Molecular-genetic analysis of natural variation in photoperiodic flowering of *Arabidopsis thaliana*

Inaugural-Dissertation zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Universität zu Köln

vorgelegt von
Antonis Giakountis

Köln, Mai 2008

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung Köln, in der Abteilung für Entwicklungsbiologie der Pflanzen (Direkotr Prof. Dr. George Coupland) angefertigt.

Berichterstatter: Prof. Dr. George Coupland
Prof. Dr. Ute Höcker

Prüfungsvorsitz: Prof. Dr. Ulf-Ingo Flügge

Tag der mündlichen Prüfung: 30 June 2008

Acknowledgements

I would like to thank my supervisor Professor Dr. George Coupland for giving me the chance to perform my PhD studies in an excellent working environment. His help, guidance and support were valuable throughout the PhD experience. Thank you George for entrusting this wondefull project to my hands.

I would also like to thank Professor Dr. Ute Höcker for being the second examiner and Professor Ulf-Ingo Flügge for accepting being the chair of the disputation.

Many thanks to Professor Dr. Maarten Koornneef who as my second advisor, helped this project with fruitfull discussions.

I could not forget to express my gratidute to Prof. Dr. Karen Hicks and Prof. Dr. Annie Schmitt for the wonderfull discussions, exchange of ideas and collaboration that we have established. It was a great aid working with you.

In addition I would like to express my appreciation the group of George Coupland at the Max-Planck Institute for Plant Breeding research and especially Stefano Torti for the work with the *in situs* and Dr. Frederic Cremer and Dr. Amaury de Montaigu for the work with the GIGANTEA project. Many thanks also to Lars Lanckaert for the wonderfull help and Sheina Sims for the collaboration in the SD mapping project. It was a pleasure meeting and working with all of you.

Finally I would like to thank my parents George and Anastasia Giakountis for their patience and support and my partner Chrysoula Kokkini for her love, motivation and support all these years.

Table of Contents

1.3 FLOWERING BEYOND CONSTANS

1.3.1 PATTERNS OF INFORMATION TRAFFICKING IN PLANTS

1.3.2 LINKING LEAVES AND SAM : THE ELUSIVE NATURE AND THE ROLE OF FLORIGEN

1.3.3 BEYOND LEAVES : THE ROLE OF THE SHOOT APICAL MERISTEM IN THE CONTROL OF FLOWERING

1.4 NATURAL VARIATION AS A TOOL IN GENETICS

1.4.1 DESCRIBING NATURAL VARIATION IN ARABIDOPSIS

1.4.2 NATURAL VARIATION AND MAP-BASED CLONING IN THE POST-GENOME ERA

2. MATERIALS AND METHODS

2.1 MATERIALS

2.2 METHODS

3. RESULTS AND DISCUSSION

3.1 NATURAL VARIATION IN FLOWERING TIME RESPONSES

3.1.1 PHYSIOLOGICAL DESCRIPTION

3.1.2 GENETIC ANALYSIS

3.2 MOLECULAR ANALYSIS OF NATURAL VARIATION IN DAY LENGTH PERCEPTION FOR FLOWERING TIME

List of Figures

List of Tables

ABSTRACT

In *Arabidopsis thaliana*, the focus of my research, three developmental switches controlling the life cycle can be recognised. The first is germination that separates embryonic from post-embryonic development. The second signals the transition from the juvenile to the adult vegetative phase while the third, flowering, marks the initiation of the reproductive phase (Isabel Baurle and Caroline Dean, Cell 2006). All three exhibit both external (environmental) and endogenous (hormones) regulation. Natural genetic variation, namely phenotypic diversity due to genetic differences between individuals of the same species, has been reported both for germination and flowering initiation (Bentsink et al., PNAS 2006; O Neill et al., TAG 2008). Since individuals of *Arabidopsis,* commonly referred to as accessions, are collected from a variety of locations, it is believed that this genetic diversity reflects differences in the seasonal oscillations of environmental cues among the collection sites leading to local adaptation.

Although natural genetic variation as a tool has been used in the study of flowering initiation in *Arabidopsis* (Alonso-Blanco and Maarten Koornneef, Trends in Plant Science 2000) a systematic survey that focuses mainly on the photoperiodic aspect of this regulation has been lacking. In order to expand the current knowledge two approaches were designed. First a survey for natural genetic variation in the flowering responses of phylogenetically distant *Arabidopsis* accessions under six different photoperiods was made. In parallel the transgenic equivalents of the same accessions, carrying a promoter fusion of the flowering time and circadian clock gene *GIGANTEA* (GI) were screened in the same photoperiods as for flowering time in order to detect for the first time trans-specific natural variation in the circadian regulation of an evening gene.

Here I present evidence that natural genetic variation is present in a wide range of photoperiods both for the circadian clock and for flowering initiation *per se*. The flowering time responses are compared with the ones of mutants and transgenic lines of previously identified flowering time genes and I show that the affected known genes cannot fully cover the different patterns of day length discrimination that the natural accessions exhibit. Five different mapping populations were constructed by selecting interesting accessions from both screens, which led to the identification of new as well as known QTL, which alter various circadian and flowering responses between short and long days of similar duration. Generating advanced genetic material allows fine mapping and eventually cloning of some of the loci, while identification of genome-wide patterns of genetic interactions reveals additional loci that classical QTL mapping approaches cannot detect. Using RT-PCR and *in*

situ hybridisation, I link this novel natural genetic variation between similar long day lengths with molecular variability in the temporal and spatial expression of flowering time genes *FT* and *SOC1* thereby also demonstrating the tight dependence of the SAM floral commitment on the FT florigen. Finally I show that in nature, genetic variability in the property of enhanced photoperiod discrimination under similar long days, is enough to prevent winter flowering in a plant without any requirements for vernalization.

Cologne, 2008

ZUSAMMENFASSUNG

Der Lebenszyklus von *Arabidopsis thaliana* kann in verschiedene Entwicklungsphasen eingeteilt werden, deren Übergänge durch externe (Umwelteinflüsse) und interne (Hormone) Signale reguliert werden. Die Keimung beendet die Embryonalentwicklung und leitet die post-embryonale Entwicklung ein. Während dieser vegetativen Wachstumsphase findet der zweite Entwicklungswechsel, der Übergang von der jugendlichen zur erwachsenen Phase, statt. Mit der Blühinduktion beginnt zuletzt die Fortpflanzungsperiode (Isabel Baurle and Caroline Dean, Cell 2006). Natürliche Varianz zwischen Individuen der gleichen Pflanzenspezies wurden für den Keimungs- und Blühprozess beschrieben (Bentsink et al., PNAS 2006; O Neill et al., TAG 2008). Verschiedenen *Arabidopsis* Individuen, auch Ökotypen genannt, wurden aus verschieden Gebieten der Welt gesammelt. Ihre genetische Diversität reflektiert die Anpassung an jahreszeitliche Änderungen an ihrem Herkunftsort.

Während genetische Varianz bereits genutzt wurde, um die Induktion der Blütenbildung in *Arabidopsis* zu untersuchen (Alonso-Blanco and Maarten Koornneef, Trends in Plant Science 2000), wurde der Aspekt der photoperiodischen Regulation bisher nicht analysiert. Um diesen Aspekt zu untersuchen, wurden phylogenetisch unterschiedliche *Arabidopsis* Ökotypen in sechs verschiedenen Tageslängen auf den Zeitpunkt ihrer Blühinduktion untersucht. Parallel zu diesem Ansatz wurde die Regulation des Gens *GIGANTEA (GI)*, das bei der Regulation der zirkadianen Uhr und der Blühinduktion eine Rolle spielt, in verschiedenen Ökotypen untersucht. Hierfür wurden die Pflanzen mit einem Promoter-Reportergen-Konstrukt transformiert und unter verschiedenen Photoperioden getestet.

In meiner Arbeit zeige ich, dass es genetische Varianz bei der Regulation der zirkadianen Uhr und der Induktion der Blütenbildung gibt. Diese natürlichen Unterschiede wurden mit der von bereits charakterisierten mutanten und transgenen Pflanzen, die einen Defekt in Blühinduktionsignalweg tragen, verglichen. Ich zeige, dass die die genetische Varianz der bereits identifizierten Blühzeitpunktgene, die der natürlichen Ökotypen nicht abdecken konnten. Zur Identifizierung von OTLs wurden fünf verschiedene Kartierungspopulationen aus unterschiedlichen Ökotypen erstellt. Hierbei konnten neue, wie auch bereits bekannte QTLs identifiziert werden, die verantwortlich für verschiedene Reaktionen auf sehr ähnliche Tageslängen sind. Eine weitere Analyse des genetischen Materials wird Feinkartierung und eventuell Klonierung der verantwortlichen Gene erlauben, sowie die Identifizierung von genomischen Interaktionen verschiedener Loci. Mit Hilfe von RT-PCR und *in situ* Hybridisierung konnte die genetische Varianz mit Änderungen der zeitlichen und räumlichen

Expression der Blühzeitpunktgene *FT* und *SOC1* in Zusammenhang gebracht werden. Dadurch wird die enge Bindung von Blühinduktion und dem Florigen FT deutlich. Zudem zeige ich, dass die Fähigkeit, sehr ähnliche Tageslängen zu unterscheiden, ausreicht, um Blühinduktion im Winter bei Ökotypen, die keine Vernalisierung erfordern, zu verhindern.

Köln, 2008

INTRODUCTION

1.1 INTRODUCING FLOWERING TIME

1.1.1 Diversity in life strategies and flowering time responses

A broad diversity in the regulation of flowering time gave rise to many different flowering strategies in the plant kingdom. Plants that complete their life cycle within one year are called annuals, those which grow vegetatively the first year and complete their life cycle by flowering and senescence the second are called biennials, while the ones that will flower repeatedly in the following years are called perennials (Ausin et al., 2005). Regardless of the different life cycle models, diversity also exists by different responses to environmental stimuli. Even though many environmental factors contribute in a moderate fashion to the environmental control of flowering time (Bernier and Perilleux, 2005) light and temperature are the most important ones due to their repeatable fluctuation in both diurnal and annual basis.

Several properties of light such as duration (photoperiod or day length), intensity and quality (wavelength) regulate the onset of floral transition, all of which show not only temporal fluctuation (both diurnal and annual) but also spatial variation (latitudinal clines). Regarding the different strategies that plants exhibit so as to utilize light duration as a regulatory stimulus of flowering time, three classes of responses have been reported. The first class consists of plants such as the model species *Arabidopsis thaliana* that flower earlier under long days (LD plants). The second class includes species in which flowering time is accelerated under short days (SD plants) such as rice while the last one refers to plant organisms in which flowering time is not affected by photoperiod (day-neutral plants, tomato).

The exposure to changing photoperiods can also control flowering time either in a qualitative or a quantitative manner. For example *Arabidopsis*, which is a quantitative species, flowers both in short and long days, but flowering is greatly accelerated under long days. On the contrary species that follow the qualitative habit will flower only when photoperiod exceeds a critical value (so-called critical photoperiod) of light or darkness (for LD or SD plants respectively).

1.1.2 General aspects of flowering initiation

Arabidopsis thaliana, the model annual species, begins its development with a program set into vegetative growth. During this phase, the apical meristem is committed with a vegetative identity, thus producing leaf primordia, which form a rosette. Leaves constantly perceive

environmental cues such as light (An and Roussot, 2004; Searle 2004) and secondary temperature (Searle et al., 2006). At the proper time of the year, usually spring, inductive long days for *Arabidopsis*, will activate a signalling cascade in the leaves (Koorneef et al. 1991., Putterill et al., 1995). A graft transmittable activating signal, the so-called florigen, will then migrate through the phloem towards the apical meristem (Zeevaart et al., 1976) and will eventually commit the shoot apical meristem (SAM) to floral identity.

Temperature is also monitored throughout the year, mainly by the apical meristem (Lang et al, 1965). There, certain repressors maintain vegetative growth during the low temperatures of winter (Johanson et al., 2000, Sung and Amasino, 2003). Once the elevated temperatures of spring after the exposure to the prolonged cold of winter, are experienced this repression is released. Then, and in combination with the activator(s) that arrives from the leaves [florigen(s)]), the transition to flowering begins. Once a set of factors called floral homeotic genes, are activated the apical meristem will be committed to floral identity and the production of leaf primordia will terminate abd be succeded by floral primordial. The plant will eventually bolt and give rise to inflorescences. Certain endogenous stimuli such as hormones and sucrose also have a role in flowering initiation, especially under non-inductive short days (Eriksson et al., 2006).

Flowering therefore can be considered as the default developmental program of plants. Consistent with this aspect, the role of vegetative growth is to delay the onset of flowering, ensuring that once it occurs, the organism will have sufficient biomass to support reproductive development. In agreement with this is initiation of leaf senescence that occurs synchronously with flowering. Senescence provides the developing siliques with nutrients through the phloem. From this point of view, environmental cues such as photoperiod, light quality and intensity as well as temperature are responsible for regulating a wide range of floral repressors and activators. A sensitive equilibrium between them, fine-tunes development by expanding or reducing the duration of the vegetative growth.

1.1.3 Regulatory pathways of flowering time

A successful transition to the reproductive phase ensures that the plant will develop and consequently set seeds in favourable environmental conditions. A tight regulation of this transition will therefore provide the possibility of simultaneous flowering within a population, a necessity especially for cross-pollinated species and agricultural crops. To ensure both, plant evolution developed complex regulatory networks governing flowering time· networks that involve both external stimuli and endogenous factors such as hormones.

Genetic studies in *Arabidopsis thaliana* revealed many genes that are involved in flowering regulation (reviewed in Koornneef et al., 1991; Coupland 1995; Simpson et al., 1999). Four pathways have been proposed to regulate flowering time, the photoperiod pathway, the vernalization pathway, the autonomous pathway, and the gibberellin pathway (Fig 1 Ausin et al., 2005). The autonomous pathway includes genes, mutations of which cause late flowering regardless of photoperiod (Simpson et al., 1999). These genes regulate the mRNA abundance of the floral repressor FLC (see below) and vernalization treatment can overcome the effect of their mutations. Recently two of these genes *FY* and *FCA* have been shown to regulate *FLC* mRNA levels via endogenous post-transcriptional gene silencing or PTGS (Swiezewski et al., 2007). The gibberellin pathway reflects the endogenous control in flowering initiation that is mediated by plant hormones (Moon et al., 2003). The role of this pathway is mainly to promote flowering under non-inductive short days or when the proper function of the photoperiod pathway is attenuated. Since this work is mainly focused on photoperiod and secondarily on vernalization I will refer in more details only to these two pathways.

1.2 THE MOLECULAR ASPECTS OF FLOWERING REGULATION

1.2.1 THE CIRCADIAN CLOCK

1.2.1.1 Historical background

Leaf movement, a physiological process of plants that is diurnally regulated, was the first historical observation regarding the existence of a timing mechanism (Androsthenis, 4th century BC). Carl Linnaeus, the Father of Taxonomy, in 1.500 AC confirmed the same pattern of leaf movement in many plant species.

However the first true experiment regarding the circadian clock was carried out two centuries later, in 1729 by the French astronomer Jean Jacques d'Ortous deMairan. He proved the sustainable character of leaf movements by placing *Helianthus* plants in continuous darkness and proving that this phenomenon is not simply a response to light. Darwin also tried to connect the leaf movement patterns of plant species with their evolution and he concluded that plants close their leaves at dusk in order to minimise the surface that is exposed to low night temperatures while at dawn the leaves open in order to maximise the exposed to sunlight area. He published his observations in his book "The power of movement of plants".

Apart from plants, experiments concerning the circadian clock were made also in insects. Just in the beginning of the 20th century, Karl von Frisch and Ingeborg Beling conducted experiments on bees after observing that the insects visit the flowers only at specific times of the day. Even after the removal of nectar, the bees continued to visit the flowers that specific time of the day, following the periodic pattern of nectar production. The first however who proposed the existence of the circadian clock was Gustav Krammer and Klaus Hoffmann in the middle of the 20th century (Introduction to Plant Physiology, William G. Hopkins, Second edition). They proposed the endogenous character of the clock as well as the ability of entrainment by the environment. Around the same time another scientist Colin Pittendrigh studied the effect of temperature on the circadian rhythms of *Drosophila pseudoobscura*. He concluded that the period of the clock remains largely unaffected over a range of different temperatures, establishing the term Temperature Compensation. In recent years however the progress in molecular biology provided the first molecular insights into the circadian clock not only of plants but also other organisms.

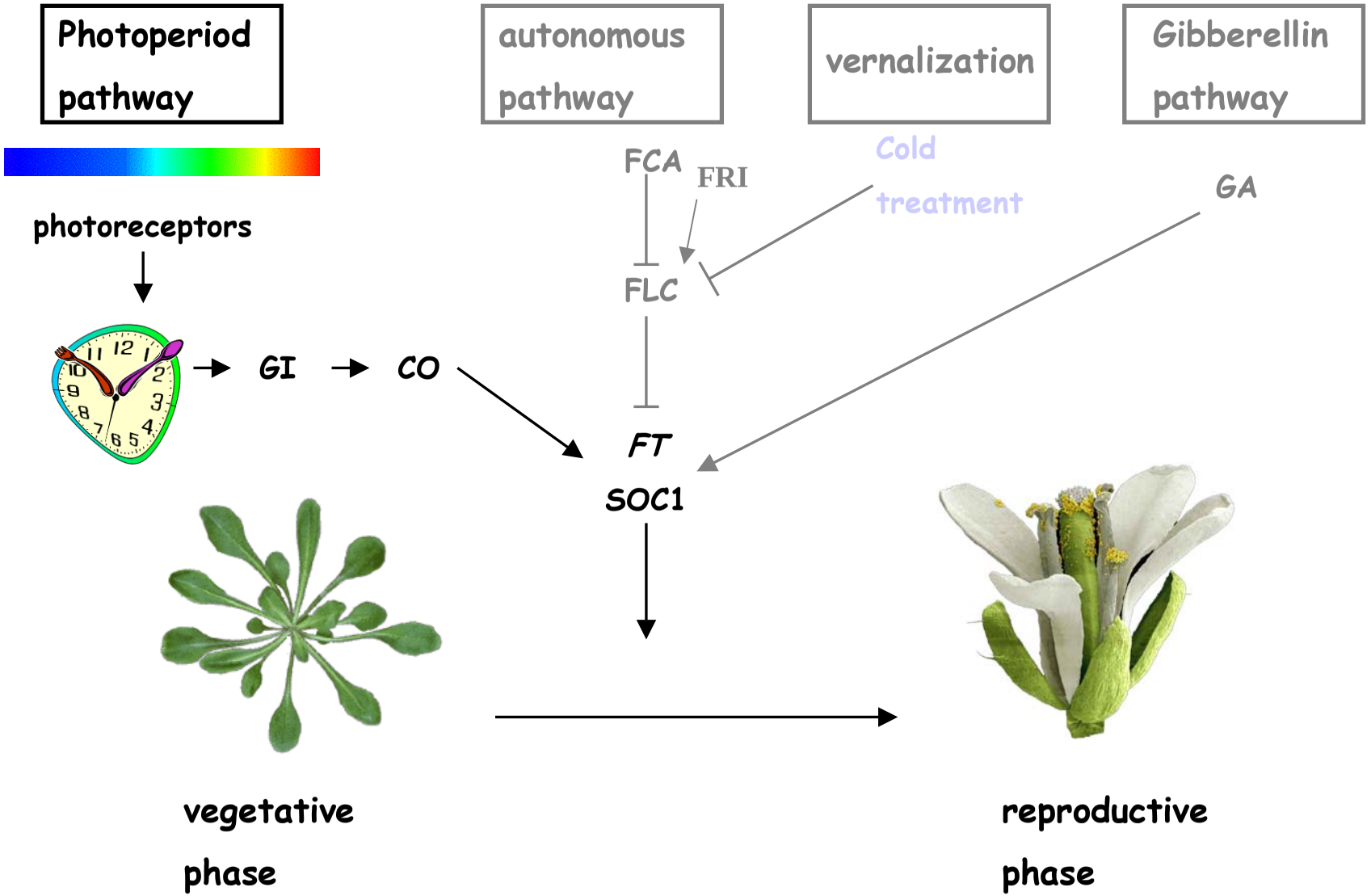


Figure 1 : Genetic pathways regulating flowering time. The photoperiod pathway controls floral induction in response to photoperiod. The vernalization pathway regulates flowering in response to prolonged exposure to low temperatures. The autonomous pathway regulates the mRNA levels of the floral repressor FLC, via association with chromatin remodelling and endogenous RNAi cascades. The gibberellin pathway adds a hormonal aspect in the regulation of flowering time. All these pathways eventually converge on the so called floral integrators FT and SOC1, activation of which will initiate the reproductive fate.

1.2.1.2 Modules and components of the plant circadian clock

Plants measure day length with the aid of the circadian clock, an endogenous timing mechanism that is found from cyanobacteria to mammals. In most organisms the circadian clock consists of three parts: the input pathways, the central oscillator and the output pathways (Fig 2). Each part has a different role and several molecular components. However some of these components are shared by more than one part. The role of the input pathways is to connect the circadian clock with the environment. Light input is mediated by photoreceptors, namely the phytochromes and the cryptochromes, that detect different aspects of light such as intensity and quality (wavelength) and signal to the components of the central oscillator. The central oscillator is the heart of the system. It consists of a negative feedback loop that is synchronised to the environment with the aid of the input pathways (simply the photoreceptors set the clock to the local time). The components of the central oscillator control the expression of many genes that consist of the output pathways, the information that derives from the clock. The parts or modules of the circadian clock differ not only in their components but also in their intra-cellular positions. Phytochromes exist in the cytoplasm from which under specific conditions (see below) they traffic into the nucleus in order to signal to the components of the central oscillator. The output pathways however can eventually express their effect not only in the cytoplasm of the same cell, but also in different cells as in the case of CONSTANS (CO).

1.2.1.3 Input pathways of the plant circadian clock - The phytochromes

Although the circadian rhythms are endogenous and genetically regulated, they must be reset to the local time. The basic characteristic of a functional circadian clock is the ability to monitor the 24h periodic light/dark cycle of the planet. In order to do so, the clock is synchronised (entrained) to the environment via the components of the input pathways. Today two major classes of photoreceptors are known in plants: the family of phytochromes and the family of cryptochromes. In *Arabidopsis* five members of the phytochromes and two members of cryptochromes have been reported.

Phytochromes are the most extensively studied photoreceptors of plants. Their discovery traces back to 1950 which are reported as the most important factors controlling flowering in short day (SD) plants. Borthwick proved that red light irradiation activates flowering in cabbage, while infra-red represses it (Borthwick et al., 1952). Based on protein stability, the five phytochromes can be classified in two categories. Category I consists of PHYA which is light labile and its protein degrades fast under light while category II consists of the other

phytochromes (PHYB-E) the proteins of which are stable under light (Quail 1997; Kozma-Bognar et al., 1997; Kay and Nagatani 1999).

All phytochromes are receptors in the visible spectrum of light and consist of two forms. The Pr (Phytochrome red form) is the biologically inactive form in which the molecule is produced. The Pfr (Phytochrome Far-Red) is the biologically active form (Quail 1997). Upon exposure to the red wavelengths of light (667 nm) the Pr form is converted to the Pfr biologically active form, accompanied with structural differences in the molecule as well as differences in the absorption spectrum. Only in this form does the phytochrome enter the nucleus (Quail, 1997). Upon infra-red irradiation (730 nm) the Pfr from is converted back to the Pr form and the molecule shuffles back to the cytoplasm (Quail, 1997). Among the five members, differences exist also in the intensity of red light that can be absorbed. Based on the intensity of light, three categories can be made, namely the Low Fluence Responses (LFRs), the Very Low Fluence Responses (VLFRs) and the High Irradiation Responses (HIRs). Phytochrome B mediates LFRs and HIRs, phytochromes D and E mediate the LFRs and PHYA the VLFR.

Studies of the structure of the apoprotein, revealed that phytochromes are homodimers consisting of two subgroups of 125 kDa each (Quail, 1991). The N-terminal domain binds covalently the chromophore (factor responsible for light absorbance) while the C-terninal domain is responsible for dimerization. This domain also contains two PAS motifs (**P**eriod circadian protein, **A**h receptor nuclear translocator protein and **S**ingle-minded protein), which are responsible for the interaction with transcription factors such as the bHLH PIF3 (**b**asic **H**elix **L**oop **H**elix **P**hytochrome **I**nteracting Factor **3**). Phytochromes are considered to be light dependent kinases due to the existence of the HKRD domain (**H**istidine-**K**inase-**R**elated-**D**omain). A direct outcome of this is the ability for autophosphorylation as well as for phosphorylation of nuclear target proteins (Fankhauser et al., 1999; Quail, 1997).

1.2.1.4 Input pathways of the plant circadian clock - The cryptochromes

Cryptochromes mediate light responses in the blue spectrum of light. After the isolation of cryptocrome 1 (CRY1) in *Arabidopsis* (Ahmand and Cashmore, 1993) orthologs were identified in almost every multicellular organism that was examined (Lin and Shalitin, 2003).

Cryptochromes are flavoproteins that have flavin and pterin as chromophores. They show similarity to DNA photolyases however lack the ability of DNA repair. The light absorbance spectrum consists of two peaks, one at the UV-A range (320-400nm) and one at the violet-

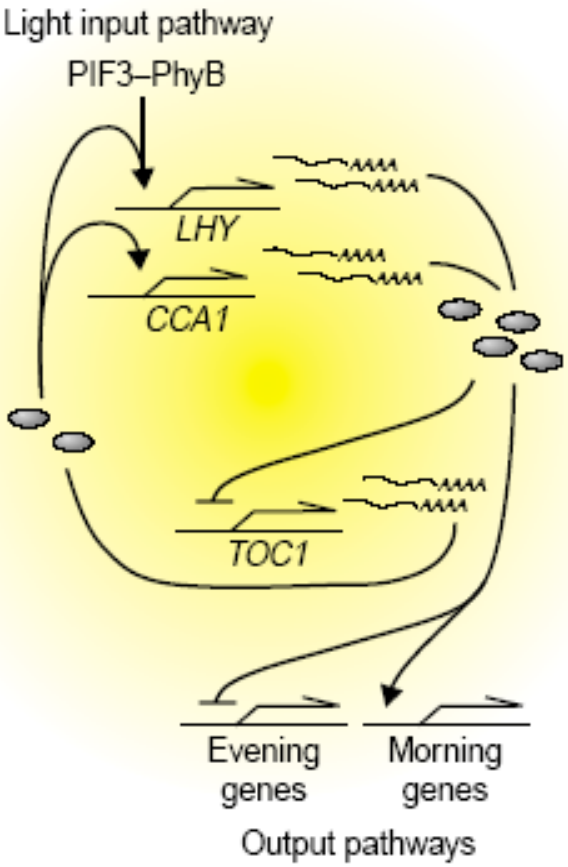


Figure 2 : Classical model of the circadian clock of *Arabidopsis*. The three components, namely the input pathways, the central oscillator and the output pathways are shown. (Hayama και Coupland, 2003).

blue range (400-500 nm). The *Arabidopsis* genome encodes two cryptochromes CRY1 and CRY2 (Ahmad 1993; Hoffman, 1996) while in tomato three have been reported (CRY1A, CRY1B and CRY2, Pergotta, 2000). The molecular weight of the apoprotein is 70-80 kDa with the N-terminal domain hosting the **P**hotolyase-**r**elated **d**omain (PHR), which as the name indicates shares homology with the DNA photolyases (Lin and Shalitin, 2003). In the N terminal domain the two molecules of flavin and pterin also bind covalently (Yanovsky and Kay, 2003). The C terminal part does not share homology with the photolyases and could mediate interactions with other proteins (Lin and Shalitin, 2003).

Both CRY1 and 2 are located in the nucleus and the cytoplasm (Cashmore et al., 1999; Guo et al., 1999; Kleiner et al., 1999) with CRY2 shuffling into the nucleus independently of light conditions. It seems that the C-terminal part is sufficient to drive the molecule into the nucleus (Cutler et al., 2000; Wang et al., 2001; Yang, 2000). Regarding the protein stability CRY2, like PHYA is light labile. Exposing etiolated seedlings for 15 min to violet light is enough to reduce the protein abundance by 90 % (Shalitin et al., 2002) and both the PHR domain, as well as the C-terminus seem to be important for the light-dependant degradation of the protein.

1.2.1.5 The combined signalling of the central oscillator

Both the red absorbing phytochromes and the blue absorbing cryptochromes entrain the plant central oscillator. Mutating members from both classes, alters the circadian period and/or phase of many circadian marker genes such as CAB. Interactions between and within the PHY and CRY families have been shown with Y2H and co-immunoprecipitation assays. A direct physical interaction between PHYB and CRY2 and were confirmed with FRET (Mas, 2000). In addition both proteins co-localise in the nucleus but such co-localisation was not observed for CRY1. This interaction leads to a C-terminal phosphorylation of the former by the latter (Mas, 2000). Phosphorylation was also observed in PHYB mutants, suggesting that PHYB is not the sole kinase that targets CRY2. Another candidate can be Casein Kinase 2 (CK2, Lin and Shalitin, 2003). According to *in vitro* results, both CRY1 and PHYB also interact with members of the ZTL/LKP1/FKF1 family and their mutations alter the mRNA expression levels of FKF1 (unpublished personal data).

Combining the above results into a model regarding the synergistic action of phytochromes and cryptochromes in signalling to the central oscillator in a range of light intensities and qualities. LFRs in the red wavelengths are controlled by PHYA and in the violet-blue by CRY1 with the latter acting downstream of the former. Increased light intensities result in the

degradation of the light labile PHYA and PHYB-E taking over the red light responses while CRY1 and 2 synergistically mediate blue light responses. HIFs in the violet-blue range are mediated solely by CRY1 since CRY2 protein is degraded. This plasticity in the light responses regarding the light intensities and qualitites provides higher flexibility to the plant in perceiving and responding to a variety of light signals and modulate it circadian rhythms accordingly.

1.2.1.6 The central oscillator of the plant circadian clock

From plants to humans, central circadian oscillators are nothing else but auto regulatory transcriptional feedback loops. In Arabidopsis the classical model of the circadian clock places two MYB transcription factors as negative elements of the system and a single protein as the positive element. The two MYB transcription factors are the LHY (Late Elongated Hypocotyl 1, Schaffer et al., 1998) and the CCA1 (Circadian Clock Associated 1, Wang and Tobin, 1998) proteins while the PRR1/TOC1 (Pseudo-Response Regulator1 / Timing of CAB expression 1, Strayer et al., 2000) is the positive element of the loop. Northern analysis showed that both *LHY* and *CCA1* transcripts peak at dawn in contrast to the *TOC1* transcript that peaks at dusk. *LHY* and *CCA1* belong to the category of the so-called morning genes while *TOC1* to the evening genes. Another difference between the components of the central oscillation is their acute response to light. Both *CCA1* and *LHY* transcripts increase upon exposure to light at dawn while *TOC1* mRNA does not show any response to light.

1.2.1.7 The central oscillator of the plant circadian clock–The TOC1 family and the PRR quintet.

TOC1 belong to the COL (**CO**NSTANS-**L**ike) family with the N-terminal part containing a motif that shares similarity to the receiver domain of the Two-component signal transduction response regulators.The typical function of this signalling system includes as a first factor one kinase which is upon activation via autophosphorylation of a conserved histidine, transmits the signal by phosphorylating an asparagine in the receiver domain of response regulator. The final outcome of the system is the activation of the transcription of target genes (Stock et al., 1991). However in the TOC1 family, two out of three conserved, functional aminoacids in the receiver domain are replaced. However the characteristic domain of this family lies in the C-terminus and consists of the highly conserved CCT domain among the members. The CCT domain of TOC1 consists of 45 aminoacids and contains one NLS signal. Immediately after that the AD domain, characteristic of transcriptional activators, follows (Cress et al., 1991).

The NLS calls for nuclear localization and indeed the TOC1-YFP nuclear speckles have been observed (Strayer et al., 2000).

Apart from TOC1, four additional pseudo-response regulators have been characterised in *Arabidopsis* (Matsuskika et al., 2002). PRR9, PPR7, PRR5, PRR3 and PRR1/TOC1 therefore form what is known as the PRR quintet of *Arabidopsis*. The Ppd-H1 gene, a PRR7 homolog has been reported also in barley and there it acts as a major target of natural variation for flowering time (Turner et al., 2005). Although *TOC1* is an evening gene, the remaining members exhibit a variety of transcriptional profiles, some of which are circadian regulated via *CCA1* under continuous light (Matsushika et al., 2002). In wild type *Arabidopsis* plants, *PPR9* behaves as a morning gene, which peaks shortly after dawn and with an acute response to light, similar to the *LHY/CCA1*. Two hours after the peak of *PPR9* mRNA, the *PPR7* transcript peaks and shortly afterwards PRR3 follows. This transcriptional quintet is complete with the peak of the *PRR1/TOC1* at dusk, thus covering the entire light part of the day (Matsushika et al., 2002).

1.2.1.8 The central oscillator of the plant circadian clock–The LHY/CCA1 components

The MYB transcription factor LHY (**L**ate **E**longated **H**ypocotyl) was isolated in 1998 in a screen for genes controlling photoperiodic flowering in *Arabidopsis* (Schaffer et al., 1998). LHY overexpressors are late flowering under long days, a sign that the activity of the photoperiod pathway is compromised due to the mis-regulation of the plant circadian clock. Indeed in these plants, the circadian regulation of many genes such as CAB2 is disrupted (Schafer et al. 1998). LHY was reported as a component of the central oscillator and not the input pathway since the circadian effects were independent of the external conditions. At the same time CCA1 was shown to regulate the expression of the genes of the light-harvesting complex of photosystem II (**C**hlorophyll **a**/b **B**inding Proteins, Wang et al., 1997).

After the morning peak, the expression of the two genes gradually decreases during the day releasing the repression of *TOC1* mRNA expression (Alabadi et al., 2001). The function of the two genes is redundant and overexpression reduces the expression of the endogenous genes. Since LHY-ox and CCA1-ox plants show reduced levels of *TOC1* mRNA, the reduced expression of the two endogenous MYB genes is due to the repression of their positive regulator TOC1 (Wang and Tobin, 1998; Fowler et al., 1999; Matsushika et al., 2002). A direct involvement of LHY/CCA1 in the transcriptional repression of TOC1 would require evidence for promoter binding. Promoter deletions and EMSA provided this evidence and

indicated that the region –734 until –687 contains the cis regulatory elements that LHY/CAA1 bind to (Alabadi et al., 2001). The domain contains the AAAATATCT motif known as the "evening element (EE)" since it is present in the promoters of 31 evening genes in *Arabidopsis*. As mentioned this regulation needs to be reciprocal and indeed in strong *toc1-2* mutants the expression of *LHY*/*CCA1* mRNA is almost eliminated in free running conditions (Alabadi et al., 2001). Interestingly apart from the evening element, these two MYB proteins bind another very closely element AAAAATCT. This element is found in the promoter region of morning instead of evening genes and was therefore called the morning element. Note that the two motifs have only one nucleotide difference (the inserted thymine marked with red).

Interestingly the accumulated LHY/CCA1 proteins promote the expression of genes, with morning element such as *CAB* early in the day while at the same time repress the expression of the evening genes such as *GIGANTEA* and *TOC1* via binding of the evening element (Wang et al., 1997). Other genes that are involved in the proper function of the circadian clock and also affect the photoperiodic control of flowering time are *ELF4* that acts as a positive element within the loop and *ELF3* (Hicks et al., 2001) that controls light input to the central oscillator in a phase dependent manner. To date many genes that affect the circadian clock in Arabidopsis have been mapped in all five chromosomes (Fig 3).

1.2.1.9 Output pathways of the circadian clock

The third module of the clock is the output pathway that simply refers to the sum of genes that are regulated by the clock. According to microarray data, 10-12 % of the *Arabidopsis* transcriptome is under circadian regulation (Harmer et al., 2000). These genes include four LHCA genes (**L**ight **H**arvesting **C**omplex **A**) of PS I and seven LHCB genes of PS II. These genes together with the PXC IX (Protoporphyrin IX Magnesium Chelatase), an enzyme responsible for the production of chlorophyll, and 22 genes that are tightly connected with photosynthesis, indicates the importance of the circadian clock in adjusting the physiology of the plant to the environment. All these genes are expressed in the morning in order to facilitate photosynthesis. In addition, members of the input pathways such as the PHYB and CRY1 and 2 are themselves regulated by the clock. Secondary metabolites such as phenyl-propanoids are also regulated by the clock and are responsible for protecting the leaves from intense light (Harmer et al., 2000).

In contrast to the morning genes, other output genes are expressed at dusk. Apart from *TOC1* and *GI*, both of which contain EEs in their promoter regions, they include transcription factors responsible for cold acclimation. DREB1a/CBF3 is a transcription factor, which

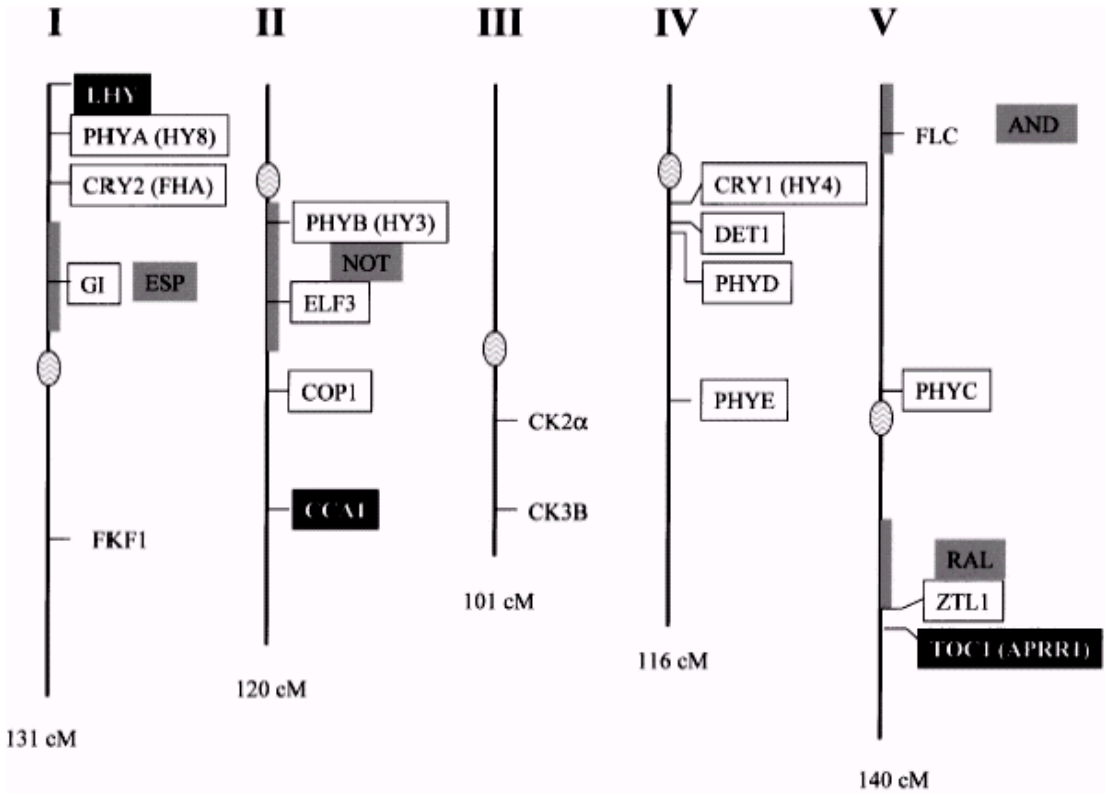


Figure 3 : Physical positions of the most important elements of the circadian clock in *Arabidopsis thaliana* (Roden και Carré, 2001).

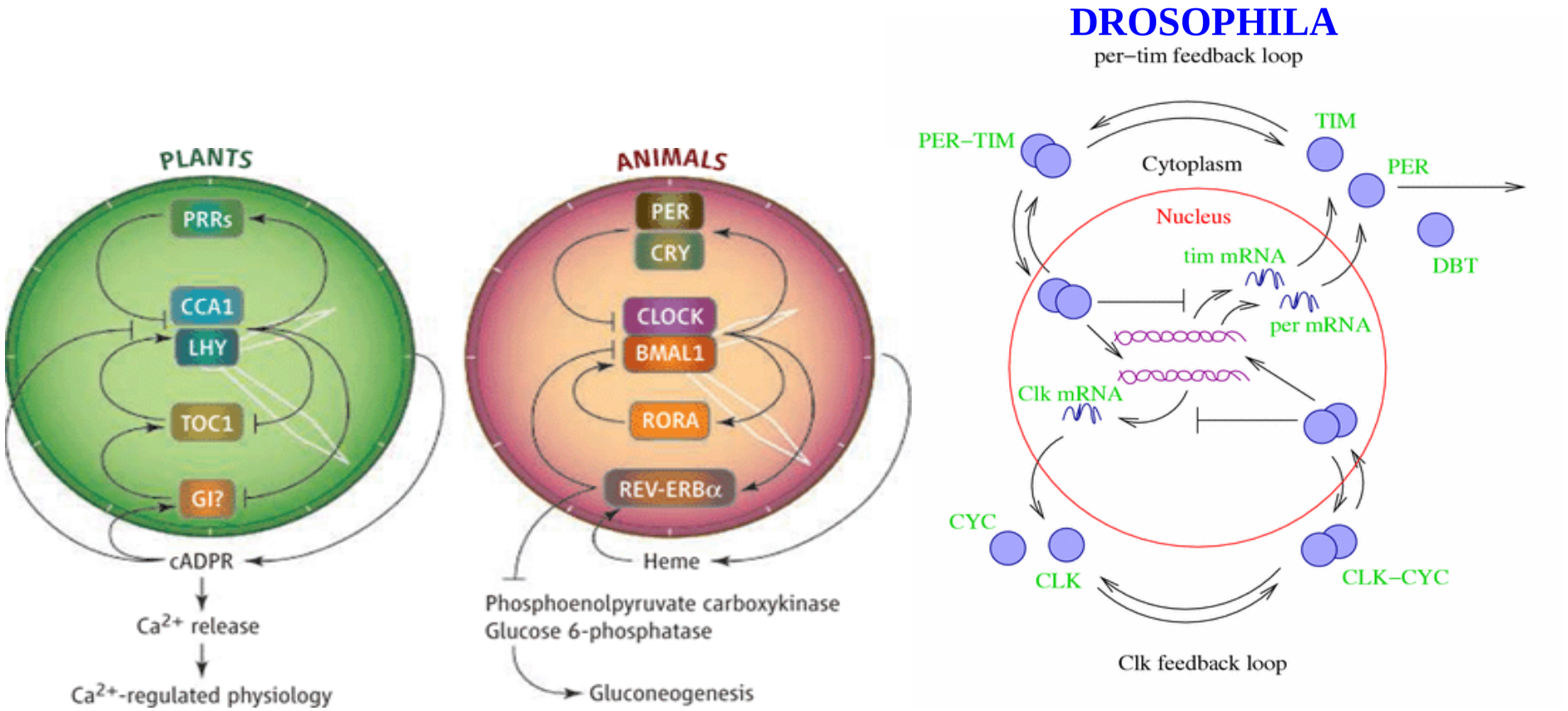


Figure 4 : Form of the circadian clock in various organisms. In plants the interlocked model predicts the existance of three loops, a morning one consisting of members of the PRR familly, the central oscillator consisting of LHY/CCA1 and TOC1 and the evening loop consisting of GI and ZTL. In Animals the morning loop consists of PER and CRY, the central oscillator consists of CLOCK and BMAL1 while the evening loop of RORA and REV-ERBa. In the *Drosophila* clock two loops, the PER-TIM and the CLK-CYC, can be identified.

regulates the expression of genes that induce tolerance to the low temperatures of night (Shinozaki et al., 2000). The energy production cascades are also regulated by the clock with six genes involved in glycolysis as well as transporters of sugar in the roots and starch accumulation being among them. Hormone signalling is also circadian regulated since two auxin transporters PIN3 and PIN7 are clock outputs. The circadian clock also controls the concentrations of Ca^{++} and K^{+} in the guard cells, regulating thus the opening of the stomata. Other physiological processes that are circadian regulated are leaf movement as mentioned, again by influencing the amount of water in the cells at the base of the stem and of course flowering initiation details of which will be presented below.

1.2.1.10 The interlocked model of the plant circadian clock.

According to the classical model, dawn resets the clock through *LHY* and *CCA1*, two morning genes with an acute response to light that is more pronounced in the morning. Apart from *TOC1*, mutations in *GI* also affect the expression of *LHY*/*CCA1*. Computer simulations predicted a more complex regulatory network than the originally proposed, containing three auto-regulatory loops forming the central oscillator. In this model the one-loop central oscillator of the classical model is complemented by an additional morning loop that consists of PRRs 9 and 7. The role of this loop is to promote the photoreceptor dependent light input and to facilitate the acute response of *LHY* and *CCA1* to light. The interlocked model also predicts the existence of an evening loop through which, dusk and not only dawn, resets the clock (Fig 4, Locke et al., 2005). Interestingly this model resembles the proposed models for the Drosophila circadian clock (see later).

1.2.1.11 GIGANTEA as a central oscillator component.

Approximately 70 % of the activity of the evening loop is proposed to be mediated by GI (Locke et al., 2005). *GIGANTEA* (Fowler et al 1999), is shown to be involved in the photoperiodic control of flowering time, circadian rhythms (Park and Somers, 1999) and also red light responses (Huq et al., 2000). *GI* is a single gene in *Arabidopsis*, it is plant specific, and highly conserved homologs have been identified also in monocots (rice) and trees (poplar), indicating the evolutionary significance of the protein.

Experimental results confirm that *GI* indeed is part of the evening loop and one of the roles of the gene is to reset the clock by dusk (Locke et al., 2005; Toth et al., in preparation). But how exactly does *GI* feed back to the clock? In the plant photoreceptors paragraph (§ 1.2.1.5), a sort reference to the ZEITLUPE (ZTL) gene was made. The N-terminal of ZTL has a

PAS/LOV domain, which can bind a flavin mononucleotide (FMN) like the cryptochromes. Mutations of this domain affect circadian rhythms in expression of genes such as *CAB2* in a light dependent manner probably reflecting the inability for covalent binding of the chromophore (Sommers et al., 2000). The second very interesting motif of ZTL is an F-Box, which lies approximately in the centre of the protein. This motif is common in proteins that mediate poly-ubiquitinylation and subsequent signalling of protein targets to the 26S proteasome. Last but not least six Kelch repeats, reported to mediate protein-protein interactions also exist in the C-terminal part.

Mutations in *ZTL* affect both evening and morning gene expression and the *ZTL* mRNA itself is not circadian regulated but is expressed constitutively. On the contrary the ZTL protein oscillates with maximum levels occurring shortly before dusk resembling the *TOC1* and *GI* profiles (Kim et al., 2003 and 2007). Together these data indicate that ZTL regulates the core of the clock and that another gene is responsible for the post-transcriptional regulation of ZTL. The first connection with the central oscillator was made with Y2H data, demonstrating that ZTL and TOC1 physically interact (Mas et al., 2003). TOC1-RNAi plants showed the same phenotype as *ztl1-1* mutant, indicating that the function of ZTL could be expressed via TOC1. Increasing allele copies of TOC1 via insertion of a minigene (TMG) demonstrated circadian phenotypes that were enhanced in a *ztl1-1* background. In addition, despite the fact that the mRNA and the protein were oscillating in TMG plants with functional ZTL, in the *ztl1-1* background, the TOC1 protein was stabilised, indicating a post transcriptional pattern of regulation. This type of regulation would be consistent with the role of ZTL as an E3-ligase. Indeed although the TOC1 protein could be rapidly turned over within 30 min, proteasome inhibitors repressed the degradation like in the *ztl1-1* mutation.

Interestingly the TOC protein was degraded only in the maximum of the rhythm but not in the minimum. How is this achieved since ZTL itself is not circadian regulated? *GI* mRNA expression pattern is almost identical with *TOC1*, exhibiting a peak shortly before dusk (Fowler et al., 1999; Park et al., 1999). At the protein level, light regulates the stability of the GI protein in a manner similar to that of CO (see below), causing a translational profile that is similar to the mRNA profile (David et al., 2006). GI therefore could be the additional gene that is required for the regulation of TOC1 by ZTL.

Evidence came first from the study of ZTL protein abundance in gi mutants and GI overexpressors. Mutations of *GI* strongly reduced ZTL level while GI-ox plants retain a constitutively high level of expression. At the same time no marked changes were observed at the *ZTL* mRNA profile indicating that the effect of GI is post-transcriptional. Additional

evidence about the putative GI-ZTL interaction were provided by Y2H analysis, which confirmed a physical interaction between the two proteins, with the LOV^{ZTL} mediating the interaction with GI and the $Kelch^{ZTL}$ repeats mediating the TOC1 interaction. Co-immunoprecipitation assays showed an enhanced GI-ZTL interaction under blue light compared to dark and red light, in contrast to the ZTL-TOC1 interactions that are light independent. Mutations in *GI* also lead to an increase in the protein level of TOC1 during the day and a decrease during the night. Since TOC1 is post-transcriptionally regulated by ZTL and ZTL post-transcriptionally regulated by GI these results point to the model that the evening loop of the plant circadian clock consists of both GI and ZTL which synergistically act in order to degrade TOC1 protein. During the day blue light stabilises the ZTL-GI interactions, marking the TOC1 protein for degradation via the proteasome while in the night the ZTL-GI interaction is weak allowing the TOC1 levels to rise (Kim et al., 2007). The mechanism not only validates the predictions of the interlocked model but demonstrates how dusk can also reset the clock via GI.

1.2.2 THE PHOTOPERIOD PATHWAY

1.2.2.1 Generally about the photoperiod pathway

The observation that in many plants photoperiod regulates the initiation of flowering suggests the existence of a molecular mechanism that mediates such photoperiodic responses. The photoperiod or long day pathway consists of genes whose mutations cause a late flowering phenotype under LD conditions (Koornneef et al., 1991). Its role is to promote flowering under inductive long day conditions for *Arabidopsis*. Basic components of the photoperiod pathway are the circadian clock, CONSTANS (CO) and Flowering Locus T (FT). GI connects the CO to the circadian clock and CO activates the expression of *FT* in the leaves.

1.2.2.2 CONSTANS – Transcriptional pattern of regulation

A key element of the photoperiod pathway is CONSTANS (CO), a zinc finger transcription factor (Putterill et al., 1995). *CO* belongs to a family of 17 genes in *Arabidopsis*, some of which are conserved in both monocots and dicots (Griffiths et al., 2003). *CO* transcript exhibits a diurnal pattern of expression under both short (SD, 8 h of light) and long days (LD, 16 h of light) This pattern of diurnal expression is mainly controlled by the circadian clock via GI (Fowler et al 1999 Suarez-lopez et al., 2001). However, only under inductive long day conditions do *CO* transcripts accumulate shortly before dusk, thus exhibiting a first peak which overlaps with light (Fig 5, Suarez-Lopez et al., 2001).

FKF1 was shown to regulate the long day exclusive biphasic pattern of *CO* expression (Imaizumi et al., 2003). *CO* transcripts in *fkf1* mutants fail to rise shortly before dusk under inductive long days, leading to a monophasic pattern of expression exhibiting only the circadian night peak. The overall levels of the *CO* mRNA however are identical to WT and are not reduced like as in gi mutants (Suarez Lopez et al., 2001). Nevertheless *fkf1* mutants, like gi mutations, cause late flowering with the *FT* transcript under normally inductive long days being reduced to almost undetectable levels found in non-inductive short days. In conclusion these data support the external coincidence model, according to which the long day specific and FKF1 regulated coincidence of the CO transcript with light is critical for the upregulation of *FT* transcription and the promotion of flowering.

FKF1 belongs to the same family as ZTL and the two proteins share similar motifs. So like ZTL, FKF1 is an E3 ligase with the F-box marking protein targets form degradation via the proteasome and the PAS/LOV domain being responsible for the binding of FMN. Indeed the

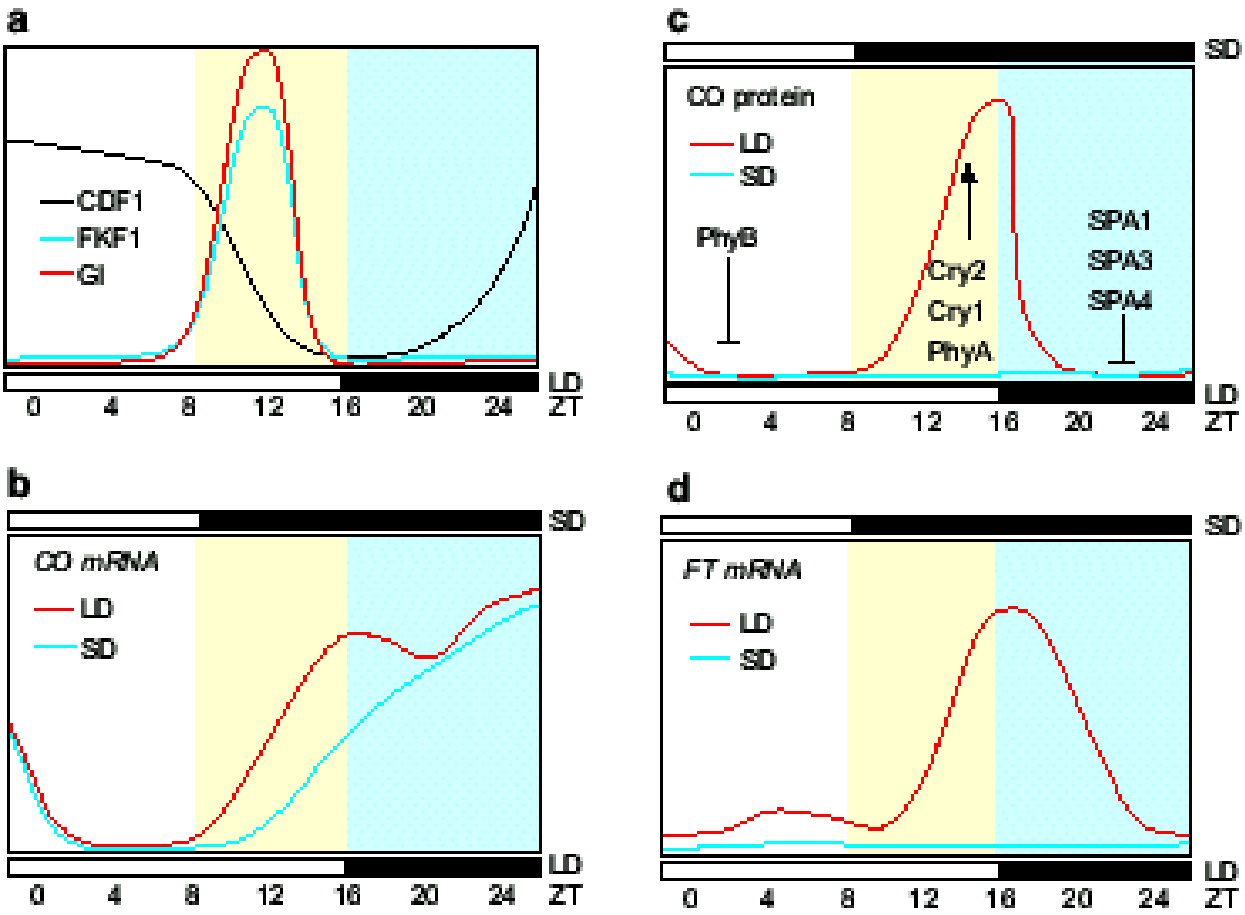


Figure 5. : The external coincidence model for flowering time initiation. a) Upon inductive long day conditions GI and FKF1 protein profiles overlap. The two proteins interact and degrade CDF1, a *CO* repressor. CDF1 is expressed in the morning and prevents *CO* mRNA from being induced during the day. b) Upon non inductive short days *CO* mRNA is monophasic exhibiting only one circadian peak during the night. Under LDs the degradation of CDF1 allows the *CO* transcript to rise before dusk and to coincide with light. c) Stabilized by the photoreceptors the CO protein rises before dusk under LDs. *PHYB* blocks the accumulation of CO protein early in the day while in the night the COP1-SPA complex degrades CO. d) The accumulated CO protein activates *FT* mRNA expression at dusk only under LDs inducing flowering. (Turck et al., 2008).

FKF1 absorption spectrum shows a peak in the violet-blue wavelengths. In contrast to *ZTL*, *FKF1* transcripts oscillate, indicating that the gene is a clock output. *FKF1* is an evening gene and its regulation by the circadian clock requires the binding of LHY/CCA1 proteins to the EE of the *FKF1* promoter, similarly to GI and TOC1. Under short days however the peak in the expression of the gene occurs approximately three hours earlier compared to the long day peak (Imaizumi et al., 2003).

FKF1 regulates *CO* expression by controlling the stability of CDF1, a DOF transcription factor. CDF1 (**C**ycling **D**OF **F**actor **1**) belongs to a family with several members in *Arabidopsis*. Missexpression of transcription factors under the control of the phloem specific promoter (REGIA project) indicated that not only CDF1 but also other members of the DOF family regulate flowering time. CDF1 mutants are slightly earlier flowering under long days compared to wild type. In addition the mRNA profile of the gene exhibits a morning peak and the expression is gradually decreasing during the day. CDF1 therefore is a repressor of photoperiodic induction of flowering by affecting the *CO* transcriptional pattern. CDF1 also binds directly to the *CO* promoter and represses its activity early in the day (Imaizumi et al., 2005). Fusions of the promoters of *CO*, *CDF1* and *FKF1* with *GUS* demonstrated that their expression is localized in the phloem, a result that is consistent with the fact that leaves are the site of photoperiod perception (An and Roussot, 2004; Imaizumi et al., 2005).

1.2.2.3 GIGANTEA as a clock output and flowering-time gene.

The fact the FKF1 is an E3 ligase indicates that FKF1 might regulate *CO* expression by degrading a repressor of CONSTANS, such as CDF1. This possibility is further supported by the fact that the phenotype of *fkf1-1* mutants resembles the late flowering phenotype of *CDF1* overexpressors and at the same time the *CO* transcriptional profile of *fkf1-1* phenocopies the *CO* mRNA profile of *CDF1*-ox plants. Indeed FKF1 and CDF1 proteins interact physically in Y2H and in *fkf1-1* mutants the abundance of the CDF1 protein is increased (Imaizumi et al., 2005) but the *CDF1* mRNA is not significantly affected. The results proved that FKF1 promotes flowering by degrading CDF1 protein and this degradation is light dependent since it is enhanced by blue light.

These observations resemble remarkably the interplay between ZTL and GI for the degradation of TOC1. Since FKF1 belongs to the same family as ZTL and the two proteins share high levels of conservation, it is conceivable that FKF1 might also interact with GI. Indeed Y2H data confirm this interaction and as in the case of ZTL-GI the PAS/LOV domain of FKF1 and the N-terminal domain of GI are essential (Sawa et al., 2007). In vivo interaction

was confirmed by co-immunoprecipitation and again as with ZTL, the interaction was enhanced by blue light and reduced by red light or dark. Interestingly only under long days the protein profiles of GI and FKF1 overlapped with the FKF1 profile being similar between the two day lengths. However the peak of GI protein occurred approximately three hours earlier under short days vs long days and thus the coincidence with FKF1 protein was reduced.

Genetically *GI* is epistatic to *FKF1* since late flowering *GI* mutations introduced into a normally early flowering *FKF1*-ox background, phenocopies the *GI* mutant phenotype, suggesting the FKF1 function is dependent on GI. Since GI interacts with FKF1 and FKF1 degrades CDF1 a connection between the regulation of CDF1 by FKF1 and GI was hypothesised. Indeed in *gi* mutants CDF1 protein levels were more abundant compared to WT despite the presence of a functional FKF1 protein. These data are consistent with the genetic data and suggest that FKF1 requires GI for the degradation of CDF1 exactly as ZTL requires GI for the degradation of TOC1. CHIP indicated that GI associates with the *CO* promoter suggesting that the GI-FKF1 complex degrades CDF1 directly on the promoter of *CO*. In conclusion these data prove that FKF1 requires GI for the degradation of CDF1 and that internal coincidence between FKF1 and GI via the day length dependent shifts in GI protein profile is critical for the interaction with FKF1 and subsequently for the long day specific external coincidence of *CO* mRNA with light.

1.2.2.4 CONSTANS – Post-transcriptional pattern of regulation

This external coincidence as it was mentioned is critical for the up regulation of *FT* mRNA expression by CO. Interestingly in plants that constitutively over express the *CO* transcript, *FT* expression is dependent on the quality of light. More specifically infrared and blue light promotes the expression of *FT* and red light inhibits it (Yanovsky et al., 2002). These data are consistent with the flowering phenotypes of photoreceptor mutants. Mutants of the red mediating PHYB are earlier flowering than WT under long days while mutants of the infrared PHYA are later flowering together with mutants of the blue mediating CRY1 and 2. In addition the effect of these mutations on *CO* mRNA levels is minimum and cannot explain the dramatically reduced levels of *FT* mRNA. The data indicate either that CO is post-transcriptionaly regulated in a light-dependent manner or that the FT gene is regulated in a CO-independent manner by the photoreceptors.

Indeed the importance of light quality and quantity in the photoperiodic regulation of flowering time is reflected by the multiple roles that plant photoreceptors have. The five

member family of phytochromes (PHYA-E, Quail, 1994) that mediate red/far red (R/FR) responses and the two member family of cryptochromes (CRY1-2, Lin C. et al., 1996; 1998) that mediate blue (B) responses not only entrain the circadian clock but also control the degradation of the CO protein. More specifically red light via PHYB degrades CO consistent with the early flowering of the mutant. In contrast both infra red and blue light mediated by PHYA and CRY1 and 2 respectively, stabilise the CO protein leading to the upregulation of *FT* mRNA and to the promotion of flowering (Valverde et al., 2004). Last by not least in dark the CO protein is rapidly turned over.

Dark dependent degradation requires an E3 ligase since proteasome inhibitors stabilised the CO protein. A good candidate for this is COP1, an E3 ligase with was shown to interact with HY5 (Osterlund et al., 2000) LAF1 (Seo et al., 2003) and HFR1 (Jang et al., 2005). All these proteins are positive regulators of photomorphogenesis and COP1 is responsible for their degradation in the dark, promoting skotomorphogenesis. COP1 expresses this dark-dependent function by shuffling into the nucleus while light restricts the protein to the cytoplasm. Recently it was shown that CO acts downstream of COP1 and COP1 downstream of CRY1 and 2 since the late flowering *cry1cry2* double mutant when combined with the *cop1*mutant phenocopies its very early flowering phenotype (Jun Lui et al., 2008). A physical interaction between CO and COP1 in Y2H, requiring the CCT domain of the former, as well as *in vivo* conformation by FRET further indicated that COP1 might target CO for degradation. Additionally both proteins co-localise in the nucleus forming speckles. In the *cop1* mutant CO protein was stabilised proving that indeed COP1 degrades CO protein.

In conclusion under short days the GI protein does not overlap with the FKF1 protein and thus the stabilised CDF1 protein blocks the up-regulation of *CO* mRNA during the day. Shuffling of COP1 into the nucleus degrades the CO protein that could accumulate during the night due to the circadian peak. Inductive long day conditions cause a phase delay of GI compared to the short day pattern, leading to internal coincidence between GI and FKF1. The activated by blue light GI-FKF1 complex degrades CDF1 on the promoter of CO, allowing the transcript of the former to coincide with light. Stabilised by the PHYA and CRY1 and 2 photoreceptors, CO protein then accumulates leading to the activation of the downstream target *FT*.

1.2.2.5 A proposed mechanism for the activation of *FT* by CONSTANS

Although CO belongs to the zinc finger transcription factors, evidence regarding the ability of the protein to bind DNA has been lacking. How then does the accumulated CO protein up-

regulate *FT* mRNA? An attempt to answer this question came from the observation that the CCT domain of CO interacts in yeast with the HAP3, 2 and 5 complex. FRET confirmed the *in vivo* character of this interaction (Wenkel et al., 2006). The components of the HAP complex are conserved in plants, yeast and in animals. Their role is to activate transcription by recruiting the transcriptional machinery to the CCAAT promoter motif of target genes which they bind (Maity and Crombrugghe, 1998). This binding from the complex alone is weak probably transactivators which will enhance it are required. The *Arabidopsis* CO was shown to have a transactivating activity in yeast. Interestingly two of the yeast homologs of the HAP complex, namely CBF-A and C interact and form a heterodimer that recruits CBF-B. The heterotriplet then can bind DNA. Both CBF-A and C show similarity with the histone-fold motif of histones H2B and H2A. It would be interesting to investigate whether the putative HAP-mediated transcriptional activation of *FT* by CO is associated with histone modifications in the FT protein.

1.2.3 THE VERNALIZATION PATHWAY

1.2.3.1 General aspects of vernalization

Although the topic of this thesis does not refer to vernalization directly, components of the pathway are major targets of natural variation in *Arabidopsis* while the pathway is memorised in the results. Therefore a short reference in the spatial and molecular aspects of the mechanism might be helpful.

Temperature is monitored throughout the year, mainly by the apical meristem and secondarily by the leaves (Lang et al, 1965, Searle et al., 2006). In both sites, FLC acts as a repressor in order to maintain the vegetative growth during winter (Johanson et al., 2000, Sung and Amasino, 2003). Once the elevated temperatures of spring after the exposure to the prolonged cold of winter, are experienced, the activated vernalization pathway will repress FLC mRNA levels. The vernalization pathway therefore includes genes whose mutations block the acceleration of flowering regardless of photoperiod via the upregulation of FLC mRNA.

1.2.3.2 The molecular basis of vernalization

Vernalization at the molecular level is mediated through modifications in the chromatin structure of FLC, which encodes a MADS box transcriptional repressor. These modifications eventually result in silencing of the gene and consequently into the release of its repression on floral transition (Sung et al., 2006).

Initiation of the chromatin modification occurs when a long exposure to low temperatures upregulates the expression of VIN3, a PHD transcription factor (Sung and Amasino, 2004). VIN3 protein presumably interacts with a histone deacetylase and a DNA binding protein, thus forming a complex that eventually deacetylates *FLC* chromatin. This deacetylation is considered the first step towards the establishment of the vernalised state but at this point this state is established in a transient manner. The result of the deacetylation in *FLC* chromatin is the recruitment of another complex consisting of *VRN1* (Levy et al., 2002) and *VRN2* (Gendall et al., 2001) and a methyltransferase protein that will methylate lysines 9 and 27 of histone 3. This pattern of methylation probably signals the binding of *LHP1* to the *FLC* chromatin and thus a stable vernalized state is established. Genes that form the autonomous pathway could repress FLC expression either by deacetylation (*FVE* Ausin et al., 2004; *FLD* He et al., 2003) or by post-transcriptional modifications (*FCA* Macknight et al., 1997, *FY* Simpsons et al., 2003). The levels of the *FLC* transcript are increased in the presence of active

FRIGIDA (*FRI*, Johanson et al., 2000) alleles, thus *FRI* regulates FLC expression and vernalization requirement in a positive way.

FLC together with FLM/MAF1 and MAF2 genes, are three MADs box genes, which were shown to have a role in vernalization responses in *Arabidopsis* (Michaels et al., 1999; Ratcliffe et al., 2001; Scortecci et al., 2001). Another MADs box gene that is light regulated is SVP (Hartmann et al., 2000). Genetic analysis proved that SVP and FLM/MAF1 act as co-regulated partners (Scortecci et al., 2003). In addition, the *svp* mutation was able to partially rescue the late flowering phenotype of *gi* and *co* mutants under long days of 16 h, in a dosage dependent manner. It has been proposed that different MADs box proteins are involved in floral development act through hetero- or multidimers (Goto et al., 2001; Immink et al., 2002).

1.2.3.3 Integration of the flowering time pathways.

All the above pathways converge and regulate the expression of three genes *FT* (**F**LOWERING LOCUS **T,)**, *SOC1* (**S**UPPRESSOR OF **C**ONSTANS **O**VEREXPRESSOR **1,** Lee et al., 2000) and *LFY* (**L**EA**FY,** Weigel et al., 1992), the so-called flower integrators. *CO* activation upregulates *FT* and SOC1 expression, FLC represses mainly *SOC1* and *LFY* (Lee et al., 2000) but also *FT* expression, while the GA pathway up-regulates both *SOC1* and *LFY* (Parcy et al., 2005). *FT* expression might be also regulated directly by *PHYB* through the repression of *PFT1*, a gene that activates *FT* (Cerdan and Chory, 2003). *SOC1* mRNA is detected in the leaves both under LDs and SDs (in much lower levels though) and also throughout the SAM under LDs. LFY is also expressed in the leaves and the SAM. Last but not least *FLC* mRNA is expressed strongly in the SAM but also in the leaves.

Once activated the flower integrators will eventually up-regulate the expression of meristem identity genes such as the MADs box transcription factors *AP1, FUL and CAL* (***AP****ETALA* ***1, FRUITFULL*** and ***CAULIFLOWER***; Mandel et al., 1992), and subsequently floral homeotic genes such as *AP1*, *AG* (***AG****AMOUS*), *PI* (***PI****STILLATA*) and *AP3* (***AP****ETALA****3***) that will commit the shoot apical meristem with a floral identity (see below). It is important to note genes like *AP1*, *FUL* and *LFY* mediate two different responses: these genes act as meristem identity genes, affecting the length of the vegetative phase by initiating floral primordia and at the same time as floral homeotic genes since their mutations affect floral patterning.

1.3 FLOWERING BEYOND CONSTANS

1.3.1 PATTERNS OF INFORMATION TRAFFICKING IN PLANTS

1.3.1.1 A necessity to look beyond the leaves

So far all the regulatory components that were described, express their function in the vasculature of leaves. When it comes to flowering time, even the components of the circadian clock, a mechanism which is otherwise considered as cell autonomous, restrict their function to the leaves since this is the site of expression of their output targets. This is supported by results from GUS staining and missexpression studies with phloem specific promoters, of genes from the photoperiod pathway. These results are consistent with the notion that leaves are the eyes of the plants and that this is the tissue that perceives day length.

However the site of production of all aerial parts of the plant, including flowers and inflorescences is a small pool of stem cells in the center of the rosette that form the **S**hoot **A**pical **M**eristem (SAM). SAM together with the **R**oot **A**pical **M**eristem (RAM) are the two centers of plant development. This indicates that in order to fully understand flowering control one needs to look beyond the leaves and elucidates the mechanisms by which i) leaves induce transition to the reproductive phase in the SAM and ii) which genes form the regulatory cascade of flowering in the SAM. Before that though it is useful to mention a few things regarding the mechanisms by which plants transmit information over long distances.

1.3.1.2 The information highway of plants : The vascular network

In contrast to the animal cardiovascular system, the plant vascular system is non circulatory. It consists of two distinct units, the phloem and the xylem. The main task of the phloem is the delivery of the photo assimilates from the source tissues (mainly the leaves) to the sink issues (mainly the roots). The xylem on the contrary is responsible for the delivery of water and inorganic minerals from the roots to the aerial parts. The two systems differ not only in function but also in structure and position and are separated by a one-cell meristematic layer called cambium.

The xylem consists of dead cells that are connected to each other and form long tubes that allow the passive flow of water and diluted minerals due to respiratory and capillary forces. On the contrary, the phloem consists of living yet highly differentiated cells. These cells have lost most of their intracellular organelles including the nucleus and the vacuole. Cambium is responsible for the biogenesis of the phloem. After two longitudinal divisions one cambium

cell will give the two progenitor cells of the phloem. One of these cells will gradually loose its organelles in order to provide space for undisturbed movement. This cells known as **S**ieve **E**lements (SE) depend solely in the second cell, known as the **C**ompanion **C**ell (CC) for their survival (Lucas et al., 2001).

The classical aspect that wants the vascular network to simply move sugars and water throughout the plant body today is expanded. Studies in pumpkins, in which the isolation of the phloem sap is easy, revealed a complete proteome and transcriptome associated with the phloem, consisting of more than 1500 protein, mRNA (Lucas, unpublished) and small RNA (Yoo et al., 2004) species. These observations suggest that plants use their vascular network in order to transmit information over long distances. Indeed cases of long distance signalling have been found. A classical root–shoot-root signalling includes the nodulation process and its coordination with growth demands. *Medicago* roots perceive signals regarding the availability of inorganic nitrogen on soil. Upon low availability a signal travels from the roots via the xylem and activates a receptor kinase in the leaves. According to growth demands, this initiates another signalling cascade and a feedback signal travels to the roots via the phloem in order to coordinate the development of the N_2 binding nodules (Searle et al., 2003). A classical shoot–root signalling involves short day specific fruit development in potato while shoot-shoot signalling refers to the flowering and the florigen theory which will be analysed below.

1.3.1.3 The toll points of the information highway

As it is known, plants also exhibit the cell wall, a structure which surrounds and seals the cell from the outer side of the plasma-membrane. Plasmodesmata, which in Greek means cytoplasmic arrays, are connection points between neighbouring cells through the cell wall.

For many years it was believed that plasmodesmata are nothing more that simple holes, which allow the undisturbed trafficking of molecules from both sides. However our view regarding these connection points has changed dramatically during the last few years. Plasmodesmata are much more than plain holes : they consist of a core of endoplasmatic reticulum membrane. This core is attached to the cell wall via several globular spoke-like proteins (Lucas et al., 2001; Wu et al., 2002). The gaps between these proteins, the plasmodesma core and the cell wall are the connecting pipelines between two cytoplasms on each side of the plasmodesma. These gaps define the most fundamental property of the structure, the **S**ize **E**xclusion **L**imit (SEL). SEL defines an upper size threshold below which movement through the plasmodesma is allowed and above which is blocked.

SEL is far from being characterised as a static property. On the contrary it is very dynamic both spatially but also developmentally. *Arabidopsis* embryos at the heart state form a single symplasmic domain. This term refers to groups of cells that are connected via plasmodesmata, allowing undisturbed exchange. However at latter steps of the embryo development and more specifically at the torpedo stage four distinct symplasmic domain can be recognised. These domains fit very accurately to the symplasmic domains of young *Arabidopsis* seedlings and define spatial patterns of trafficking (Kim et al., 2005).

These spatial patterns are created by the differential SEL of the plasmodesmata that connect the different tissues. GFP localisation studies (32 kDa) have proved the existence of these domains. More specifically a single GFP molecule, under the activity of the *SHOOTMERISTEMLESS* promoter that is located in the SAM and just above the RAM, was able to traffic throughout the plant. A double GSF fusion however (>60 kDa) was excluded from the cotyledons and a triple was restricted only in the stem being excluded from the root as well. Expression of the double fusion from the MSG2 promoter, the activity of which is located in the SAM and RAM only, confirmed the retention of the fusion only in the two meristems suggesting that the stem and the leaves exhibit different SELs. According to their position the plasmodesmata have three different fates. Plasmodesmata that connect the progenitor cells of two SEs will be degraded. On the contrary plasmodesmata in the SE-CC junctions will be structurally differentiated compared to plasmodesmata between two common cells (Lucas et al., 2001). Last but not least based on the SEL plasmodesmata exist in three distinct forms: the open, which allows passive flow, the closed which blocks movement and the diluted, which promotes active trafficking (Wu et al., 2002). The above suggest that plasmodesmata are important regulatory points in the exchange of information between cells and tissues. But how exactly this type of trafficking occurs?

1.3.1.4 Information over distance : Cell to cell trafficking

There are two general types of mediators regarding the information over distance : RNA and proteins. Although the literature regarding intra-cellular trafficking is really exciting I will focus here on the aspects of intercellular trafficking. Proteins and mRNAs which are about to travel through cells, reach the plasmodesmata with the aid of the cytoskeleton, which is part of the trafficking and anchoring mechanism of intra-cellular signalling. Evidence about the involvement of cytoskeleton comes from the co-localisation of viral MP (Movement Proteins) with the microtubules. Viral MPs are responsible for the spreading of the viral RNA inside the

plant and share similarity with endogenous MPs that facilitate movement of endogenous transcripts and/or proteins (Lough et al., 2006).

Certain protein motifs and cis-regulatory RNA elements known as zip codes (Jansen, 2001), allow anchoring of the target to the plasmodesmata by the cytoskeleton. In the case of mRNA, 3'UTR or even CDS zip codes carry **R**ibo-**N**uclear **P**roteins (RNPs). These proteins are recognised by spatial chaperones. Target proteins are also recognised by chaperones directly. The target-chaperone complex gets recognised by a plasmodesmata-anchoring protein, which interacts with certain receptor proteins and allows docking on the plasmodesma unit. A second motif of the chaperone will interact with the binding site of the activated SEL protein and thus will trigger two conformational changes. First the structure of the plasmodesma will change and thus will increase SEL. In parallel the chaperon will partially denature the protein targets. Both of these changes will allow the trafficking via the plasmodesma to the second cell, in which binding to the cytoskeleton can follow (Lucas et al., 2001).

1.3.1.5 Information over distance : Long distance trafficking and surveillance on the highway

A protein or a transcript destined to travel from cell-to-cell will eventually reach the CCs. There in the case of mRNA the zip codes and the specialized RNPs that carry, will mark its non-cell autonomous fate from the cell autonomous fate of other transcripts driving it to the differentiated plasmodesmata. In a similar process as cell to cell trafficking both protein and mRNA will enter through the plasmodesmata into the SEs. There the RNP will detach the transcript and the zip code will be recognised by a competent **S**upracellular **T**ransport **P**rotein (STP). STPs will recognise protein targets as well. STPs are the postmen of the vascular network. Their role is to deliver the target to the right tissue, reducing cases of miss-delivery. Binding of the target requires competence, which is achieved via phosphorylation during or after the entrance to the phloem. Inactivation occurs upon return to the CCs.

Apart from the active transport, passive transport following the phloem flow can also occur. In addition the STPs reduce but do not eliminate cases of miss-delivery. Therefore in order to protect from miss-targeted products plants have developed a surveillance mechanism called **Sur**veillance **F**ield (SurF), which filters the signals that arrive from the phloem before allowing entrance to the SAM or RAM (Lucas et al., 2001; Foster et al., 2002). Although not much are know about the SurF, it is believed that it is based on the differential SEL of the meristems from the phloem and only certain STPs can allow entrance to the meristem. The

role of SurF is not only to restrict entrance to the meristematic cells but also to direct the granted-access targets to specific cell layers. The degradation machinery, components of which are present in the phloem sap, will degrade miss-delivered products.

This protection mechanism is believed to serve primarily as a defence against viral infection. Apart from endogenous targets, viral proteins or transcripts as well as PTGS components are blocked from the meristem. Interestingly during evolution, viruses developed ways of bypassing the defence of plants. Typical example is the case of the **p**otexvirus Potato **v**irus **X** (PVX). In WT plants the viral RNA is excluded from the meristem. Plants that over express the viral MP TGB1, allow invasion of the virus into the SAM after inoculation. Interestingly non-inoculated TGM1-ox individuals show developmental abnormalities due to the inversion of the meristem by otherwise blocked endogenous signals (Foster et al., 2002)

1.3.2 LINKING LEAVES AND SAM : THE ELUCIVE NATURE AND THE ROLE OF FLORIGEN

1.3.2.1 The history of florigen

One of the first studies in the attempt to describe the nature of florigen was in Perilla. The tools included the use of radioactively labelled ^{14}C-photoassimilates and safranin as markers of phloem flow in combination with grafting experiments using induced short day leaves as donors and non-induced long day grown plants as acceptors (King and Zeevart, 1972). The connection between photo assimilates and movement of flowering related signals became apparent from studies of *Pharbitis nil* in which is was shown that the velocity of the transmissible signal associated with flowering induction, equals the velocity and range of photo assimilates suggesting that the two substances travel in the same manner through the phloem (King et al., 1968).

These experiments showed that indeed flowering formation was related in a quantitative manner with the transport between donor-acceptor. This quantitative nature has been confirmed in many species, from *Xanthium strumarium* (van De Pol, 1972) to *Impatiens balsamina* (Pouteau et al., 1998). Induced short day leaves are an absolute requirement for floral induction and the fact the living cells are important suggested phloem as the route of movement. Another indication that phloem connections are important came from the observation that the effect of donor leaves was evident in shoots of the same orthostichy. Interestingly the position and number of non-induced leaves affected the distribution of both photo assimilates and floral induction to various parts of the plant from the induced donor leaves. More specifically vegetative leaves were able to import photo assimilates from LD leaves but flowering shoots received photo assimilates only from SD induced leaves.

In addition the authors demonstrated that there is a time duration regarding the presence of the induced leaves on the plant, suggesting the existence of i) a critical threshold for the inductive signal, required to promote flowering ii) that in the phloem insignificant amounts of the florigen exist. In the opposite case removal of the induced leaves would not inhibit flowering induction. These properties of duration for floral induction together with the non-efficient stem accumulation of the florigen were contrasting the kinetics of photo assimilate movement and accumulation in the stem, separating the two substances. Interestingly although *Perilla* plants require 9 SDs for induction, leaves from plants experiencing 6 SDs were unable to trigger flowering completely and showed only position specific weak properties of floral induction. This could be interpreted as a moderate diffusion of a poorly

accumulated inductive signal. However species such as *Xanthium*, exhibit the phenomenon of indirect induction, according to which the induced state can be transferred through several rounds of successive grafting. Interestingly, although inter-specific grafting experiments suggested that the identity of the florigen is common of at least very similar between a variety of SD and LD species, not all species exhibit properties of indirect induction (Zeevaart et al., 1976)

1.3.2.2 The molecular path towards the identification of florigen

As discussed, CO protein accumulation under inductive LDs in *Arabidopsis* results in the up regulation of FT mRNA expression. GUS staining indicates that both CO and FT promoters are active in the CCs of the minor veins of the phloem (An et al., 2004). Expression was excluded from the meristem under the two native promoters, however these results must be taken with caution since the expression level of both genes is generally very low. Miss-expressing CO in the minor veins of the phloem was enough to rescue the late flowering phenotype of co mutations, suggesting that expression of CO solely from the leaves is sufficient to induce flowering. Grafting experiments suggested that this ability was indeed transmitted over long distances.

This connected CO with the florigen and suggested that either CO itself or its targets could participate as long-distance signals. Interestingly in SU2::CO:GFP plants fluorescent signal was detected only in the CCs suggesting that either the CO:GFP fusion (~60 kDa) was very large or in general incapable to enter to the major veins of the phloem or that the native CO protein does not move beyond its site of transcription. At any case the GFP fusion retained functionality and rescued the late flowering phenotype of *co* mutants. However missexpression of CO in the SAM was not able to rescue the late flowering of co mutants, suggesting that the CO-target FT could mediate the long distance response.

1.3.2.3 FT protein is a central part of the florigenic signal

FT is a 20 kDa protein which shows similarities to the **R**af-1 **k**inase **i**nhibitors (RKI) of mammals. The *Arabidopsis* genome encodes apart from FT, five more similar proteins, namely TSF (**T**win **S**ister of **F**T), MFT (**M**other of **F**T and **T**FL1), BFT (**B**rother of **FT**), TFL1 (**T**erminal **Fl**ower **1**) and ATC (***A**rabidopsis **t**haliana* relative of **C**entroradialis, Turck et al., 2008). Of these proteins TSF shares redundant temporal and spatial pattern of expression but much lower level of expression while evidence for functional redundancy were also provided by the analysis of double mutants. TFL1 on the contrary, although it is very

similar to the FT protein, it acts as a floral inhibitor and it is expressed in the apical meristem. In addition FT like genes have been identified in other Brassicaceae species, in *Pharbitis*, in monocots (rice), trees (poplar, citrus). This conservation together with the small size of the protein makes it partially capable of fulfilling the requirements of a florigen candidate.

Miss-expressing FT both in the CCs of the minor veins and in the apical meristem rescued the late flowering of *co* and *ft* mutations, indicating that the protein promotes flowering in both sites. This suggested that FT itself or an FT-target this time could mediate the response. Compromising the ability of the FT protein to move beyond the CCs resulted into late flowering of SUC2::FT:2xGFP plants, at the same time that single FT:GFP fusions were restoring early flowering and allowing fluorescence signal to be detected in the SAM (Corbesier et al., 2007).

In addition grafting experiments both in *Arabidopsis* (Corbesier et al., 2007) and tomato (Lifschitz et al., 2006) proved that the FT mRNA could not move through the graft junction, however the acceptor plant was induced to flowering. Similar results were obtained with the rice honolog of FT (Tamaki et al., 2007), proving that the FT protein is indeed a large part of the florigenic signal across diverse species. The fact that *Arabidopsis* eventually will flower also under short days, suggests that other molecules could also contribute to the long-distance induction of flowering, with hormones and specifically gibberellins being attractive candidates.

1.3.3 BEYOND LEAVES : THE ROLE OF THE SHOOT APICAL MERISTEM IN THE CONTROL OF FLOWERING

1.3.3.1 Introducing the shoot apical meristem

Plants retain a post-embryonic ability for development, which in the case of the aerial parts, requires a functional **S**hoot **A**pical **M**eristem (SAM). In *Arabidopsi*s the SAM consists of stem cells that are positioned in the center of the rosette, surrounded by young leaves that protect them. In terms of direction of cell division two different forms can be identified: the tunica and the corpus. The tunica that is located in the periphery of the meristem consists in many species of two layers, namely the L1 and 2. In both layers, cells divisions take place vertically to the surface of the SAM, thus expanding the length of the tunica but not its width. On the contrary the corpus cells can divide in any direction.

Judging from the mitotic activity and fate of the meristematic cells, two zones can also be distinguished. The peripheral zone consists of cells that divide vividly while the central zone consists of cells that are mitotically less active in vegetative apices. Cells of the peripheral zone are destined to give all the new organs while cells of the central zone are pluripotent and retain a stem character (Brand et al., 2001). Since the meristem needs to create new organs but at the same time must retain a pool of stem cells, a cross talk between the peripheral and the central zone is necessary. Floral meristems of *Arabidopsis* differ from the inflorescence meristem, since the former exhibit complete floral differentiation and thus determinate growth (see below).

Interestingly, upon floral induction several physiological changes occur in the meristem. Plasmodesmata density increases but the meristem at the same time divides into a peripheral and a central symplasmic field, restricting communication between the two domains (Kwiatkowska et al., 2008). This restriction is essential for the subsequent differentiation of the floral primordia and the maintenance of the inflorescence meristem. Cell division rates in the central zone equalise the rates of the peripheral zone in order to coordinate organ growth and meristem. Subsequent steps of floral induction include separation of the floral meristems via the formation of a boundary zone between the SAM and the floral primordium. This zone consists of a layer of slowly dividing cells and overlaps with the expression domain of CUC2 (**Cu**p-shaped **C**otyledon **2**) shape. Floral induction leads also to alterations in the geometry of the SAM with a marked increase in the height and shape leading to meristem doming. However the changes in the geometry are also related to changes in the size of the cells due to vacuolation rather than strictly to cell division. Last but not least, certain early changes occur

in the meristem also at the molecular level. A general RNA increase has been observed in meristems of *Sinapis alba* and *Pharbitis* 16 h after photoperiodic induction (Zeevaart et al., 1976), consistent with the idea that the first steps of floral induction alter the transcriptional profile of the SAM, marking the transition to the reproductive phase.

1.3.3.2 Perpetuation of the apical meristem

Maintaining the stem cell pool of the meristem is vital for the continious productions of new organs. This maintenance must be controlled with the rhythm of organ production and the central zone must supply with soon-to-be differentiated cells the peripheral zone. Four genes, namely CLAVATA 1, 2 and 3 (CLV) and WUSCHEL (WUS) play central roles in the whole process. Both *CLV1* and *2* genes encode leucine-rich repeat transmembrane (LRR) proteins with the CLV1 having a C-terminal serine/threonine kinase domain (Clark et al., 1997). It is was shown that both proteins form an heterodimer that is anchored in the plasmamembrane with the LRR domain facing the outer part and the kinase domain of CLV1 being cytoplasmic. CLV1 is expressed in the L3 layer, below the L1 and L2 of the peripheral zone above the central zone. Members of the LRR kinases express their function via the binding of a receptor that activates the cytoplasmic kinase domain and facilitates intra-cellular signalling and eventually nuclear target activation.

CLV3 seems to be the ligand that activates the CLV1 and 2 heterodimer (Brand et al., 2000 and 2001; Weites and Simon 2000; Muller et al., 2006). The gene encodes a small protein and its mRNA is found in the L1 and L2 layers. CLV3 is secreted from the L1 and L2 cells and binds the CLV1 and 2 complex in the L3 forming a 450 kDa complex which is believe to be biologically active. Protein interactions were verified in Y2H and demonstrated that the CLV1/2 complex is stabilised in the presence of CLV3. A phosphatase (KAPP) and a GTPase (ROP) also participate in the 450 kDA complex. Upon CLV3 binding a signalling cascade including MAPKs is initiated and eventually the expression of WUS is turned off inside the nucleus. Therefore the CLV complex is a negative regulator of WUS.

But what is WUS? WUS is a homeodomain transcription factor, whose mRNA is first detected at the 16-cell stage of *Arabidopsis* embryos (Mayer et al., 1998). Its role is to promote a stem cell identity. Mutant analysis confirmed the negative regulation of WUS by the CLV complex, a regulation that occurs at the transcriptional level. Loss of stem cell fate in *wus* mutants not only alters the geometry of the meristem but also causes determinate growth and differentiation, as expected. Nevertheless *wus* mutants still induce axillary meristems suggesting that the activity of WUS is not to initiate the meristem but to maintain its

pluripotent nature, providing in parallel stem cells to areas that will give rise to new organs. Interestingly WUS positively regulates CLV1 and 3 expression.

In conclusion low number of stem cells producing the CLV proteins due to differentiation, release WUS repression and the gene provides new stem cells. At the same time promotion of the CLV genes by WUS will establish again the feedback negative regulation, ensuring equilibrium and avoidance of uncontrolled growth (in the case of WUS hyperactivity) or determinate growth (in the case of CLV hyperactivity) (Brand et al., 2000). This regulation of WUS by CLV however is not strictly linear since 10-fold fluctuations in CLV expression do not affect WUS expression (Muller et al., 2006).

1.3.3.3 SAM regulatory cascades : Leaves and their role in the regulation of apical transcription

One of the first studies that connected molecular cascades of the apical meristem with events in the leaves included a combination of a Dex-inducible system regulating CO and *in* situ hybridisations (Simon et al., 1996). The advantage of the approach was that it bypassed additional complexity due to the application of LD conditions that could alter the expression of many genes in the apical meristem, focusing at the same time into the CO-dependent effects.

This study showed that 24 h after application of DEX in CO:GR plants grown under SDs the expression of TFL1 (**T**ERMINAL **FL**OWER **1**), AP1 (**AP**ETALA**1**) and LFY (**L**EA**FY**) was up-regulated in levels comparable with LD grown plants. The role of each one of these flowering genes will be discussed later but in short both LFY and AP1 promote the floral fate and TFL1 promotes the inflorescence or shoot fate. These results confirmed that a major regulator of flowering expressed in the leaves, is capable to induce molecular alterations of flowering genes in the meristem. A genetic approach to start classifying the hierarchy of the whole process, included mutant combinations of LFY and flowering time genes (Nilsson et al., 1998). According to these results two general types of mutations could be identified: mutants which altered the expression of LFY and therefore could be assigned upstream of the gene and mutants which altered the response to LFY activity, the effects of which could be assigned downstream or in parallel to LFY.

Interestingly *ft* and *fwa* (a gene that regulates the expression of FT) mutants showed the biggest effects in the response to LFY activity, estimated by alterations of flower structure (an indication of abnormal LFY activity). The two genes on the other hand had minor yet detectable effects in the expression of LFY. On the contrary *co* and *gi* mutations although had

very small effects on the response to LFY activity, reduced dramatically the expression of LFY. Interestingly together with *gi* and *co* mutations, mutants of the autonomous pathway, which up-regulates FLC expression, showed similar patterns. Both CO and FLC regulate FT and SOC1 (see 1.2.3.3) and since *ft* mutants had smaller effects, a good candidate for mediating the expression specific effect of these mutations to LFY is SOC1 (see also below).

In an attempt to expand our molecular knowledge regarding SAM-specific transcriptional events in floral induction, microarray hybridisations were conducted in the two common laboratory strains of *Arabidopsis* Col-0 and L*er* and this results were compared with microarray hybridisations in *ft*, *co* and *lfy* mutants (Schmid et al., 2003). The authors induced all plants to flowering after a shift from short non-inductive days into long inductive conditions. These results confirmed that *LFY*, *AP1* and *TFL1* expression is being up-regulated in upon floral induction in WT plants with, *co* and *ft* mutants blocked this up-regulation, although some of the target genes responded differently to the different mutations. The effects of *lfy* mutations were subtle and restricted after floral induction, indicating that LFY acts further downstream in the meristem floral cascade. These results will be discussed in detail in the next paragraphs in an attempt to dissect the different processes of SAM floral induction. All these data however confirm the strong effect that the signal from the leaves exhibits on the expression of meristem specific genes associated with flowering.

1.3.3.4 SAM regulatory cascades : Molecular consequences of florigen arrival from the leaves – AP1

Since genes in the leaves regulate the expression of genes in the SAM, it is logical to assume that this control passes through florigen. The fact that FT regulators such as FWA, CO and GI, affect the expression in the SAM, further strengthens this argument. Since FT protein is the long sought florigen then FT must be able to mediate at least part of these responses biochemically. Genes like *FUL* and *CAL*, the proteins of which show similarities with the *AP1* and share partially redundant roles in the meristem, are affected in the transcriptional level in *ft* mutants in the meristem (Schmid et al., 2003). These genes are also expressed in the leaves, without affecting flowering time there. However in the leaves FT alone is not able to affect their expression since this ability required the presence of the bzip FD (**F**lowering locus **D**, Koornneef et al., 1991) protein (Teper-Bamnolker et al., 2005, Wigge et al., 2005). Mutants of FD cause late flowering, masking the extremely early flowering of FT-ox plants (Wigge et al., 2005). The gene is expressed throughout the apical meristem and to a lower extent in the leaves and upon floral induction by long days its expression is induced (Wigge et

al., 2005). This slight induction in the meristem is not due to FT (Searle et al., 2006), although in *ft* mutants FD mRNA in the leaves is slightly reduced with the difference being comparable to the SD-LD difference (Abe et al., 2005). Combined these data indicate that FD acts either synergistically or downstream of FT in the control of flowering time genes in the meristem.

Experimental evidence regarding a physical interaction of FD and FT were provided by Y2H and FRET, proving that the two genes act synergistically (Abe et al., 2005). Interestingly the FT/FD complex is able to up-regulate the expression of AP1 in the periphery of the meristem and more specifically in the floral anlagen (Wigge et al., 2005; Abe et al., 2005). Mutations in either gene fail to up-regulate AP1 expression leading to late flowering and abnormalities in the floral pattern. Direct evidence about this transcriptional regulation were provided by CHIP the results of which mapped an FD response element in a 130-bp region of the AP1 promoter. Recruitment of FD in the AP1 promoter requires FT and this is supported with enriched AP1 promoter amplicons in CHIP from LD vs SD samples (Wigge et al., 2005).

1.3.3.5 SAM regulatory cascades : Molecular consequences of florigen arrival from the leaves – SOC1 and LFY

Is AP1 the only target of FT/FD complex in the meristem? According to microarray data apart from *AP1*, *SOC1* mRNA expression is also reduced in the meristem in *ft* mutations (Schmid et al., 2003). *SOC1* is expressed also in the leaves and *SUC2::SOC1* plants can rescue the *soc1-1* late flowering phenotype, by inducing FT expression (Searle et al., 2006). In the meristem, it is possible that the FD/FT complex is required, in a similar to AP1 way, to regulate SOC1 expression. Indeed *fd* mutants, after a shift from SDs to LDs failed to induced SOC1 expression in the meristem in contrast to WT plants, in a similar way to *ft* mutants (Searle et al., 2006).

In parallel, plants which express FLC in high levels due to functional FRI alleles, failed to up-regulate SOC1 expression in the meristem. Repression of SOC1 by FLC was shown to be direct due to binding of FLC to a CarG box in SOC1 promoter (Hepworth et al., 2002). Similarly FLC binds to a CarG box interestingly in the first intron of FT, repressing the expression of the gene (Searle et al., 2006). FLC also binds to a CarG box in FD repressing the expression of this gene as well and thus explaining how strong FLC alleles delay flowering. Last but not least, strong FLC alleles down-regulate LFY in the meristem (Lee et al., 2000). Collectively these data suggest that likewise to AP1, the FT/FD complex up-regulates SOC1 expression in the meristem and that strong FLC expression repress it together with LFY.

As mentioned, microarray data indicated the LFY expression in the meristem is down regulated in *co*, *gi* or *ft* mutations. However GUS staining proved that LFY expression in the meristem is strongly down regulated only in *co* or *gi* mutants (Nilsson et al., 1998) and is less affected by *ft* mutations. Since *CO* regulates both *SOC1* and *FT* in the leaves and since FT/FD regulate *SOC1* in the meristem, it is possible that this down-regulation of *LFY* in SAM is in part, mediated by these genes. Indeed SOC1 can partially up-regulate LFY in the meristem. But is this regulation direct or not?

It seems that AGL24 (**AG**AMOUS-**L**IKE **24**) a MADs box gene, mediates this regulation at least in part. *AGL24* mRNA is expressed in the apical meristem increasingly with development (Yu et al., 2002). *agl24* mutants are late flowering and partially suppress the early flowering of FT and SOC1 overexpressors, indicating that the gene acts downstream of CO, FT and SOC1. *ft* mutations do not strongly affect *AGL24* expression but interestingly SOC1 mutations strongly affect *AGL24* expression. Surprisingly recent experiments suggest that AGL24 feeds back to SOC1 and directly up-regulates its expression via binding in the promoter region (Liu et al., 2008). Such positive regulation can occur independently of FT and in response to gibberellins, suggesting a mechanism for amplification of the floral signal in the apex in response to the GA pathway. Although *SOC1* is regulated but *FT* and the photoperiod pathway, do not forget that *SOC1* as a floral integrator is also positively regulated by the GA1 pathway. The above therefore can explain the inconsistency between the FT-mediated *SOC1* regulation and the differential *FT* and *SOC1* mutant effects on *AGL24* expression.

LFY overexpressors suppress the late flowering of *agl24* plants, placing the gene downstream of SOC1 and upstream of LFY. GUS staining indicates that in *agl24*, *LFY* mRNA expression is delayed and reduced but not eliminated, indicating that the role of AGL24 is to partially initiate LFY expression upon floral transition, without being a necessary factor for its maintenance. This is consistent with microarray in which LFY up-regulation is severely delayed in *co* and *ft* but is eventually initiated much later than WT, probably due to the positive effects of the GA pathway (Schmid et al., 2003).

Collectively all the above data suggest that upon arrival from the leaves, FT interacts with the bZIP transcription factor FD. The FT/FD complex is then competent to up-regulate first the expression of SOC1, one of the earliest targets of floral transition in the meristem. SOC1 upregulates AGL24 and AGL24 is able to at least partially to initiate LFY expression, which is restricted in the periphery of the meristem in the floral anlagen. The FT/FD complex up-regulates also AP1 expression in the floral anlagen, but the gene is expressed slightly later

than LFY (Hemplel et al 1997). This is consistent with the earlier SOC1 expression that is required for LFY initiation. These differences could reflect different requirements of FT protein accumulation in the meristem for transcriptional activation between SOC1 and AP1.

1.3.3.6 SAM regulatory cascades : Defining floral primordia

It is by now obvious that the arrival of FT from the leaves, triggers a transcriptional chain reaction that prepares the meristem to change into the floral fate. But what is the switch that will initiate eventually flowers? First floral primordia must be defined. Key role in this process have the already up-regulated *LFY* and *AP1* transcripts together with CAL and FUL. AP1 CAL and FUL are partially redundant MADs box genes, with AP1 having the primary role, CAL being able to substitute partially AP1 and FUL being a gene which in the absence of both AP1 and CAL can still weakly complement for their loss and lead to limited flower development (Ferrandiz et al., 2000). FUL acts also as a flowering time gene with *ful* mutants displaying later flowering than the WT. Interestingly this flowering time delay is more the outcome of increased cauline leave production than rosette, reflecting most probably the enlarged duration of the meristem for the commitment after the arrival of the florigen. AP1 and CAL are expressed only in the floral anlagen (see below), domains in the periphery of the meristem which are destined to become floral primordia, which will be developed into flowers. FUL on the contrary is expressed throughout the meristem after photoperiodic floral induction and its expression appears also later in the center of the flowers, suggesting a role in carpel development (Ferrandiz et al., 2000).

The function of these genes was elucidated in mutant combinations. A triple *ap1calful* mutant convers all of the floral meristems into leaf shoots, resembling strikingly the *lfy* mutant phenotype. It is logical to expect then that the role of the three MADs box genes is to either regulate the expression of LFY or to attenuate the response to it function. *In situ* hybridisations showed that although LFY mRNA was reduced in *ap1* and *cal* single mutants, meristems of *ap1cal ful* triple mutants displayed even lower levels of LFY mRNA, comparable to the initial levels after photoperiodic induction (Ferrandiz et al., 2000). Interestingly LFY is able also to boost the expression of AP1, after this has initiated (Parcy et al., 1998) suggesting that once activated, both genes act synergistically to maintain their expression and to promote floral fate. LFY regulates AP1 but also AP3 and AG expression (see below) via direct binding in a CC(A/T)(A/G)(G/T)G(G/T) site in their promoter region (Chae et al., 2008).

However is this down-regulation of LFY mRNA the sole explanation for the *ap1calful* triple phenotype? The answer is partially yes, since over-expressing LFY was able to partially rescue the late flowering and flower-to-shoot triple mutant phenotypes. It is very important to note that the flowering time of 35S::LFY*ap1calful* was almost double the flowering time of WT Ler plants, while the flowering time of 35S::LFY alone was almost half. This demonstrates that LFY can control the putative threshold for flower commitment but this control depends on either the AP1 / CAL / FUL genes or to an outcome of their function (see also below).

A number of conclusions can be made here. First these results explain why the triple mutant does not create flowers compared to the single and double AP1 / CAL mutants. They also prove that LFY acts in parallel of AP1, CAL and FUL and that one of the roles of the these genes is to spatially define meristematic regions, the floral anlagen, which have a strict floral destiny. In addition these three genes are not responsible for initiating LFY expression but for boosting and maintaining it after floral induction. This separates their role from the role of AGL24 in the control of LFY. LFY can feedback on AP1 and boost its expression as well, unravelling a loop which is responsible for the maintenance of the floral fate.

1.3.3.7 SAM regulatory cascades : Transcriptional events in the development of floral primordia

Flower primordia have entirely different properties than inflorescence meristems. The flower organs, namely the four sepals and four petals, the eight stamens and the two fused carpels in *Arabidopsis*, are not produced in a spiral way like the leaves but in four whorls. The outer whorl 1 has the sepals, whorl 2 the petals, whorl 3 the stamens and in the central whorl 4 is one fused carpel. In addition, although the growth of inflorescence meristem is indeterminate, flower primordia by definition will have a determinate growth. What is the mechanism that causes all these differences in the floral primordia and eventually how flowers are formed? The already expressed LFY and AP1 genes posses key roles in the process.

In an attempt to understand the transcriptional events that determine floral meristem fate in a genome-wide level, microarray analysis was performed (Wellmer et al., 2006). In contrast to the already mentioned study of Schmid and collegues, who were interested in the early transcriptional events in the SAM upon floral photoperiodic induction, this second study focuses on more downstream events. The authors used *ap1cal* mutant plants, which normally do not produce flower primordia. In these plants they inserted an 35S::AP1:GR construct,

which upon DEX application allowed them to initiate uniformly and at will floral primordia and subsequently normal flowers and at the same time identify many known and unknown genes which were up- or down-regulated. These results suggest that the role of AP1 in determining floral primordia includes both activation and a repression identity depending on the target.

The microarray data confirmed the role of AP1 in the up-regulation LFY but interestingly FUL was slightly reduced confirming previous observations about a role for AP1 in repressing FUL (Ferrandiz et al., 2000). However the expression of the MADs box homeotic genes AP3 and 2, SEP3, PI and AG, known also as ABC genes, was up-regulated again confirming previous studies. According to the standing theory, flowers are developed by the combined action of these distinct functions. The A function alone is responsible for the development of sepals. The combined action of A and B is responsible for the development of stamens. B and C together lead to the development of stamens and C alone allows carpels to develop (Weigel et al., 1994). These and also morphological results demonstrate that upon DEX treatment floral development of 35S::AP1:GR *ap1cal* plants is similar to wt plants.

During the first day after DEX application, a large number of genes were repressed while at the same time few targets were up regulated. Between 3 and 5 days after treatment, the up vs down-regulation ratio was more equalised due to the activation of genes responsible for floral patterning. This result indicates that two different phases mark the floral transition: an early phase during which many genes are repressed and a latter phase during which genes that regulate floral patterning are activated. These two phases are separated by a few days. TCP transcription factors and MADs box genes were among the families of genes, the expression of which was affected during both phases. TCP transcription factors regulate cell division and proliferation while many MADs box genes are involved in floral patterning. Members from both families are therefore needed to ensure the determinate growth of the floral meristem and the production of floral organs.

1.3.3.8 On the way to flower creation : Sepals

AP1 mRNA is expressed throughout the floral primordia early in their development. Latter it is confined only to the outer regions destined to create sepals and petals (Gustafson-Brown et al., 1994). This observation suggests that AP1 may express the A function. Indeed *ap1* mutants lack both sepals and petals while carpeloid structures are found in the second whorl. The fact that in *lfy* mutants, the sporadically developed flowers lack sepals and petals, further supports this view since LFY as mentioned positively feeds back on AP1 expression. When

mutations in AP2, another MADs-box factor, are combined with AP1, floral organs arise in a spiral fashion resembling inflorescence meristem organisation. Strong AP2 mutants in combination with AP1 mutants, seem to phenocopy LFY mutations since many flowers are completely transformed into shoots (Weigel, 1995).

1.3.3.9 On the way to flower creation : Petals and stamens

Two MADs box transcription factors, *AP3* and *PI*, mediate the B function. Mutations in either gene result in ectopic appearance of sepals and carpels in the second and third whorl respectively. This indicates that the combined action of both is required for the production of sepals and stamens (Goto et al., 1994; Jack et al., 1992; Samach et al., 1999). Indeed both proteins form an heterodimer (Riechmann and Meyerowit, 1997). Overexpression of AP3 causes ectopic appearance of staminoid carpels in the inner whorl while PI-ox plants have petalloid sepals in the outer whorl. These observations indicate that the B function has increased.

But how AP3 and PI are expressed in such defined regions? UFO (**U**NUSUAL **F**LORAL **O**RGANS) an F-box gene and LFY regulate this expression pattern. Weak *lfy* mutants are lacking petals and stamens, suggesting that LFY is required for the B function (Weigel et al., 1992). It was shown that LFY is able to regulate AP3 expression via direct binding in its promoter (Lamb et al., 2002). *ufo* mutants lack normal petals and stamens while overexpressors have many. This effect of UFO overexpression is lost in *lfy* mutants and since LFY expression is not affected, a synergistic role of the two genes is suggested. Overexpression of both PI and AP3 rescues the loss of petals and stamens in *ufo* mutants (Krizek and Meyerowitz, 1996), and this establishes that UFO positively regulates the expression of the B genes. Its activity is required for the maintenance of PI and AP3 expression but not initiation, with the control of AP3 expression being direct and PI indirect. PI maintains its expression in an AP3 dependent manner (Jack et al., 1994) therefore the indirect effect of UFO on PI expression is most probably via AP3 (Samach et a., 1999).

UFO in addition to these roles, seems to have a more general role in the normal growth of the floral meristem (Samach et al., 1999). In the inflorescence meristem UFO seems to be required for difining the position of the floral primordia (Tepper-Bamnolker et al., 2002). In addition UFO is responsible for the suppression of the bract in *Arabidopsis* flowers (Tepper-Bamnolker et al., 2002). Although UFO expression is detected in all meristems, in floral primordia its expression is eventually restricted in the second and third whorl (Samach et al., 1999)

LFY and UFO were recently shown to interact physically based on Y2H and pull-down assays, with the C-terminal region of LFY mediating the interaction and the N- terminal stabilising it (Chae et al., 2008). This interaction recruits UFO in the AP3 promoter, since LFY can bind the CC(A/T)(A/G)(G/T)G(G/T) element present in the promoter of the latter, and there UFO acts as a co-activator of LFY. The fact that UFO is an E3 ligase and that proteasome inhibitors reduce induction of AP3, suggests that the LFY/UFO up-regulates AP3 by degrading putative repressors (Samach et al., 1999; Chae et al., 2008). This pattern of transcriptional activation is not unusual and reminds the cases of TOC1 and CDF1 regulation by GI/ZTL and GI/FKF1 complexes respectively (see previous chapter).

1.3.3.10 On the way to flower creation : Carpels

AG mRNA is detected in the central whorl where carpels will be developed (Parcy et al., 1998). This observation suggests that AG may mediate the C function, responsible for stamen development in whorl 3 together with PI, and carpel development in whorl four. In flowers of *ag* mutants, normal carpels are absent from the central whorl while stamens are replaced by sepals in the third whorl proving that indeed AG is the C gene (Gustafson-Brown et al., 1994). At the time that *AG* expression is initiated in the central whorl *AP1* expression, although uniform initially, becomes restricted to the first two whorls (Gustafson-Brown et al., 1994; Parcy et al., 1998). This suggests that AG may negatively regulate AP1. Indeed in *ag* mutants *AP1* mRNA can expand to the central regions of the flower and convert stamens into sepals (Gustafson-Brown et al., 1994). Flowers of *ap1* mutants do not have sepals while carpeloid structures are ectopically found in the outer whorls. In *ap2* AG mRNA is ectopically expressed in the first two whorls and AP1 mRNA levels are reduced (Gustafson-Brown et al., 1994). This effect AP2 on AP1 expression is indirect through AG since in *ap2ag* double mutants AP1 levels are not affected (Gustafson-Brown et al., 1994).

But AP1 can also repress AG, thus preventing it from expanding towards the outer whorls (Weigel et al., 1994) and thus establishing a mutual antagonism between the inner and outer whorls that separates the A from the C function. The mechanism of AP1 mediated repression of AG includes the recruitment of the SEU (**SEU**SS), LUG (**L**E**U**NI**G**) repressor complex (Sridhar et al., 2004) in the second intron of AG (Gregis et al., 2006). Two additional MADs box genes SVP (**S**HORT **V**EGETATIVE **P**HASE, Hartmann et al., 2000) and AGL24 are also involved (Gregis et al., 2006). Both genes have dosage-dependent yet opposite effects on flowering time with SVP being a repressor and epistatic to the AGL24 activator (Liu et al., 2007). Interestingly in high temperatures the double mutant shows homeotic transformations

in all floral whorls. Since after the transition to flowering their mRNAs overlap in the inflorescence meristem while the two proteins are very similar, it was postulated that the two genes have an additional role in floral patterning.

In situ hybridisations showed that AP3 mRNA, which is normally restricted in the second whorl, is ectopically expressed in all parts of the floral meristem. In addition AG mRNA that is normally restricted in whorls 3 and 4, in the double agl24svp is expressed earlier and in all whorls. Both genes do not affect the expression of LUG or SEU and do not interact on them suggesting that their effect on AG in independent of the LUG-SEU complex. Interestingly AP1 interacts with the LUG-SEU complex as well as SVP and AGL24 (Pelaz et al., 2001; de Folter et al., 2005). This suggested that the effect of AGL24 and SVP on AG could be in mediated through AP1. Testing AP1 interactions with SEU in Y2H in the absence of functional SVP and/or AGL24 proved that AP1 does not interact. On the contrary yeast three hybrid showed that AP1-SVP and AP1-AGL24 dimers interact with SEU. Finally in a Y4H assay it was shown that the LUG requires all three proteins to bind. Collectively these data suggested that the AP1-SVP and AP1-AGL24 dimers allow SEU binding and this complex recruits the LUG repressor (Gregis et al., 2006). Once these interactions are established AP1 recruits the SEU-LUG repressor complex to the binding sites of AG second intron, repressing its expression.

1.3.3.11 On the way to flower creation : Ensuring determinate growth

In *lfy* AG expression is not up regulated. LFY therefore plays significant role in the up-regulation AG, while AP1 and AP2 represses it. Both LFY and AP1/2 are expressed throughout the floral meristem while AG only in the centre. Additionally both LFY and AP1/2 mRNAs precede that of AG. Since LFY is expressed earlier than AP1, it can up-regulate AG. But how this up-regulation is restricted in the central whorl, especially since LFY is expressed throughout the floral primordium?

Like in the case of AP3, LFY requires a co-regulator to express regulate AG expression specifically in the center of the leaves. WUS was identified as this target, however this identification was not straightforward since most *wus* do not have obvious floral homeotic transformations. However a strong *wus* mutant allele occasionally produced flowers that lacked carpels and most stamens (Laux et al., 1996). Based on these observations and on the fact that WUS mRNA can be detected in the center of floral primordial in subsequent steps of their development, the possibility of WUS being a regulator of AG expression was further tested (Lohmann et al., 2001). It was found that WUS indeed regulates AG expression

through regulatory sequences in the second intron of AG gene. Interestingly the binding sites of WUS and LFY on AG are different and mutating site one does not abolish binding of the other factor. Therefore WUS and LFY can bind DNA independently and not cooperatively. In addition there are no evidence that the two proteins interact physically.

Since WUS is expressed in the floral primordial, one could expect that it must also promote stem cell fate and thus indeterminate growth. However basic principal of flower primordial is that they determine their growth producing carpels in the center of the primordium. Therefore there must be a mechanism according to which WUS expression is turn off or its activity reduced in the center of the flowers. It was shown that after WUS activates AG, the latter negatively feeds back and switches of W US expression leading the primordium to determinate growth (Lohmann et al., 2001). This negative feedback loop resembles the CL3/WUS mechanism in the inflorescence meristem, however in the AG/WUS loop the regulation is temporal since both genes are expressed in whorl 4 with WUS first and AG second.

1.3.3.12 Maintaining the inflorescence fate and indeterminate growth of the shoot meristem

By now it is perhaps obvious that after the floral induction, in the SAM there are two different structures, the floral primordia that give flowers and the inflorescence or shoot meristem, responsible for the production of the inflorescence. The floral primordia will be sequentially determined in the flanks of the inflorescence meristem, however the two structures have very different roles and properties. How is then possible to maintain both without compromising spatial and functionally the role of each other?

In paragraph 1.3.3.7, a microarray study focusing on the transcriptional events that define the early steps of floral primordia development was mentioned (Wellmer et al., 2006). According to these data during the first two days of AP1-induced floral development, the majority of the genes are down regulated, while afterwards this is equalised with transcriptional activation of other genes. One could therefore speculate that floral primordia initiate with the repression of genes that define inflorescence fate. After this point, a large number of genes that define the floral fate initiate. A basic view regarding the second part was given in the previous paragraphs. But what are the inflorescence specific genes and how do they guard the shoot meristem from floral invasion?

As mentioned in 1.3.3.3, one of the early studies regarding transcriptional changes in the SAM in response to CO, *TFL1* mRNA was up regulated. TFL1 is a repressor of flowering and

a TFL1 like gene have been identified in other species, including the *Antirrhinum* CEN (**CEN**TRORADIALIS) gene (Bradley et al., 1997). However CEN does not control flowering time only floral determinacy. Shift experiments, in which plants were initially grown in SD, were then transiently exposed into LD and then experienced again SD, showed that *tfl1* mutants responded earlier to the inductive conditions requiring two to three days less exposure to inductive LDs compared to WT plants (Bradley et al., 1997). This suggests that the *tfl1* mutation allowed earlier commitment of the SAM to the floral identity.

Interestingly in *tfl1* mutants apart from being early flowering, the inflorescence also exhibits a terminal flower (Bradley et al., 1997, Conti et al., 2007). All this indicate that *tfl1* mutants are compromised in the ability of restricting the floral identity out of the inflorescence meristem leading to a pattern of determinate growth. It is therefore possible to expect that TFL1 either repress the expression of floral meristem identity genes such as *AP1* and *LFY* or that compromises their ability to express their function. Before describing how this is achieved, it is important to enter into details regarding the expression of TFL1, LFY and AP1 mRNA and proteins after floral induction.

In WT plants TFL1 mRNA is expressed in low levels throughout the vegetative meristem but upon floral induction the expression increases and at the same time is restricted to the very center of the meristem (Bradley et al., 1997). In contrast to the *TFL1* mRNA the TFL1 protein is diffused throughout the IM (**I**nflorescence **M**eristem, Conti et al., 2007). TFL1 protein like FT, is small (20 kDa) and since no diffused pattern was observed the movement is most probably active through the plasmodesmata. This is further supported by the cytoplasmic nature of the protein. However exactly like the mRNA TFL1 protein was excluded from the floral meristems. This observation suggests obstacles of movement and it is fully consisted with the observation that young floral primordia are separated from the SAM (see 1.3.3.1).

This expression pattern is totally complementary to the expression of *LFY*, which at the same time is expressed in floral primordia, being excluded from the center of the SAM. In *tfl1* both LFY mutants demonstrate a shoot conversion of floral primordia. Since TFL1 promotes shoot identity it is possible that in *lfy* mutants TFL1 expression is altered. In *lfy* mutants TFL1 mRNA is de-repressed and although TFL1 has some putative LFY binding sites, direct binding was not shown (Parcy et al., 2002). Interestingly miss-expression of *LFY* mRNA in the L1 zone completely complements the *lfy* phenotype suggesting that the protein of LFY can move inside the floral primordium. Protein immunolocalization proved this hypothesis and showed that LFY protein shows a gradient towards the center of the floral primordium in plants missexpressing LFY in the L1 zone (Sessions et al., 2000). These data are fully in

aggrement with patterns of trafficking through plasmodesmata in young floral primordia upon floral induction (Gisel et al., 1999), indicating that plasmodesmata have active control in determining floral and inflorescence fate.

Microarray data support that TFL1 together with SOC1 and AGL24 is one of the early targets of AP1, which are gradually down-regulated upon floral induction (Wellmer et al., 2006). However this effect of AP1 on TFL1 could be due to LFY since AP1 together with CAL boost LFY expression. We have seen how SOC1 mediates LFY up-regulation via AGL24 and how interactions between AP1 and AGL24 and SVP, recruit the SEU-LUG co-repressor complex that negatively regulates AG expression in the outer whorls where AP1 is expressed. However both AGL24 and SVP promote inflorescence identity therefore their expression eventually must be eliminated from the floral primordia.

Over-expression of SOC1, SVP and AGL24 suggest that AGL24 and SVP show flower-to-shoot transformations with AGL24 having the most severe flower abnormalities (Liu et al., 2007). In addition, flower-to-shoot transformations in *ap1* mutants were reduced by *svp*, *soc1* and *agl24* mutants suggesting that AP1 controls the shoot promoting activity of these genes. AGL24, like SOC1 is expressed in the L1 zone of the floral tunica and throughout the IM. In *ap1* and *ap1cal* plants, AGL24, SOC1 and SVP mRNA are ectopically expressed in the center of the floral meristem (Liu et al., 2007, Yu et al., 2004). Therefore AP1 controls the expression of AGL24, SOC1 and SVP. Indeed this regulation is direct since AP1 binds CarG boxes in all three genes (CHIP; Liu et al., 2007), proving that apart from activator (etc. LFY, AP3) AP1 can be a repressor consistent with the events of early repression and subsequent activation of target genes during floral development.

Summarizing, upon floral induction both LFY, AP1 and TFL1 expression is induced. TFL1 mRNA becomes restricted to the center of the SAM but the protein is present throughout IM. LFY mRNA is present throughout the floral meristem and the protein diffuses exhibiting a gradient towards the IM. Therefore the role of TFL1 in *Arabidopsis* can be summarised in two parts : during vegetative phase TFL1 prevents floral commitment, regulating thus flowering time. After floral induction TFL1 role is to prevent the inflorescence meristem from the floral fate. Mutual exclusion prevents overlap of LFY and AP1 expression domain with TFL1, separating the inflorescence from the floral fate. Since co-regulators mediate LFY action, it is possible that another factor recruits LFY to TFL1. Alternatively LFY might regulate the expression of a TFL1 activator. The fact that LFY is a nuclear protein and TFL1 cytoplasmic supports the second notion. Recently TFL1 was shown to localise intra-cellularely in the vacuole membrane (Sohn Jo et al., 2007). It is therefore possible that TFL1 express its role on

flowering time and floral patterning by leading for storage to the vacuole a LFY activator. AP1 promoters floral fate by directly repressing additional shoot promoter genes such as AGL24, SOC1 and SVP.

1.3.3.13 Summarising

Photoperiodic regulation of flowering is a complicated regulatory cascade that involves an interplay between the leaves and the SAM. Upon inductive long day conditions, the endogenous timing mechanism known as circadian clock shifts the mRNA expression of GI later during the day. This results in an overlap between GI and FKF1 proteins, which interact. The FKF1/GI complex is stabilised by blue light and leads to the degradation of the CO repressor CDF1 on the CO promoter. Upon release of the CDF1 mediated repression, CO mRNA raises and this leads to a biphasic pattern expression, allowing coincidence with light. PHYA and CRY1 and 2 stabilise CO protein and this leads to the up-regulation to downstream targets FT, TSF and SOC1.

The FT protein then becomes the signal that travels through the phloem from the leaves to the SAM. There it interacts with FD and the FT/FD complex activates SOC1. SOC1 activates AGL24 in the IM and AGL24 activates first LFY expression in the floral primordia. The FT/FD complex activates also AP1 expression a bit later. LFY further boosts AP1 expression and vice versa leading to a positive feedback loop, which maintains floral identity. The protein of TFL1, an inflorescence meristem specific gene, is excluded from the floral primordia due to LFY and AP1 prevents invasion of additional inflorescence specific genes SOC1, AGL24 and SVP. TFL1 prevents expansion of the LFY and AP1 expression into the IM, thus maintaining together with SVP and AGL24 its shoot characteristics and indeterminate patterns of growth.

AP1 promotes then the creation of four sepals in the first outer whorl of flowers. LFY interacts with UFO and this allows the expression of AP3 in the second whorl. AP1 and AP3 together form four petals in the second whorl, while AP3 activates PI in the third whorl. In parallel WUS is expressed in the central whorl and with LFY they activate AG expression. AG and PI form eight stamens in the third whorl while AG forms one fused carpel in the center of the flower. The recruitment of the LUG-SEU co-repressor complex by the API-AGL24 and AP1-SVP represses AG in the outer whorls while AG represses AP1 in the central whorl, separating the two domains of the floral primordium. Subsequently the restricted AG negatively regulates WUS expression and eventually terminates flower growth.

1.4 NATURAL VARIATION AS A TOOL IN GENETICS

1.4.1 DESCRIBING NATURAL VARIATION IN ARABIDOPSIS

1.4.1.1 Introduction in natural variation and biogeography of *Arabidopsis*

Natural genetic variation can be a powerful tool when the regulation of a specific trait is coupled with adaptation to the natural world, such as flowering and/or circadian clock (Michael et al., 2003; Dodd et al., 2005). The worldwide spreading of *Arabidopsis thaliana* in the north hemisphere (from the Cape Verde islands near the equator to northern Finland close to the arctic circle regarding latitude and from USA to Japan regarding longitude) coupled with a wide range of habitats that the weed is found to grow (from sandy soils near the see level to frozen steppes in altitudes of 3000m; Hoffmann et al., 2002) and the extensive existing genetic and molecular information about the species, provide a unique opportunity to further unveil the genetic mechanisms regulating various traits. The plant community is fully aware of all this and supports the rapid growth of the field. More than one third of all publications that utilize natural genetic variation in *Arabidopsis* have been published only recently (PubMed data 2007; Weigel and Nordborg, 2005).

1.4.1.2 The use of natural variation in functional genomics

The use of artificially mutagenised methods such EMS and irradiation treatment or T-DNA insertion, greatly facilitates our knowledge regarding the genetic and subsequently molecular regulatory components of the trait of interest. However these approaches, although powerful in detecting major loci controlling the phenotype under study, are less capable to provide insights about additional loci with minor contribution. In *Arabidopsis* estimates of spontaneous mutation rates range from 10^{-4} to 10^{-6} per locus (Byers et al., 2000) leading to an average difference of seven nucleotide per kbp (Mitchell-Olds and Johanna Schmidt, 2006). *Drosophila* exhibits similar rates while humans have ten times less diversity between allele pairs. The use of this naturally occurring variation in the level of the single gene, in combination with QTL mapping largely complements for this loss of knowledge since such approaches not only can detect major regulators but depending on the experimental design, they can also separate their effects from secondary players leading to more complete gene discovery.

In addition natural variation can been seen in the context of the background. From this point of view characterising phenotypically and genetically a large number of additional

genetic backgrounds, multiplies our power of detecting genetic similarities in comparisons between different strains and treatments. Therefore we do not anymore characterise major regulators but the whole genetic architecture of the trait of interest as nature shapes it, no matter how complex that might be (Tonsor et al., 2005). This of course means that the use of artificially induced mutations should be complemented of even better combined rather than replaced. Indeed using these classical approaches in new backgrounds alleviates the problem of phenotypic loss of a mutated gene due to epistasis in the laboratory strains Ler and Col-0.

It is therefore essential to separate the terms of gene function and gene effect. The function will be the same in every genetic background but the effect might be masked (Tonsor et al., 2005). This separation is due to the multilocus genetic control of most traits in combination to the complicated relationships of genes. Issues such as dominance (the intra-genic cases of interactions) and epistasis (the inter-genic aspects of genic function), together with pleotropy (the multiple regulatory potential of genes) can change between different backgrounds and that is why natural variation can help describing the genetic architecture of complex traits

Both the genic and background views of natural variation are connected through natural selection, with fitness. And that is the real power of the approach since it conditionally allows conclusions to be made regarding the adaptive significance of the results. This ecological touch is not relevant only in the native environment of each ecotype but also and perhaps more interestingly in novel niches. Expanding and sharing the above knowledge also to the molecular level, allows steps towards an integrated understand of plant biology to be made, consistent with the concepts of the *Arabidopsis* 2010 project.

1.4.1.3 Structure of molecular diversity underlying natural variation

Several mechanisms can influence nucleotide diversity of a genome. Recombination directly affects it and a positive correlation has been observed not only in *Arabidopsis* but also in other species (Nordborg et al., 2005; Mitchell-Olds and Johanna Schmidt, 2006; Borevitz et al., 2007). Areas of low recombination rates for example pericentromeric regions, are not severely influenced by strong selection and allow large haplotypes to be fixed. This allows a large portion of molecular diversity to be identified. In contrast, “genic” regions are more prone to selection forces reducing diversity.

However that is not a rule without exceptions since the type of selection greatly determines the pattern of diversity. For example “genic” sites under balancing selection are more diverse by definition compared to “ genic” sites, gone through selective sweeps. Typical examples of this are the R genes (genes conferring resistance to pathogens). Although these genes are

advantageous for the carrier their activity leads to a higher energy cost. Inactive haplotypes show approximately 8% larger biomass compared to active ones. A direct consequence of that in *Arabidopsis* is a clear pattern of balancing selection in these genes. Interestingly there is a clear cut positive correlation between the occurrence of R clusters in the genome and the overall levels of nucleotide diversity. This relationship is obvious in the lower arm of chromosome 5, around the PHYC region. This area has a big cluster of R genes and its levels of nucleotide diversity among the most elevated, compared to the rest of the "genic" regions (Borevitz et al., 2007, Clark et al., 2007).

The above introduce some important features of nucleotide diversity. There is a structure regarding the distribution of it in the genome. Natural selection strongly influences natural diversity both by increasing and decreasing it. Therefore common haplotype frequencies of specific mutations can indicate their adaptive importance. Low frequencies suggest either that the region has high recombination rates (resulting in high linkage equilibrium, disrupting haplotype preservation) or that the region under goes balancing selection, with the opposite suggesting selective sweeps of advantageous mutations (Borevitz et al., 2007). These properties of natural variation, better known as the analysis of Linkage Disequilibrium, are the hallmark of genetic mapping in human population genetics and recently attempts were made to introduce this approach into *Arabidopsis*. With an average LD decay of roughly 50 kb in the species (Nordborg et al., 2005), LD has the potential of fine-mapping mutations relatively quickly, reducing the labour. However the effects, advantages and caveats of LD are not the scope of this thesis and will not be analysed further.

Apart from the genome wide pattern, intra-and inter-genic structure of nucleotide diversity, expressed as SNPs (**S**ingle **N**ucleotide **P**olymorphisms) can also be identified, again as an outcome of selection. At relatively moderate false discovery rates, coding and inter-genic regions share almost identical densities, with the inter-genic increasing with false discovery rate, and thus being gradually separated. This suggests that mutations in cis regulatory elements, could be more frequent than coding regions with interesting consequences (see below, Clark et al., 2007). Diversity per accession within introns and UTRs is lower than diversity in intergenic and CDS regions, while overall diversity in the genome is much higher, consistent with the "genic" recombination rates and molecular diversity properties of these sites. In between-accessions comparisons however, there is a significantly much higher diversity in coding regions versus inter-genic and intronic, suggesting that major source of functional natural variation could be activity vs expression attenuation (Clark et al., 2007). Again here R genes show by far the biggest rates of large effect diversity and transcription

factors such a bHLH and MYB the lowest (comparable to the highly conserved ribosomal sequences). Comparisons of ESTs between different accessions reveal that there is a higher ratio of silent to replacement polymorphisms when compared to randomly chosen loci, consistent with the strong selective pressure on the first (Schmidt et al., 2003).

1.4.1.4 Occurrence of environmental clines ? A view through population structure and colonization of *Arabidopsis*

Molecular diversity in coding regions may underlie functional attenuation. It was mentioned in the beginning that natural genetic variation, expressed as molecular sequence diversity, could lead to adaptation in specific niches characterised by strong environmental influence. Such adaptation can create environmental clines, arranging haplotypes according to geography. In *Arabidopis*, clines in the vernalization requirement (Stinchcombe et al., 2004) and light sensitivity, expressed as hypocotyls lengths (Maloof et al., 2001; Stenoen et al., 2002) have been identified however the strength of these correlations expressed as R^2, is relatively weak. In *Drosophila*, a typical example of latitudinal clines is the circadian clock (more below).

There are certain criteria that must be fulfilled in order to observe such correlations. First the geographical distribution must exposure the species under investigation in a wide range of habitats. This is the case both in *Arabidopsis* and *Drosophila* (Fig 6 and 7). However, complex genetic architectures with pleotropic and epistatic components complicate a clear-cut observation. The pattern of selection on each trait is also important in determining such associations. Traits undergoing balancing selection have a higher potential of such clinal appearance compared to traits that experienced selective sweeps in the recent past. The pattern of sexual reproduction is also of importance with populations of selfing species have a higher chance a being genetically distinct between different locations versus populations of out crossers. Although a selfing species, *Arabidopsis* has extensive variability within populations that equals variability between populations from different regions (Nordborg et al., 2005).

Paradoxically such admixture does not eliminate population structure. Accessions from Asia are considerably distinct genetically from European accessions with many studies confirming this (Sharbel et al., 2000; Nordberg et al., 2005; Schmid et al., 2006). Within Europe the picture is very complicated. In the south, namely Iberian Peninsula, Italy and the Balkans, there seems to be distinct patterns of variability compared to the rest of Europe. The Spanish accessions are significantly different not only from the Asian but also the central and

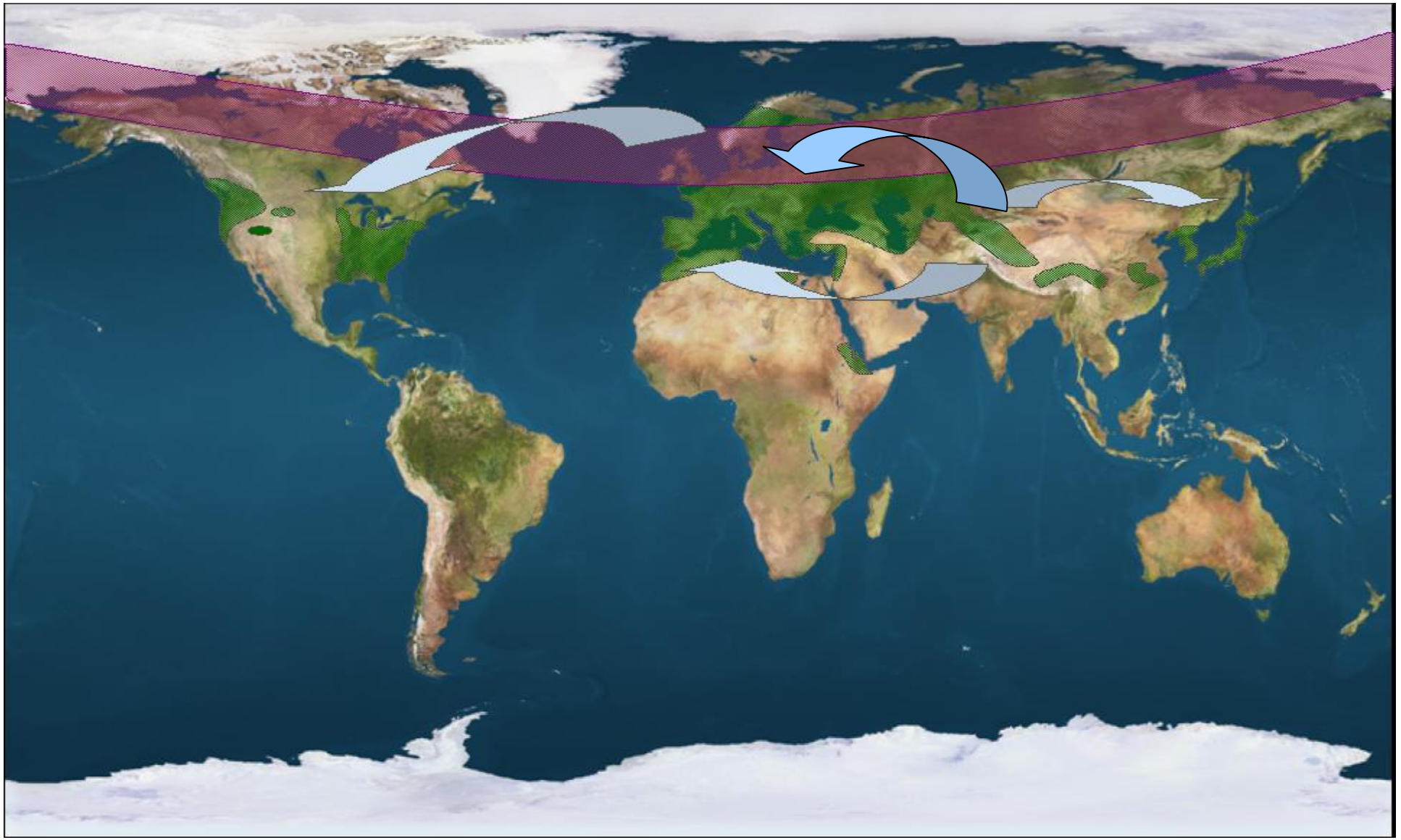


Figure 6 : Migration and geographical distribution of *Arabidopsis*. The species originates from central west Asia. From there it expanded towards Europe following two independent routes, one via the south and a second via the east. From there migration eventually occurred towards USA and Japan. Green areas indicate the current geographical distribution of the species, while red shaded areas indicate the geographical zone, which experiences LD 16 h in nature.

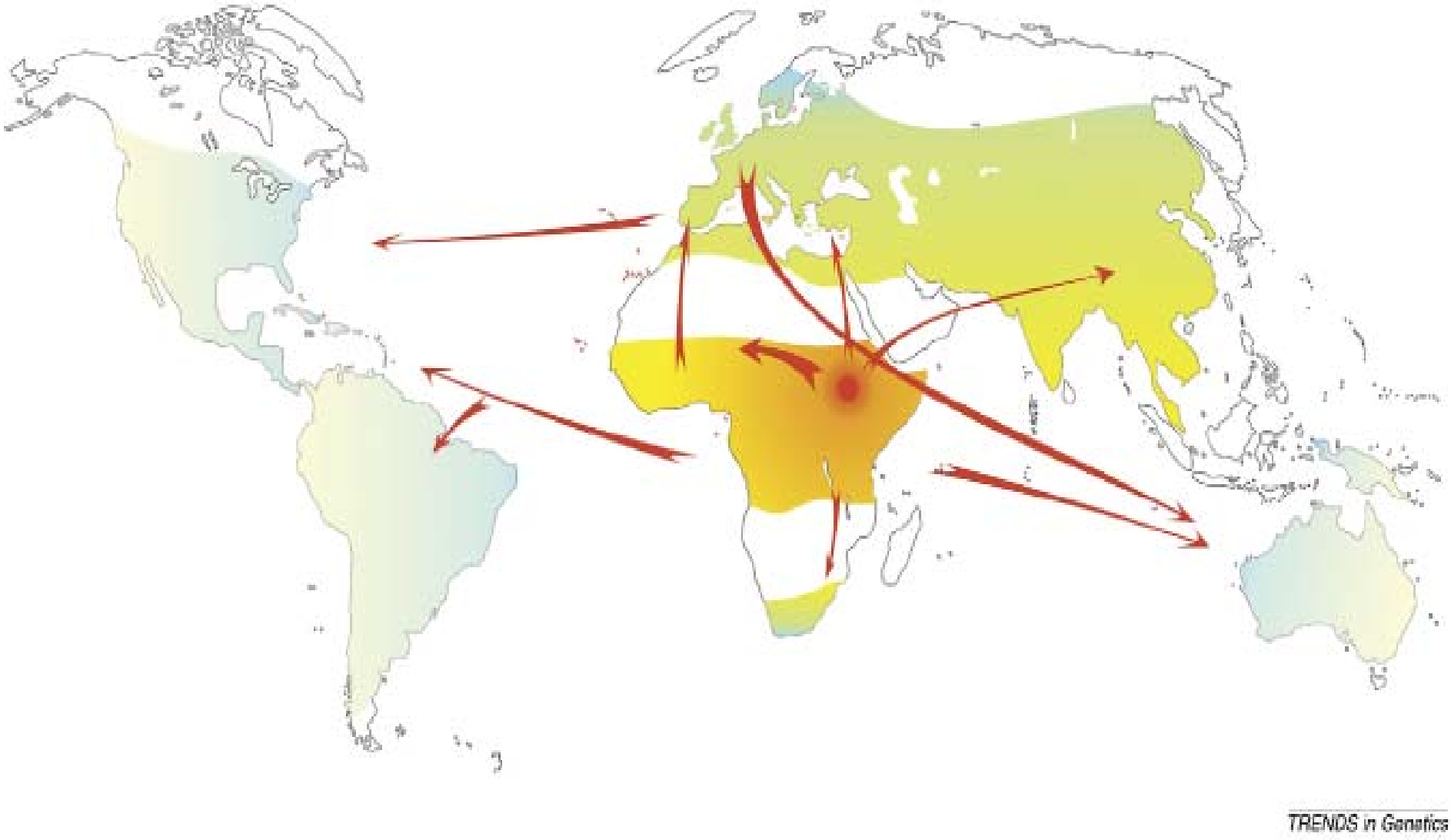


Figure 7 : Migration and geographical distribution of *Drosophila*. The species originates from central east Africa. From there it expanded towards Europe and Asia. Both European and African migration waves colonised USA and Australia. (Kyriacou et al., 2007).

Northeast European, but this does not mean that the rest of the south does not follow this trend. Most probably this is an outcome of accessions sampling, which is extensive in Spain and limited in Balkans. In addition, even within Spain, large variability can be observed (Carlos Alonso-Blanco, personal communication). Scandinavian accessions are also different from the central European collections, with extensive haplotype sharing (Nordborg et al., 2005) being evident within Sweden and Norway (Nordborg et al., 2005; Stenoen et al., 2005). The recently introduced to the USA, accessions exhibit haplotype sharing between populations, and overall similarities to British accessions (most similar to Spanish), suggesting that colonization of the New World from a few pioneer individuals of *Arabidopsis*, was mainly an outcome of human migration. In the central Europe, admixture is more than anywhere else evident, but small evidence about a Central-west and Spanish like to Central-east and Asian like trend can be supported (Nordborg et al., 2005; Sharbel et al., 2000).

According to the standing theory, *Arabidopsis* originates from the Southeast Asia. From there is expanded to Europe following two routes, one from the East through Rusia and one from the South. This created a first diversity gradient that eventually separated the Asian accessions from the Europeans. During the last glacial period in the Pleistocene, ice covered the Northern and central Europe, forcing many species including *Arabidopsis* to retreat to southern refugia such as the Iberian Peninsula. In the post-glacial period, recolonization of central Europe took place both from the south Iberian Peninsula, to the central west and from Russia to central east. These migration waves in combination with human interference created significant admixture in the central part of Europe. Much later, human migration from the British Isles allowed colonization of the new world.

1.4.2.5 Natural variation and clines in the circadian clock: *Drosophila* as a case study

Like in *Arabidopsis* and the animals, the circadian clock of *Drosophila* consists of a series of positive and negative feedback loops (Kyriacou et al., 2007). The expression of the negative transcriptional regulators TIMELESS (TIM) and PERIOD (PER), which make up the PER–TIM loop, and CLOCKWORK ORANGE (CWO) is mediated through binding of their E-boxes (hexameric sequence CACGTG) by the positive transcription factors CLOCK (CLK) and CYCLE (CYC), which form a heterodimer. The negative regulators feed back, after a time delay, to inhibit CLK and CYC activity. These delays are imposed by the kinases SHAGGY (SGG), casein kinase 2 (CK2) and DOUBLETIME (DBT) and by phosphatases including protein phosphatases 2A (PP2A) and 1 (PP1), which modulate the stability of the

negative regulators. A second feedback loop interlocks with the PER–TIM loop. Negative regulation of CLK occurs via by VRILLE (VRI) and positive regulation by PAR domain protein 1e (PDP1e). However, more recent data have shed some doubt on the role of PDP1e in this process.

Cytoplasmic PER and TIM levels start to rise at the beginning of the night, and late at night PER and TIM translocate to the nucleus where they negatively regulate (through CLOCK-CYCLE) their own genes. At dawn, the blue-light circadian photoreceptor cryptochrome (CRY) is activated and interacts with TIM. This interaction stimulates TIM phosphorylation, which recruits the F-box protein JETLAG (JET), a component of a SKP1-CULLIN1-F-box (SCF)-E3 ubiquitin ligase, which polyubiquitylates TIM, thereby marking it for degradation. Under conditions of constant darkness, TIM is degraded by the proteasome at subjective dawn by a CRY-independent pathway. PER monomers are strong transcriptional repressors, but in the absence of TIM, PER is vulnerable to phosphorylation by DBT kinase, and phosphorylated PER is recognized by the F-box protein SUPERNUMARY-LIMBS (SLIMB) and marked for proteasomal degradation. The negative regulation of per and tim is now lifted, and CLK–CYC begins the cycle again by binding the per and tim E-boxes. In the brain, the PER and TIM proteins are expressed in 75 neurons on each side of the brain, as well as in glia and photoreceptors. Of the two main groups of neurons, the dorsal and the lateral neurons, the lateral neurons are indispensable for rhythmic behavioral expression. However, many peripheral tissues of the fly are also rhythmic, and in peripheral tissues, CRY seems to function as a clock component and a photoreceptor.

There are two different types of robust clines that can be identified in the *Drosophila* clock An uninterrupted stretch of alternating threonine-glycine (Thr-Gly) pairs is encoded within the per gene. *D. melanogaster* natural populations in Europe and North Africa differ in the length of Thr-Gly region: two major alleles, *per* $(Thr\text{-}Gly)_{17}$ and *per* $(Thr\text{-}Gly)_{20}$, account for 90% of the variation, *per* $(Thr\text{-}Gly)_{23}$ making up to 8%, and *per* $(Thr\text{-}Gly)_{14}$ at 1%. All variants coalescing back to a *per* $(Thr\text{-}Gly)_{23}$ sequence, while European and North African *per* sequences suggested that balancing selection acts on the Thr-Gly region The northern *per* $(Thr\text{-}Gly)_{20}$ allele provided the most thermally stable periods, with an endogenous period that barely deviated from 15–20 min less than 24 h. The *per* $(Thr\text{-}Gly)_{17}$ allele provided a more precise 24-h free-running period at 29 ºC, but the period shortened significantly in cooler temperatures. Thus, the 'temperature compensated' *per* $(Thr\text{-}Gly)_{20}$ allele is probably better adapted in the colder and more thermally variable higher latitudes, whereas *per* $(Thr\text{-}Gly)_{17}$ is

suited to the warmer Mediterranean region: a potential outcome of balancing selection scenario that is consistent with the statistical analysis of Thr-Gly sequences

A more photoperiodic aspect of clines refers to polymorphisms in the TIM gene. The 5' coding region of tim in *D. melanogaster* is characterized by a single nucleotide insertion. The insertion haplotype (*ls-tim*) has two in-frame initiation codons, producing a long TIM isoform from the upstream methionine (L-TIM_{1421}) and a shorter isoform from the downstream ATG. By contrast, the major products of the deletion haplotype are the short isoform, S-TIM_{1398}, and a putative truncated 19-residue peptide, S-TIM_{19}, from the upstream ATG. Natural European *D. melanogaster* populations show a latitudinal cline for this polymorphism, with *ls-tim* being most frequent in southeastern Italy, whereas *s-tim* prevails in northern Europe Unlike the per polymorphism, the TIM variants do not show any systematic differences in temperature compensation. Because TIM is the light-sensitive clock molecule, seasonal day length, which co-varies with latitude, is a potential selective agent. In light pulse experiments, circadian photosensitivity is compromised in homozygous *ls-tim* flies, and this finding was recapitulated using transformants that mimicked the natural variants In a yeast two-hybrid assay, CRY interacted more strongly with S-TIM than L-TIM. If this change in TIM–CRY dynamics is reflected in vivo, *ls-tim* flies should have more stable TIM products. Thus, the L-TIM isoform likely provides greater overall TIM stability and causes the attenuated circadian and seasonal photoresponses of the ls-tim genotype.

1.4.2 NATURAL VARIATION AND MAP-BASED CLONING IN THE POST-GENOME ERA

1.4.2.1 Historical aspects

Great advance has been conducted the last years in the available tools for the isolation of genes controlling complex traits in *Arabidopsis*. Historically the first attempts to link phenotypes with genetic characters involved the use of morphological markers. The limited availability and application of such markers for various traits together with their reduced objectiveness quickly became problems. With the break of the molecular era, it became obvious that other, DNA based approaches should be used for direct phenotype-genotype linking. The first approaches included southern based techniques such as RFLPs since PCR was not invented yet. After the discovery of PCR, amplification based approaches such as AFLPs flourished. Sequencing and better understanding of the genome structure and nature allowed microsatelites to enter into the picture. However all these molecular markers were not abundant in high densities or in genomic areas of will and also were not anchored against any physical position.

A true revolution in biology was whole genome sequencing (The Arabidopsis genome sequencing, 2000). In functional genetics not only it allowed more and new markers to be discovered but thanks to it, a physical map was now available together with new exciting methods for genotyping. Today SNPs posses central role in the construction of dense linkage maps.

1.4.2.2 Genetic marker identification and availability in *Arabidopsis* accessions

With the complete genome of Col-0 sequenced and the position of the genetic markers known, it was now possible to simply anchor QTL regions (Quantitative Trait Loci, genomic locations harbouring genes which control quantitative traits) to the physical map of Col-0 and to scan for candidate genes. Re-sequencing of genes or genomic regions of other *Arabidopsis* accessions, although useful in a local scale was not enough when new markers were needed genome wide. In a second sequencing approach shotgun sequencing was conducted in Ler. Approximately 700.000 500 bp sequences were isolated, representing 92.1 Mbp of total assembled genomic sequence (The Cereon project, Jander et al ., 2002). By comparing the two genomes more than 37.000 SNPs, 18.000 small InDels (<100 bp) and nearly more than 1000 large InDels (>100 kbs) were discovered in a total sum of 56.000 polymorphisms.

These markers are distributed equally in the five chromosomes and within exons on average there is one SNP every 3.1 kb and within introns every 2.2. This likely reflects the higher preservation of mutations in the non-conserved introns. Regarding InDels cases up to 38 kb were identified with an average of 175 bp. Most Indels as expected were found in non coding regions. These markers can be used in order to test whether some of them are polymorphic between different accessions and subsequently used in mapping programs.

Expanding this further, other studies focused on directly identifying markers by sequencing large numbers of *Arabidopsis* accessions. A first attempt included generation of markers from 10.000 EST (**E**xpressed **S**equence **T**ags) and 600 STS (**S**equence **T**agged **S**ites) in six and 12 accessions respectively (Schmid et al., 2003). Over 4.000 SNPs were identified, with the majority spanning coding regions and one third of them resulting in aminoacid changes. And this expands further and further with the sequencing of nearly 1000 fragments from 96 accessions (Nordborg et al., 2005) resulting into 17.000 SNPs and InDels. Apart from sequencing nucleotide differences can also be identified by mis-hybridization. Here the use of high-throughput techniques such as array hybridisation is valuable and leads to the discovery of SFPs (**S**ingle **F**eature **P**olymorphisms) due to non-hybridisation on the chip due to the presence of an SNP against the Col-0 sequence. Nearly 20.000 SFPs are currently available from 14 accessions (Borevitz et al., 2005).

1.4.2.3 Methods for high-throughput genotyping

Classical methods for genotyping included a PCR reaction, followed by size separation of the amplicons on agarose gels. In the case of SNP based markers, allele specific restriction digestion is performed between the PCR and the gel separation, converting thus the DNA polymorphism into size difference. These standard approaches are reliable and easy technically but become labor intensive in the case of a modern mapping project. More specifically for an average mapping project of 200 individuals and 10 markers per chromosome, 10.000 PCRs are needed and approximately 100 gels of 100 samples each need to be casted. All this is time consuming and labor intensive, without including PCR set-up and data verification.

Luckily modern high-throughput methods are today available in a continuously decreasing cost. One such method is DHPLC (**D**enaturing **H**igh-**P**erformance **L**iquid **C**hromatography). The method allows the detection of SNPs in fragments as large as 1 kb and is highly automated (Steinmetz et al., 2000). The method is based on denaturation kinetics of amplicons based on their DNA differences. The final results are displayed in the form of

chromatograph with peaks corresponding to the DNA differences, being determined by the number, position and chemical nature of the differences. I have also developed a novel and easy to use method for rapid, high-throughput and agarose-free genotyping. The method is called HmMPP (**H**ybridization-**m**ismatch **M**elting **P**oint **P**rofiling) and it has been tested successfully in verifying genotyping results and in fine-mapping in the projects of the current manuscript. At the moment is it applicable to InDels and SSRs (microsatelites) but it will be further modified with the hope to cover SNPs as well.

Another method was already discussed in the previous paragraph. Tiling or expression arrays can be used in order to hybridise genomic DNA Signatures from mis-hybridisation can be interpreted as DNA variations. Moderate levels of false positive identification (FPI) due to poor hybridisation regardless of DNA sequence, can be counterbalanced by the rich abundance of polymorphic signatures that can be obtained. When the SNP sequence is know, Taqman markers are also another option, although their application for mapping is limited due to high cost. Pyrosequencing instead is another method, suited more for SNPs but not large deletions at a moderate cost. For example typing 100 samples with 50 SNPs with restriction digestions made with 30 enzymes (assuming some will be designed for application between different markers) requires an average cost of 3500-4500 € only for the enzymes. With pyrosequencing the same project can be covered with 3000 € and with much less effort.

1.4.2.4 QTL and fine mapping approaches in *Arabidopsis*

QTL mapping is a method which allows the identification and subsequent isolation of genes controlling quantitative traits. A typical mapping project has three general requirements : a mapping population, a linkage map and the statistical approaches that will combine them.

The mapping population is the focus of the study. Consists of individuals which segregate for the trait/s of interest under the desired experimental conditions. The first step in creating such a material is a cross between two parents, the phenotype of which differs in the trait/s of interest. In *Arabidopsis* this is usually followed by selfing and allowing the F2 progeny to segregate. Upon segregation, the genetic mosaics of the F2 individuals will create a quantitative distribution of the trait/s reflecting the recombination and subsequent uncoupling of genes which control the trait. Phenotypic scoring of this population allows the measurement of this impact.

However it is essential to have also a molecular insight of the recombination events. This will indicate the pattern of genic uncoupling and will allow the estimation of its consequences on the observed phenotypes of the individuals. For this reason evenly distributed across the

genome, genetic markers that distinguish the alleles of the two parents across the genome are needed. One these markers are obtained genotyping of each individual will distribute them on each chromosome based on their recombination frequencies in the population. Statistically combining the phenotypic data with the genotypic information will eventually i) indicate in which areas of the genome allelic differences from the two parents have the biggest impact on the phenotype ii) measure this impact iii) estimate the statistical significance of its occurrence after comparison to a randomly generated data set, a process known as genotype permutation iv) estimate intra-genic only interactions such as dominance across the genome.

Subsequent selfing of the F2 individuals for 6-7 generations, through SSD (**S**ingle **S**eed **D**ecent) will result in the production of RILs (**R**ecombinant **I**nbreed **L**ines). The advantage of the RILs, also know as immortalised F2s, is that the subsequent selfing expanded homozygosity in 99 % of the genome and thus, once genotyped the RIL genetic mosaic is known no further genotyping is needed. This fixed genotype can therefore used in any mapping project and the only thing that is needed from now on is to phenotype the RILs against the desired trait. RIL mapping is also more powerful statistically than F2 mapping since each genotype is not represented by an individual plant but by a population. Among the caveats are the labour of creating the RIL population, and the time that is needed to reach the desired number of generations (usually 7-9). In addition, often the final genotyping is not with dense genetic maps making it inaccurate. Indeed several times partial additional genotyping is required to complement the initial genotyping into specific genomic areas of interest.

However even in mapping projects with dense genetic maps (>100 markers) and big populations (>200 individuals), the outcome of the QTL mapping places the genes of interest in large genomic regions, together with hundreds of other genes. These preliminary map positions need to be resolved further in a process known as fine mapping. This will eventually indicate the responsible genes and allow their isolation. In cases in which many QTL have been identified, splitting of the QTL will allow the independent study of their phenotypic impact. Usually the method of splitting these QTL is by backcrossing selected F2 individuals that contain alleles causing the desired phenotype around the area of the QTL, to the reciprocal parent. This genetic material, known as NILs (**N**ear **I**sogenic **L**ines) will allow background purification and will create individuals which differ from the reciprocal parent only at the positions of the QTL, thus being nearly isogenic, nearly the same.

The size of the introgressed region varies in literature, from 7-10 down to 5 cM and depends on the location of the QTL, the number of backcrosses and the size of the segregating population after each backcross. Alternatively one can keep backcrossing the F1 hybrids after

each BC independently of their segregation, but according to their genotypes across the introgressed region and in the rest of the genome. Once a NIL is available, this is used to i) confirm the QTL mapping ii) isolate the gene. This NIL will be backcrossed to the reciprocal parent and a very large population will be used in order to increase the resolution of the introgressed region. Many markers (ideally one per 100 kb or even less to start with) spanning the introgression must be identified while the size of the population depends on the recombination frequency of the region, but usually requires thousands of segregants. In humans due to ethical reasons, such genetic approaches of course cannot be applied and therefore there genetic mapping takes place by studying the genome wide haplotype structure, estimated by the association pattern of nearby markers, in global samples of populations, a method known as Linkage Disequilibrium

1.4.2.5 QTL and fine mapping approaches in *Droshophila*

In order to provide a more complete picture regarding the isolation of quantitative genes, a reference to genetic approaches used in other organisms will be made. Since *Drosophila* led the way on QTL mapping it was selected as the focus of this paragraph. A large toolbox of visible markers such as the eyFLP transgene causing eye-color mosaicism, are available in Drosophila (Berger et al., 2001). A very common method for fine-mapping in flies is whole or partial chromosome (or better autosome) substitutions, known as CSSLs (**C**hromosome **S**egment **S**ubstitution **L**ines). This approach was introduced to flies 40 years before its application to mammals (Mackay T, 2001). In addition, in *Drosophila* for the first time introgressions allowed sub-chromosomal mapping.

Like in *Arabidopsis*, also in *Droshophila* it is common to isolate the QTL by introgression to the reciprocal, to the desired phenotype, parent, producing NILs. Such cases where successfully led to the identification of QTL responsible for bristle number (Mackay T, 2001). However fine-scale mapping after this point differs dramatically in flies, following two methods deficiency and complementation mapping, which circumvent the issue of limited recombinations in a small genomic area (Mackay T., 2001). Deficiency stocks are chromosomes missing a stretch of genomic sequence. The deficiency stocks have overlapping breakpoints covering thus the genome. By testing complementation of the mutant phenotype by the series of the deficiency stocks, the location can be pinpointed quickly and without much effort. Quantitative complementation is analogous to deficiency and suggests that the alleles of the QTL are tested from complementation of a mutant series instead of complementation of absent stretches (Mackay T., 2001).

Currently large surveys of genetic markers make QTL mapping powerful in the species and cases of fine-mapping QTL down to the intra-genic resolution of 2 kb are reported (Berger et al., 2001). These studies utilise the existence of visible markers FRT and ER in order to define recombination events, by the mosaic eye, which show the desirable phenotype and therefore having the mutation. Once a genomic region is obtained, PCR screening and amplicon sequencing result in the identification of SNPs which can be used as markers. With polymorphisms being abundant in the genome, this on average results in rich yields.

1.4.2.6 eQTL : welcoming the new era of mapping

A usual caveat of any mapping approach is the isolation of the corresponding genes. Even in humans this is a very slow process and very often, identified QTL lack a molecular explanation (Glazier et al., 2002; Korstanje and Paigen, 2002). The molecular cause of the phenotypic differences is usually associated with two things: i) change of expression levels of a gene ii) change in the functionality of the gene. The change in the expression of the gene can be further divided in cis- and trans specific effects. Cis effects often include mutations residing in the promoter region of the gene of interest, while trans effect resides in the coding sequence or promoter region of a regulator of that gene. Therefore gene effect can be treated in a transcriptional level. In classical mapping approaches, phenotypes are treated without any molecular insights regarding the mechanistics of the effect. Even in metabolic QTL mapping, measurements of metabolites although molecular, are describing the outcome and not the cause. A promising bridge between the effect and the mechanism is the eQTL mapping.

In eQTL mapping the phenotype are global-wide transcriptional levels (Hansen et al., 2008). Depending on the type of the genetic material that is used eQTL mapping can be further divided into standard eQTL and introgression eQTL mapping. In standard eQTL mapping, RILs are used in order to estimate the transcriptional effects of the presence of the two parental alleles throughout the genome. Introgression eQTL mapping requires NILs, and here the global transcriptional effects from the introgression of alleles of one parent, in an otherwise homozygous genome consisting of alleles from the reciprocal parent, are studied. In other terms RIL eQTL mapping is more general throughout the genome while NIL eQTL mapping focuses in the global effect of a small region.

The advantage of eQTL mapping is that it brings cause and effect together. What is measured as a phenotype is at the same time what is causing the phenotype : changes in expression. eQTL phenotyping can be combined also with standard phenotyping to further reflect the effects of the eQTL in the physiology of the plant (assuming of course that the

physiological phenotype is caused by transcriptional differences, which is not always the case). In addition introgressed eQTL mapping is usually following a standard QTL mapping program linking genetics with transcription. According to their nature, eQTL can be divided in two parts : cis and trans. Cis-eQTL are caused but mutations in the promoter regions of the target gene, while trans-QTL are caused by mutation in the CDS or the promoter region of a trans regulator of the target gene (Hansen et al., 2008).

These two categories can be separated by the location of the polymorphism. In the polymorphism is located in the same region with the gene, the expression levels of which are changed, then most probably this is a cis eQTL or a case of a linked trans eQTL. The latter case is not unusual in mammalian eQTL programs in which regulatory sequences of genes can be distant to the gene position. On the contrary if the location of the polymorphism is located in a very distant (e.g other chromosome of over the centromer etc) to the gene whose expression levels are measured position, then the eQTL is definitely trans. An additional property of trans eQTL that separates them form cis is their ability to regulate more than one genes, resulting into trans regulatory eQTL circuits. There are two possible ways of constracting such networks : the *a priori* and the *posteriori* anylysis. The first requires previous at least partial knowledge of the hierarchy within a regulatory pathway, while the second does not and it can be applied in the identification of novel transcripionaly regulatory pathways. Nevertheless, the effect of cis QTL is much stronger than the effect of trans QTL since it is direct. Trans regulation of a target by many genes can buffer the effects caused by mutating only one of them (the observed trans eQTL).

In *Arabidopsis* two reports have been made on eQTL mapping. The first included the Bay x Sha (West et al., 2007) while the latter the Ler x Cvi RIL population (Keurentjes et al., 2007). The first approach reported the identification of more than 36.000 eQTL wile the second 4000. In the first case the majority (86 %) of the eQTL were in trans while in the second trans and cis eQTL were equally abundant. These differences could be attributed to the population size and replication number of the hybridizations, which differ between the two studies. For example the bigger population size of the Bay x Sha experiment can increase the power of detecting the usually small effect trans eQTL. Generally, the ratio of the trans to cis eQTL is related to the ratio of inter-genic vs CDS occurrence of molecular diversity (see § 1.4.1.2). eQTL mapping programs were conducted in other organisms as well. Results from budding yeast confirm the stronger effect of cis eQTL but also the widespread nature of trans eQTL suggesting the existence of transcriptional regulatory networks and confirming the complex nature of transcriptional regulation even in single cell organisms (Brem et al., 2002). In

humans interestingly most eQTL that were identified were in trans. Sever hotspots with effects spread over many different chromosomes were detected suggesting that large part of the transcriptional regulation in humans occurs via robust regulatory networks (Morley et al., 2004). However it is important to remember that mapping approaches of any type in humans are not using RILs but populations and refer more to association mapping. These pioneering studies are the first yet crucial step towards the elucidation of the transcriptional aspects of natural variation in diverse species.

Aim of this study

This work utilizes the extensive natural genetic variation in *Arabidopsis thaliana* in multiple ways in order to further elucidate the complex pattern of photoperiodic regulation of flowering time. More specifically I am interested in understanding the genetic and molecular mechanisms by which *Arabidopsis* plants discriminate between nearly identical photoperiods in terms of flowering time. Additionally, I am evaluating the level of contribution that is being conferred by the circadian clock, through natural genetic variation in the expression of the *GIGANTEA* gene. Furthermore I proceed in identifying novel loci that are responsible for a) efficient discrimination between different day lengths in terms of flowering time b) causing variation in the expression pattern of *GI* under different photoperiods.

Several reasons justify the selection of natural variation as the tool to address the above. Progress towards understanding the complex and multigenetic nature of flowering regulation, which frequently involves epistatic relationships, pleotropical effects and redundancy between the members of the five pathways, was accomplished almost exclusively by creating artificial variation through mutagenesis, T-DNA insertion events and activation tagging in two genetic backgrounds Columbia-0 (Col-0) and Landsberg erecta (La*er*, Ler). Those approaches identified the major genes that regulate flowering time, but were less powerful in detecting other genes whose contribution is masked by those key elements. Additionally the use of mainly two *Arabidopsis* accessions sets limits in the utilization of the extensive phenotypic and genetic variation, which is observed in the species, in order to understand the important epistatic and pleotropic relationships among already characterised genes.

All of the studies referring to flowering time were performed in extreme photoperiods of 16 h for long days or 8 h for short days. In nature, such photoperiodic conditions are restricted to a limited number of northern locations (Fig 6). Interestingly these locations are in the periphery of the geographical distribution of the species, suggesting that environmental conditions might shape genetic diversity and population structure differently in these niches. For example, a stronger need for a vernalization requirement and higher sensitivity to white light would be more obvious in these populations versus the southern ones (Stinchcombe et al 2002; Maloof et al., 2001). Regarding the study of vernalization, mainly genotypes with absolute and not mild requirements for vernalization were selected for further analysis. In the real world such genotypes are restricted only to northern latitudes. However, several natural strains of *Arabidopsis*, mainly from the southern latitudes, demonstrate a mild requirement for vernalization. In addition, the majority of the mutations were isolated mainly in either

Columbia or Landsberg, thus restricting the search for epistatic interactions in these two genetic backgrounds. It is therefore possible that additional players together with critical cross-talk between the two pathways were masked, since such extreme conditions would mainly benefit one of the two pathways over the other

The same set of accessions was transformed with a promoter fusion of *GIGANTEA* to the firefly luciferase gene in order to describe molecular variation in the expression of this circadian clock component. Real-time bioluminescence was then used to monitor the expression of the *GI::LUC* in the resulted transformed lines. Following the expression under the same photoperiods as for flowering time, allows us to draw conclusions regarding the effect of the circadian clock in photoperiod discrimination over a wide range of day lengths and to identify natural variation in the circadian parameters of the expression of the gene. By monitoring the expression under constant darkness, I distinguish the effect of light in the expression of the gene and detected natural variation, which is strictly under circadian control and not affected by light. Effects of the changing day length into the free running period under constant conditions are also investigated with the aim to identify and further characterise not only natural variation and the genes that control it but also to understand the mechanisms through which these different regulatory parameters shape circadian rhythmicity.

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Plant material (Table 1)

Abbreviated name	Full name	Country of origin	Habitat	latitude
		Accessions		
Aa-0	Aua/Rhon	Germany	field border	N51
Ak-1	Achkarren	Germany	vineyard	N48
An-2	Antwerpen	Belgium	Botanic Garden	N51
Ang-0	Angleur	Belgium		N50
Ang-1	Angleur	Belgium		N50
Ba-1	Blackmount	United Kingdom	roadside 300m	N52
Bay-0	Bayreuth	Germany	fallow land	N49
Bch-1	Buchen	Germany	deposited sand	N53
Bd-0	Berlin/Dahlem	Germany	Botanic Garden	N52
Be-0	Bensheim	Germany		N49
Be-1	Bensheim	Germany		N49
Bl-1	Bologna	Italy	Botanic Garden	N44
Bla-1	Blanes	Spain		N41
Bla-3	Blanes	Spain		N41
Br-0	Brunn	Czechoslovakia		N49
Bs-1	Basel	Switzerland	Botanic Garden	N47
Bs-2	Basel	Switzerland	Botanic Garden	N47
Bs-5	Basel	Switzerland	Botanic Garden	N47
Bsch-0	Buchschlag	Germany	near horticultural nursery	N50
Bsch-2	Buchschlag	Germany	near a rail line	N50
Bu-0	Burghaun	Germany		N50
Bu-2	Burghaun	Germany		N50
Bur-0	Burren	Ireland	wall by roadside	N52
Cen-0	Caen	France	Botanic Gardens	N49
Chi-0	Chisdra	Russia	river slope	N54
Cl-0	Cl			
Co	Coimbra	Portugal	Botanical Garden	N40
Co-2	Coimbra	Portugal	Botanic Garden	N40
Col-0	Columbia	USA		N38
Col-2	Columbia	USA		
Col-3	Columbia	USA		
Ct-1	Catania	Italy		N37
Cvi	Cape Verdi Islands	Cape Verdi Islands	rocky wall with moss	N15
Da(1)-12		Czechoslovakia		
Da-0	Darmstadt	Germany		N50
Db-0	Dombachtal	Germany	stony, sunny road	N50
Dijon-G	Dijon	France		

Dijon-M	Dijon	Russia		
Dr-0	Dresden	Germany	Botanic Garden	N51
Dra-0	Drahonin	Czechoslovakia		
Edi-0	Edinburgh	United Kingdom	Botanic Garden	N56
Ei	Eifel	Germany		N50
Eil-0	Eilenburg	Germany		N51
El-0	Ellershausen	Germany	limestone, south side	N51
En-2	Enkheim	Germany	field border	N50
En-D	Enkheim	Ukraine		
Enkheim-T	Enkheim	Tadjikistan		
Ep-0	Eppenheim	Germany		N50
Er-0	Erlangen	Germany	dry, sandy way	N49
Est	Estland	Germany		
Fi-0	Frickhofen	Germany		N50
Fr-4	Frankfurt	Germany	Botanic Garden	N50
Gr	Graz	Austria		N47
H55		Czechoslovakia		
Hi	Hilversum	Netherlands		N52
H-O-G	Hodja-Obi-Garm	Tadjikistan		
Je54		Czechoslovakia		
Ler	Landsberg erecta	Germany		
Ler-1	Landsberg erecta	Germany		
Lip-0	Lipowiec	Poland	loamy soil/limestone	N50
Lm	Le Mans	France	wheat field	N48
Nd	Niederzenz	Germany		
Oy-0	Oystese	Norway		
Petergof	Petergof	Russia		
RLD1				
Rsch	Rschew	Russia		N56
Rubenz	Rubezhnoe	Ukraine	near lake	
S96		Netherlands		
Sha	Shakdara	Tadjikistan	mountains	
Sn(5)-1		Czechoslovakia		
Ta	Tabor	Czechoslovakia		N49
Ws	Wassilewskija	Russia		N52
Yo	Yosemite Nat. Park	USA	granite sand valley	

Mutants and trangenic lines

35S::cocry1cry2
35S::cophyAcry1cry2
ado1-1
ado3-1
cca11
cca1-1
co-2

co-3
co-7
cop1-4
cry1-1
cry2-1
edi1-1
elf3-1
fca1
fcaco-2
fcaco-3
fcalhy1
ft-7
lhy-11
lhy1-1
lhycca1-1
phyA
phyB
prr5prr71
prr5prr72
soc1-1
suc2::coco-2
suc2::soc1soc1-1

2.1.2 Kits

RNease kit for isolation of total RNA :	Qiagen, Hilden Germany
DNease kit for isolation of total RNA :	Qiagen, Hilden Germany
Biosprint kit for DNA extraction :	Qiagen, Hilden Germany
Turbo DNase :	Ambion biosciences
PCR purification kit :	Qiagen, Hilden Germany
Qiagen plasmid preparation kit :	Qiagen, Hilden Germany
TOPO TA cloning kit :	Invitrogen, Karlsruhe, Germany

2.1.3 Enzymes

Restriction enzymes :	New England Biolabs, Frankfurt am Man
Expand High Fidelity Taq polymerase :	Roche, Mannheim, Germany
Taq polymerase :	Invitrogen, Karlsruhe, Germany
Superscript Reverse transcriptase : :	Invitrogen, Karlsruhe, Germany
RNAse inhibitor : :	Invitrogen, Karlsruhe, Germany

2.1.4 SSR markers (Table 2)

Alias	Chromosome	Position	Forward primer	Reverse primer
F19P19-75410	1	1,16	CCACGTAGGTCAAGAAGAAGAAG	TGTCTGCTGCGATAGAGAGAG
NT1G11	1	1,24	AATTGCATAAGGCACTTGAAAG	GAAGACAAAGCTCTGCAGTAATG
T1G11-89778	1	1,24	GAAGACAAAGCTCTGCAGTAATG	AATTGCATAAGGCACTTGAAAG
NF7G19	1	2,85	TCGTTGAAAACGATTAGATTGG	TTCAAAAATCGTGAGATGAAATG
F7G19-92589	1	2,85	TTCAAAAATCGTGAGATGAAATG	TCGTTGAAAACGATTAGATTGG
NF21M12	1	3,21	TTACTTTTTGCCTCTTGTCATTG	GGCTTTCTCGAAATCTGTCC

CXC750	1	11,31	ATAATAGATAAAGAGCCCCACAC	CGCCAAAGACTACGAAATGATC
T27K12-SP6	1	15,93	GGAGGCTATACGAATCTTGACA	GGACAACGTCTCAAACGGTT
CIW1	1	18,37	ACATTTTCTCAATCCTTACTC	GAGAGCTTCTTTATTTGTGAT
NGA111	1	27,36	TGTTTTTTAGGACAAATGGCG	CTCCAGTTGGAAGCTAAAGGG
NGA1145	2	0,68	GCACATACCCACAACCAGAA	CCTTCACATCCAAAACCCAC
NGA114	2	0,68	CCTTCACATCCAAAACCCAC	GCACATACCCACAACCAGAA
CIW2	2	1,19	CCCAAAAGTTAATTATACTGT	CCGGGTTAATAATAAATGT
F7B19.22	2	3,77	CCCCGCTGAACTGACTGACTACGAG	TCCGCCCACCGATAAGATACGAC
T13H18-T7	2	4,49	GGTAACAGCCTTCACTCGTC	AAAGACTTGTATTTGGGATTTG
CIW3	2	6,41	GAAACTCAATGAAATCCACTT	TGAACTTGTTGTGAGCTTTGA
PLS1	2	8,22	TATGTTGCCTAAGTGCTGTA	TATGGTAGATGGTAGATTATGTG
PLS2	2	8,37	TACGCGAATTATTTTTAGGAGA	AATTTATTTTGAGTCGGATGC
PLS3	2	9,00	TAGTCGTTTCTCTGGTTGTAG	TTGCCTGTCGATGTAGATTTGT
PLS4	2	9,02	CCACCCTCTAGACTTATTACG	AAGAGGCAAAAACTGTCATTAG
PLS5	2	9,18	GATGCCTTTCTCCTGGTTG	AATATAGCCGTCGTCTTCATCA
PLS6	2	9,28	ATAGTGGGTGATCTTTGAATGTA	ACTTGTCTCGTCGCACTGTT
CHIB	3	3,96	ATGAGAAGCTATAATTTTTTCAATA	CTCATATATACAAAGAACTACTATAC
NGA162	3	4,61	CTCTGTCACTCTTTTCCTCTGG	CATGCAATTTGCATCTGAGG
T27C7-SP6	3	13,05	ATGCCTAACTATTCGCTGAC	TTCTGTAGTTCTTTGTGAGTGC
F1P2-TGF	3	17,55	TTTGTCTGAAGATGTGGAGAGAGAG	CAAAACCCCACTCTTCATTATTGTT
T16K5-TGF	3	18,44	TTGTCGAAATAAAAATTGACCGTTA	TGGATGTGGATTCTATTGTTTCTCA
CIW4	3	18,90	GTTCATTAAACTTGCGTGTGT	TACGGTCAGATTGAGTGATTC
F24M12-TGF	3	19,07	GTTCTCTGCATTCCACACATACTCT	CTTGGGTATTCTGAAGAGCATAAAT
CDC2BG	3	20,07	ATTGAACTGTGTTGGTTTCTGG	GGGAAAAACGAAGTGACGTG
NGA707	3	21,76	CTCTCTGCCTCTCGCTGG	TGAATGCGTCCAGTGAGAAG
FUS6.2	3	22,64	TTCCTTGATCAGATTTGGTCG	TCGTTACACTGGCTTGCTTG
NGA6	3	23,04	ATGGAGAAGCTTACACTGATC	TGGATTTCTTCCTCTCTTCAC
F14G16-T7	4	4,19	TCGACTAGATTTATTATTTCTCTCAG	TTTGGCTTGACTCTGTGAAC
CA72	5	4,25	CCCAGTCTAACCACGACCAC	AATCCCAGTAACCAAACACACA
NGA151	5	4,67	CAGTCTAAAAGCGAGAGTATGATG	GTTTTGGGAAGTTTTGCTGG
PAT1.2	5	5,96	CATGCTTCATCATTGCCC	AGCTGAAGCTCTGCCACC

2.1.5 RT-PCR primers (Table 3)

Gene	Forward primer	Reverse primer
Act	GGTAACATTGTGCTCAGTGGTGG	AACGACCTTAATCTTCATGCTGC
Act2	AACGACCTTAATCTTCATGCTGC	GGTAACATTGTGCTCAGTGGTGG
AP1	CTTCTTGATACAGACCACCCA	GCACATCCGCACTAGAAAAAA
COFran	TAAGGATGCCAAGGAGGTTG	CCCTGAGGAGCCATATTTGA
CRY1	CTTTCACAGATGTGGCAACT	CATTTTCTTCACTAGGGAAC
CRY1n2	GCCACATCTGTGAAAGCGCT	TCTTGATGTTGTTAAATCCA
CRY2	AGCTCGTGAGCTACTAGCTA	CTTCTTCCTCAGGTTTCACT
CRY2n2	CTAGTAGCTCACGAGCTGTG	TCAGATTACTACAAGTTTGG
FKF1c	ATCTCCAGTGTTCCAGTTATCT	GTCGTAACTGTCGATTCCTACA
FLC	ACGCATCCGTCGCTCTTCT	GCATGCTGTTTCCCATATCGA
FRL1	ATGCAGGGAAATCGTCGTC	CCGATGATGACGGTGTGTAG
FT_new	ATGGTGGATCCAGATGTTCC	ATTGCCAAAGGTTGTTCCAG
FT_Fran	CGAGTAACGAACGGTGATGA	CGCATCACACACTATATAAGTAAAACA
GIc	CTGTCTTTCTCCGTTGTTTCACTGT	TCATTCCGTTCTTCTCTGTTGTTGG

HUA2	ACGGCGATACAGAAATGAGG	AGCATCTGGGGAGCAGTATG
HUA2n2	ACAGAACCATCGGTTTCAGC	TCTAGGAAGGCCTGGATGTG
LFY_Fran	GGTACGCGAAGAAATCAGGA	ATGACGACAAGCGATGTTCA
PHYA	CAGAAGTCATGTTGTCGTCT	GTTAGCTGACCTCCGGATGG
PHYB	TGGAAGTGTCATAAACGCGA	CTAATATGGCATCATCAGCA
SOC1	GGATCGAGTCAGCACCAAACCG	CTTGAAGAACAAGGTAACCCAATGAACAAT
SOC1N2	CTTGAAGAACAAGGTAACCCAATGAACAAT	GGATCGAGTCAGCACCAAACCG
SVP	GCTCGTTCTCTTCCGTTAGTTGC	CGCTCTCATCATCTTCTCTTCCAC
TFL1	TTCAACTCATCTTTGGCAGT	TATCTTTCCTAATATCCCTTCG
TSF	ACCTGCCACCACTGGAAAT	CTACGTTCTTCTTCCCCCACA
TSFn2	CTACGTTCTTCTTCCCCCACA	ACCTGCCACCACTGGAAAT
Tub	ACACCAGACATAGTAGCAGAAATCAAG	ACTCGTTGGGAGGAGGAACT
APRR9	CAAGATTCGATAAAAATGGTGTTG	TCTCTGGCTTCTACTCCAACG
APRR7	ATGTCATCTCATGACTCAATGGGG	AGAGATCTTATTTTCATCCGC
APRR5	ACACAGGACTCGGTGAATACTGTG	TTCTCGTTGAAGAGATTGCTG
APRR3	TCACATGATTCTATGGTTCTGGTC	TTGTGCCCATCGGTCACTTCC
LHY	GCGTTTCCAGCTTGTCATTCAC	GTCGGCATTTAGCTTGGTTACTCC
CCA1	AACCAGTTTGTCATGAGCAGCCTTCT	TGCCATTGGATATCTTTGTTCTTGTTG
TOC1	CAGCACCGAAGAAAAGTAGATTGAAGA	ATGTGGGAAGCGTGCGTGACC

2.2 METHODS

2.2.1 Flowering time analysis

For the flowering time analysis 75 accessions from the Altmann and the Nordborg collection and 30 mutants and transgenic lines were selected. The accessions have been through single seed selection to reduce heterozygosity. The flowering time was scored in six different day lengths three of which are short days (SD) of 6, 8 and 10 hours of light and the remaining ones are long days (LD) of 12, 14, 16 hours of light respectively. The analysis was performed in controlled environment growth chambers at 22 °C, after stratification for 3 days at 4 °C. A population of 18 individuals represented each accession. Flowering time parameters such as a) number of total leaves b) days until bolting at the stage of 1 cm c) days until anthesis were monitored. The experimental design was a Randomized Complete Block (RCB). For the vernalization treatment, 10 days old seedlings, grown under SD of 8 h at 22 °C, were exposed to 4 °C for one month under the same photoperiod. For the genetic analysis, 150 to 200 F2 plants were used for each population and 50 individuals were used for F3 populations.

2.2.2 GIGANTEA real time bioluminescence monitoring

A transcriptional fusion of 2.5 kbp of the promoter of *GIGANTEA* from Ler to the firefly luciferase gene was made. The resulting construct was used to transform the above accessions. For the top count imagining seeds, were stratified for 3 days at 4 °C and then sterilized with 95 % ethanol and 10 % bleaching solution. The seeds were then plated in Petri dishes containing appropriate medium with antibiotic (PPT) and allowed to germinate. Twenty-four, ten days old seedlings represented each accession. During the first day of the analysis, the seedlings remained in the growth chambers under specific photoperiodic conditions and manual measurements of luminescence were taken every 20 min for the duration of each day length respectively. At dusk, the plates were transferred to constant dark to monitor the free running period under that condition. An FFT NLLS (Fast Fourier Transform Non Linear Least Squares) analysis was performed using the BRASS program and parameters such as the phase under each LD entrainment and the free running period and the RAE (Relative Amplitude Error) under constant dark were calculated. For the analysis of the data the TopTempII program was also used

2.2.3 mRNA extraction and cDNA synthesis

For the mRNA expression analysis of the accessions, total RNA was extracted with the RNease kit (Qiagen) from 10 days old whole seedlings grown on soil under a specific day length. Since the growth rate is different under LD 16, 14 and 12 h, the all seedlings were selected to be at the stage of the first two true leaves. 3 mg of RNA were tested in formaldehyde gel in order to evaluate the quality of the RNA. 5 mg of total RNA were DNase treated with the DNase kit (Ambion). For large batches of RNA extraction the Biospint protocol for DNA extraction was modified for RNA preps. The optimum conditions of the PCR for each pair of primers were evaluated after a gradient Real time PCR.

cDNA synthesis mix :

5 x first strand buffer (Invitrogen): 8 µl

oligodT_{15} 10 mM : 4 µl
dNTPs 2,5 mM : 2 µl
DTT (Invitrogen): 4 µl
Superscript reverse transcriptase II : 1 µl
RNase inhibitor : 1 µl
RNA : 20 µl (5 µg).

Denaturation of RNA alone for 5 min in 65ºC
Incubation at 42ºC for 2 h
Heat inactivation of reverse transcriptase : 75ºC 5 min
Coll on ice and add 110 µl of ddH_2O.

A typical PCR program for RT-PCR is shown below :

Initial denaturation (hotstart) : 95ºC, 3 min
Denaturation : 95ºC, 1 min
Anealing : Primer defined
Extension : 72ºC, 45 sec
Repeat 30-40 times depending on gene
Final extension : 72ºC for 10 min.
Short term storage : 4ºC

2.2.4 Shift experiment conditions

For the first shift experiment, Col-0 seeds were stratified and sterilized as mentioned above and then plated on agar plates containing appropriate medium without sucrose. 10 day-old seedlings grown under SD of 8 h were used. Sampling began at day 11. At day 12, plates were transferred to LD 16 h for two days and then back to SD 8 h for another two days. Plates that remained under SD of 8 h throughout the experiment, served as negative control. Sampling was performed for all cases at ZT 4, 12, 16, and 20.

For the second shift experiment, Ler seeds were stratified for 3 days at 4 °C and then were sown on soil under SD 8 h. 14 days old seedlings at the stage of the first two true leaves were selected for sampling and only the two true leaves and the two cotyledons were harvested every 4 hours and at ZT 1. At day 15 seedlings were transferred in SD 8 h to serve as negative control and the remaining plants were exposed to LD 16 h for three days. After the induction, all plants returned to SD 8 h. As an additional negative control, induced and non-induced co-2 seedlings in Ler background were harvested with the same way and at the same age as for the WT only during the three days of the induction. For the shift experiments of the accessions the same experimental setup was used. mRNA extraction, DNase treatment and cDNA synthesis were performed in all cases as described above.

2.2.5 Statistical analysis

For the statistical analysis SigmaStat v.3 was used. A two way ANOVA was performed with accessions and day lengths as the two factors and for parameters such as TLN and BT for flowering time and phase and free running period for the *GI::LUC* expression. Demonstration of the results was performed with SigmaPlot v. 10. For the hierarchical clustering, Cluster v. 3 was used with the same parameters as for the two way ANOVA. The data were mean-centered in an accession depended way and SOMs were calculated for both the accessions (100.000 iterations) and the day length (20.000 iterations). Both factors were then clustered according to the single linkage clustering method using uncentered correlation as similarity metric. TreeView v. 1.6 was used to demonstrate the results. PCA was performed using Cluster. QTL mapping was performed using the MapQTL program. Linkage maps were created using JoinMap. A permutation test defined the LOD threshold for each population. First an interval mapping was performed and subsequently MQM mapping with automatically selected cofactors was used.

2.2.6 DNA isolation and genotyping

DNA isolation was performed in 100 mg of fresh tissue with either the CTAB method (for small number of samples) or with the semi-automated method for DNA extraction using the Biosprint robot (QIAGEN) for the mapping populations. For genotyping PCRs a standard protocol was used :

Initial denaturation (hotstart) : 95ºC, 5 min
Denaturation : 95ºC, 1 min
Anealing : 55ºC, 1 min
Extension : 72ºC, 1 min
Repeat 30 times
Final extension : 72ºC for 10 min.
Short term storage : 4ºC

The PCR fragments were then analysed in 3 % agarose gels after electrophoresis in 100 V for 40 min approximately.

2.2.7 *In situ* hybridization

Meristems were harvested at ZT 8 in all conditions in a solution of 4% (w/v) formaldehyde, with 0.1% tween-20 and 0.1% triton-x-100. Fixation was achieved with vacuum infiltration for approximately 10-20 min depending on the accessions and the age, followed by O/N. filtration in new solution. The next two days ethanol exchanges were performed according to the following scheme :

30% Ethanol	1 Hr	on ice
40% Ethanol	1 Hr	on ice
50% Ethanol	1 Hr	on ice
60% Ethanol	1 Hr	on ice
70% Ethanol	2 Hr	on ice

85% Ethanol	OVERNIGHT	4°C

Day 3

95% Ethanol	4 Hr	4°C
100% Ethanol	4 Hr	4°C
100% Ethanol	OVERNIGHT	4°C

At this points samples were stored at 4°C until embedding with wax. Prior to embedding staining of the meristems in an 0.2 % Eosin ethanol solution was performed for 40 min. The meristems were then transferred in the embedding machine for one O/N and the next day each meristem was transferred separately in peel-away molds (Polysciences Inc. Cat #18985 or 18646A) with additional wax. The embedded meristem blocks were stored at 4 °C prior to sectioning. Sectioning was performed in 7 μm thickness. *In situ* hybridisation and signal development was performed with digoxigenin antibodies and was performed according to the protocol described in Searle et al., 2006 with minor modifications.

2.2.8 Gating experiment

For the gating experiment, three accessions, namely Dijon-G, Lip-0 and Dijon-M were selected based on their pattern of GI phase reposne over the different day lengths. Plants were initially grown in Petri dishes with GM growing medium for ten days under LD of 16 h and SD of 8 h. On the last day they were transferred to 96 well microtiter plates with GM medium. Luciferin 5mM was added and plants were transferred to dark. In total 12 identical plates were prepared for each day length, giving a resolution of 2 h in the gating experiment. A additional plate was also prepared serving as a referance. On the day of the measurement the referance plate was kept continuously in the dark while measurements were recorded. A single plate from the rest was exposued to with light (70 μE) for half an hour and then measurements of the luciferase intensity were taken for 2 hours. At the end of the 2h interval another plate, kept in the dark, was exposed to light in a similar way until the end of the day. Each accession was represented by 24 individuals and all were included in a single 96-microtiter plate.

RESULTS AND DISCUSSION

3.1 NATURAL VARIATION IN FLOWERING TIME RESPONSES

3.1.1 PHYSIOLOGICAL DESCRIPTION

3.1.1.1 Introduction

Seventy five, geographically diverse *Arabidopsis* accessions from the Altmann and the Nordberg collection were selected (Table 1). The flowering time of these accessions was monitored in controlled environmental conditions, under six different day lengths three of which were long days of 16, 14 and 12 hours of light and the remaining ones were short days of 10, 8 and 6 h. A large number (>30) of known mutants and transgenic lines of genes from several pathways was also included in the analysis. Comparing the accessions with the mutants allows the selection and subsequent testing of candidate genes that might explain interesting phenotypes observed in accessions.

A series of questions was raised. i) How does a quantitative, long day species such as *Arabidopsis* respond in a wide range of different photoperiods ii) Although *Arabidopsis* is a long day plant, can we detect natural variation under short days or is it only restricted to long days? iii) How sensitive is the equilibrium between photoperiod and vernalization over a range of different photoperiods? iv) With both photoperiod and temperature changing along with latitude can we identify a pattern in the already mentioned day length responses that correlates robustly with geographical and climatological information? Finally v) can we identify additional players in the genetic framework of flowering time that control this natural variation and if yes, which are the epistatic interactions between them and which are the spatial and temporal consequences of this genetic variability at the molecular level? Here the selected approaches to answer these questions are presented.

3.1.1.2 Describing photoperiod discrimination under a range of day lengths – Accessions.

The flowering time analysis revealed extensive natural variation in the ability of the accessions to discriminate between the selected day lengths. A general aspect of this variation among accessions, expressed as Total Leaf Number (TLN) is shown in Figure 8. A general pattern of day length discrimination was observed with Day length Response Curves (DRCs) of flowering time following the sigmoid curve. In order to further support this observation, three different types of theoretical curves were selected for correlations with the experimental data. The strength of the correlation was then evaluated by the distribution of the Rsquare

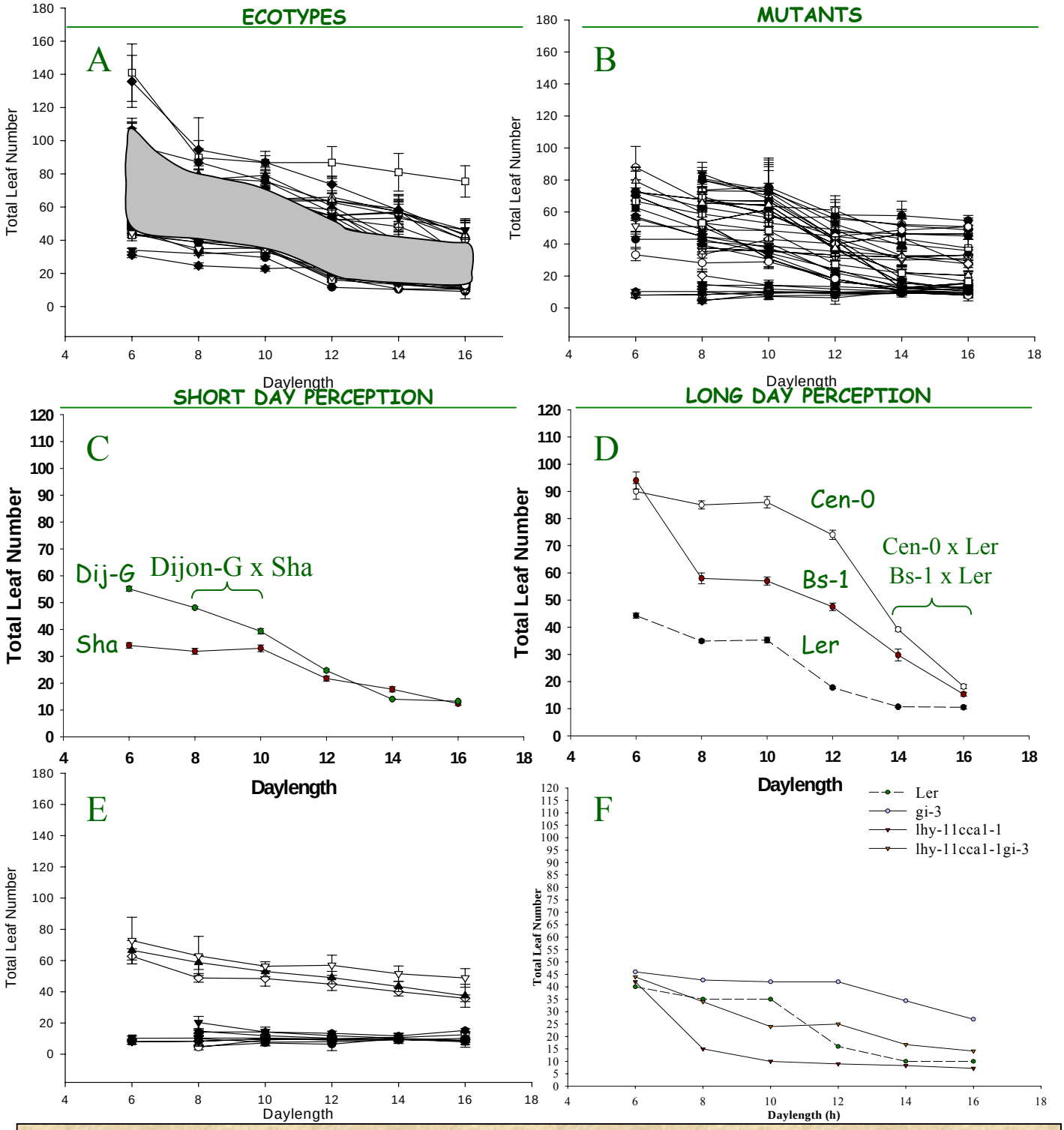


Figure 8 : Identified natural variation in the flowering time responses across different day lengths. A) The general pattern of the response to day length for flowering time. The shaded part highlights the general trend of sigmoid response. B) **D**aylength **R**esponse **C**urves (DRCs) of mutants and transgenic lines used in this project C) Natural variation under nearly exact long days. Two accessions Cen-0 and Bs-1 showed differences of 20 leaves between LD 16 h and LD 14 h. Both accessions were crossed to Ler (which does not respond to these two daylengths) in order to identify QTLs which control the response under nearly exact long days. D) Natural variation under nearly identical short days. Generally a strong response was observed between LD 12 h and SD 10 h, marking the transition from LDs to SDs. One cross of Dijon-G with Sha was used in order to detect QTLs which discriminate between SDs of 8 and 10 h. E) Cases of mutants and transgenes, which result in daylength insensitivity with either ealry or late flowering. F) Disruption of the normal function of the circadian clock affects daylength discrimination. Interestingly all genotypes converge at SD 6 h.

values of all accessions. The highest strength with Rsqr values close to 1 was for the sigmoid curve while the linear together with the quadratic polynomial curve showed weaker correlations. Another type of sigmoid curve, the logistic was also tested and verified the above observation (data not shown).

LD 16 h as expected result in early flowering in almost all accessions (
9). However already under LD 14 h a group of accessions is separated from the main group, forming a late flowering tail. As the day becomes shorter, this tail becomes more obvious and eventually forms a second peak in the histogram, even though the flowering times of the major group are shifted towards later flowering. Regression analysis was also performed between Days to Bolting (BT) and Total Leaf Number (TLN). In general a varying correlation between the total leaf number and the days until bolting across different day lengths was observed, with R^2 values ranging from 0.44 to 0.85. Higher correlation between the two parameters was observed under LDs compared to SDs. As expected the three long days were clustered together and independently of the SD of 8 and 10 hours. SDs of 6 hours clustered independently of all. These results are based on three different types of influence diagnostics that were selected in order to identify accessions the flowering time of which deviates significantly from the rest.

Interestingly a significant transition in both the leaf number and the days to bolting between SD of 10 h and LD 12 h was observed, which is more evident for the leaf data. A high level of correlation was observed between the days until bolting and days until anthesis across all day lengths for both groups (R^2 values above 0.90 in all cases, data not shown), suggesting that the two estimations of flowering time are consistent. Regarding correlations of flowering time with the origin of the accessions and more specifically with latitude, no statistically significant associations were observed in any of the analysed day lengths, both for the case of TLN and BT.

3.1.1.3 Describing photoperiod discrimination under a range of day lengths – analysis of mutants.

A set of thirty mutants and transgenic lines affecting known genes influencing flowering time were also selected and their flowering times monitored over the six day lengths (Table 1). The mutant DRC (**D**ay length **R**esponse **C**urve) range was comparable to the accessions (Fig 10), however there where many cases in which day length response was totally abolished resulting in a horizontal DRC both under early flowering and late flowering ranges (Fig 8E). Apart from these cases, which indicate that the function of major genes of the photoperiod

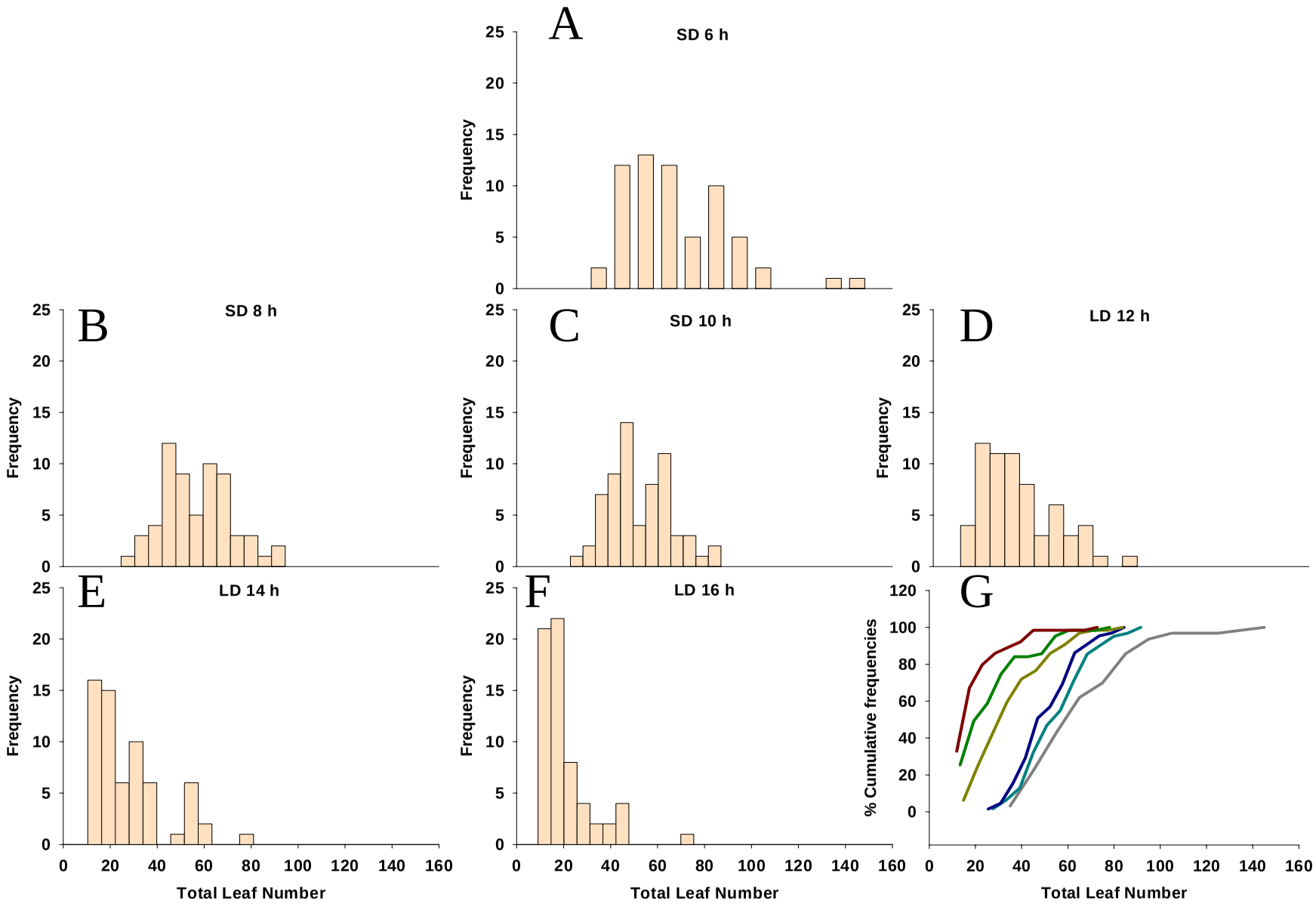


Figure 9 : Distribution of flowering times of accessions under six different daylengths based on TLN data. A) SD 6 h. B) SD 8 h. C) SD 10 h. D) LD 12 h . E) LD 14 h. F) LD 16 h. G) Combined cumulative frequencies of the same plots, allow direct comparisons. As the day becomes shorter the distribution as expected is shifted in earlier flowering times but also the shape of the distribution is more sharp indicating more uniform flowering time responses between the accessions. Note the three clusters (LD-SD and SD 6) that are formed confirming the results of the BT-TLN correlation and the Hierarchical clustering.

pathway is either masked by epistatic mutations or abolished directly, attenuation in the function of other genes resulted in more subtle changes affecting day length perception under LD only. Such examples are the clock genes (Fig 8F). A double *lhycca1* mutant is very early flowering between LD of 16 h and SD of 8 h. By contrast *gi-3* mutant is late flowering between LD 12 h and SD 6 h. Between LD 12 h and LD 16 h this particular mutant retains a reduced ability for day length discrimination. Therefore the property of late or early flowering is not always connected to the property of increased or decreased response to day length. The *gi-3* mutant although late flowering generally, can still respond to day length however this response is greatly reduced compared to the WT.

Another interesting aspect of *gi* is that the critical day length for the response is shifted towards higher values. In the WT, SD of 10 h are the point from which the response initiates. In the *gi-3* mutant however this is shifted to LD 12 h. But also the kinetics of the transition itself are changed. In WT the response is nearly saturated at LD 12 h. In *gi-3* however gradually earlier flowering but not saturation can be observed between LD 12 h and LD 16 h, suggesting that the saturation point for this mutant is placed in even longer day lengths. Interestingly a combination of an *lhycca1* double and early flowering mutant with a late flowering *gi-3* is close to the WT in terms of overall lateness in flowering time, suggesting that the loss of function of the two MYB genes is epistatic to loss of GI in the control of flowering time. Since photoreceptors are controlled by the clock and in turn control CO protein stability, this class of genes could be an interesting target for this regulation.

Since in the double *lhycca1* mutant *CO* mRNA shows a phase advance, a scenario in which external coincidence with light is achieved in much shorter day lengths compared to WT leading to early flowering is highly likely to occur. Interestingly all these mutations have such a diverse effect on flowering time over many day lengths, but converge under SD 6 h. This indicates that such photoperiodic conditions are too short to allow any type of external coincidence even though ectopic CO phase advance occurs.

3.1.1.4 Statistical approaches identifying natural variation in the day length Response Curves for flowering time.

To further elucidate the observed natural variation in photoperiod discrimination, three different statistical approaches were selected. First a Two-Way ANOVA approach was selected. This type of bi-factorial analysis allowed the identification of statistically significant differences both among different accessions in the same day length (Factor A, first plane of analysis) and across different day lengths for the same accession (Factor B, second plane of

analysis), both in terms of bolting time and leaf number. In general, more statistically significant differences were detected with the leaf number data compared to the days until bolting.

Although the two-way ANOVA approach helped us in identifying natural variation in the flowering time responses in a statistically significant manner, the classification of the accessions was based on pair wise comparisons of the flowering time means between different day lengths. This approach has the advantage of being a high-resolution analysis, however it does not classify the accessions based on differences in the general shape of the response. Therefore an additional approach, which takes into account all day lengths simultaneously in order to detect different patterns, is required. For that reason a Hierarchical Clustering approach was selected (Fig 10). Instead of genes, the analysis had accessions and instead of expression levels, flowering times. The arrays were the six different day lengths.

Generally two major clusters were identified with overall late flowering accessions in the first and overall early flowering ones in the second (Fig 10). Within these two major clusters, several minor ones were formed according to day length discrimination properties. Several groups of responses were identified. Some cases include gradually increasing day length discrimination properties as the day becomes shorter, or almost horizontal lines indicating a moderate ability to respond to the changing day length. Apart from the clustering of the accessions, the six different day lengths were also grouped in two major clusters. The first cluster consists of LDs with 16 h and 14 h being almost identical and 12 h being slightly different. The second cluster consists of SDs with 8 h and 10 h being again identical and 6 h slightly different. This clustering confirms the regression analysis between days to bolting and total leaf number.

Although the hierarchical clustering allowed classification of the different accessions based on their overall performance in the day length range, the resulting dataset is complex and limits subsequent analysis and correlation with other traits such as latitude or allelic properties of candidate genes for each accession. In addition hierarchical clustering cannot separate easily the two major properties of flowering time in this project, namely overall lateness and day length discrimination. To overcome these complications, a third approach namely a Principal Component Analysis (PCA) was combined with the Two-way ANOVA and the Hierarchical Clustering. The philosophy of PCA is simplification of the dataset without losing accuracy in the description. The conditions under investigation (in this project the six day lengths) are replaced by components, each one of which concentrates one general property (e.g overall lateness, concavity, sharpness of the response etc). These components then are

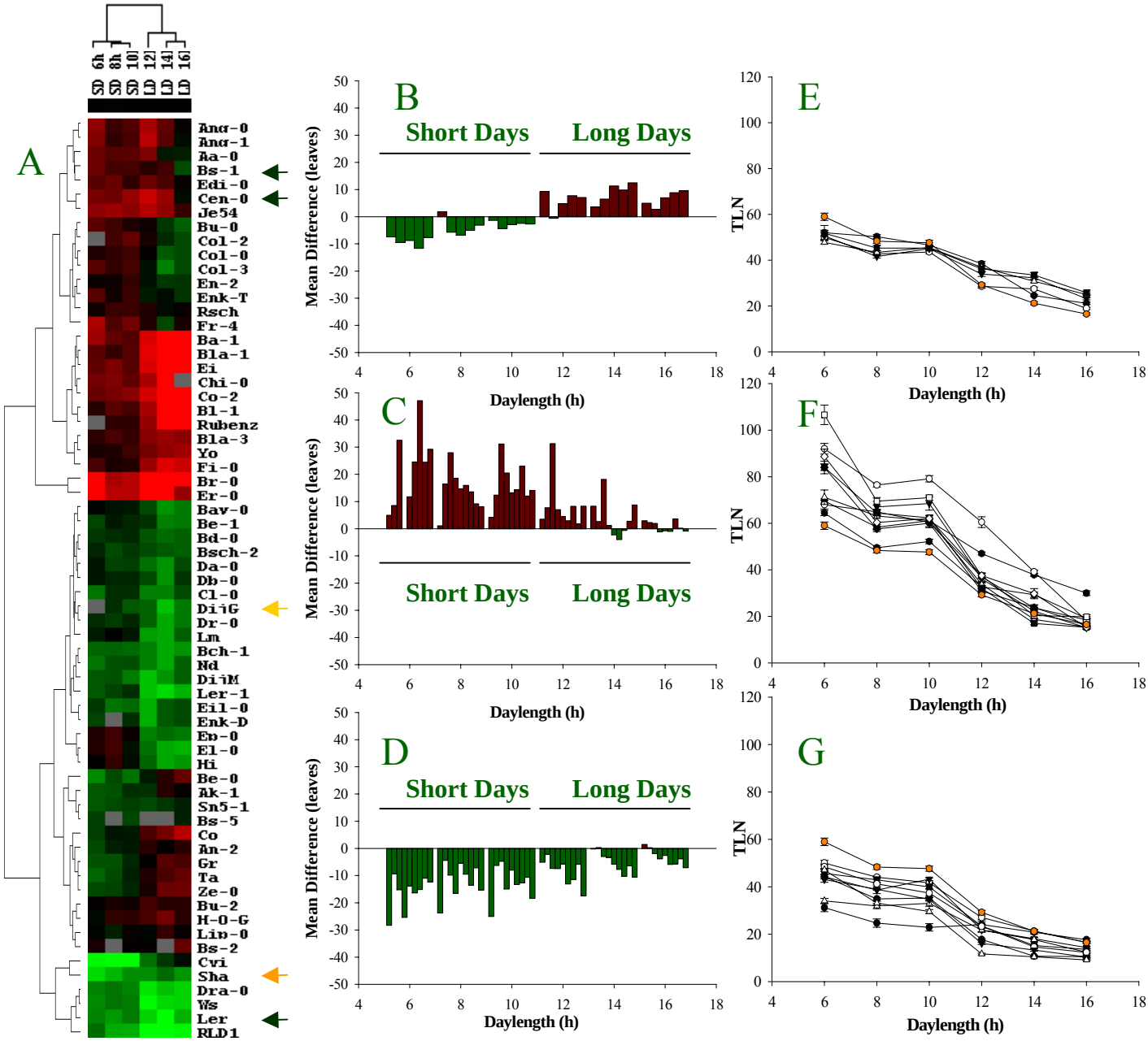


Figure 10 : The hierarchical clustering approach identifies natural variation in the flowering time responses across the six day lengths. The experimental flowering time data are log transformed to become symmetric around zero and then mean centered in an ecotype dependent manner. A self organized map with 100.000 iterations was then calculated to give a uniform shape in the final tree. Clustering was performed with an uncentered correlation similarity metric and a complete linkage was performed. A) The output tree for the clustering of the flowering times across different day lengths. Green corresponds to negative centered values and red to positive. In terms of flowering time, green corresponds to earlier and red to latter flowering compared to the average dataset response (orange curve). The numbers correspond to the correlation coefficients of the three major clusters. Several cases of daylength discrimination were identified. Some accessions demonstrate early flowering under short days and later flowering under long days, resulting in almost a straight curve, while in others late flowering times throughout the six daylengths, resulting I a gradually increasing strong response to daylength as the days become shorter. Finally some accessions demonstrate early flowering in all daylengths, resulting into a moderate response to the changing photoperiod. B) C) D) Distance from the mean of representative accessions from the three groups. E) F) G) Flowering time curves of representative accessions from three different groups, across the different day lengths. The average response is shown in orange. Green arrows indicate the selected accessions for the LD crosses, while the yellow arrows the selected accessions for the SD crosses. Note that in both cases the parental responses are extremely diverse.

classified according to the overall variability in the dataset that each one explains. A successful PCA concentrates as much of the variability as possible (<70 %) in the first two components. In such cases a simple scatter plot between these two components can reflect accurately the behaviour of each line in the initial conditions of the experimental design.

PCA analysis for flowering time was successful both for the accessions and the mutants (Fig 11). Component ranking revealed very similar properties between the two datasets, suggesting that a direct comparison is possible. Both datasets concentrated 70 % of the variability within the first two components, namely overall lateness (45 %) and day length discrimination (25 %). Interestingly a direct comparison of the mutant and the accession dataset revealed the presence of a group of accessions, the flowering properties of which cannot be predicted by the selected mutants and transgenic lines of the project. This comparison indicates that flowering time regulation over a range of day lengths in these lines is determined either by unknown genes, or unknown alleles of known genes or by conditional epistasis between such alleles.

3.1.1.5 Determination of the vernalization properties of the accessions under study.

In *Arabidopsis*, vernalization is often a mechanism that affects flowering time. Major gene of the vernalization pathway is the FLC MADs box gene while another gene, FRI acts as positive regulator of FLC elevating the mRNA levels of the latter. Both genes are major targets of natural variation for flowering time in the species with active alleles in both loci leading to extremely late flowering, while inactive alleles in FRI lead to moderate lateness and response to vernalization. FLC is located at the top of chromosome 5 while FRI at the very top of chromosome 4. Another gene similar in sequence and role with FLC is FLM/MAF1. This gene belong to a family with 5 member and together with MAF2 cause a weak response to vernalization, which is more obvious under short days. One case of natural variation has been reported also for MAF1 (see introduction).

Since the FRI/FLC genes are major targets of natural variation it is important to investigate the vernalization properties of the accessions involved this study and the putative role of these genes in the regulation of day length perception. Phenotyping of Bs-1 with and without vernalization revealed that this accession shows a moderate response to vernalization (Fig 12A). This response was confirmed also with days to bolting. By contrast vernalization did not rescue the late flowering of Cen-0 under LD 14 h, both in terms of leaf data and days to bolting. In summary the Bs-1 accession responds to vernalization while the Cen-0 does not.

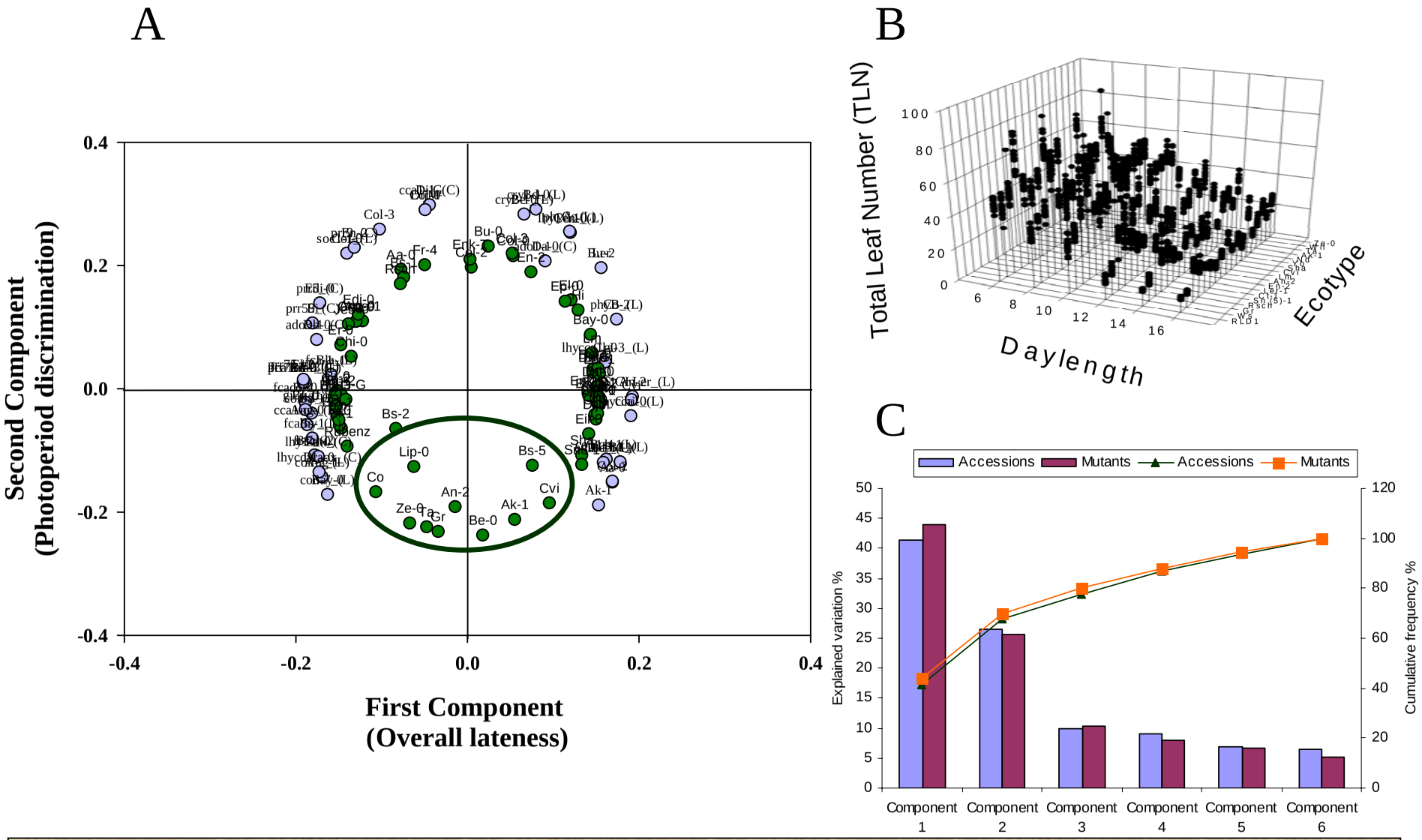


Figure 11 : PCA and Two way ANOVA analysis for flowering time. (a) Combined **P**rincipal **C**omponent **A**nalysis (PCA) between the flowering time responses of the accessions (green) and the mutants and transgenic lines (blue) of known genes affecting flowering time. Note the appearance of an accession cluster, the properties of which cannot be explained by the properties of mutants and transgenic lines of known genes. b) 3D plot of the distribution of flowering times of accessions under six different daylengths c) Component ranking of the PCA analyses for the accessions and the mutants. Left Y axis shows explained variability of each component while the right axis cumulative frequencies of the PCA components. The PCA properties of the accessions and the mutants are almost identical and therefore comparable with the first two components explaining most of the observed variability of the data set.

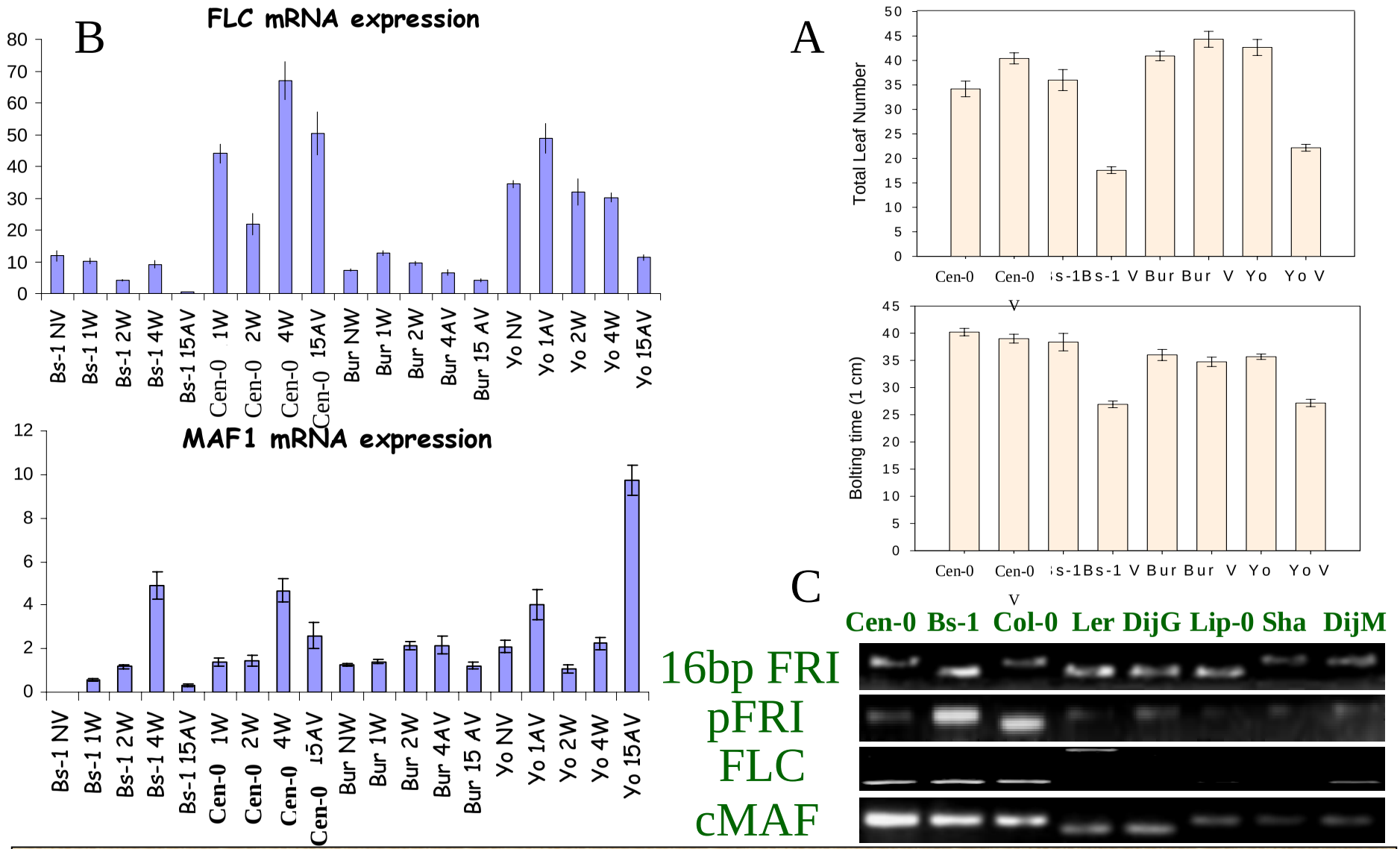


Figure 12 : Vernalization response and properties of selected accessions. A) Vernalization responses of selected accessions under one month. Only Bs-1 and Yo respond to vernalization. B) Comparison of *FLC* and *MAF1* levels during a 30 days vernalization time course for various accessions. NV : No vernalization, 1WV : 1 week of vernalization, 2 WV : 2 weeks of vernalization, 3WV : 3 weeks of vernalization, 4WV : 4 weeks of vernalization, 15AV : 15 days after vernalization. Note the high levels of FLC and the slight reduction in FLC expression after vernalization in the Cen-0 accession. C) Genotyping FRI, FLC and the MAFs in selected accessions. 16 bp FRI : 16bp Insertion in the first intron. pFRI : promoter deletion in FRI. FLC : presence of the transposon in the first intron of FLC. cMAF : dCAP marker. Cen-0 has an active FRI but Bs-1 has the inactive 16bp insertion.

Ler also does not respond to vernalization. Ragarding Sha and Dijon-G both do not respond to vernalization.

The next step was to determine the FLC and MAF1 mRNA levels in a time course during a one-month exposure to low temperatures. As shown in Fig 12B FLC levels are moderate and decrease during vernalization in the Bs-1 accession. Two weeks after the treatment, FLC mRNA is almost undetectable confirming the vernalization response. Interestingly FLC mRNA in the Cen-0 accession is very high throughout the experiment and after vernalization only a moderate reduction takes place. In the case of MAF1 both Bs-1 and Cen-0 have comparable overall levels, which initially increase during vernalization, and subsequently decrease rapidly afterwards to almost undetectable levels.

The very high levels of FLC in the Cen-0 accession suggest that this accession carries a functional FRI allele. Indeed genotyping with markers that distinguish between the reported inactive from the active alleles of FRI proved that the Cen-0 FRI does not host the promoter deletion of the Col-0 weak allele or the 16 bp insertion of the Ler weak allele, confirming its full functionality (Fig 12C). Bs-1 in contrast has the 16 bp deletion similar to Ler and therefore is inactive. This explains why the levels of FLC mRNA are low compared to the Cen-0 levels. Dijon-G carries also the Ler mutation and therefore has an inactive FRI allele, while Sha, like Cen-0 carries non of the known mutantions and FRI in that accession is reported as active.

These results confirm published datasets of large vernalization screens between *Arabidopsis* accessions (Fig 13). According to these screens Sha, Lip-0, Ler and Dijon-G have low levels of FLC and do not respond to vernalization, Cen-0 has high levels but does not respond to vernalization as well and Bs-1 has moderate levels yet responds to vernalization. How is it then possible to have accessions such as Cen-0 with high levels of FLC and active FRI alleles, without a response to vernalization? The secret is mutations in the CDS of FLC, which render the gene inactive. Indeed the FLC of Cen-0 carries a rare mutation, which disrupts a splicing site in the first intron of the gene, leading to a misplaced and inactive larger protein product (Lempe et al., 2005).

3.1.1.6 Identifying natural variation in the flowering time responses under natural conditions – Connection to adaptation in the wild?

The Holy Grail in most studies that include natural variation, is a connection of the phenotype with fitness, which could suggest importance of the variation to particular environments. Unfortunately a usual caveat in such associations are the conditions of the

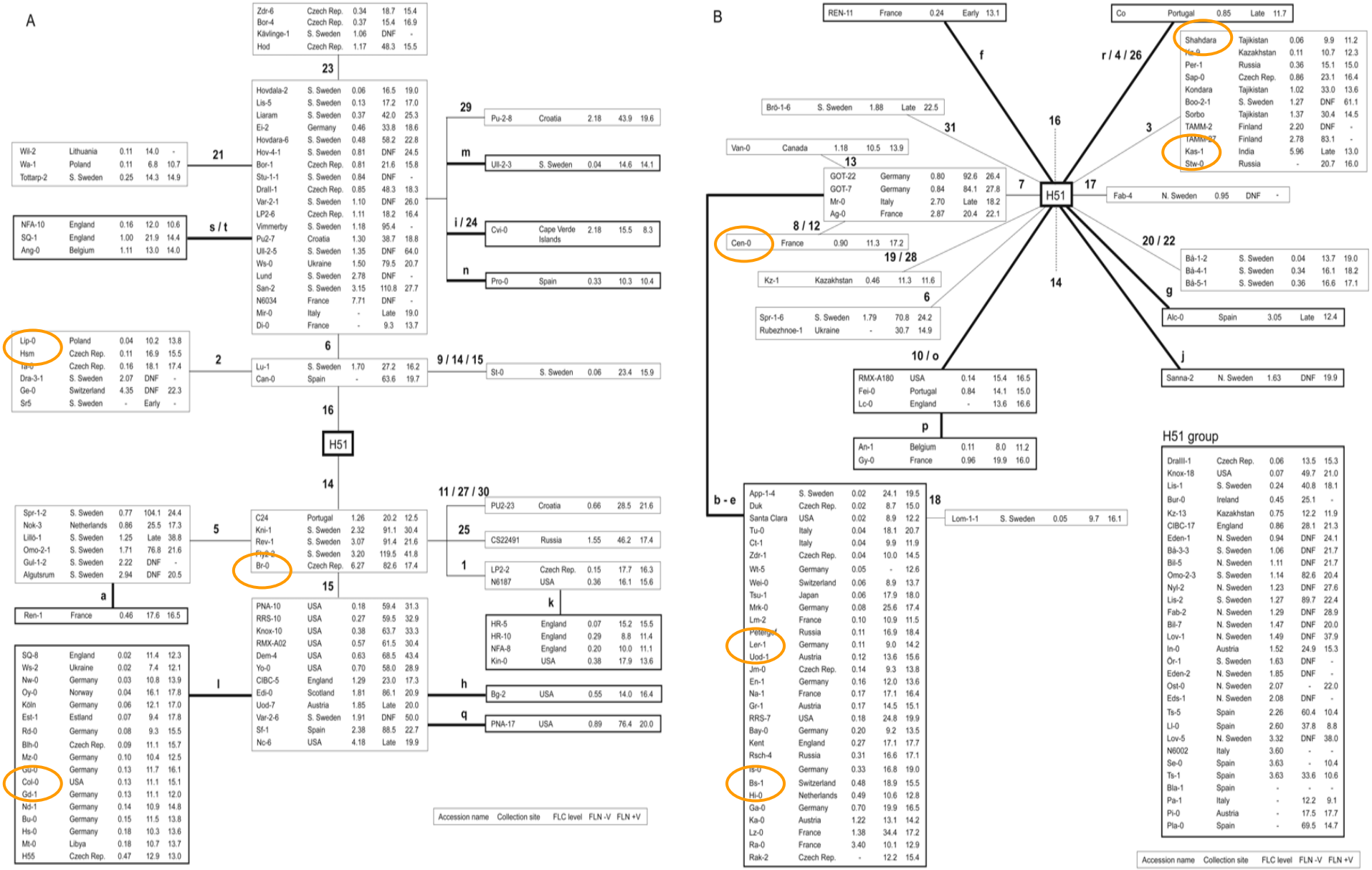


Figure 13 : Vernalization responses and FLC levels among *Arabidopsis* accessions. Note that Cen-0 *FLC* mRNA levels are high compared to the high-*FLC* Kas-1 accession from Kasmir, but there is no vernalization response. On the contrary FLC mRNA in Bs-1 is on average levels yet this ecotype can respond to vernalization. Other accessions, which are interesting for this project are also indicated with orange cycles. (Sup fig, Shindo et al. 2005)

experiment that do not fully reflect the environmental conditions that plants experience in the wild. In this project, the controlled environmental conditions were excellent in isolating day length as the sole parameter of the experiment, however it would be interesting to expand the results of day length perception under controlled environment to the real world.

In collaboration with Brown University and Prof. Annie Schmitt, the flowering times of selected accessions from our screen were included in a large project aiming to characterise the flowering behaviour of accessions, mutants and NILs in five different sites across Europe. These sites were Norwich (UK), Oulu (Finland), Koeln and Halle (Germany) and Valencia (Spain), ensuring a large coverage of the *Arabidopsis* native latitudinal range. The flowering behaviour was evaluated after sowing in the spring, the summer and the fall.

According to the results of this screen, natural variation as expected was detected among the accessions, with larger effects in sites and seasons where day length was below 12 h (Fig 14A). This day length marks the transition to short days and higher overall differences were observed below that threshold also in the controlled environment conditions. The general pattern of the response in the wild was similar to the one that was obtained in controlled environment conditions, with an early plateau in day lengths above 14 h, an enhanced transition phase in intermediate day lengths between 14 and 10 and a late plateau in shorter photoperiods. Nevertheless natural variation defined deviations from this general trend also in the wild. Interestingly the two hyper-responsive accessions Bs-1 and Cen-0, which under controlled environments discriminated LD of 16 from LD of 14 h, showed again an increased ability to respond in the same range. For example, in the fall planting of Norwich the longest day length was LD of 14 and gradually declined to SD of 8 h. Both Cen-0 and Bs-1 flowered very late and their pattern was similar to photoperiod pathway mutants such as *gi* and *ft* consistent with the idea that the long day pathway contributes very little in the determination of flowering under these conditions in both accessions. In contrast in the same conditions Ler, which was identified in our screen as an accession that does not detect any difference between LD of 16 and 14 h, flowered early and with almost no difference in the flowering times during spring and summer. Under SDs Cen-0 shows a higher response compared to Ler, yet the differences are less obvious compared to LDs, just like in the controlled environment experiments. Collectively these data indicate that the difference between Cen-0 or Bs-1 and Ler is more obvious under long days also in the wild and fully confirm the analysis in the cabinets.

Comparing only Norwich spring (day length max : 16 h and min : 10 h) with fall (day length max : 14 h and min : 8 h, Fig 14B) the Ler DRC is almost a horizontal line suggesting

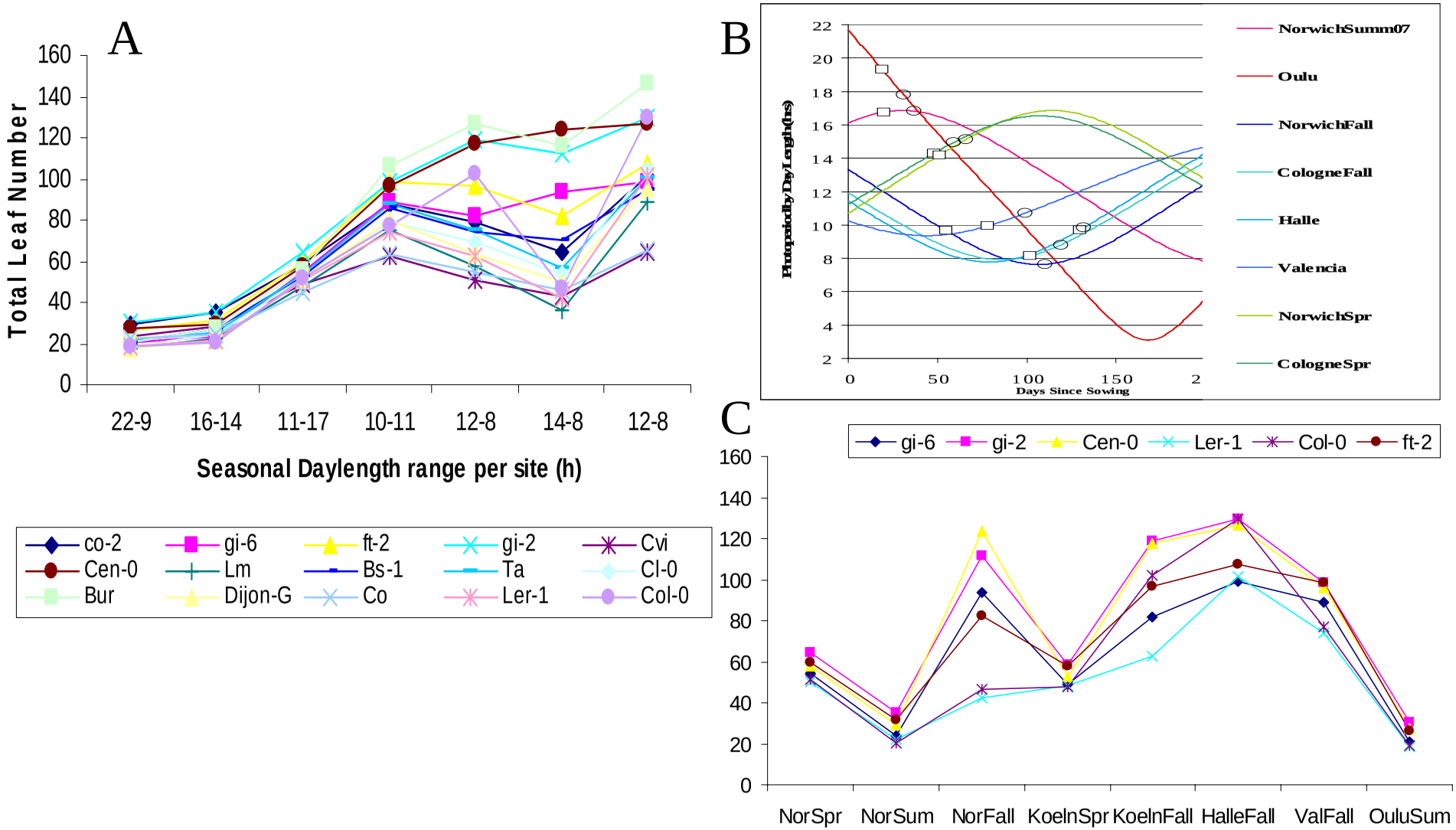


Figure 14 : Comparative flowering time analysis in natural cohorts across Europe. Five sites, namely Norwich (Nor, UK), Koeln (DE), Halle (DE), Oulu (FI), Valencia (Val, ES) were selected and plants were sown in three seasons Spring (Spr), Summer (Sum) and Fall. A) Bolting times of selected accessions and mutants according to sorted daylength range per site and season. B) Photoperiod measurements per site and season during the experiment. Note that in Norwich during fall, photoperiod had a maximum daylength of 14 h, which is already short days for Cen-0 but not Ler. C) Bolting times of selected accessions and mutants per site and season. Note that Cen-0 retains a very strong ability of discriminating between Summer and Fall in Norwich in contrast to Ler. This ability delays flowering for three months, allowing this vernalization insensitive accession to over winter in the vegetative phase.

no response, while in Cen-0 a response above 60 days is evident (Fig 14C). Furthermore when sown in fall Cen-0 requires more than three months to flower while accessions like Ler flower in 40 days. Since both Ler and Cen-0 do not respond to vernalization, this suggests that the enhanced response of Cen-0 to day length alone is enough to prevent flowering during winter, thus protecting the individuals of this accession in a vernalization independent manner.

3.1.1.7 Conclusions

Natural variation was detected in day length perception for flowering time in *Arabidopsis*. The patterns of the Daylength Response Curves were sigmoid, with early plateaus under long and short days and a strong transition phase under intermediate photoperiods. Although the species flowers under long days, significant natural variation was detected also under short days. Interestingly accessions with similar flowering times in commonly used yet physiologically extreme photoperiods of LD 16 h and SD 8 h, revealed significant differences in intermediate photoperiods. Patterns of enhanced response to the changing day length were identified yet no clear association with latitude or climatological parameters was seen. Day length perception and overall lateness are two parameters of flowering time that are not connected. Interestingly comparative PCA analysis between accessions as well as mutants and transgenic lines of known flowering time genes, revealed that the properties of day length discrimination in a group of accessions cannot be explained by the known genes and most probably is the outcome of complex genetic interactions and/or novel regulatory loci.

Surprisingly accessions with enhanced response to almost identical long and short days were identified. Under long days, Cen-0 and Bs-1 discriminate efficiently LD of 14 h as a shorter photoperiod compared to LD of 16 h. Such trends were also observed under short days, with Dijon-G being more efficient in discriminating between SD of 10, 8 and 6, compared to Sha. Both accessions however, shared identical flowering times under LDs, suggesting a SD dependent effect in this discrimination. All these accessions were tested for a response to vernalization, with only the Bs-1 parent showing a moderate response.

The results of the controlled environmental conditions reflect very accurately the pattern of day length discrimination, which is observed in natural conditions. In natural conditions Cen-0 and Bs-1 are able to discriminate more efficiently the changing day length of the fall in northern sites (Norwich) and if sown in the fall delay their flowering in order to overwinter, in contrast to the day length responsive Ler. Interestingly in southern sites (Valencia) the flowering times of these accessions converge, indicating a very small window of day lengths that makes the difference in the wild. These data are in agreement with the LD of 16-14 h

differences that the cabinet data indicated for these accessions. Such results are both unexpected and exciting since it was believed that vernalization was the sole mechanism of winter avoidance regarding flowering in *Arabidopsis*. The flowering time data need to be supplemented with measurement of fitness such as seed yields. Such experiments are in progress. Collectively these data also indicate that standard controlled cabinet experiments although they do not reflect 100 % the natural conditions, are sufficient in describing accurately the behaviour of different genotypes in the wild.

3.1.2 GENETIC ANALYSIS

3.1.2.1 Introduction

One of the main targets of this project was to characterise genetically the observed natural variation in day length discrimination, to describe the genetic architecture of the trait and eventually to isolate the new genes/alleles that underlie it. For this reason several mapping populations were created and QTL mapping was used as a first step toward the genetic characterisation of day length discrimination both under short and long days. Genome wide genetic interactions in the level of F2 populations were determined in order to identify regulatory loci which are usually masked by standard QTL approaches. The results of the QTL mapping were compared with results from many published QTL projects both under standard long and short days, after anchoring them to the *Arabidopsis* physical map. Last but not least advanced genetic material was created in order to validate the QTL mapping, to isolate the responsible genes after fine mapping and to link the genetic analysis with molecular consequences in a spatial and temporal manner. A general scheme of the standard mapping design for the Bs-1 x Ler population is shown in Fig 15.

3.1.2.2 Selecting accessions for the genetic analysis

Based on the Two Way ANOVA, the Hierarchical Clustering and the Principal Component analyses, natural variation was detected both under long days and short days. Interestingly, accessions that had very similar flowering times in extreme day lengths, showed radically different responses to intermediate photoperiods. An example is Cen-0 and Cvi both of which flower with 17 TLN under LD of 16 h. Cvi however, shows almost no response in the remaining day lengths. By sharp contrast, Cen-0 detects LD of 14 h as short days and flowers with 20 more leaves compared to 16 h (Fig 8D). Similarly strong differences are detected between LD of 14 and LD of 12 as well as LD of 12 and SD of 10 h. However this ability in response to day length of Cen-0 is greatly restricted under LDs. No difference was detected between SD of 10 h and SD of 8 h while reduced response was observed between SD of 8 h and SD of 6 h. Another accession, Bs-1 showed a similar pattern to Cen-0. Bs-1 shows a mild response to vernalization and Cen-0 does not. Crosses of these two accessions with Ler, which does not respond to vernalization or photoperiod between LD of 16 and 14 h can help identifying QTL that control the responses between nearly exact LDs and also to evaluate the role of vernalization in such responses.

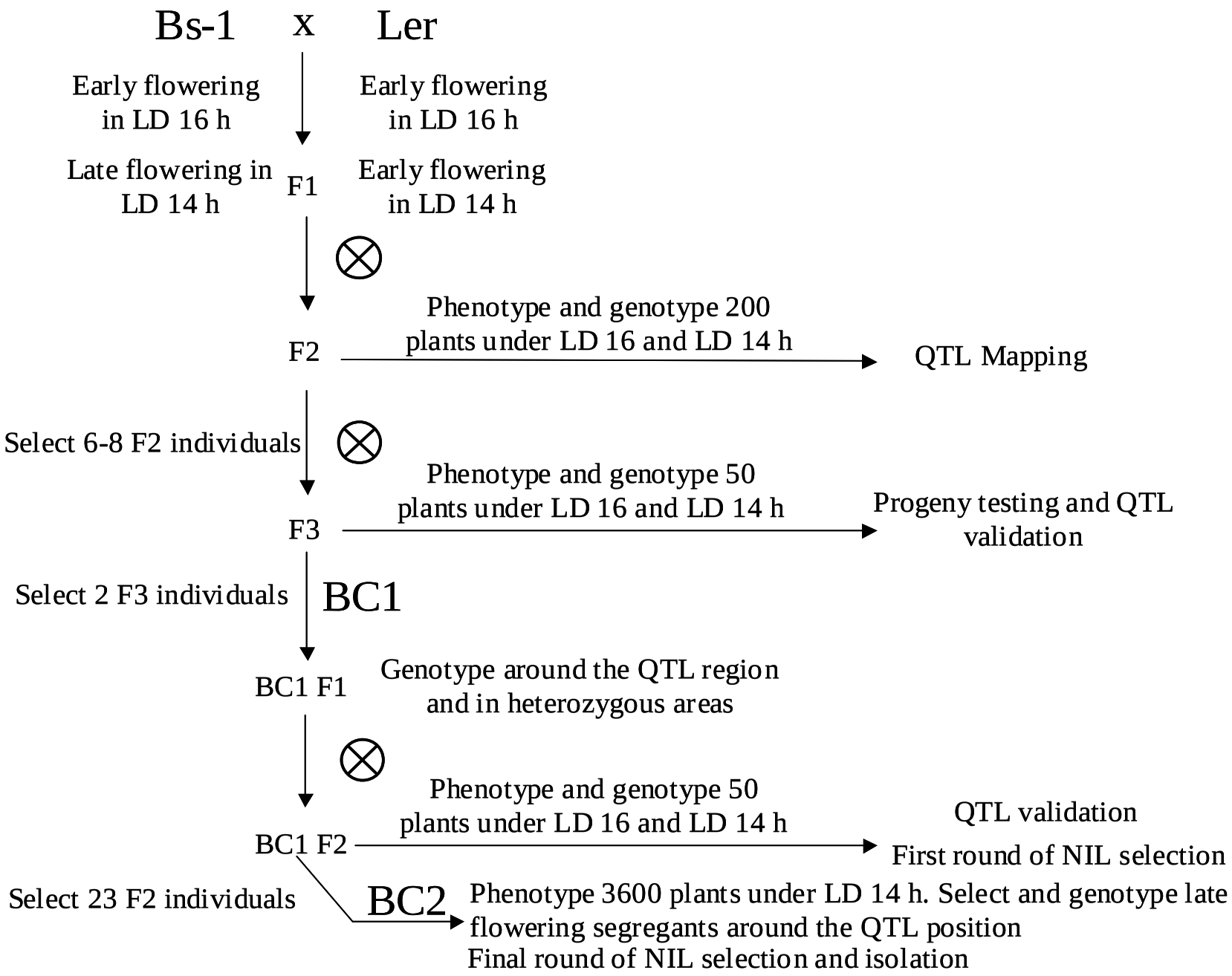


Figure 15 : Scheme of experimental design for the isolation and cloning of the large effect QTL in the Bs-1 x Ler population

Apart from Cen-0 and Bs-1 that show increased day length discrimination specifically under long days, other accessions exhibit different properties of discrimination preferentially under SDs. Sha and Dijon-G share very common flowering times under LDs. However under SDs, Sha exhibits a very weak response while Dijon-G retains ability for discrimination (Fig 8C). The two accessions have 6 leaves of difference under 10 h, 16 leaves under 8 h and 21 leaves under 6 h. Since Sha shows no response from 10 to 6 h, and both accessions share very similar flowering times under LDs this cross can be used to identify QTL that control responses exclusively under all three SDs. RILs derived from such a cross are not available, however genotyped RILs derived from crosses of both accessions with other ecotypes, exist and therefore specific polymorphic markers from both accessions can be tested.

Bay also exhibits similar pattern with Dijon-G. Bay and Sha share a difference of 5 leaves under 12 h (22 and 27 respectively). Sha however flowers with 33 leaves under 10 h, while Bay flowers with 47. Sha is so far the only accession after Cvi, which exhibits a reduced response throughout the SD range. However a Dijon-G x Sha cross will serve as a better alternative of the existing Bay x Sha RILs for this purpose, because the difference of the Dijon / Sha parents under SD 10 h is more than half the one of the Bay / Sha (6 leaves compared to the 14 leaves of the Bay / Sha). With these approaches, the whole range of day lengths can be covered and QTL with general or very specific roles in day length discrimination identified. The use of one common parent under LD (Cen-0 x Ler and Bs-1 x Ler) allows the use of non-linear mixed models and increases the probability of identifying epistasis.

3.1.2.3 Genetic dissection of natural variation in day length perception – Discrimination under similar long days.

In order to identify QTL that are responsible for the late flowering under LD of 14 h compared to 16 h, Bs-1 and Cen-0 were crossed to Ler. The resulting two F2 populations were phenotyped under LD of 14 h and LD of 16 h. Since Ler shows no difference between these two day lengths, late flowering plants under LD 14 h in the F2 progeny would be due to QTL from the Bs-1 or Cen-0 genome or due to epistatic events between elements of the two parental genomes. In the latter case, transgressive variation can be expected.

As shown in Fig 16, under LD of 14 h without vernalization, the Bs-1 x Ler population revealed a late flowering tail starting from approximately 20 leaves and extending up to 35 leaves. This late flowering tail overlaps with the flowering time of Bs-1 under LD 14 h (32 leaves). Both the Bs-1 x Ler and the Ler x Bs-1 F1 hybrids showed intermediate flowering

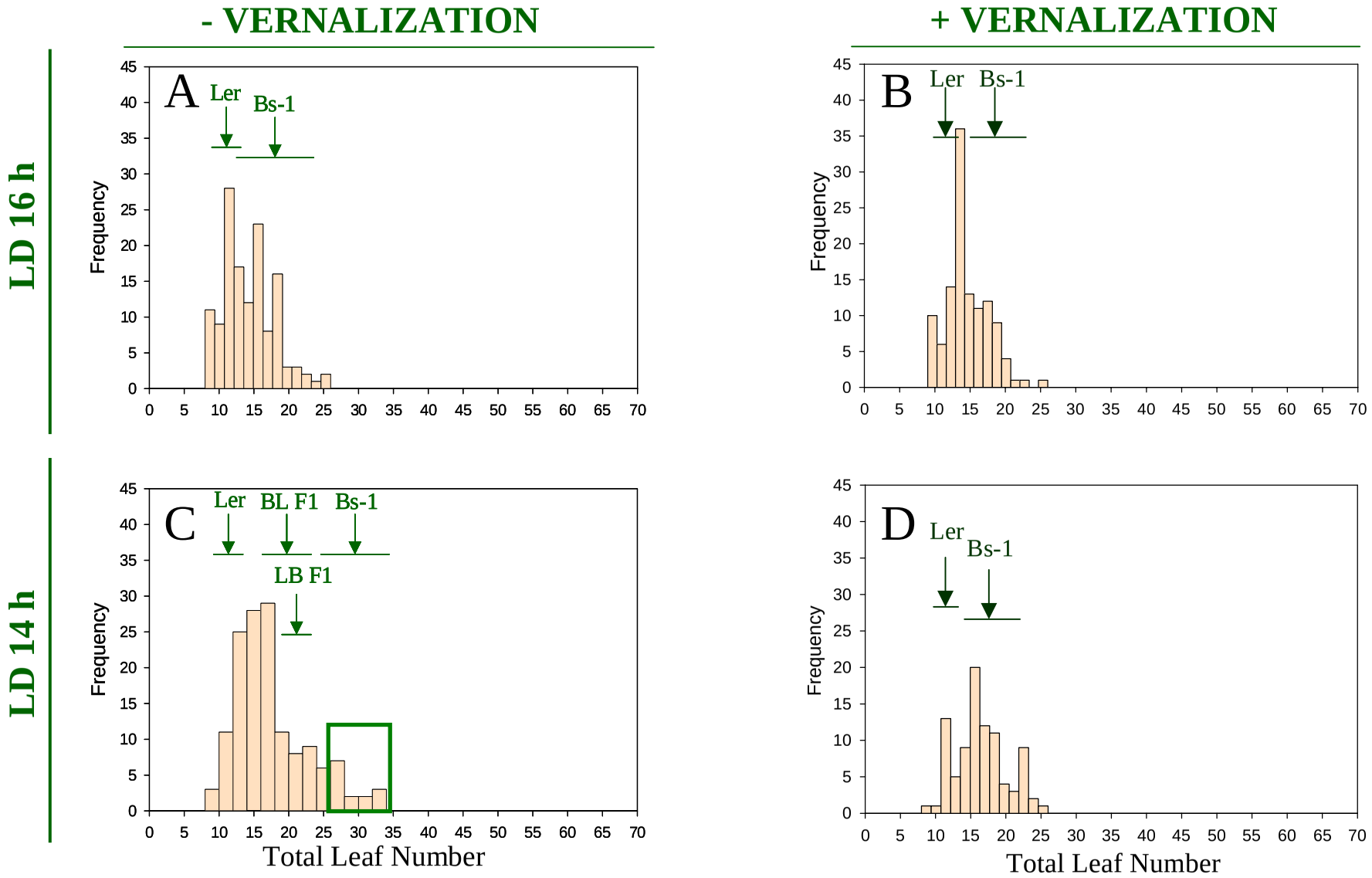


Figure 16 : Frequency distributions of the F2 population for the Bs-1 x Ler cross under various conditions. A) LD 16 h, –Vernalization B) LD 16 h, +Vernalization C) LD 14 h, -Vernalization D) LD 14 h, +Vernalization. The green box indicates the LD 14 h-specific late flowering. Although vernalization results in identical range under both daylengths, late flowering bins can be identified under LD 14 h compared to LD 16 h. Transgressive variation was not detected while the trait segregates as semi-dominant based on the F1 phenotypes under LD 14 h.

times compared to the two parents, which approach the peak of the segregation (20 leaves), suggesting semi-dominance. Interestingly, the segregation under LD of 16 h is almost identical to the one under 14 h up to approximately 20 leaves. However, the lateness that was observed under 14 h between 20 and 35 leaves is significantly reduced. Therefore it is most probably caused by elements of the Bs-1 genome that are responsible for detecting the changing day length, and not for causing late flowering in general. In contrast, since the two segregation patterns are almost identical up to 20 leaves, this part of the segregation is likely to be controlled either by QTL, that are not regulated by photoperiod or by elements of the Ler genome exclusively.

Since Bs-1 responds to vernalization, the effect of exposure to cold temperatures for one month in the segregation both under LD 14 and 16 h was tested. Vernalization reduced the lateness that was observed between the bins of 20 and 35 leaves under LD 14 h (Fig 16C). The resulting segregation pattern was similar to the one under LDs of 16 h without vernalization. Therefore vernalization under 14 h was sufficient to phenocopy the effect of LDs of 16 h and to reduce photoperiod discrimination between the two day lengths. In order to test whether, this effect was restricted to 14 h, the segregation was tested under 16 h after vernalization. No difference was observed in the segregation between long days of 16 h with or without vernalization (Fig 16 B and D). To summarize, long days of 14 h cause specific late flowering between 20 and 35 leaves. Longer days of 16 h or vernalization act specifically to reduce this observed late flowering. However under LD of 16 h, vernalization fails to trigger earlier flowering. The above suggests that the vernalization pathway could have a role in the mechanism that enables the observed response of Bs-1 between LD of 16 and 14 h, and suggests the existence of a sensitive equilibrium between the two pathways in natural populations, as the day becomes longer.

The segregation of an F2 population derived from a Cen-0 x Ler cross was also tested under LD 14 h. Although the flowering times of the parents Bs-1 and Cen-0 are similar both under 14 and 16 h, the two segregation patterns are different (compare Fig 16 and 17). Similarities were restricted only up to the bin of 20 leaves, confirming that these flowering time responses are caused most probably by the Ler genome and thus are independent of the two accessions, photoperiods and vernalization treatments. This part of the segregation includes also the main peak. However, compared to Bs-1 x Ler segregation, the Cen-0 x Ler pattern shows a large amount of lateness under 14 h. This lateness forms a second peak, which fits to the flowering time of the Cen-0 parent. In addition, extensive transgressive variation was detected around this peak. Interestingly both the Cen-0 x Ler and the Ler x Cen-

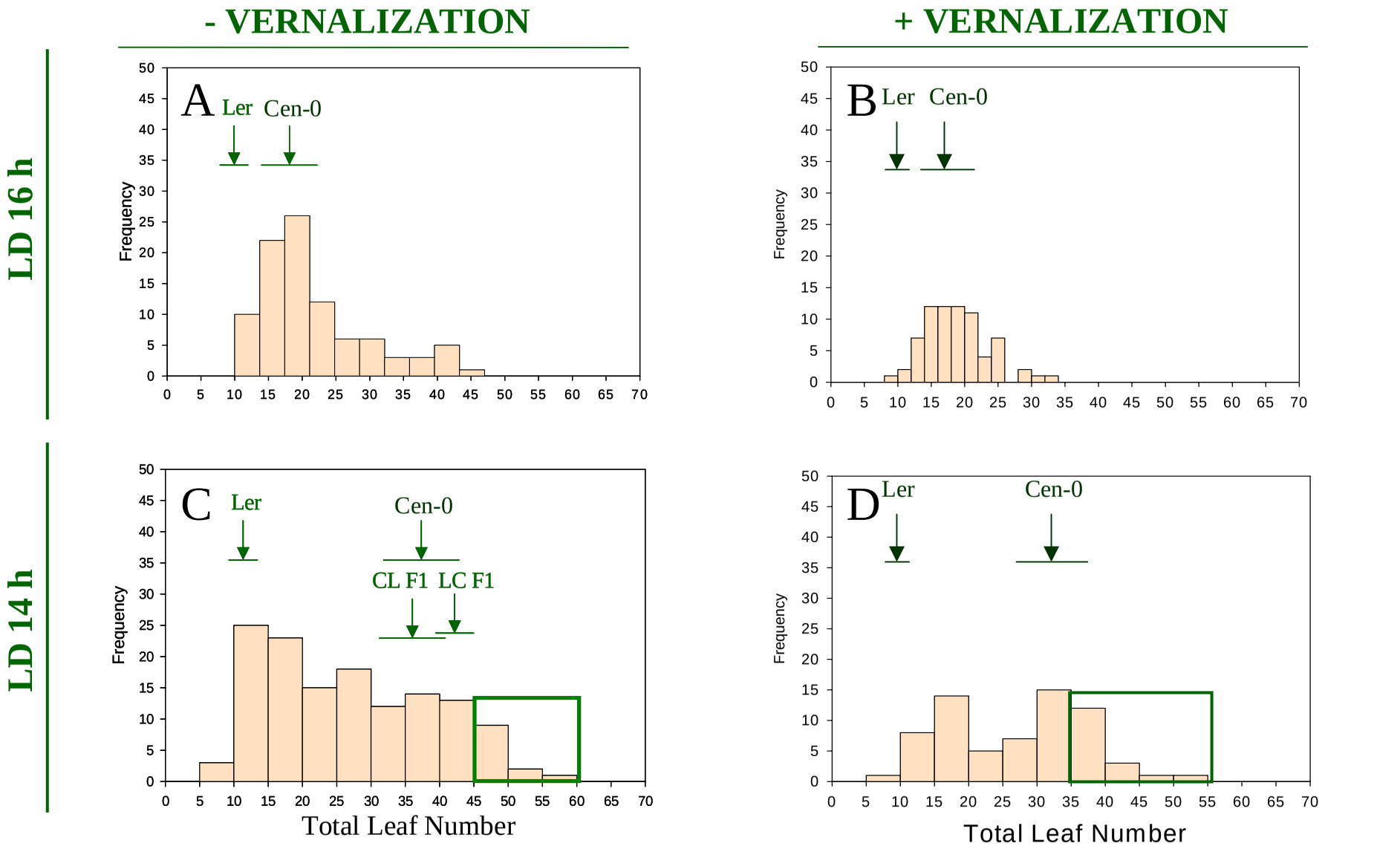


Figure 17 : Frequency distributions of the F2 population for the Cen-0 x Ler cross under various conditions. A) LD 16 h, –Vernalization B) LD 16 h, +Vernalization C) LD 14 h, -Vernalization D) LD 14 h, +Vernalization. The green box indicates the LD 14 h-specific late flowering. Although vernalization results in earlier flowering under both daylengths, late flowering specific under LD 14 h, can be identified indicating that vernalization is not responsible for the LD 14 h late flowering. Transgressive variation towards later flowering was detected under both daylegths and was rescued by vernalization indicating that restoration of the vernalization response occurred after the cross and does not occur in the parents. The trait segregates as dominant for Cen-0 based on the F1 phenotypes under LD 14 h.

0 F1 hybrids were as late as the Cen-0 parent suggesting the existence of at least one major dominant QTL to be among the causes of the 14 h lateness.

Under LD of 16 h the segregation was again similar up to the bin of 20 leaves. Lateness caused by transgressive variation was also observed with two recombinants flowering with 47 leaves. However the frequency of the late flowering recombinants was reduced compared to the 14 h while there was no plant within the range of 50 to 60 leaves, in contrast to LD 14 h. Together these data suggest that a portion of the transgressive variation under LD 14 h up to the bin of 40 leaves was caused by epistatic interactions between Ler and Cen-0 loci,and this causes lateness under both 14h and 16 h. Since the number of the recombinants within this range was reduced under LD 16 h, the remaining variation is controlled by QTL that discriminate between the two conditions. Finally, the LD 14 h specific transgressive variation within the range of 50 to 60 leaves is controlled exclusively by a combination of QTL that cause extreme late flowering and respond to LDs 14 h.

Although the Cen-0 parent did not show a response to vernalization, the pattern of segregation was tested in both conditions after one month of exposure in cold temperatures. Interestingly in the level of the F2 population a response to vernalization was restored. In contrast to the Bs-1 population, the vernalization response of the Cen-0 population was not conditional and affected the segregation both under LD of 16 and 14 h. However under LD of 14 h, vernalization failed to rescue the late flowering below 35 leaves (Cen-0 parental range) and instead a second peak in the segregation was observed at this flowering bin. This treatment rescued the transgressive variation between 40 and 60 leaves. In addition between 14 h and 16 h after vernalization differences were evident (in contrast to the Bs-1 population) collectively suggesting that vernalization is not responsible for the discrimination of the two day lengths in the case of Cen-0. This observation is consistent with the fact that Cen-0 does not respond to vernalization.

3.1.2.4 Genetic dissection of natural variation in day length perception – Discrimination under similar short days.

Under SD two accessions, Sha and Dijon-G were selected and the resulting mapping population was phenotyped under SD of 8 and 10 h without vernalization (Fig 18). Flowering time was estimated both as number of leaves and days to flower since under SD the two flowering time estimates deviate more compared to LD. For leaf number under SD 10 h a segregation pattern with a single peak was observed. The same was evident for SD of 8 h. However in both cases extensive transgressive variation was observed and the resulting

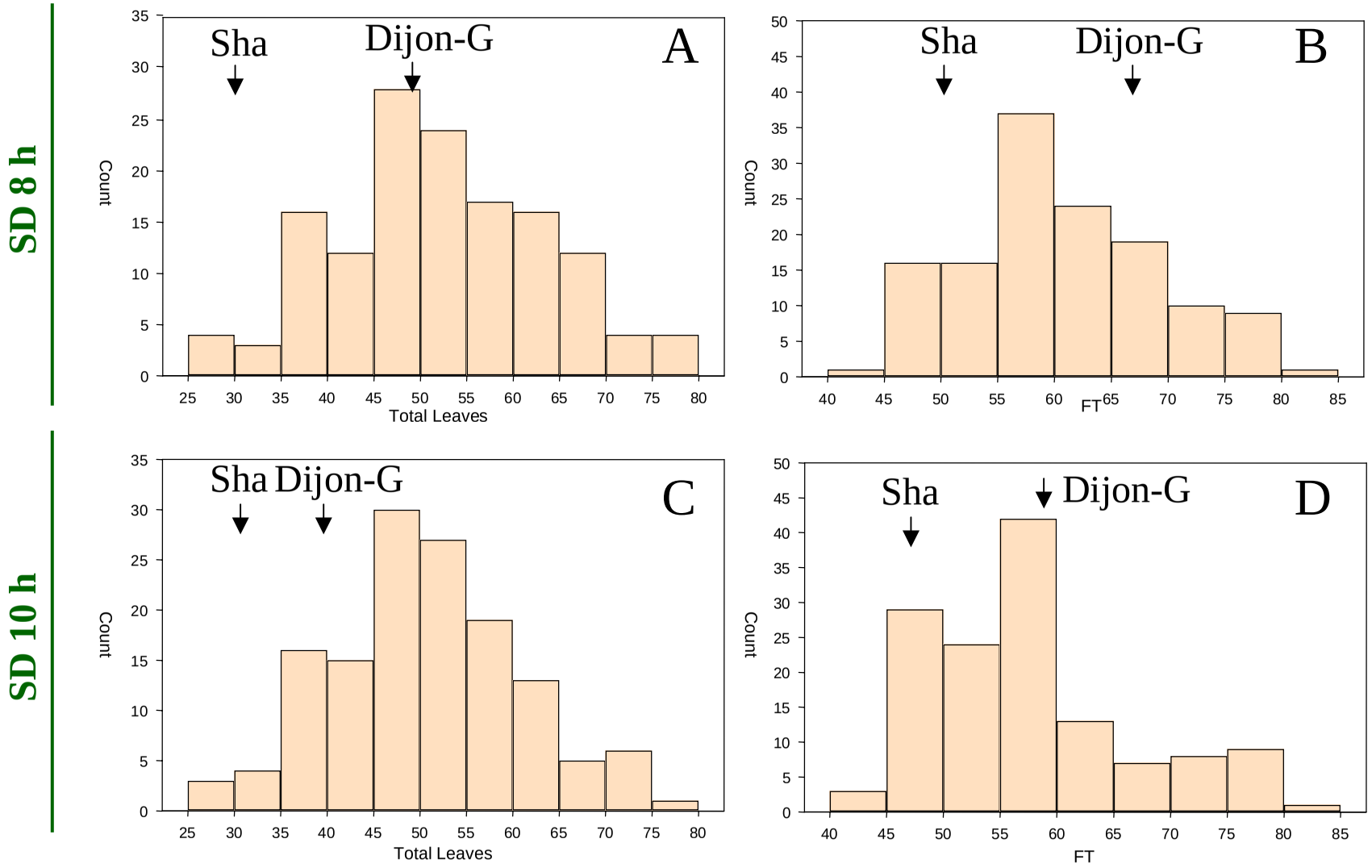


Figure 18 : Frequency distributions of the F2 population for the Dijon-G x Sha cross under various conditions. A) SD 8 h, TLN B) SD 10 h, TLN C) SD 10 h, BT D) SD 8 h, BT. Transgressive variation towards later flowering was detected under both daylegths but it was more obvious under SD 10 h.

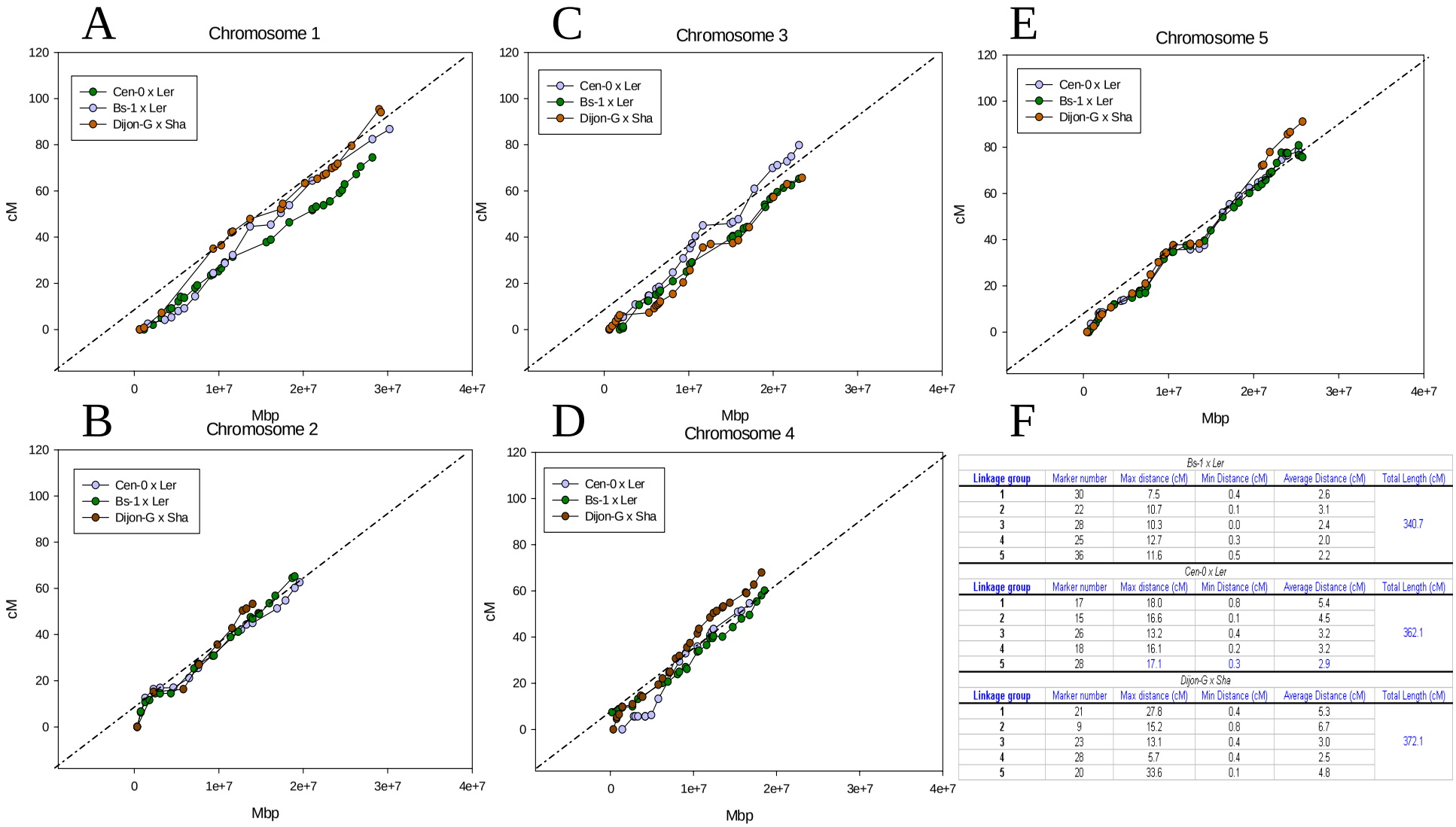
histograms were nearly identical. Such trangressive variation is often indicative of epistasis between multiple loci that regulate the trait. Interestingly between the flowering time range of 60 to 80 leaves differences could be detected between the two conditions with SD 8h having higher frequencies. This result indicates that at least part of this transgressive variation is conditional.

Regarding days to bolting, the resulting histograms were very similar to the leaf number data. Again a monophasic segregation was observed in both conditions, while transgressive variation was extensive. The observed pattern of segregation was similar between the two day lengths but again differences were evident in the transgressive range of 60 to 85 days to bolting. Short days of 10 h were able to significantly rescue the late flowering of that range while under SD of 8 h higher frequencies were observed. Collectively these data indicate that extensive transgressive variation masks the differences between the two days lengths and indicate the presence of multiple loci segregating in the population. Conditional epistasis might also regulate the observed differences in the transgressive range both for leaf number and days to bolting.

3.1.2.5 Genetic dissection of natural variation in day length perception – Construction and properties of linkage maps.

The next step after phenotyping the three mapping populations under nearly identical long and short days was genotyping. For that reason three linkage maps were constructed using markers that were polymorphic in specific crosses. DNA was extracted from populations growing under LD of 14 h for the Bs-1 and Cen-0 populations and under SD of 8 and 10 h for the Dijon-G x Sha population. The plants were scored for flowering time and genotypied in collaboration with Sequenom Inc. (San Diego), a genotyping service provider. A correlation between the physical and the genetic distance per chromosome for each linkage map, revealed absence of large segregation distortions while local distortions where observed as expected around the centromere on each chromosome (Fig 19). In fact comparing the linkage map properties of this project with previously published maps (data not shown), one can see a much better analogy between the physical and genetic linkage, most probably due to the presence of many molecular markers on each chromosome.

In the Bs-1 population the linkage map consists of 141 markers while in the Cen-0 population and in the Dijon-G x Sha 104 and 101 markers respectively form the linkage map. The average genetic distance between the markers in the Bs-1 population is between 2 and 3.6 cM across the five chromosomes while the total length of the map is 340.7 cM, values



F

Bs-1 x Ler

Linkage group	Marker number	Max distance (cM)	Min Distance (cM)	Average Distance (cM)	Total Length (cM)
1	30	7.5	0.4	2.6	
2	22	10.7	0.1	3.1	
3	28	10.3	0.0	2.4	340.7
4	25	12.7	0.3	2.0	
5	36	11.6	0.5	2.2	

Cen-0 x Ler

Linkage group	Marker number	Max distance (cM)	Min Distance (cM)	Average Distance (cM)	Total Length (cM)
1	17	18.0	0.8	5.4	
2	15	16.6	0.1	4.5	
3	26	13.2	0.4	3.2	362.1
4	18	16.1	0.2	3.2	
5	28	17.1	0.3	2.9	

Dijon-G x Sha

Linkage group	Marker number	Max distance (cM)	Min Distance (cM)	Average Distance (cM)	Total Length (cM)
1	21	27.8	0.4	5.3	
2	9	15.2	0.8	6.7	
3	23	13.1	0.4	3.0	372.1
4	28	5.7	0.4	2.5	
5	20	33.6	0.1	4.8	

Figure 19 : Properties of the constructed linkage maps. A) to E) Alignment of the genetic vs the physical position of the markers for chromosomes 1 to 5 for the flowering time mapping populations. No severe segregation distortions were observed. F) Table with the descriptive statistics of each linkage map.

comparable to the published data. In the Cen-0 population the average distance is between 2.9 and 5.4 cM across the five chromosomes and the total genetic distance is 362.1 cM. Finally in the Dijon-G x Sha population the average distance is between 2.5 and 6.7 cM while the map spans 372.1 cM. Collectively the above data indicate that highly saturated genetic maps without significant segregation distortions are used in this project, with descriptive properties similar to what has been published in *Arabidopsis*.

3.1.2.6 Genetic dissection of natural variation in day length perception – QTL detection under long days.

The highly saturated genetic map of the Bs-1 x Ler population was used in order to map one major QTL on the short arm of Chromosome 5, just upstream of the centromere (Fig 20). Interval mapping resulted in a preliminary map position spanning 10 cM while MQM mapping restricted this area to approximately 3 cM, between 6.5 and 7.5 Mbps. However, in the initial linkage map there was lack of markers between 7.5 MBps and 10 Mbps and therefore the QTL could be located further down on the chromosome. This QTL accounts for more than 30 % of the observed variability in that population under LD 14 h. Additional minor QTL were also detected on Chromosome 5, 1 and 4. In total QTL mapping in that population could explain 50 % of the observed variability.

In the case of Cen-0, interval mapping resulted into two QTL at the top of chromosomes 4 and 5 (Fig 21). These preliminary map positions spanned nearly half of each of the chromosomes and since the genetic map is saturated, the probability of having more than one QTL in these regions was high. Indeed MQM mapping dissected these large QTL into two smaller QTL on chromosome 5 and three QTL on chromosome 4. The QTL on top of chromosome 4 extends from the beginning of the chrosomome until 3 Mbps and accounts for 15 % of the observed variability in the population. The additional two QTL on the same chromosome extend from 9 to 12 Mbps and from 12.5 to 15.7 Mbps and each accounts for approximately 5 % of the variability. On chromosome 5 the first QTL extends from 3.6 to 5.7 Mbps and accounts for 30 % of the observed variability while the second one extends from 13.6 to 16 Mbps and accounts for 7 %. Collectively the identified QTL in the Cen-0 population additively account for 40 % of the total variability. Note that approximately 1 Mbps separates the two QTL on top of chromosome 5 in the two populations (Bs-1 and Cen-0). Since the two linkage maps consist of common markers in that region, this difference indicates that the QTL present in different populations are not identical.

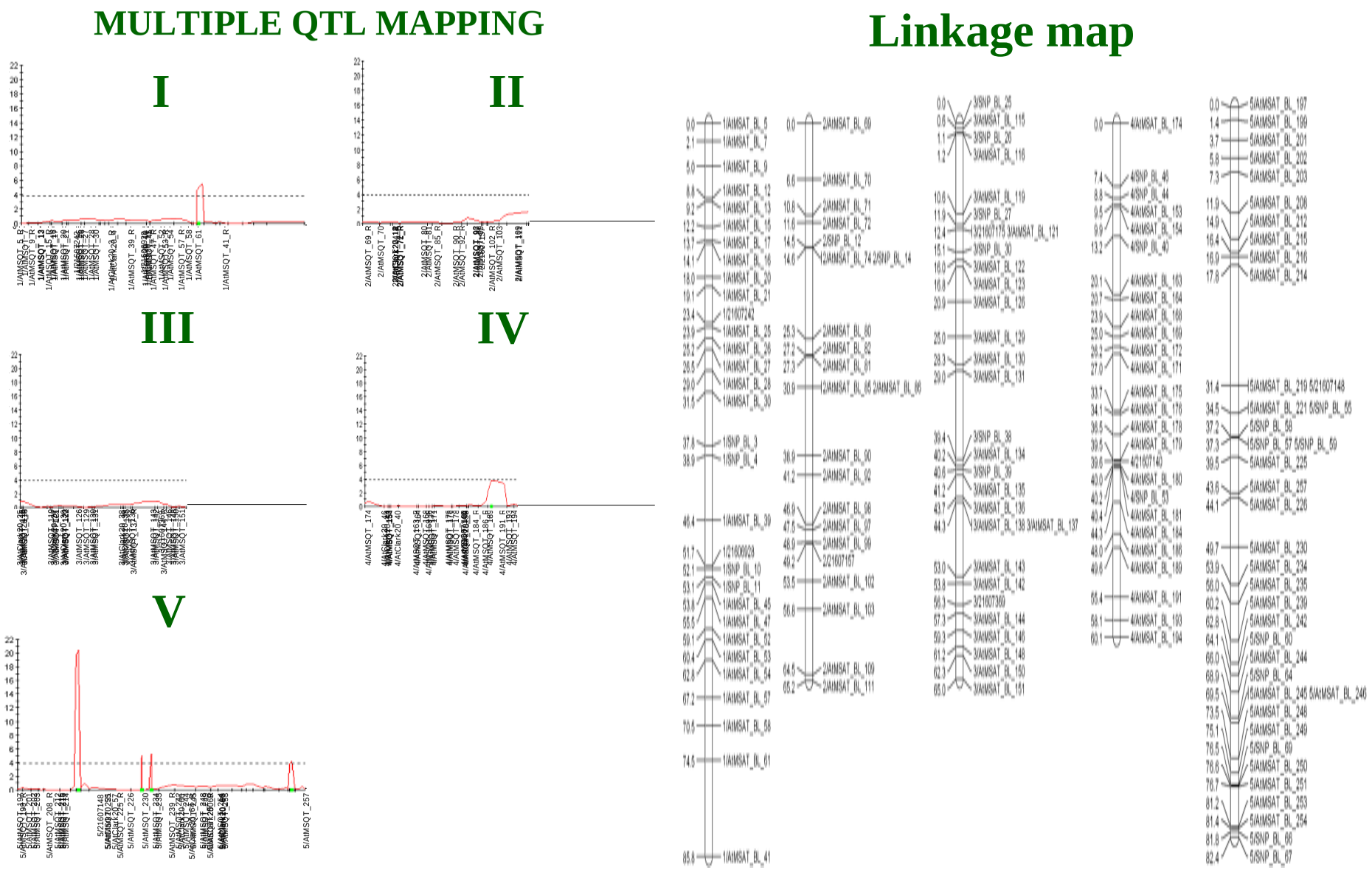


Figure 20 : QTL and Linkage map for the Bs-1 x Ler population under LD 14 h.

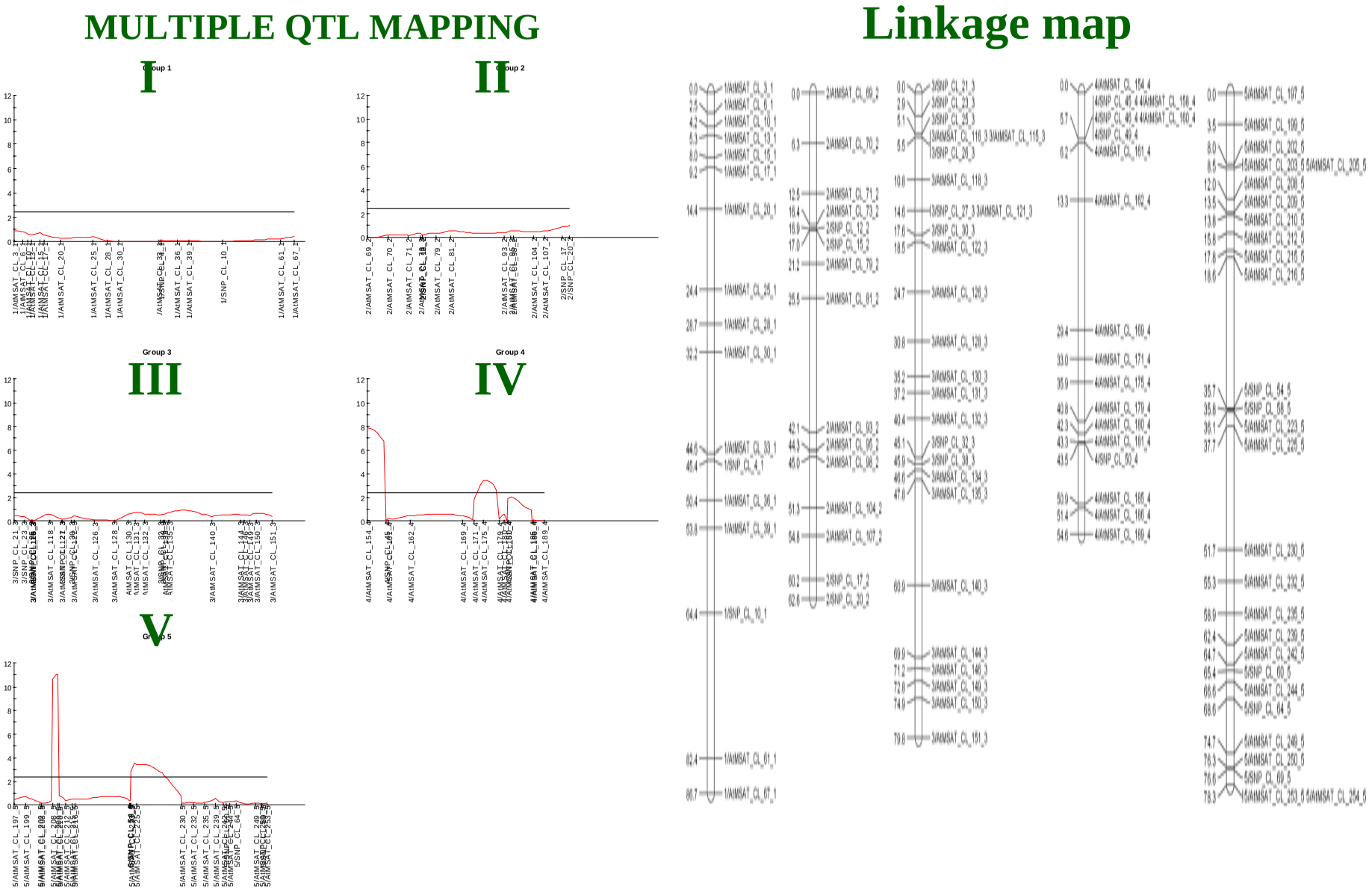


Figure 21 : QTL and Linkage map for the Cen-0 x Ler population under LD 14 h.

SD 8 h MULTIPLE QTL MAPPING

Linkage map

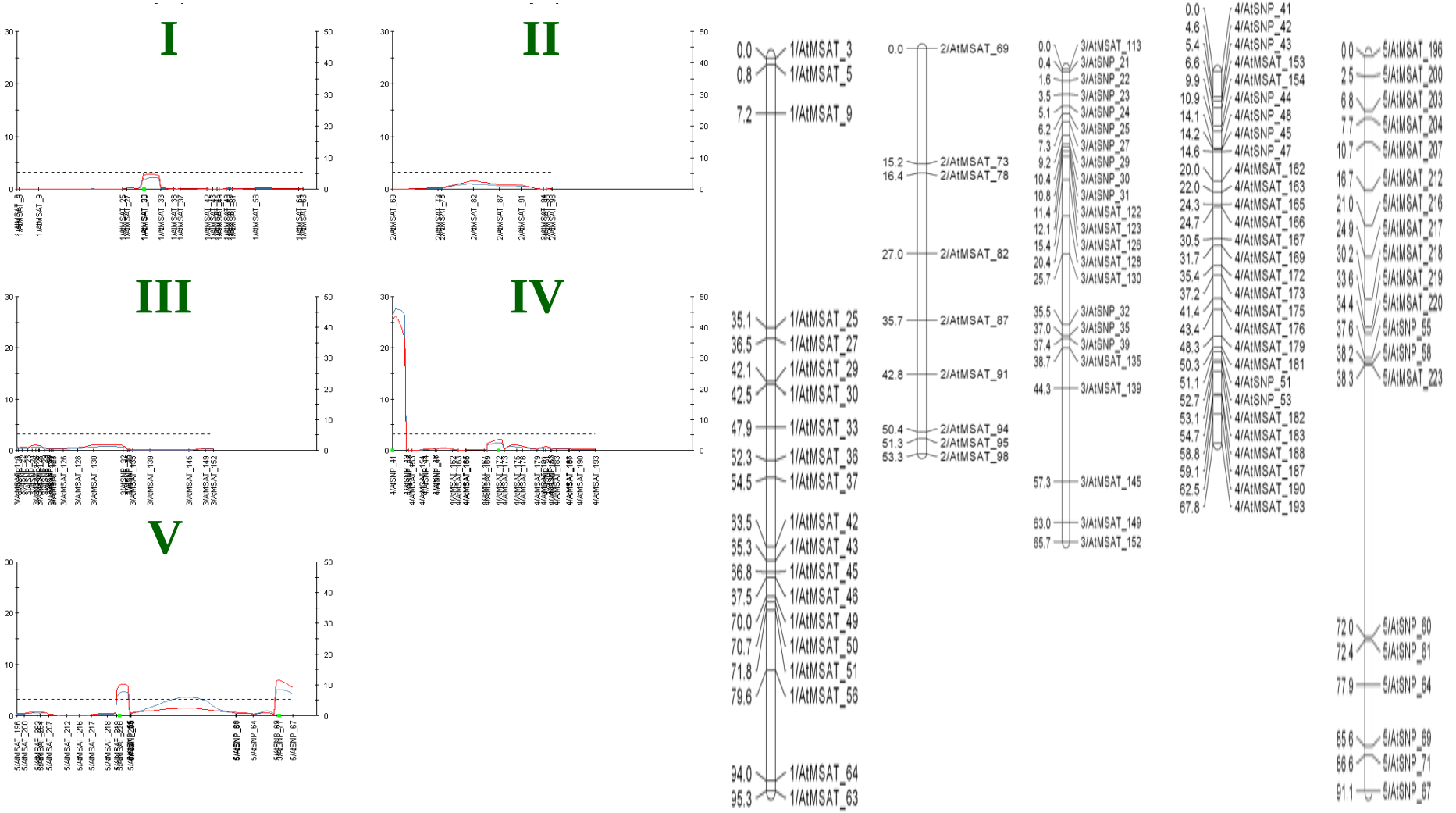


Figure 22 : QTL and Linkage map for the Dijon-G x Sha population under SD 8 h.

SD 10 h MULTIPLE QTL MAPPING

Linkage map

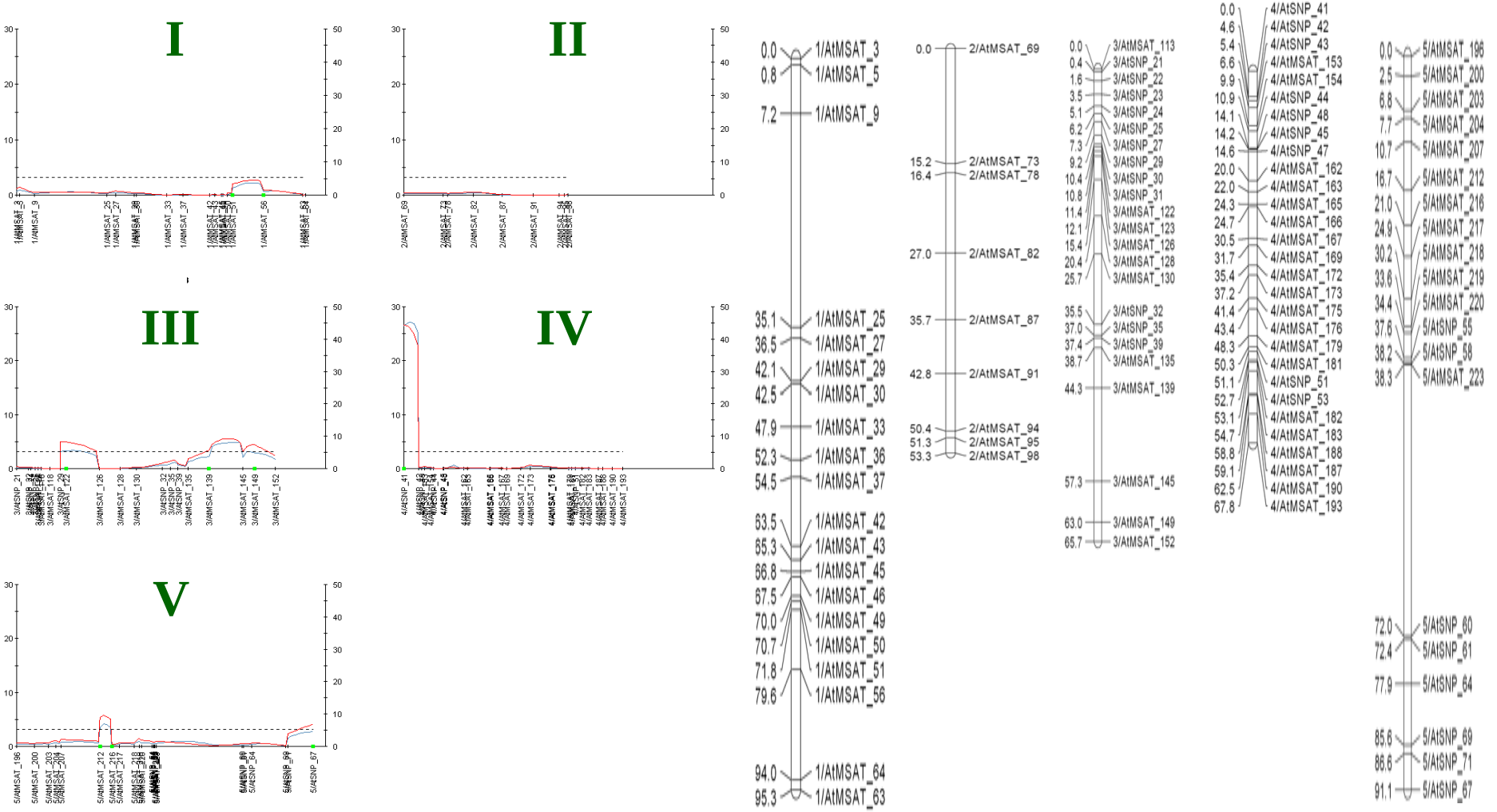


Figure 23 : QTL and Linkage map for the Dijon-G x Sha population under SD of 10 h. Note the existence of the conditional QTL on chromosome 3 , which appear only in SD 10 h in contrast to the rest of the QTL that are identical between the two conditions.

3.1.2.7 Genetic dissection of natural variation in day length perception – QTL detection under short days.

The Dijon-G x Sha population was grown and scored for flowering time under SD 8 and 10 h. DNA was extracted for genotyping from plants grown under both conditions. Under SD of 8 h four QTL influencing leaf number on chromosomes 1, 4 and 5 where detected (Fig 22). For days to bolting similar results were obtained (data not shown). The QTL on chromosome 1 is located between 11.4 and 13.7 Mbps and accounts for 5 % of the observed variability. In addition the QTL at the top of chromosome 4 is located between 0.3 and 1 Mbps and accounts for 45 % of the variability. The second QTL on the same chromosome is located between 8.3 to 9.5 Mbps and accounts for 3 % of the variability. On chromosome 5 the first QTL is located between 8 and 12 Mbps and accounts for 7 % of the observed variability. The second QTL on the same chromosome is located between 23 and 25.7 Mbps and accounts for 10 % of the observed variability. Additively under SD of 8 h the identified QTL account for 70 % of the observed variability.

Under SD of 10 h the same QTL on chromosomes 1, 4 and 5 as for SD of 8 h were detected (Fig 23). However two conditional QTL on chromosome 3 appeared specifically under SD 10 h. The first QTL is located between 5.9 and 8.1 Mbps and accounts for 5 % of the observed variability, while the second QTL is located between 17 and 23.4 Mbps and accounts for 7 % of the observed variability. Additively the identified QTL under SD of 10 h explain 75 % of the observed variability.

3.1.2.8 Genetic architecture of natural variation in day length perception –Genome wide genetic interactions.

Traits such as flowering time are under multilocus control and very often inter-allelic interactions (interactions between genes) are important for shaping the phenotypic character. These sites are often under the control of natural selection and therefore genetic variation is evident. Classical QTL mapping projects cannot detect such interactions directly since the statistics on which such tests are based are similar to one-way ANOVA and therefore the effect of a single factor (gene) with many characters (alleles) can be tested. Genetic interactions require a two-way ANOVA approach since pairwise comparisons of two factors (genes) with many characters (alleles) are needed to determine the phenotypic effect. Usually after a standard QTL mapping project, individuals the alleles of which differ in the identified QTL, are selected and their phenotypes are investigated for the presence of significant

interactions. If such cases of interactions are identified the next step is to confirm them genetically by creating NILs and crossing them.

However three caveats lie in this approach. First proper survey of such types requires many individuals with allelic differences and markers spanning the QTL. With many QTL under question, constructing such files is time consuming. Having many individuals with low variability within the categories increases statistical power of the detection and many markers ensure accuracy, validate the results and increase resolution in the region. Therefore such analysis is suited for RIL populations but often in F2 it is difficult to apply. Last and most significant of all, this type of analysis is totally biased since the hypothesis that only the identified QTL regions interact with each other is made.

In this project a method for genome-wide search for genetic interactions was applied at the level of the F2 population (Fig 24). The whole genome of each individual in the population was interrogated using the full linkage maps of each population as anchors. The benefits of this analysis are its speed and accuracy even in the level of F2, together with the reliability and confirmation of the standard QTL mapping. A disadvantage is the high computation power that is needed for the analysis. All populations were analysed with this method and results for selected cases are shown (Fig 25). However additional computational power is need for determining the significance threshold in some of them. Interestingly by this approach conditional epiQTL can be detected in some crosses. Such a case is one highly significant epiQTL in case of Dijon-G x Sha cross. This epiQTL is located between two loci, one at the top of chromosome 5 (1.2 to 2.1 Mbps) and a second one at the middle of the same chromosome (13.6 to 17.1 Mbps). Interestingly this epiQTL appears only under SD of 8 h while under SD 10 h it is absent. Its phenotypic effect is significant since it delays flowering time by approximately ten leaves. The epiQTL was confirmed also with the days to bolting (data not shown). Smaller effect epiQTL have been identified also in the Cen-0 x Ler population. Apart from conditional epiQTL, differences may exist also in epiQTL between different traits.

3.1.2.9 Genetic architecture of natural variation in day length perception –Comparative QTL analysis.

So far two approaches have been selected for the analysis of the genetic architecture of day length perception in flowering time. First classical mapping projects were performed in mapping populations under various day lengths. When possible these populations had a common parent in order to discriminate the effect of the uncommon reciprocal parental

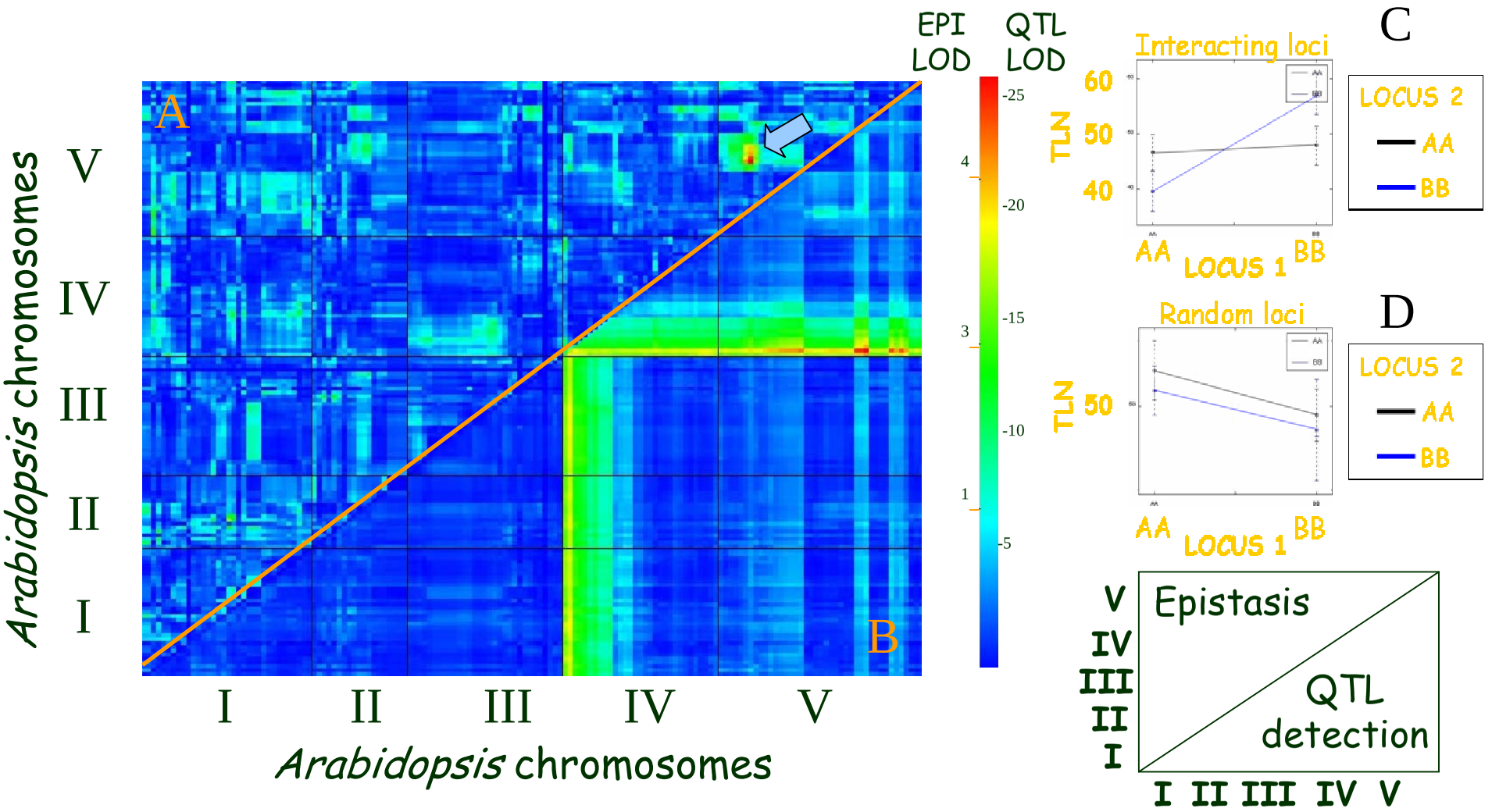


Figure 24. Identifying genetic interactions in a genome wide range in F2 populations. The knowledge of genetic interactions is based on genetic studies between known genes or NILs. However it is not always feasible to isolate the genes while the production of NILs is time-consuming. This lack of knowledge in the genetic architecture of flowering time is expanded, utilizing the highly saturated linkage maps of this project. A) Upper triangle showing a genome (Y axis) by genome (X axis) search for two-loci interactions affecting the trait of interest. In this particular case an interaction between a locus at the top of Chromosome 5 and a locus in the middle of chromosome 5 was statistically significant (arrow), according to the EPI LOD score (left scale). The interaction between alleles from both loci delays flowering by 10 leaves (C), while interaction between two randomly peaked loci has no effect on flowering time (D). B) Triangle showing a standard QTL mapping represented in a heat-mapped way. Note that no standard QTL is identified in the position of the interaction, indicating the limitations of QTL mapping.

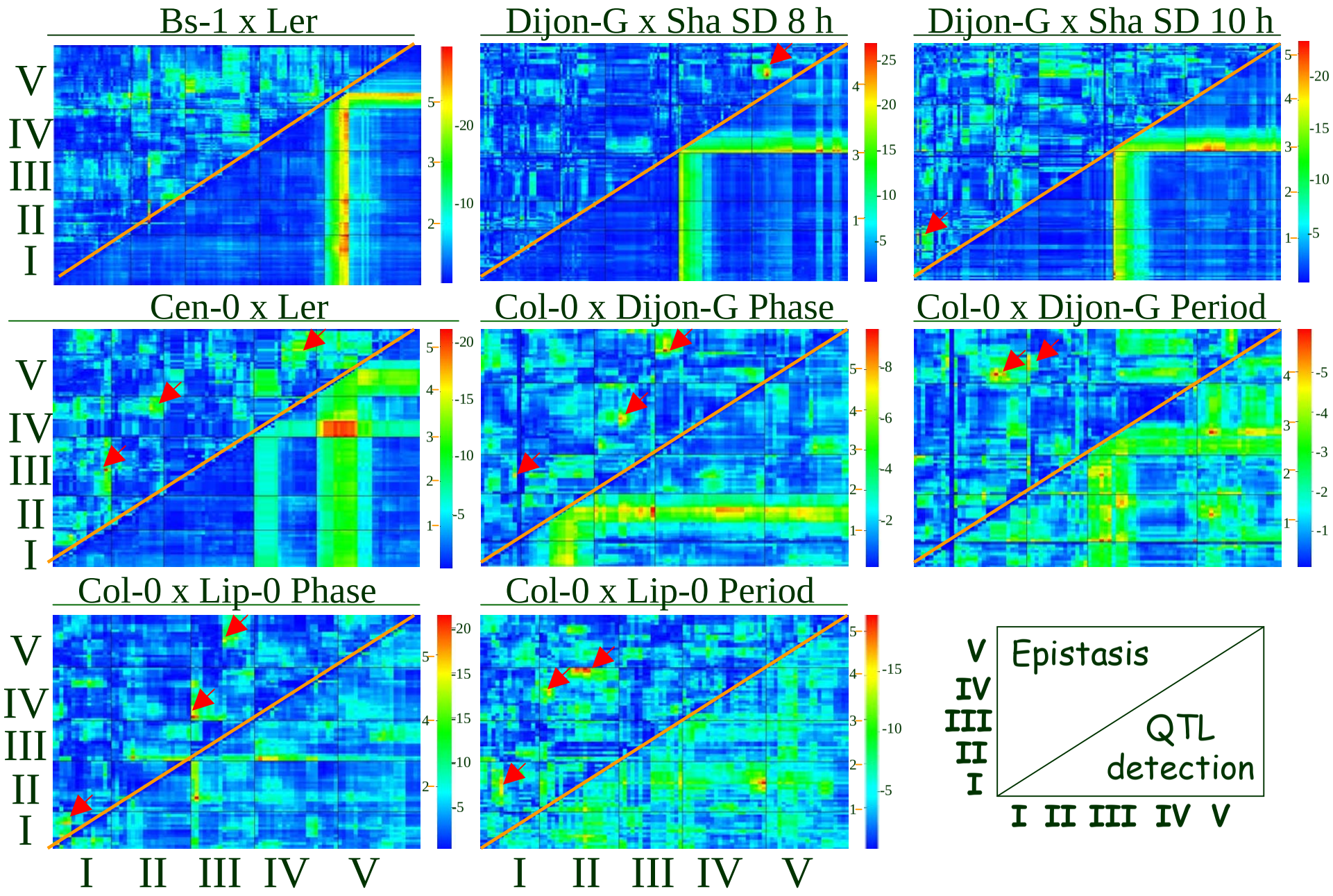


Figure 25 : Genetic interactomes for some populations and traits of this project. Red arrows indicate the presence of significant interactions (see text for details).

genomes on the trait. The second approach was to detect genetic interactions on a global, unbiased scale and to identify conditional and between traits epiQTL together with the classical QTL. These approaches helped in fully describing the genetic architecture of the day length perception under the unusual conditions of LD of 14 h and SD of 10 h of this project. However in order to approach a better understanding of the genetic architecture of day length perception for flowering time, insights must be gained by comparative analysis between the results of this project and published results between various accessions and conditions.

For that reason 9 mapping populations together with the 2 of this project were anchored in the physical map of *Arabidopsis* for long days. In addition two additional mapping populations together with the Dijon-G x Sha population were anchored as representatives of SD analysis. The eleven populations of long days are balanced since approximately half of them (5) have Ler as a common parent. Therefore such populations can be used to discriminate Ler specific alleles that act as source of natural variation. The remaining populations (6) are between random accessions and can be used in order to detect similarities between many different accessions. Approximately half of the populations are phenotyped after vernalization and this can buffer the effects of the vernalization pathway in the QTL detection. Last but not least all of the selected populations apart from the Bs-1 x Ler and the Cen-0 x Ler are phenotyped under LD of 16 h and this can help distinguishing QTL, which are specific for the LD of 14 h condition. Under SD all of the populations have Sha as a common parent, while phenotyping apart from SD of 8 and 10 h also includes intermediate LD of 12 h but also LD of 16 h in order to bridge day length perception under SDs with day length perception under LDs.

Under long days low accumulation of QTL is detected in the first three chromosomes of *Arabidopsis* (Fig 26,27,28). Exception is the appearance of four QTLs at the bottom of chromosome 1 mainly in Ler crosses (Fig 26) and three QTL at the top of chromosome 3 again in Ler crosses (Fig 28). These three QTL might be either conditional since they appear only under LD of 16 h or the outcome of special conditions that the corresponding study applied. By sharp contrast the picture changes dramatically in the remaining two chromosomes, which accumulate large numbers of QTL from various populations. Intra-chromosomal structure seems to be present since on chromosome 4 two regions, the very top and the middle seem to mainly accumulate the QTL (Fig 29). The same applies also for chromosome 5 where the bottom and not the middle, together with the top accumulate almost exclusively the identified QTL. Interestingly the two mapping populations used in this project share common QTL only in the middle of chromosome 4, in contrast to the rest 9 populations.

Regarding chromosome 5 only the Cen-0 population shares one QTL at the top of this linkage group with almost all the rest populations (Fig 30). Both Bs-1 and Cen-0 appear QTL in the middle of the same chromosome, in a region that only two more populations appear QTL sporadically. Both of these populations do not have Ler as a parent and therefore such QTL cannot be attributed to a Ler contribution. Regarding the bottom of chromosome 5 only Bs-1 shares a minor QTL with the rest populations. Collectively these data suggest that chromosomes 4 and 5 hyper-accumulate QTL for flowering time and that the Bs-1 and Cen-0 major QTL are largely unique.

The picture under SD for the Sha crosses is similar to LDs in that in the first three chromosomes QTL appear sporadically and without significant overlapping between the populations (Fig 31,32,33). In the Ler x Sha population such QTL are not conditional since they appear both under LD of 16 h and intermediate days of 12 h. However in the Bay x Sha population the QTL at the top of chromosome 3 do not appear in the Bay or Dijon-G x Sha population under LD or SD respectively suggesting a Ler specific effect. Interestingly the two conditional QTL, which appeared in the Dijon-G x Sha population of this project, are unique in that population.

Similar cases of accession specific QTL appear also on chromosome 1 in the Dijon-G x Sha and Ler x Sha (Fig 31). Dijon-G and Ler are two accessions with little difference in their genomes. Out of a pool of 360 non-redundant yet randomly selected highly polymorphic markers, spanning the whole genome of *Arabidopsis*, any two average accessions will have 50-70 % percent chance for a hit. None of these markers was polymorphic between Dijon-G and Ler (data not shown). In addition a cross between Dijon-G and Ler phenotyped under SD of 10 and 8 h, showed very sharp segregation, fitting 100 % to the two parental phenotypes (which are very similar) and not transgressive variation whatsoever, suggesting absence of major sources of genetic variability between the two genomes (data not shown). Collectively these data suggest that the absence of co localisation of the QTL on chromosome 1 in the Dijon x Sha and the Ler x Sha could be due to their conditional nature since one population is analysed under short and the other under long days.

Again chromosomes 4 and 5 collect many common QTL (Fig 34,35). On chromosome 4 all populations have one common QTL at the top suggesting that this affects flowering independently of photoperiod and that is common source of natural variation. At the bottom of chromosome 5, common QTL exist in the Dijon-G and Ler populations but are absent from the Bay population. In the middle of the same chromosome QTL specific only for the Dijon-G population are identified but these are not conditional. Not conditional and specific for the

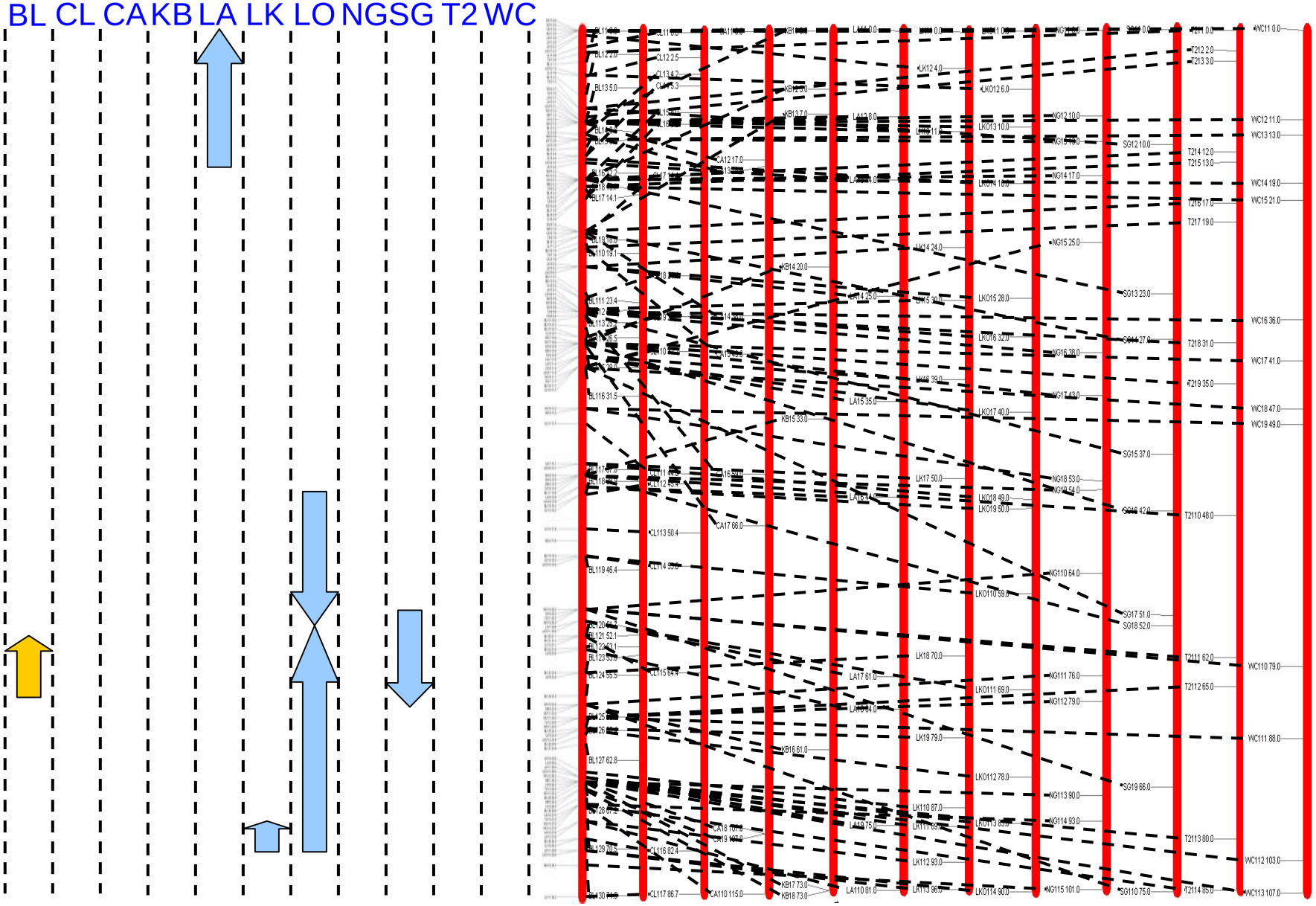


Figure 26 : Genetic anchoring of flowering time QTL, mainly from Ler crosses, on chromosome 1 to the *Arabidopsis* physical map. These flowering time QTL were detected under LDs. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : BL Bs-1 x Ler, CL Cen-0 x Ler, CA Cvi x Ag-0, KB Kondara Br-0, LA Ler x An, LK Ler x Kas, LO Ler x Kondara, NG Nok-3 x Ga-0, SG Sorbo x Gy-0, TS5 x 24014, WC Wt-5 x Ct-1

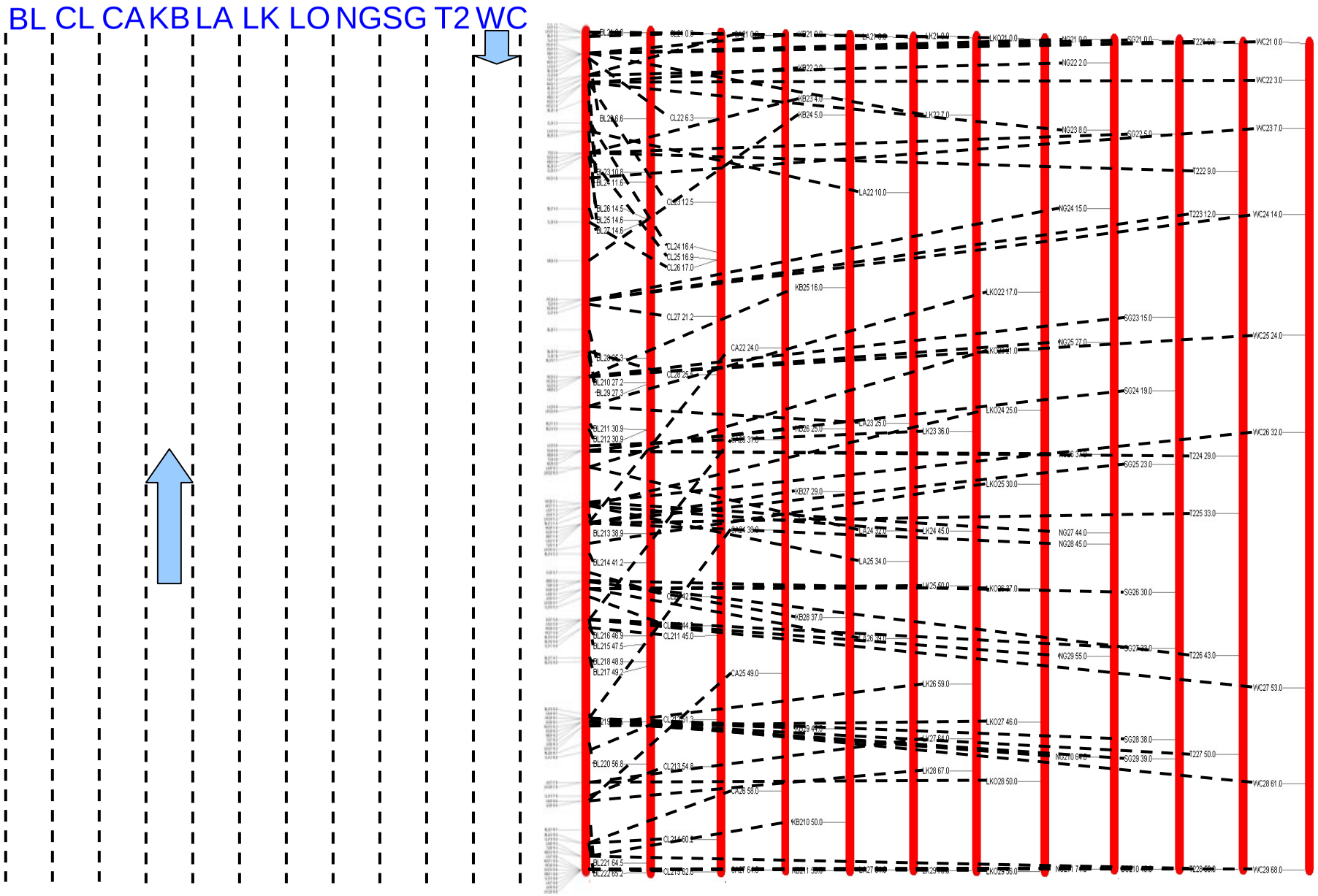


Figure 27 : Genetic anchoring of flowering time QTL, mainly from Ler crosses, on chromosome 2 to the *Arabidopsis* physical map. These flowering time QTL were detected under LDs. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : BL Bs-1 x Ler, CL Cen-0 x Ler, CA Cvi x Ag-0, KB Kondara Br-0, LA Ler x An, LK Ler x Kas, LO Ler x Kondara, NG Nok-3 x Ga-0, SG Sorbo x Gy-0, TS5 x 24014, WC Wt-5 x Ct-1

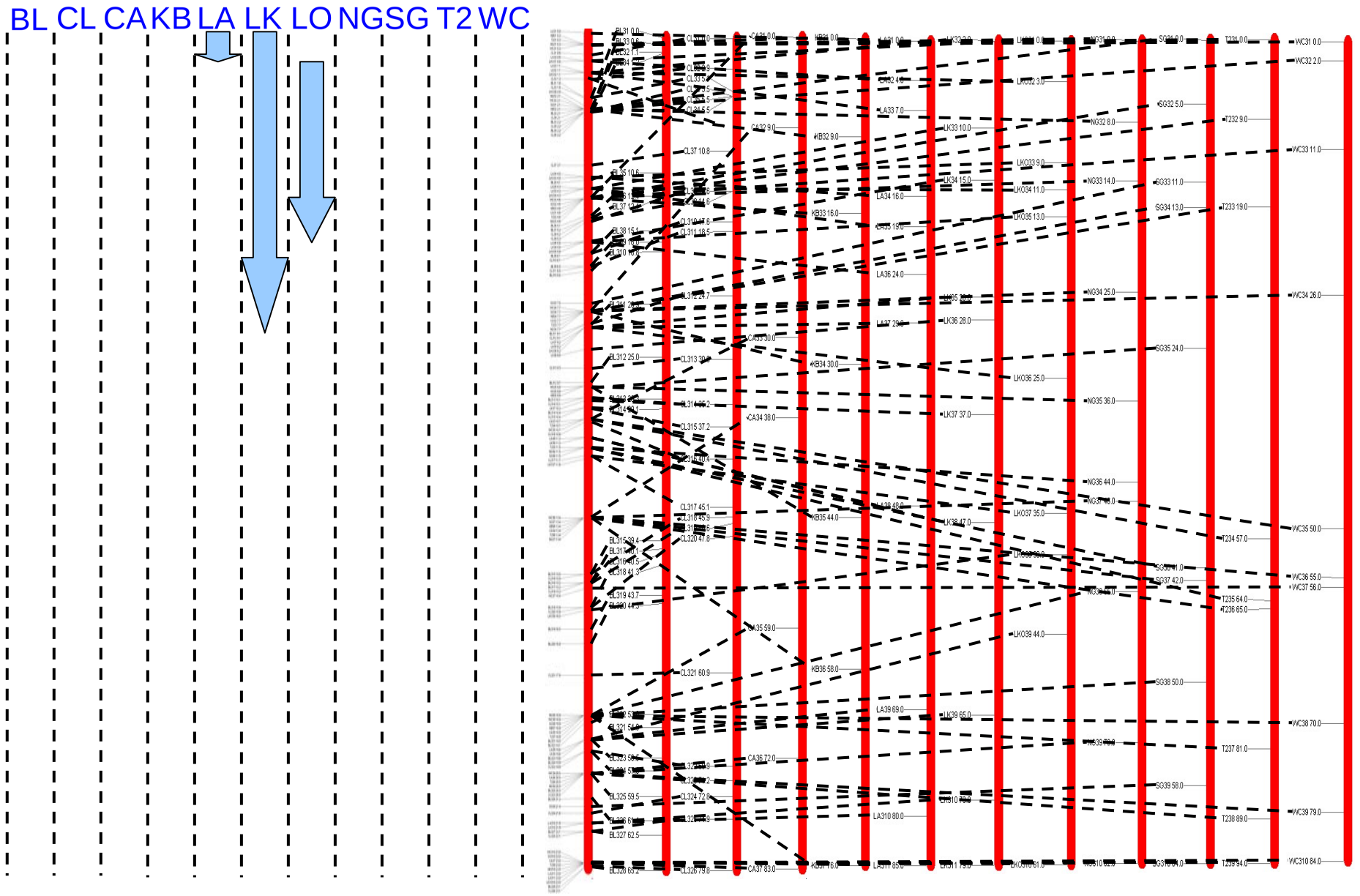


Figure 28 : Genetic anchoring of flowering time QTL, mainly from Ler crosses, on chromosome 3 to the *Arabidopsis* physical map. These flowering time QTL were detected under LDs. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : BL Bs-1 x Ler, CL Cen-0 x Ler, CA Cvi x Ag-0, KB Kondara Br-0, LA Ler x An, LK Ler x Kas, LO Ler x Kondara, NG Nok-3 x Ga-0, SG Sorbo x Gy-0, TS5 x 24014, WC Wt-5 x Ct-1

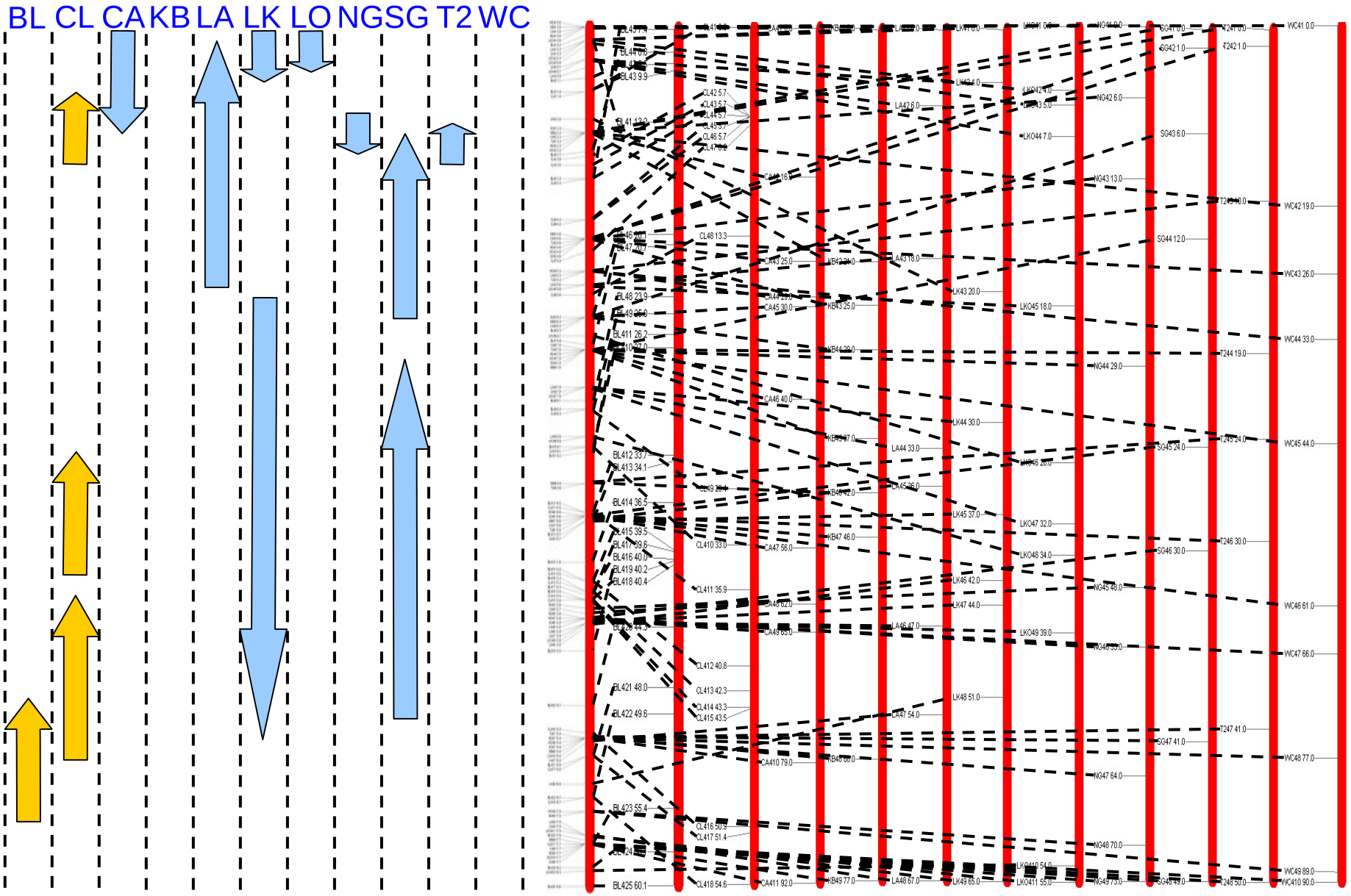


Figure 29 : Genetic anchoring of flowering time QTL, mainly from Ler crosses, on chromosome 4 to the *Arabidopsis* physical map. These flowering time QTL were detected under LDs. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : BL Bs-1 x Ler, CL Cen-0 x Ler, CA Cvi x Ag-0, KB Kondara Br-0, LA Ler x An, LK Ler x Kas, LO Ler x Kondara, NG Nok-3 x Ga-0, SG Sorbo x Gy-0, TS5 x 24014, WC Wt-5 x Ct-1

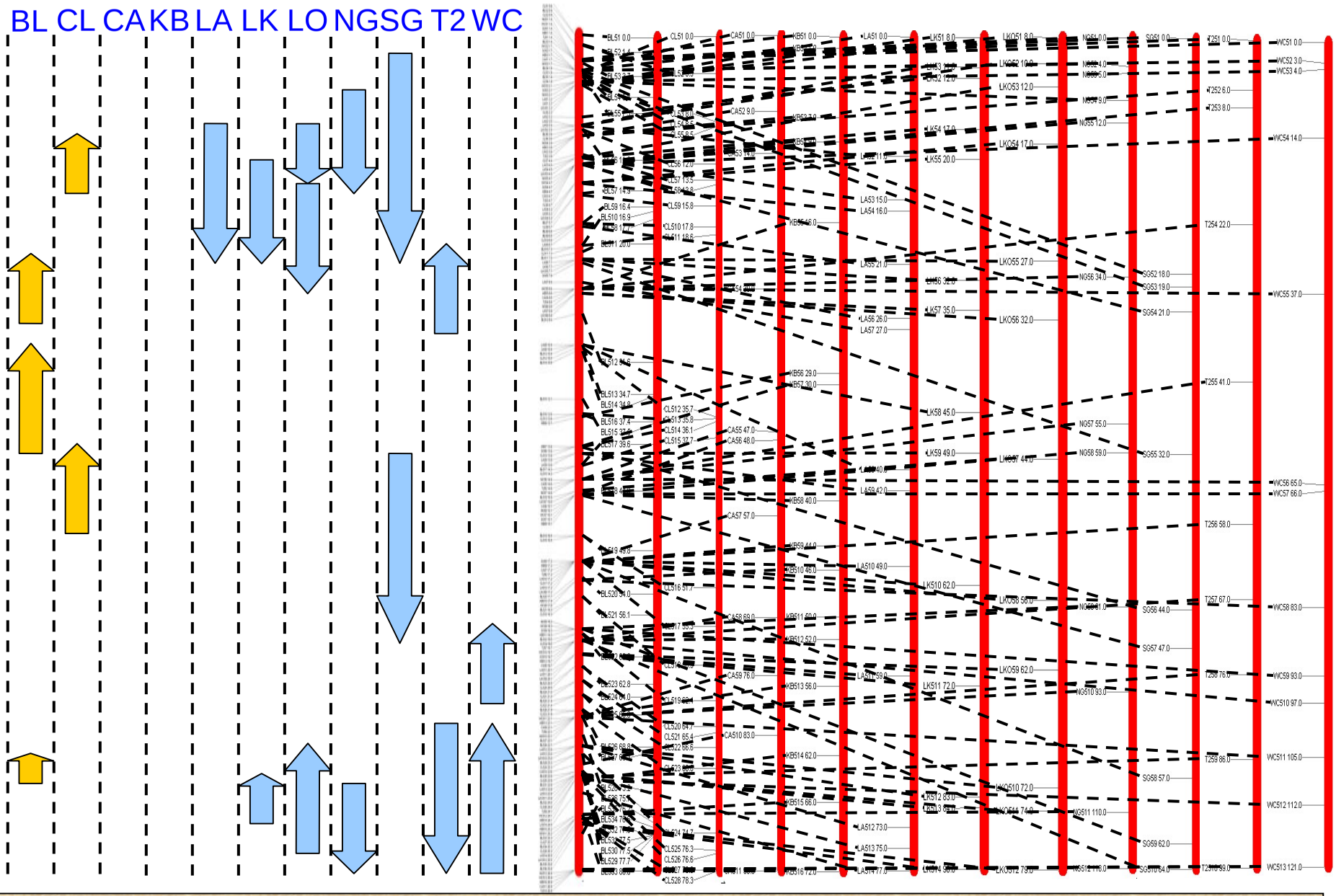


Figure 30 : Genetic anchoring of flowering time QTL, mainly from Ler crosses, on chromosome 5 to the *Arabidopsis* physical map. These flowering time QTL were detected under LDs. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : BL Bs-1 x Ler, CL Cen-0 x Ler, CA Cvi x Ag-0, KB Kondara Br-0, LA Ler x An, LK Ler x Kas, LO Ler x Kondara, NG Nok-3 x Ga-0, SG Sorbo x Gy-0, TS5 x 24014, WC Wt-5 x Ct-1

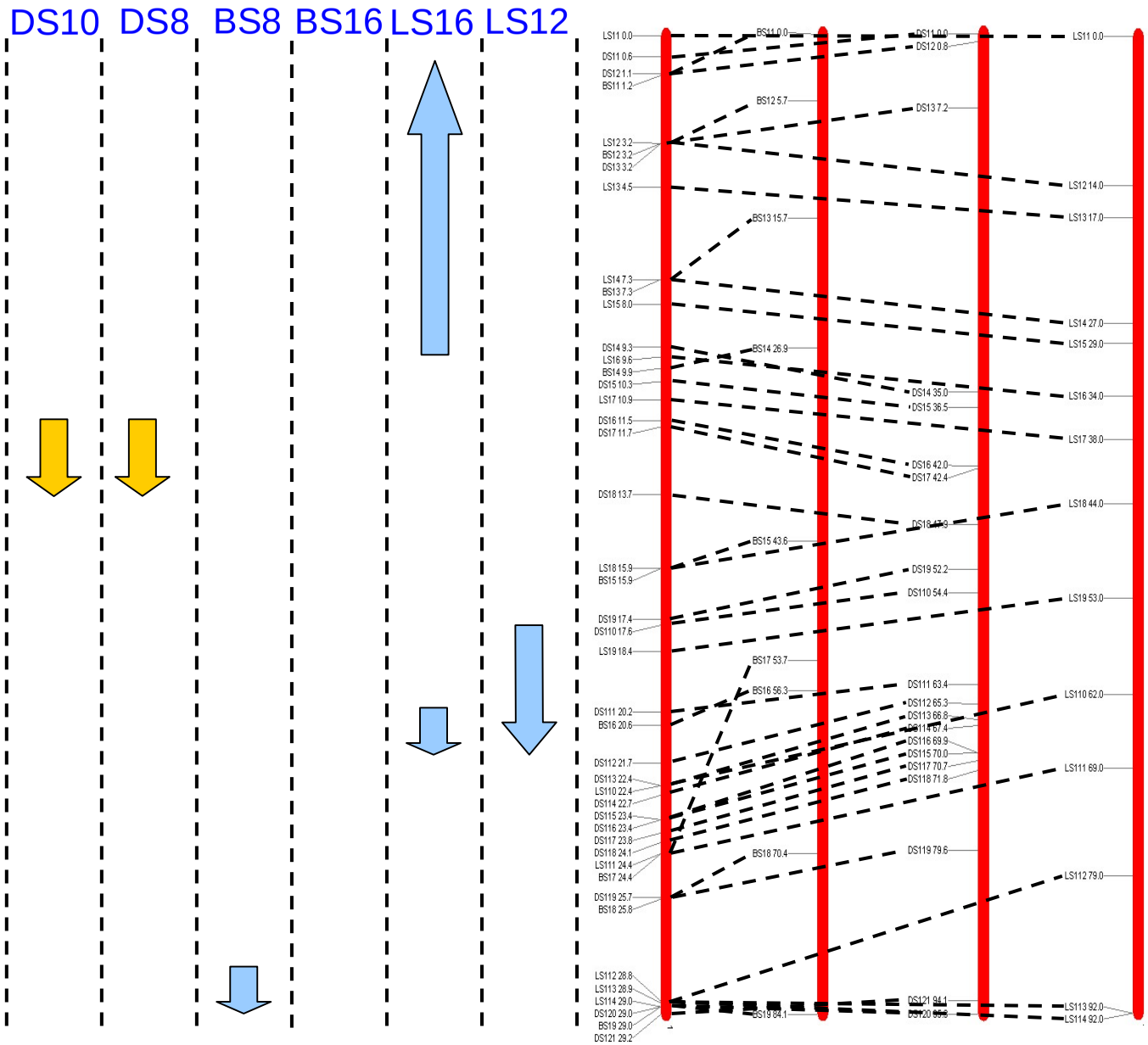


Figure 31 : Genetic anchoring of flowering time QTL from Sha crosses, on chromosome 1 to the *Arabidopsis* physical map. These flowering time QTL were detected under various conditions. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : DS 10 Dijon-G x Sha SD 10 h, DS 8 Dijon-G x Sha SD 8 h, BS8 Bay x Sha SD 8 h BS16 Bay x Sha LD 16 h, LS16 Ler x Sha LD 16 h, LS12 Ler x Sha LD 12 h.

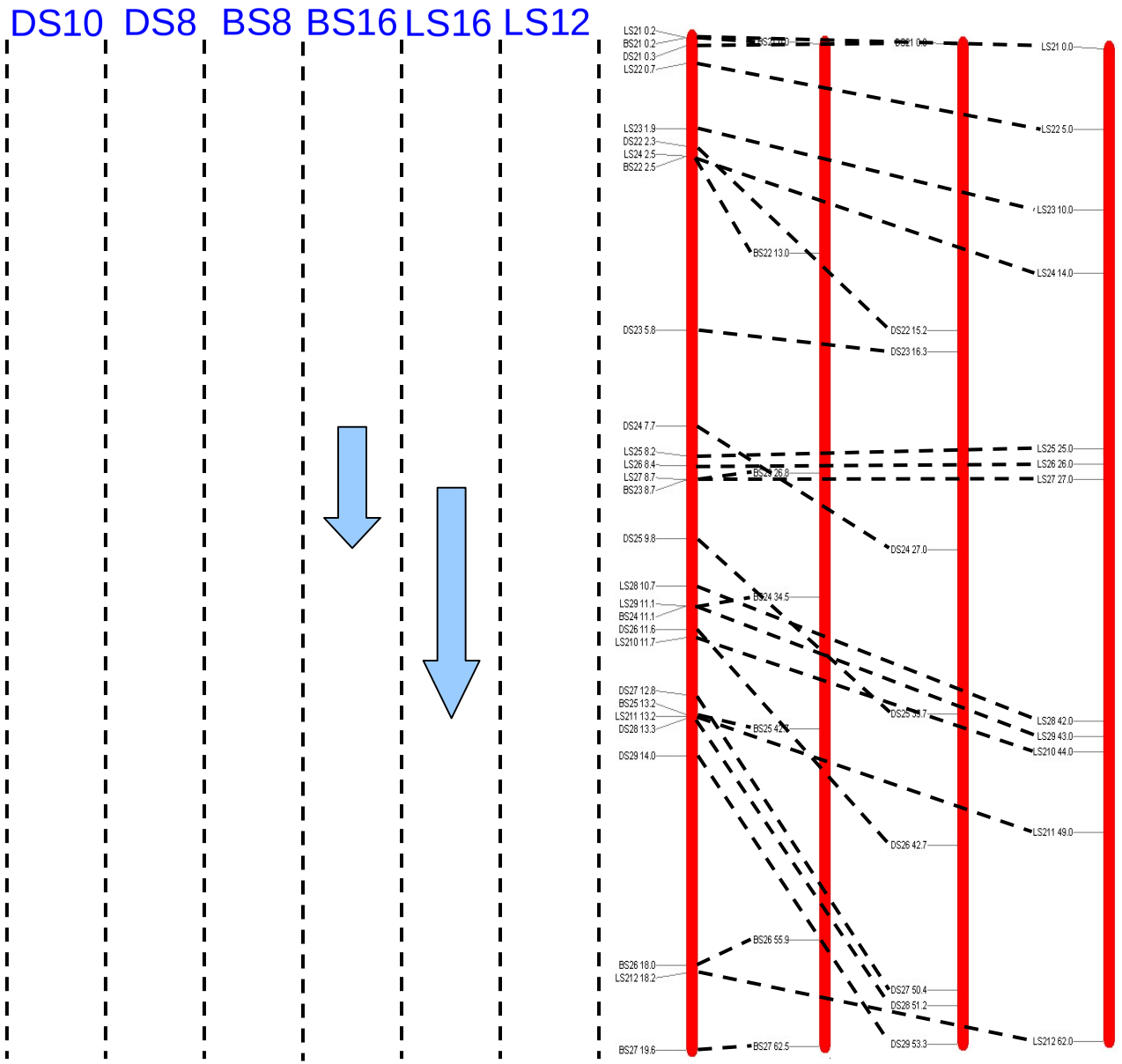


Figure 32 : Genetic anchoring of flowering time QTL from Sha crosses, on chromosome 2 to the *Arabidopsis* physical map. These flowering time QTL were detected under various conditions. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : DS 10 Dijon-G x Sha SD 10 h, DS 8 Dijon-G x Sha SD 8 h, BS8 Bay x Sha SD 8 h BS16 Bay x Sha LD 16 h, LS16 Ler x Sha LD 16 h, LS12 Ler x Sha LD 12 h.

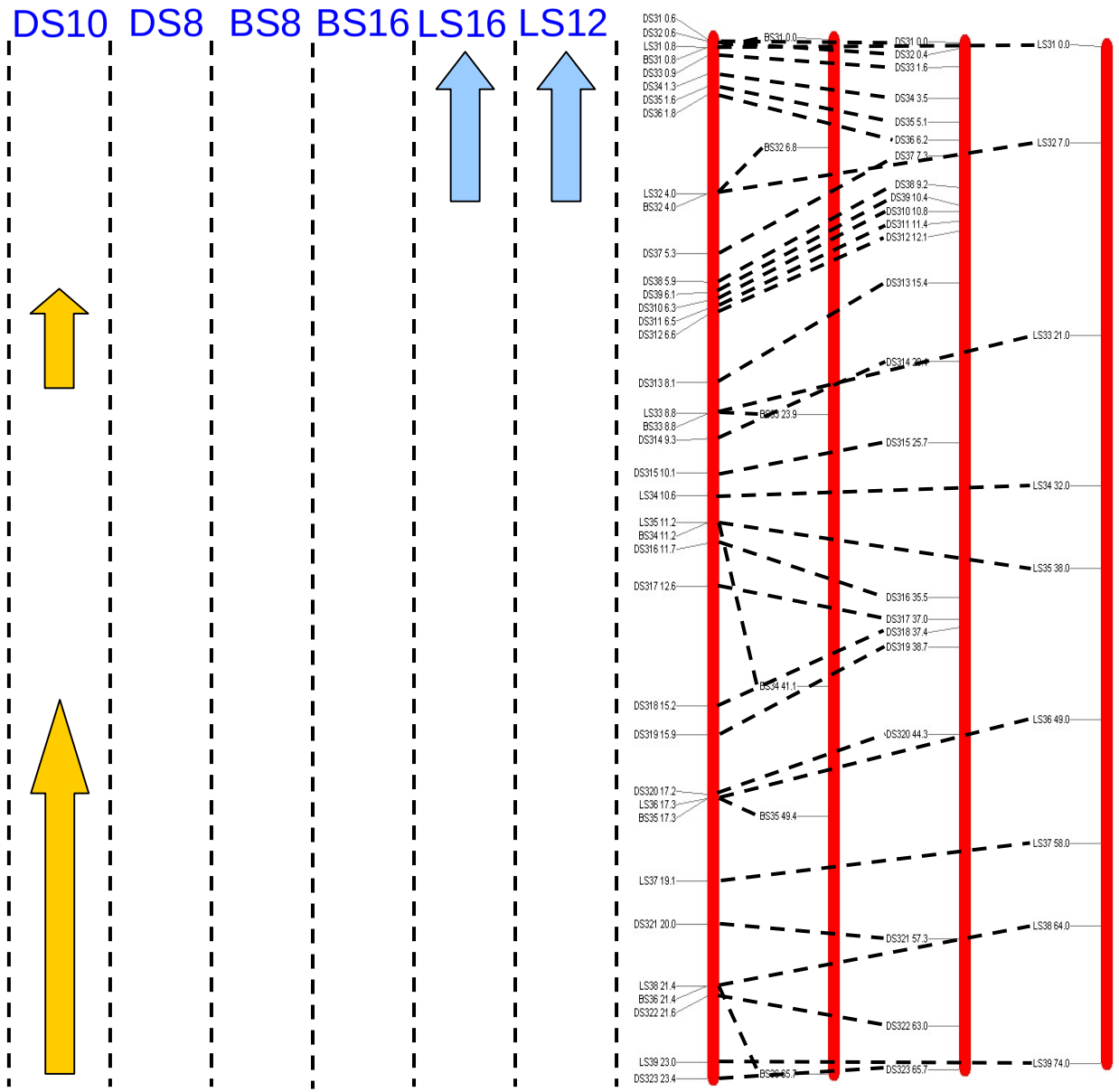


Figure 33 : Genetic anchoring of flowering time QTL from Sha crosses, on chromosome 3 to the *Arabidopsis* physical map. These flowering time QTL were detected under various conditions. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : DS 10 Dijon-G x Sha SD 10 h, DS 8 Dijon-G x Sha SD 8 h, BS8 Bay x Sha SD 8 h BS16 Bay x Sha LD 16 h, LS16 Ler x Sha LD 16 h, LS12 Ler x Sha LD 12 h.

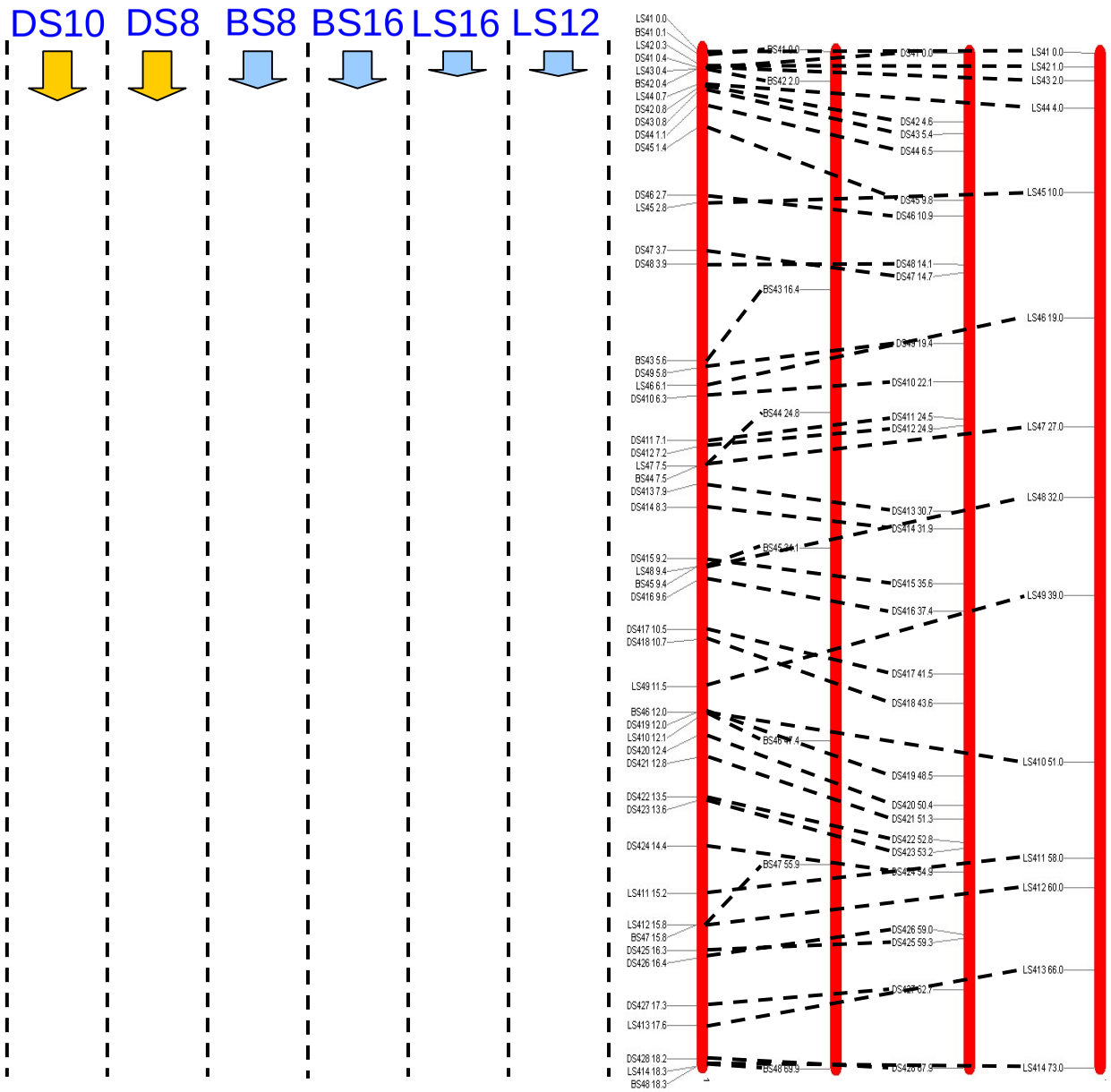


Figure 34 : Genetic anchoring of flowering time QTL from Sha crosses, on chromosome 4 to the *Arabidopsis* physical map. These flowering time QTL were detected under various conditions. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : DS 10 Dijon-G x Sha SD 10 h, DS 8 Dijon-G x Sha SD 8 h, BS8 Bay x Sha SD 8 h BS16 Bay x Sha LD 16 h, LS16 Ler x Sha LD 16 h, LS12 Ler x Sha LD 12 h.

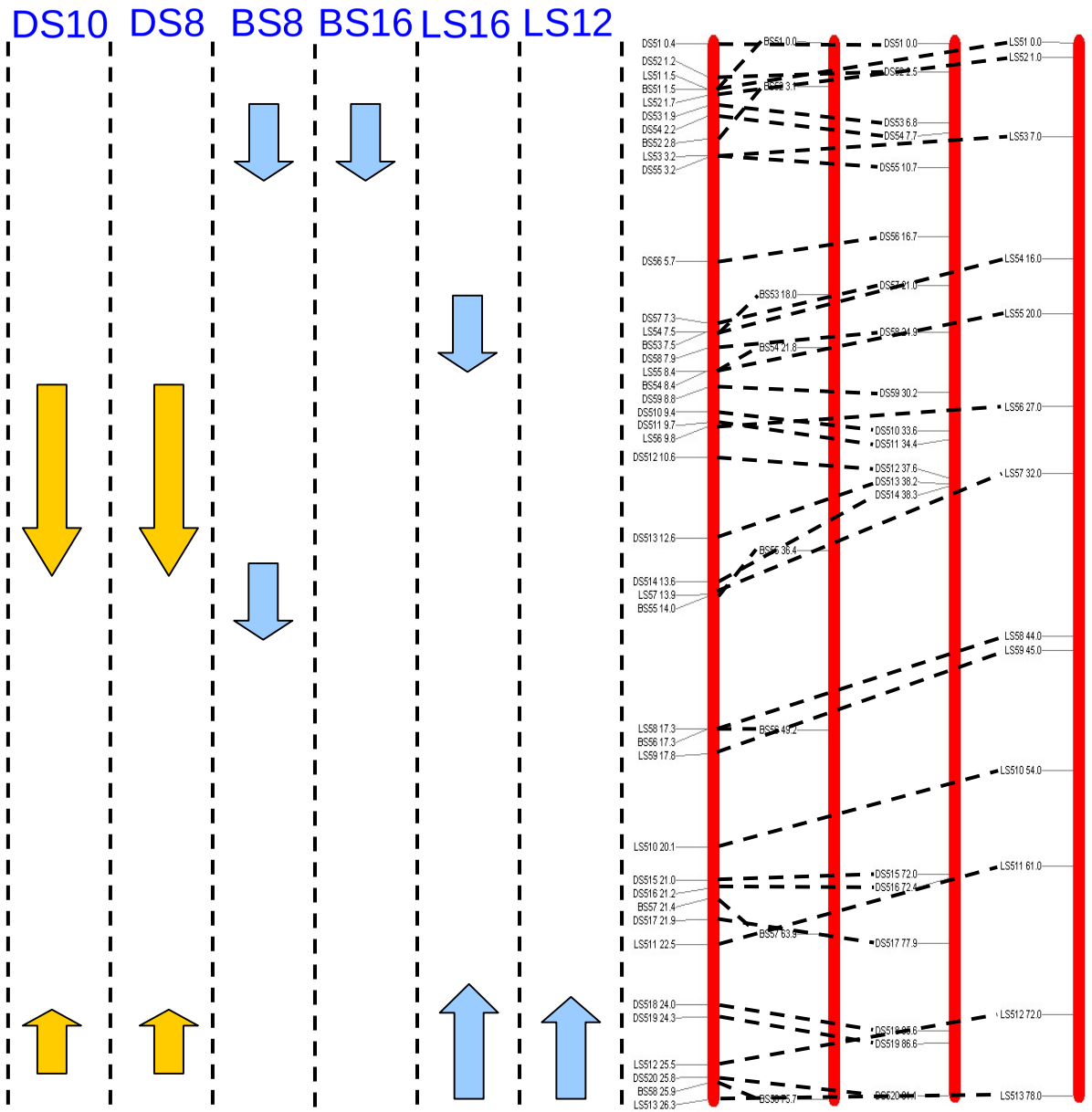


Figure 35 : Genetic anchoring of flowering time QTL from Sha crosses, on chromosome 5 to the *Arabidopsis* physical map. These flowering time QTL were detected under various conditions. Blue arrows indicate the existence of a QTL in a published population and the yellow arrows are QTL detected in this project. Populations : DS 10 Dijon-G x Sha SD 10 h, DS 8 Dijon-G x Sha SD 8 h, BS8 Bay x Sha SD 8 h BS16 Bay x Sha LD 16 h, LS16 Ler x Sha LD 16 h, LS12 Ler x Sha LD 12 h.

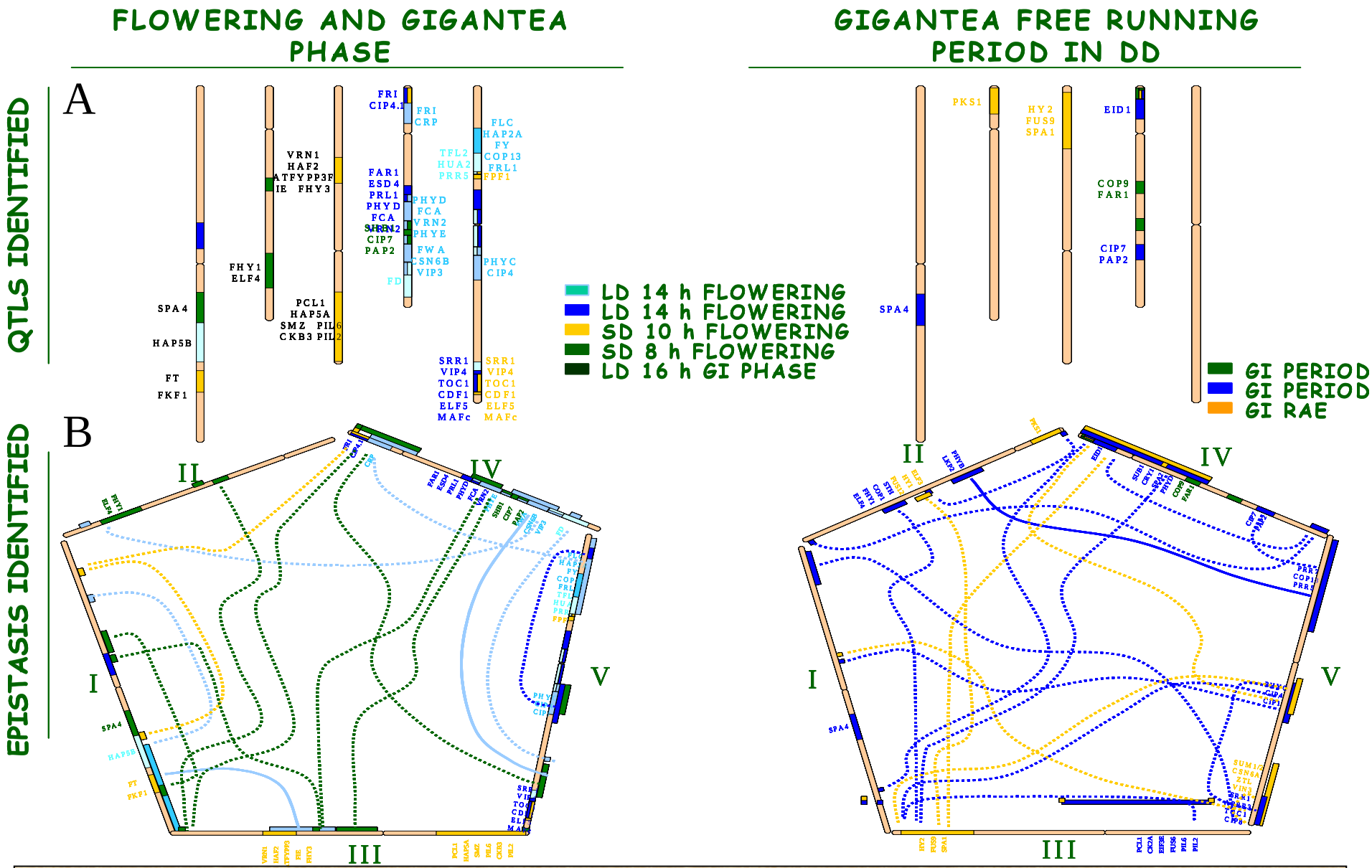


Figure 36 : Summary of the genetic analysis. A) Using multiple populations allowed the mapping of more than 20 QTL for several circadian and flowering parameters. This analysis spans a large number of different day lengths covering the geographical range that the species experiences in nature. Clusters of QTL can be identified in the genome highlighting the importance of the corresponding genetic areas. Several candidate genes are present in these regions and are reported on the graph. B) The standard QTL analysis is accompanied by an extensive network of genetic interaction, some of which are between QTL not identified by standard mapping approaches.

Bay parent is one QTL at the top of the same chromosome, which fits the position of many QTL identified there in the LD populations. Collectively these data indicate that chromosomes 4 and 5 hyper-accumulate QTL for flowering time with one region at the top and another in the middle for chromosome 4 and one region at the top and another at the bottom for chromosome 5 being over-represented suggesting the existence of common targets of natural variation. In mapping populations of this project QTL, which are conditional and unique have been identified together with common QTL. A summary of the identified QTL and genetic interactions together with candidate genes is shown in Fig 36.

3.1.2.10 Towards the isolation of QTL affecting day length perception – Progeny testing and validation of the QTL mapping.

QTL mapping is a useful approach that guides the attempt to describe genetically a phenotype. However this approach, although based on real data analysed with powerful statistical tests, is theoretical and additional genetics are required for validation. Usually the first step after on QTL mapping is to proceed in progeny testing of selected F2 individuals in order to confirm the phenotype. In the current project the phenotype is complex since the trait under investigation is not overall flowering time but day length perception. Therefore the primary interest of the project is to detect QTL that make the differencde between similar day lengths and not QTL which cause lateness or earliness in many conditions in parallel. Progeny testing based on F3 populations, serves both the need for validation of the QTL positions and the need for conformation of the specificity and conditional character of the QTL effects.

Various types of information were used in order to properly select F2 individuals from the results of the QTL mapping for both the Bs-1 and the Cen-0 populations (Fig 37). Perhaps the most important character is the genotype of the F2 individual. Plants, which were fixed in a large percentage in their genome as Ler homozygous, were advantageous since most of the genome was purified from background effects. In addition such genotypes are a good first step for the subsequent production of NILs, reducing the number of backcrosses that are required. However such individuals should be heterozygous in the QTL regions because this allows i) subsequent selection of smaller QTL regions carrying the character in the next generation ii) separation of the different QTL and thus investigation of confirmed genetic interactions among them. Having QTL at the heterozygous state can be complicated since it requires knowledge of the allelic hierarchy of each QTL position. Therefore another property of the F2 mapping that is useful in the selection is presence of dominance for each QTL separately.

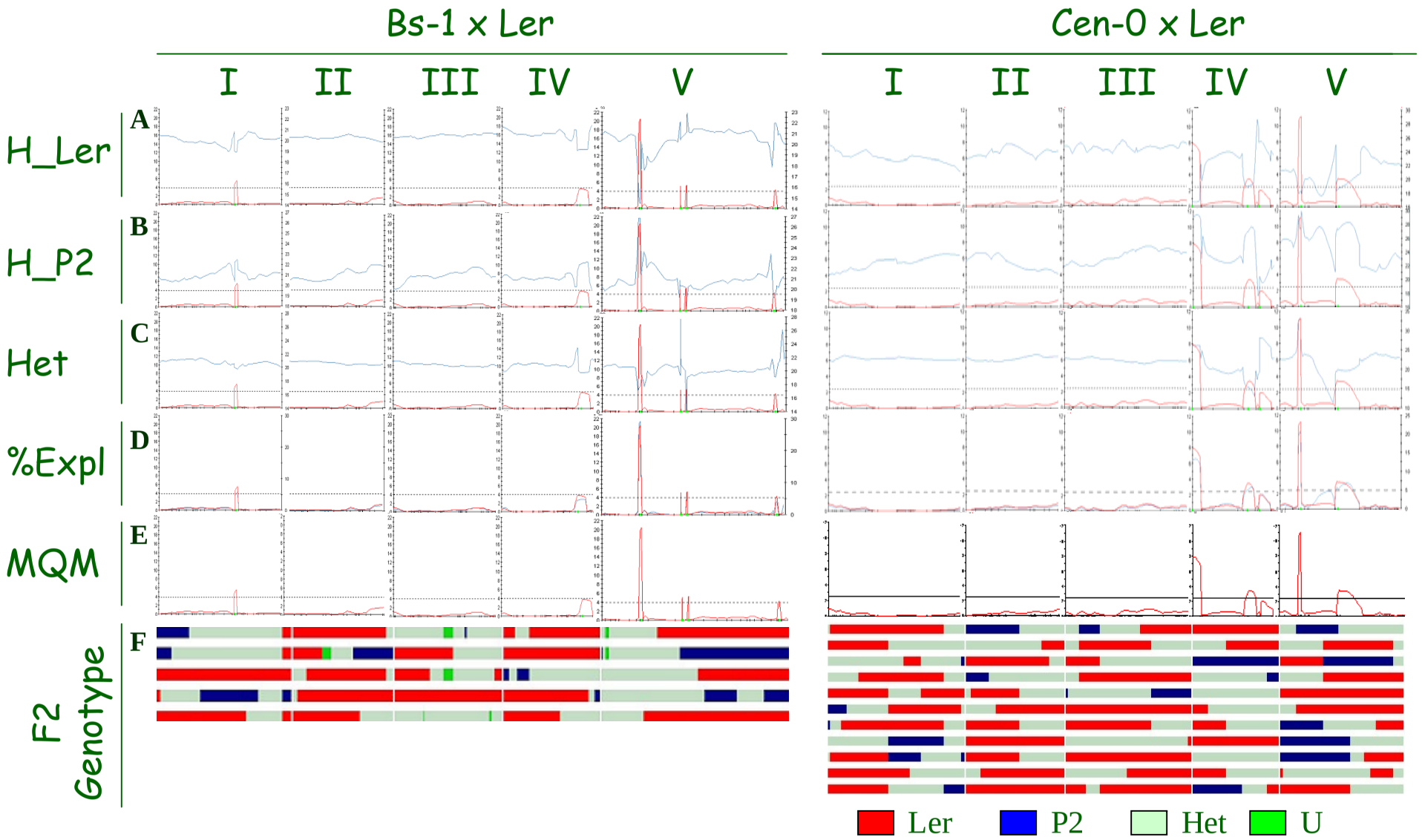


Figure 37 : Allelic properties in the Cen-0 x Ler and Bs-1 x Ler populations. a) Ler effect. The blue line on the right Y axis indicates the effect of Ler homozygous alleles on flowering time. b) Bs-1 or Cen-0 effect. The blue line of the right Y axis indicates the effect of Bs-1 or Cen-0 homozygous alleles on flowering time c) Heterozygocity effect. The blue line on the right Y axis indicates the effect of heretozygous alleles on flowering time. It also shows the presence of dominance. D) The blue line on the right Y axis indicates percentage of variability that each QTL explains. E) The red line on the left axis indicates the results of the MQM mapping and the position of the QTL. F) Genotypes of selected F2 individuals for progeny testing and subsequent NIL generation. Red : Ler homozygous, Blue : Bs-1 or Cen-0 homozygous, Light green : Heterozygous, Green : Undefined region.

All these properties combined with the QTL position, allowed the selection of several F2 individuals from the Bs-1 and Cen-0 populations, the progeny of which was phenotyped both under LD of 14 and 16 h. Since the individuals were selected as heterozygous in the QTL region segregation was expected in the F3 generation. Indeed in the case of the Cen-0 population, segregation was observed with late flowering specifically under LD of 14 h being reported (Fig 38). However not all populations discriminated between the two day lengths successfully. For example the CL 20 and the CL97 populations showed the highest frequencies of LD 14 h specific late flowering while by contrast the CL 111 and 91 populations segregated almost identically between the two conditions. The CL 20 and 97 populations appeared with transgression suggesting that in the two populations genetic interactions between the QTL might take place. CL 20 is heterozygous at the top two QTL of chromosome 4 and 5 and Ler homozygous in the rest, indicating that these two QTL are largely responsible for the trangressive variation observed in the population. CL 97 is heterozygous in the whole chromosome 4 and Cen-0 homozygous in the upper half. This combination detects day length strongly however when Cen-0 alleles at the top of chromosome 5 are combined with Ler homozygous alleles on chromosome 4 (CL 91) day length perception is abolished but late flowering remains. The inverted allelic combination (CL 111) gives the same phenotype however when the whole chromosome 5 and the top of chromosome 4 are Ler homozygous day length perception is totally abolished and early flowering similar to Ler parental phenotype is restored (data not shown). These data indicate that a complex network of interactions between the QTL regulates day length perception and late flowering in this population and that the two QTL at the top of chromosomes 4 and 5 are responsible for the observed transgressive variation. Note that this variation was abolished after vernalization treatment and therefore these QTL might be associated with the vernalization pathway. However through interactions with the remaining QTL enhanced day length perception is achieved independently of overall late flowering.

In the progeny of the Bs-1 cross all populations discriminated between the two-day lengths (Fig 39) but some were more efficient. Two interesting populations are the BL 72 and the BL 4. The two parental phenotypes are very similar with the exception that the lower end of the QTL in the BL 72 genotype is in a Ler homozygous state. Both populations discriminate well the two day lengths, confirming thus the QTL mapping. Interestingly in the BL 4 population the frequencies both under LD of 14 and 16 h are shifted towards later flowering times suggesting the existence of additional loci, which cause non conditional lateness. This lateness was rescued by vernalization and the BL 4 population behaved similarly to the BL

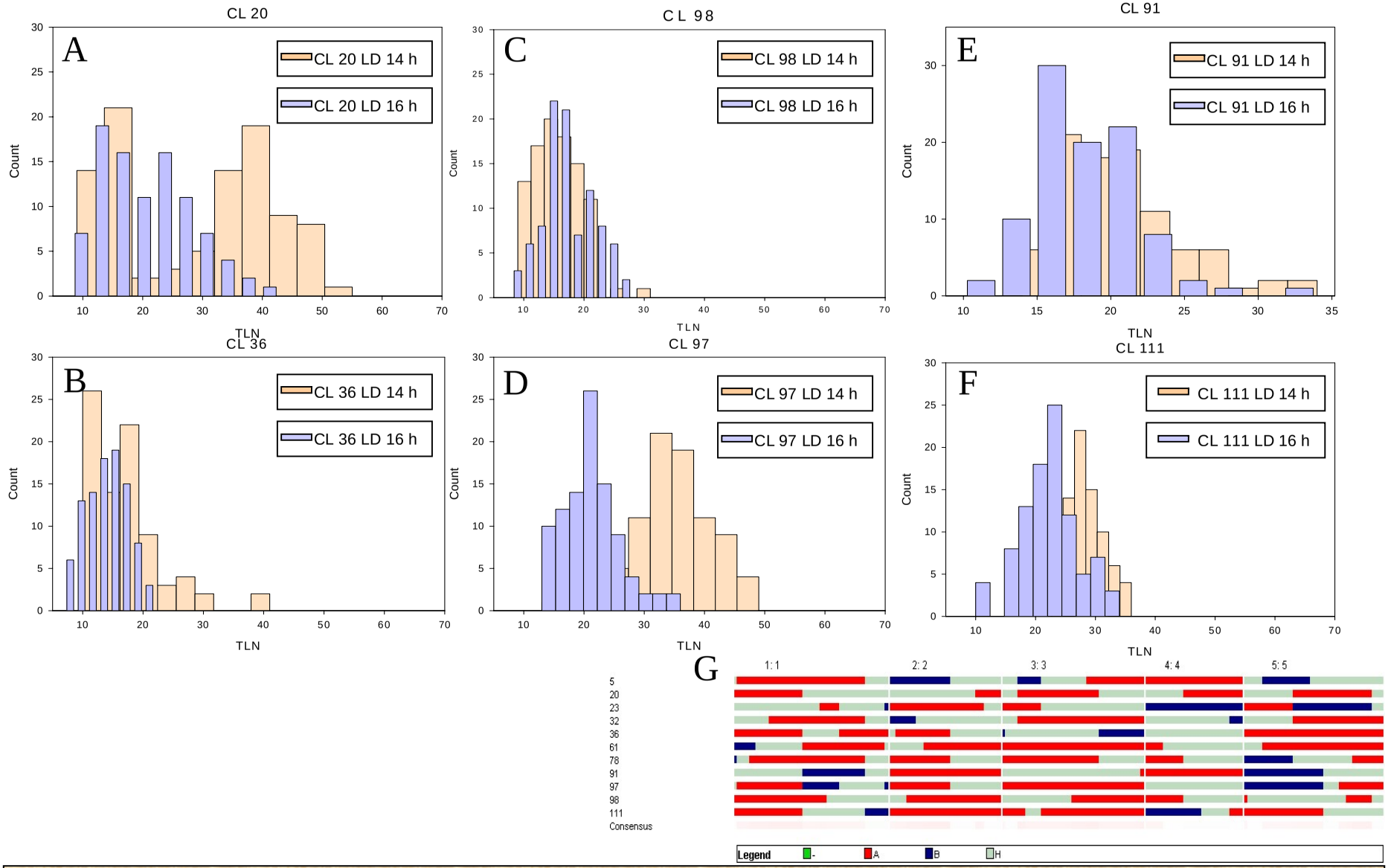


Figure 38 : Progeny testing in selected F2 genotypes of the Bs-1 x Ler cross. A) Segregation of 50 individuals of the CL 20 F2 progenitor B) Segregation of 50 individuals of the CL 36 F2 progenitor C) Segregation of 50 individuals of the CL 98 F2 progenitor D) Segregation of 50 individuals of the CL 97 F2 progenitor E) Segregation of 50 individuals of the CL 91 F2 progenitor. F) Segregation of 50 individuals of the CL 111 F2 progenitor Blue : LD 16 h, yellow : LD 14 h. G) Genotypes of all F2 individuals Red : Ler homozygous, Blue : Cen-0 homozygous, Light green : heterozygous, Green : unknown. Note the later flowering of all populations under LD 14 h with the exceptions of CL 98 and 111 . The CL 20 and 97 are the two populations with the most marked differences between the two daylengths.

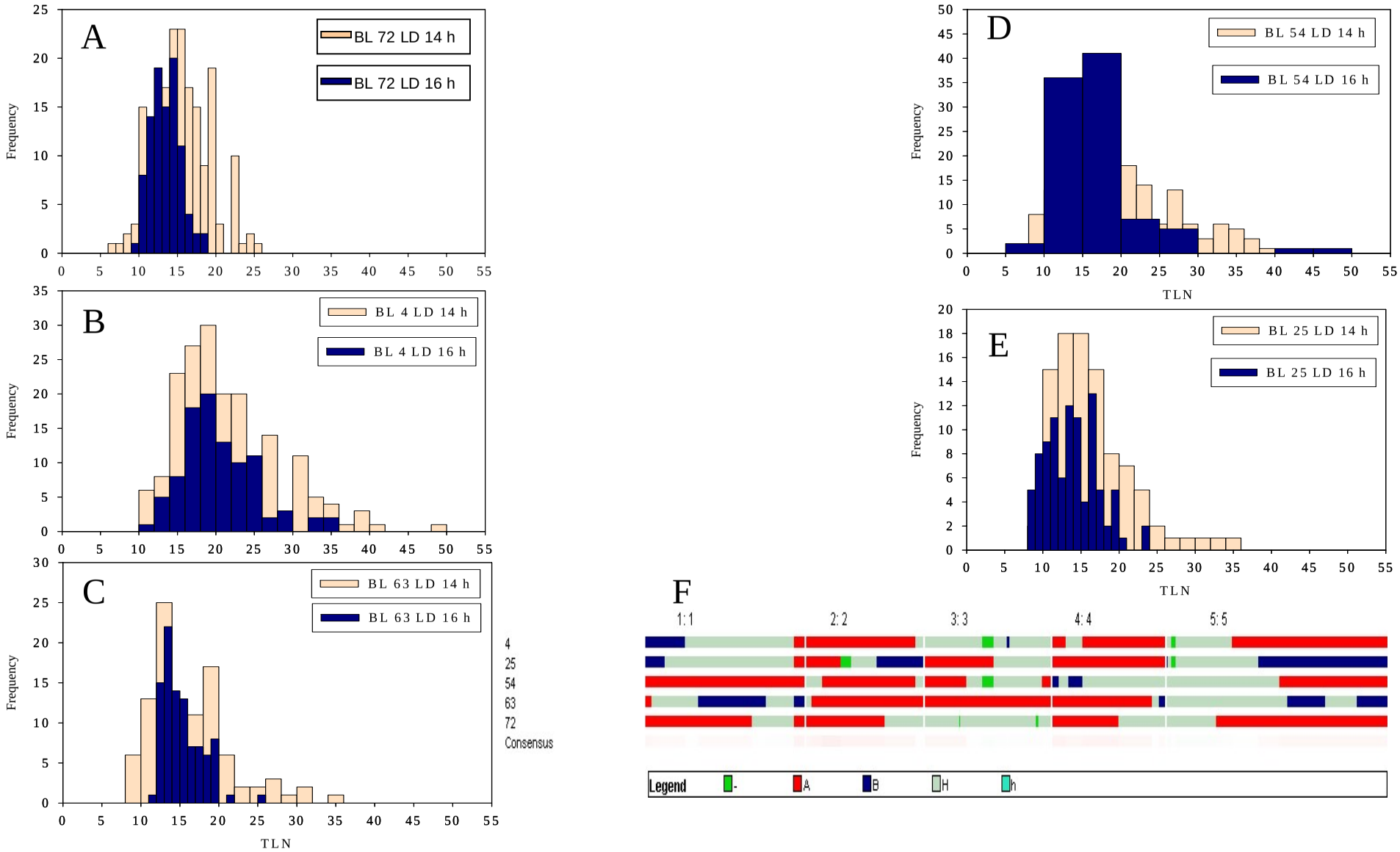


Figure 39 : Progeny testing in selected F2 genotypes of the Bs-1 x Ler cross. A) Segregation of 50 individuals of the BL 72 F2 progenitor B) Segregation of 50 individuals of the BL 4 F2 progenitor C) Segregation of 50 individuals of the BL 63 F2 progenitor D) Segregation of 50 individuals of the BL 54 F2 progenitor E) Segregation of 50 individuals of the BL 25 F2 progenitor. Blue : LD 16 h, yellow : LD 14 h. F) Genotypes of all F2 individuals Red : Ler homozygous, Blue : Bs-1 homozygous, Light green : heterozygous, Green : unknown. Note the later flowering of all segregants under LD 14 h. The BL 72 and 4 are the two populations with the most marked differences between the two daylengths.

72, suggesting that the observed non-conditional late flowering could be associated with the vernalization pathway. The BL 72 population did not respond to vernalization, yet retained the ability of day length perception suggesting that vernalization is not responsible for that trait in this population.

3.1.2.11 Towards the isolation of QTL affecting day length perception – Fine mapping in the Bs-1 x Ler population and production of NIL

From the Bs-1 x Ler population F3 individuals from the BL 72 population were selected and backcrossed to Ler. The resulting BC1 F1 individual plants were selfed and approximately 20 populations and 500 individuals in total were screened both under LD 16 and 14 h. In all cases late flowering specifically associated with LD of 14 h was detected (data not shown) confirming for a second time the results of the QTL mapping. In the initial map, the QTL region contained no markers very close to the peak of the QTL. Additional markers were found based on the existing resources available for mapping between accessions in *Arabidopsis* (see introduction and Appendix). Both SNP-based dCAP markers and copy-based SSR markers were used with the HmMPP (see introduction) method being applied successfully in the latter. These combined efforts allowed initial fine mapping of the QTL into a region of 1 Mbp or 3 cM from the initial 10 Mbp of the interval mapping results (Fig 40).

This region is located slightly downstream compared to the indicative region of the MQM mapping, mainly because of the marker gap in the initial linkage map. At any case the fine mapped region fits perfectly to the peak of the QTL in the interval map, even though the region is ten times smaller. Bs-1 homozygous alleles cause late flowering under LD of 14 h in that region while Ler homozygous alleles are early flowering and heterozygous individuals have an intermediate phenotype consistent with the semi-dominant character of the QTL. This initial fine mapping allowed also the exclusion of candidate genes that are close in the region. *CO*, a gene mutations of which cause late flowering, is located upstream of the fine mapped interval, in an area were all individuals are Ler homozygous. Since Ler has a fully functional *CO* gene and the BC1 F2 individuals differ in their phenotypes, the possibility of having *CO* as the causal gene is eliminated. With the same concept the *FLC* gene can be excluded. *FLC* is located higher on chromosome 5 than *CO* and therefore even further from the fine mapped interval, again within the region of the Ler homozygous alleles. Functional *FLC* alleles cause late flowering independently of the photoperiod, yet Ler has a weak *FLC* allele due to a transposon insertion in the first intron of the gene. This together with the fact that the parental population of the BC1 plants (BL 72) did not respond to vernalization excludes *FLC* from the

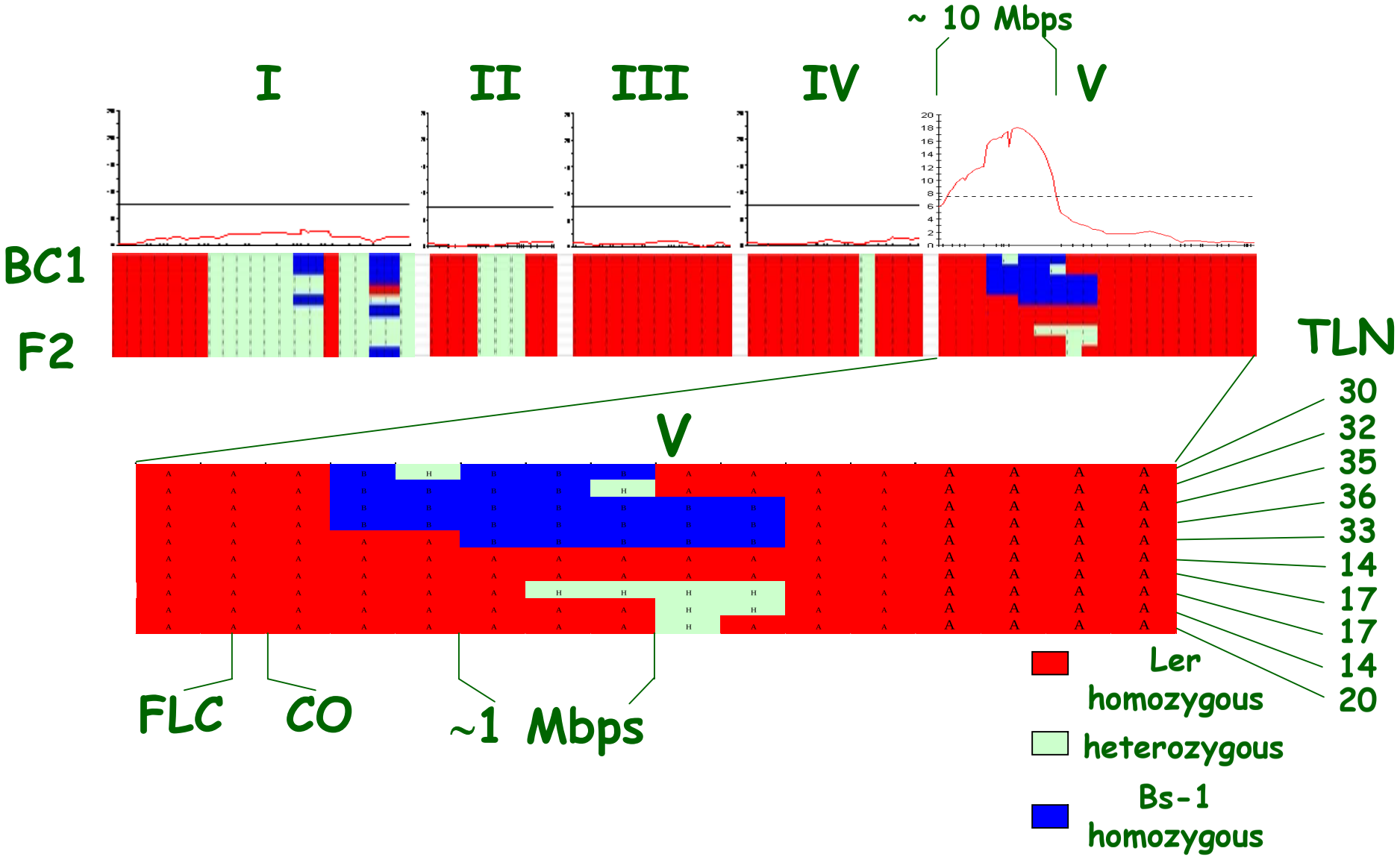


Figure 40 : Genotyping of selected BC1 plants from the Bs-1 x Le population. The QTL region in the initial QTL mapping spans 10 Mbps. After one BC, the QTL region was reduced ten times (0.9 Mbps) with Bs-1 homozygous alleles casing late flowering and Ler homozygous alleles causing early flowering. The rest of the genome is nearly complete Ler homozygous and therefore these lines can be also used as initial NILs. Genes such as FLC and CO can be excluded from the candidate list since Ler homozygous alleles span these regions in both early and late flowering individuals. The flowering times of the BC1 F2 individuals are shown next to their genotypes.

candidate list of day length perception. However since the late flowering of the BL 4 population compared to the BL 72 population was rescued by vernalization, it is possible that *FLC* has an effect in the regulation of flowering in general in the population. This is consistent with the vernalization response of the Bs-1 parent and the effect of vernalization in the F2 population, and provides additional evidence that day length perception and overall lateness in flowering time are not interconnected traits.

From the total BC1 F2 population several individuals were selected as putative NILs and their flowering times were scored both under long days of 14 and 16 h. These individuals were Bs-1 homozygous in the fine-mapped region and Ler homozygous in the rest of the genome with the exception of the bottom of chromosome 1 which remained largely heterozygous after the first backcross. Therefore Ler alleles span approximately 85 % of the genome in those individuals, heterozygous alleles approximately 14 % while Bs-1 homozygous alleles consist of less than 1 % in total. Phenotyping confirmed the overall late flowering of the putative NILs compared to the Ler flowering time both under LD of 16 and 14 h (Fig 41A). Most importantly the NILs behaved like the Bs-1 parent, flowering later under LD of 14 h compared to 16 h in steady-state conditions. These NIL will be a valuable tool both in the next level of fine mapping and also for the molecular analysis.

3.1.2.12 Discussion

Targets of natural variation affecting flowering time in Arabidopsis

Changes in climatological parameters such as light duration and temperature occur between different geographical locations. Since plants adjust their physiology to these variables it is logical to expect that natural variation is evident for flowering time. Indeed there are many studies supporting the presence of flowering time specific natural variation (reviewed in Koornneef et al., 2004; Tonsor et al., 2004). QTL identified in these studies show statistically significant interactions with the environment, suggesting that this genetic variation is largely in response to environmental changes and could reflect adaptation. In parallel, genetic studies that simulate natural conditions in controlled environments, confirm that in nature conditional QTL are accompanied with genetic interactions, reflecting a complex genetic architecture that is under strong environmental influence (Li et al., 2006). Interestingly in these studies QTL at the top of chromosome 4 and 5 appear often, just like in standard cabinet conditions. *FRI* and *FLC*, two major targets of natural variation for flowering time in *Arabidopsis* are located in these positions respectively and both are very good candidates for these QTL (Johanson et al., 2000; Lempe et al. 2005; Gazzani et al., Shindo et al., 2005). Local adaptation has shaped

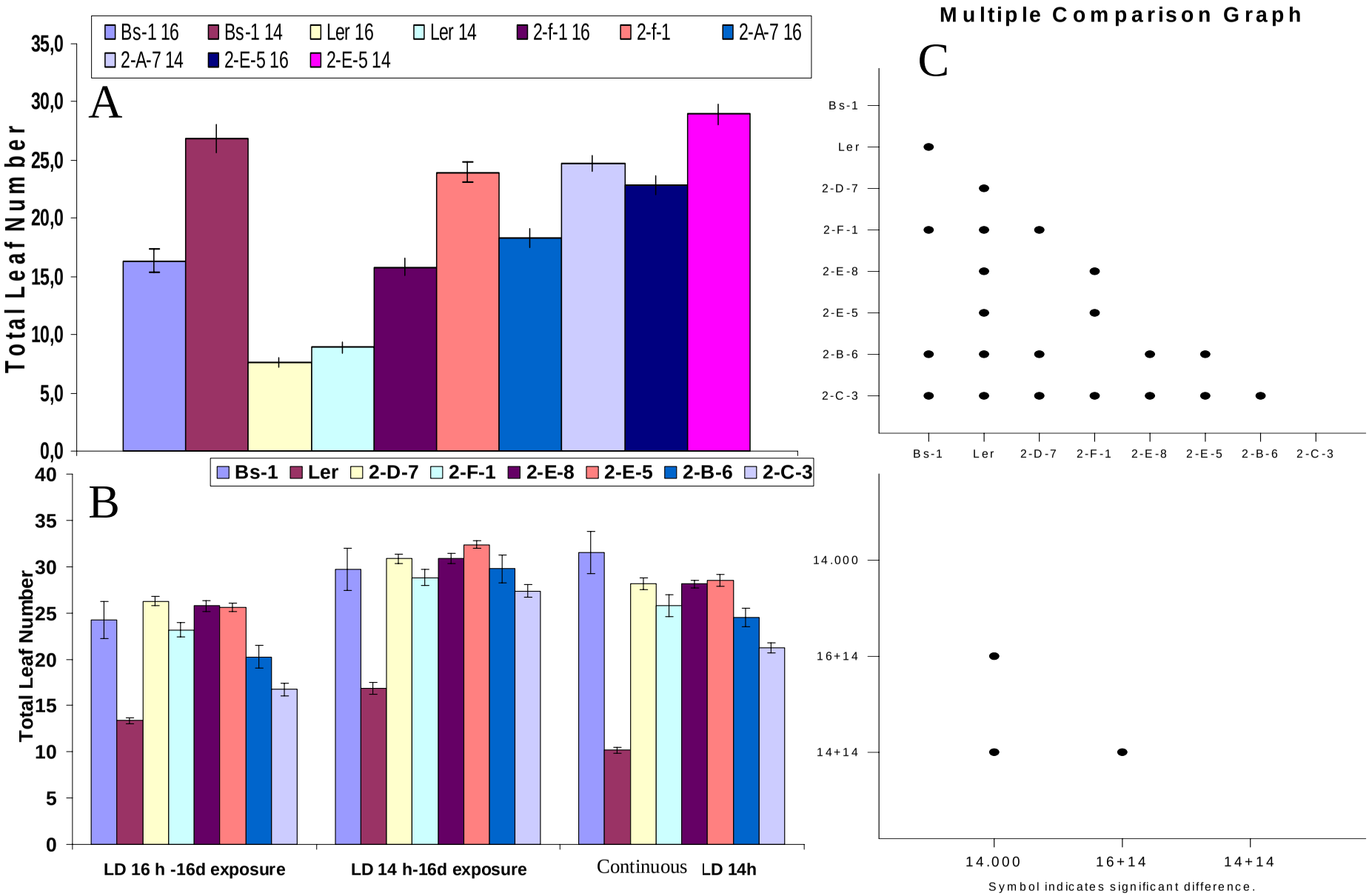


Figure 41 : Isolating NILs for the Bs-1 x Ler population. A) Flowering times of NILs 2-F-1, 2-E-5 and 2-A-7 under LD 16 h and LD 14 h. All lines behave like Bs-1 flowering later under LD 14 h vs LD 16 h in contrast to Ler, which sees no difference. B) Flowering times of NILs in a trasient exposure to inductive LD of 16 and 14 h for 16 days. All lines behave similarely to Bs-1, responding weaker to LD 14 h. C) ANOVA result from comparisons between the NILs and the parents. Dots indicate the presence of a statistically significant pair wise comparison.

patterns of haplotype shaping around both loci, sufficient for successful Linkage Disequilibrium and eventually fine-mapping in a set of 196 accessions. Strong population structure and appearance of latitudinal clines have been also reported in these studies (Hagenblad et al., 2004; Stinchcombe et al., 2004).

It is interesting to note that although both FRI and FLC are targets of natural variation, the allelic properties of this variability are largely different between the two loci. In FRI, transposon insertions in the first intron and the promoter of the gene, result in knock out mutations due to failure in expression (Johanson et al., 2001; Shindo et al., 2005). The different types of deletions suggest that the loss-of-function character of FRI was evolved several times independently. In FLC however, weak alleles exist and are expressed although no function is being mediated (Fig 12, Michaels et al., 2003; Shindo et al., 2005). That is the situation with the Ler allele of FLC. Although largely capable of reducing the late flowering of an active FRI allele, the Ler FLC allele still retains partial activity (Michaels et al., 2001 and 2003; Caroline Dean personal communication). This pattern of allelic variability between the two loci is beneficial in an evolutionary context because it allows re-appearance of the vernalization sensitive character to occur even when vernalization insensitive populations are mixed.

Strong environmental interactions have been identified also for light responses and latitudinal clines have been identified (Maloof et al., 2001; Stenoien et al., 2002). Yet such clines are detected for hypocotyls lengths and not flowering time. Nevertheless several photoreceptors have been reported as targets of natural variation for flowering time in *Arabidopsis*. CRY2 represents such a target from cryptochromes, identified in Cvi. This accession is unique since it shows a day length insensitive profile with early flowering across all the examined day lengths of this project. The $CRY2^{Cvi}$ was identified as the major source of early flowering under short days in this accession. The CRY^{Cvi} allele carries 12 SNPs of which, 4 are located in the promoter, one in the 3' UTR and seven in coding regions resulting in 4 aminoacid changes (El-Din El-Assal et al., 2001). One highly conserved valine was among the affected aminoacids and as a result the CRY2 protein was stabilized under SDs, promoting early flowering. Interestingly the CRY2 protein profile was not affected under LDs, providing a molecular mechanism for conditional variability in flowering time. The allele of CRY2 is not found in other accessions and one reason for it could be its special adaptive significance. Cvi originates from Cape Verde Islands near the equator and an early flowering and day length insensitive phenotype is best adapted there compared to northern sites.

Mutations that affect protein stability have been selected in other photoreceptors by nature. The PHYA allele in Lm accession is one such case. Lm is approximately 100 times less sensitive to far-red light than Col-0 (Maloof et al., 2001). Usually light causes degradation of PHYA protein. However in the $PHYA^{Lm}$ allele a conserved methionine is changed to threonine. This sequence change stabilises the PHYA protein in the light. In addition autophosphorylation of the $PHYA^{Lm}$ was reduced by approximately 60 %. Collectively the authors suggest that the affected methionine is responsible for proper phosphorylation of the PHYA allele and subsequent degradation of the protein in a light dependent manner.

However not all photoreceptors are affected naturally at the level of protein or mRNA abundance. Natural variation resulting into aminoacid changes consisting of 65 synonymous and 14 non-synonymous polymorphisms among 14 accessions has been reported also for PHYB (Filiault et al., 2008). Evidence for population structure and strong LD are also found in the locus. Among 140 accessions, which were genotyped for these polymorphisms on PHYB, differences in hypocotyls length under red light exists. However the abundance of PHYB is not affected both at the mRNA and the protein level, suggesting differences in the functionality rather than stability of the PHYB message. This was confirmed with complementation test of the a *phyb* mutant by overexpressors of PHYB from Ler and Cvi (which differ in the aminoacid level). Generally $PHYB^{Cvi}$ conferred less responsiveness to the mutation than the $PHYB^{Ler}$, proving that indeed the aminoacid changes in PHYB are responsible for functionality rather than abundance of the protein or the transcript. In conclusion photoreceptors are also major targets of natural variation in *Arabidopsis*, which can affect flowering time. Unlike the FRI deleterious mutations, photoreceptor variants in nature are expressed and result in changes at the protein level and therefore resemble more the FLC pattern of natural variability. In CRY2 and PHYA the protein abundance is affected while in the case of PHYB the functionality of the protein is changed. Last but not least the cases of $PHYA^{Lm}$ and $CRY2^{Cvi}$ alleles appear to be unique in these accessions in contrast to the changes in the PHYB locus that are spread over many natural variants.

Cen-0 x Ler population

In the Cen-0 population two QTL have been identified at the top of chromosome 4 (0-3 Mbps) and 5 (3,0-5.7 Mbps, Fig 21). FRI (Chrom 4 - 0,26 Mbps) and FLC (Chrom 5 - 3,1 Mbps) are located in both regions. The fact that the F2 population shows non-conditional transgressive variation, which can be rescued by vernalization (Fig 17), strengthens the argument that these two QTL are indeed FRI and FLC. But how a response to vernalization

occurred since both Cen-0 and Ler do not respond to vernalization (Fig 12A and 13)? As mentioned Cen-0 has an active FRI allele since it can complement a *fri* null mutant (Fig 12C, Lempe et al., 2005). FLC levels are high in this accession (Fig 12B and 13) yet it does not respond to vernalization because of a novel weak allele of FLC (Lempe et al., 2005). More specifically an alternative acceptor splice site is being used, leading to partial deletion of coding sequences and a frame shift. Collectively the above data indicate that in the Cen-0 x Ler F2 population the FLC^{Ler} allele responds to the FRI^{Cen-0} allele leading to a late flowering transgressive segregation and to a vernalization response. This late flowering is greatly rescued when Ler alleles are inherited in the top of Chrom 4 in a homozygous state (CL 98, Fig 38) and is eliminated when Ler homozygous alleles are inherited in both the top of Chrom 4 and 5 (F3 segregation data not shown). FRI and FLC therefore cannot be the main genes that are responsible for the ability of Cen-0 to discriminate between the two day lengths.

PHYC is another photoreceptor together with PHYA and B and CRY2, which was identified as a target of natural variation affecting flowering time in *Arabidopsis* (Balasubramanian et al., 2006). In particular, one accession Fr-2 flowers early under short days, thus exhibiting a reduced response to day length. QTL mapping indicated a region in the middle of Chromosome 5 and microarray analysis in that QTL showed that among 34 accessions only the PHYC transcript showed reduced levels specifically in Fr-2. Sequencing of the gene in different accessions revealed many polymorphisms and a haplotype sharing pattern with a Col-0 type and a Ler type cluster. This indicates that diversity in PHYC alleles, like for PHYB and unlike PHYA and CRY2 spreads over many accessions.

$PHYC^{Fr-2}$ failed to complement the *phyc* mutant in Col-0 suggesting that this particular allele is not functional. Interestingly the $PHYC^{Ler}$ is also not a strong allele, unlike the $PHYC^{Col-0}$ since it failed to complement the *phyc* mutants as well (Supp Fig 2C, Balasubramanian et al., 2006). PHYC mediates red light responses like PHYB and the two sequences are very similar. In fact the PHYC and PHYB genes are considered to have originated from a duplication of a common ancestor. In *Arabidopsis* strong PHYB alleles cause late flowering due to the degradation of the CO protein (see introduction, Valverde et al., 2004). Fr-2 carries a weak allele of PHYC and shows an early flowering phenotype under SD, while the *phyC* mutant under SD has approximately 20 leaves of difference compared to the later flowering Col-0 WT. These data are consistent with the function of PHYC as a floral repressor.

QTL at the region of PHYC are expected in populations in which the two parents differ in the activity of this locus. Such QTL have been reported in Col-0 x Ler and Col-0 x Fr-2

populations. In the Cen-0 x Ler population, apart from FRI and FLC, another major QTL that segregates is located on chromosome 5 between 13.6 to 16 Mbps. In that position the PHYC photoreceptor is located (14 Mbps). Since Ler has a weak and therefore early flowering allele of PHYC, Cen-0 might carry a functional PHYC allele. This allele could be stronger than Col-0 and therefore be responsible in part for the later flowering under 14 h but not 16 h that is observed in Cen-0. The flowering time difference of 20 leaves is well within the effect that the PHYC allele has been reported to confer in SDs. In addition microarray data confirm that the PHYC transcript in Cen-0 does not differ from the rest of the accessions and therefore the differences are at the level of protein functionality (Lempe et al., 2005, AtGenExpress project). The above collectively suggest PHYC as a strong candidate for that QTL.

Bs-1 x Ler population

Unlike the Cen-0 population, in the Bs-1 case the NILs indicate that the major QTL on top of chromosome 5 (7,5 - 8,8 Mbps) regulates the character of day length discrimination between LD of 14 and 16 h. Within that region three candidate genes are located, namely HUA2 (7,7 Mbps), PRR5 (8,3 Mbps) and FPF1 (8,5 Mbps). HUA2 was identified as an enhancer of the *ag-4* phenotype together with HUA1 (Chen et al., 1999). The protein sequence contains a RPR domain, which is involved in mRNA metabolism. Therefore it is believe that HUA2 is involved in mRNA processing of certain targets. Apart from enhancing the homeotic phenotypes of AG mutants, HUA2 has also a flowering time phenotype since it acts as a repressor of flowering by enhancing the expression of certain floral repressors such as FLC (Doyle et al., 2005). Mutations in the gene affect daylength perception between classical LD of 16 h and SD of 8 h mainly by causing early flowering under SDs (Doyle et al., 2005).

This gene was recently shown to be a target of natural variation both for flowering time and shoot architecture (Wang et al., 2007). More specifically a tyrosine substitution results in a gain-of-function unique allele in Sy-0 ecotype, causing elevated levels of FLC in the SAM, late flowering in combination with active alleles of FRI and eventually altered shoot morphology through AG depended effects. Apart from FLC the effect on other MADs box repressors is questionable. Wang and colleagues report no effect on the MAF and SVP mRNA expression by the Sy-0^{HUA2} allele, yet in the study of Doyle and colleagues there is an effect. These differences could be due to the special nature of the Sy-0 allele that was used in the first study. Interestingly some but not all strains of Ler (Landsberg) carry a weak HUA2 allele. WT Ler plants seem to have normal HUA2 alleles, which can complement *hua2*

mutants. Nevertheless L*er* (Landsberg *erecta*) plants that are the most commonly used laboratory strain of Ler, seem to be defected in HUA2, probably due to a mutation from the mutagenesis screen that resulted in the *erecta* phenotype (Maarten Koornneef personal communication, Doyle et al., 2005). The L*er* strain that is used in the current project falls into that category. Collectively the above suggest that a second gain-of-function $HUA2^{Bs\text{-}1}$ allele might cause the difference between LD of 14 and 16 h by elevating the levels of FLC. Ler would have a weak allele and therefore would not show this difference. This would also fit to the observation that in the level of the F2 population, vernalization rescued late flowering under 14 h and phenocopied the 16 h pattern of segregation (Fig 16). However several arguments suggest also against of HUA2 as a candidate. In the F3 validation, vernalization of the BL 72 population, failed to show any response, yet a day length response was observed. In addition the effect of vernalization on the BL 4 population was limited only in causing earlier flowering both under LD of 14 and 16 h, without affecting day length response (data not shown). In addition the gain-of-function allele of Sy-0 is unique and although Bs-1 was not studied, such alleles were not found in 150 other accessions, suggesting that are rare. Even in such a case an explanation must be given as to why the elevated FLC levels would affect flowering specifically under 14 and not 16 h.

Another very promising candidate is PRR5. The gene is located exactly in the middle of the fine mapped region, and usually genes in such positions are the causes of the observed phenotype. Generally the PRR genes apart from clock specific effects have a role in flowering time. This role was largely neglected until now. Several combinations of PRR mutations were included in the mutant screen of this project. A synergistic role for PRR5 and 7 was identified with mutations in both genes causing late flowering and changing the properties of day length perception stronger under LDs compared to SDs (Fig 42A). An effect of *prr* mutations on flowering time was recently reported also in another study (Nakamichi et al., 2007). Double mutants of *prr5prr9* and *prr5prr7* result in late and very late flowering respectively and suggest functional redundancy between these genes (Fig 42A). Interestingly CO mRNA levels were not affected by the PRR mutations but on the contrary, *FT* and *TSF* transcript levels were significantly reduced (Fig 42B). Such observations implicate an effect of the PRR genes at the post-transcriptional level, locating them between *CO* and *FT*.

An effect at the protein level of CO, should involve a connection to the plant photoreceptors, which are known to control CO protein stability and activity (Valverde et al., 2004). In order to establish such a connection molecularly, the effect of the different PRR mutant combinations is the photoperiod pathway, was compared to the effect on

photoreceptor mutants. Several genes with diverse functions and hierarchy within the pathway were selected and their mRNA profiles were tested with RT-PCR in ten mutant combinations under LD of 16 h. The results of this transcriptional analysis were included in a hierarchical clustering after analogous to microarray data transformations. The hierarchical clustering analysis showed that the PRR mutations cluster very closely with CRY mutants based on the effect on all tested genes (Fig 42C). More specifically genes such as *FKF1*, *GI* and *TOC1*, which are evening, *LHY* which is morning and *CO*, *FT* and *TSF* which are further downstream in the hierarchy of the pathway show almost identical pattern of expression between the *prr5* and the double *cry1cry2* mutant (Fig 42B).

In order to further establish a role with cryptochromes, a genetic approach was followed. An early flowering overexpressor of CRY2 was crossed independently to four single and double late flowering *prr* mutants. The segregation of each cross was determined under LD of 16 h. The *prr* mutants overcame the effect of the CRY overexpression in a dosage dependent manner with the double *prr7prr5* mutation having the strongest effect and the single *prr9* mutant the weakest exactly as it was observed for the flowering times of the *prr* mutants alone (Fig 43). Collectively these data indicate that the role of the PRR genes is to promote flowering at the post-transcriptional level of *CO*. The *PRR* genes express their role by acting as positive elements of the cryptochrome signalling, at least downstream of *CRY2*. These phenotypes together with the dosage-dependent effect of the *prr* mutations make the *PRR5* a strong candidate for the Bs-1 x Ler population.

So far the previous candidates express their function mainly in the vasculature. The last candidate for this QTL is FPF1, a gene that is expressed in the SAM. The full-length cDNA of FPF1 was isolated in cDNA library from meristematic RNA while the protein encodes a small peptide of only 12 kDa (Kania et al., 1997). The transcript of this gene is not detected in the leaves, while its expression is up regulated in the SAM upon floral induction. *FPF1* is one of the earliest flowering time genes the expression of which changes upon floral induction (Kania et al., 1997). The gene is associated with a GA response, since overexpressors increase the response to applied GAs. Last but not least *FPF1* seems to affect juvenile phase in *Arabidopsis* with FPF1 overexpressors shortening it (Melzer et al., 1999).

Several mutations were combined with the FPF1 overexpressor and the flowering times of these lines were tested under long and short days. The *FCA* and *FVE* autonomous pathway genes showed the biggest shortening in the flowering time when combined with the FPF1 mutant, while *CO* and *GI* mutations responded weaker. This indicated that *FPF1* cannot fully complement the function of *CO* and *GI*. In addition the early flowering of FPF1

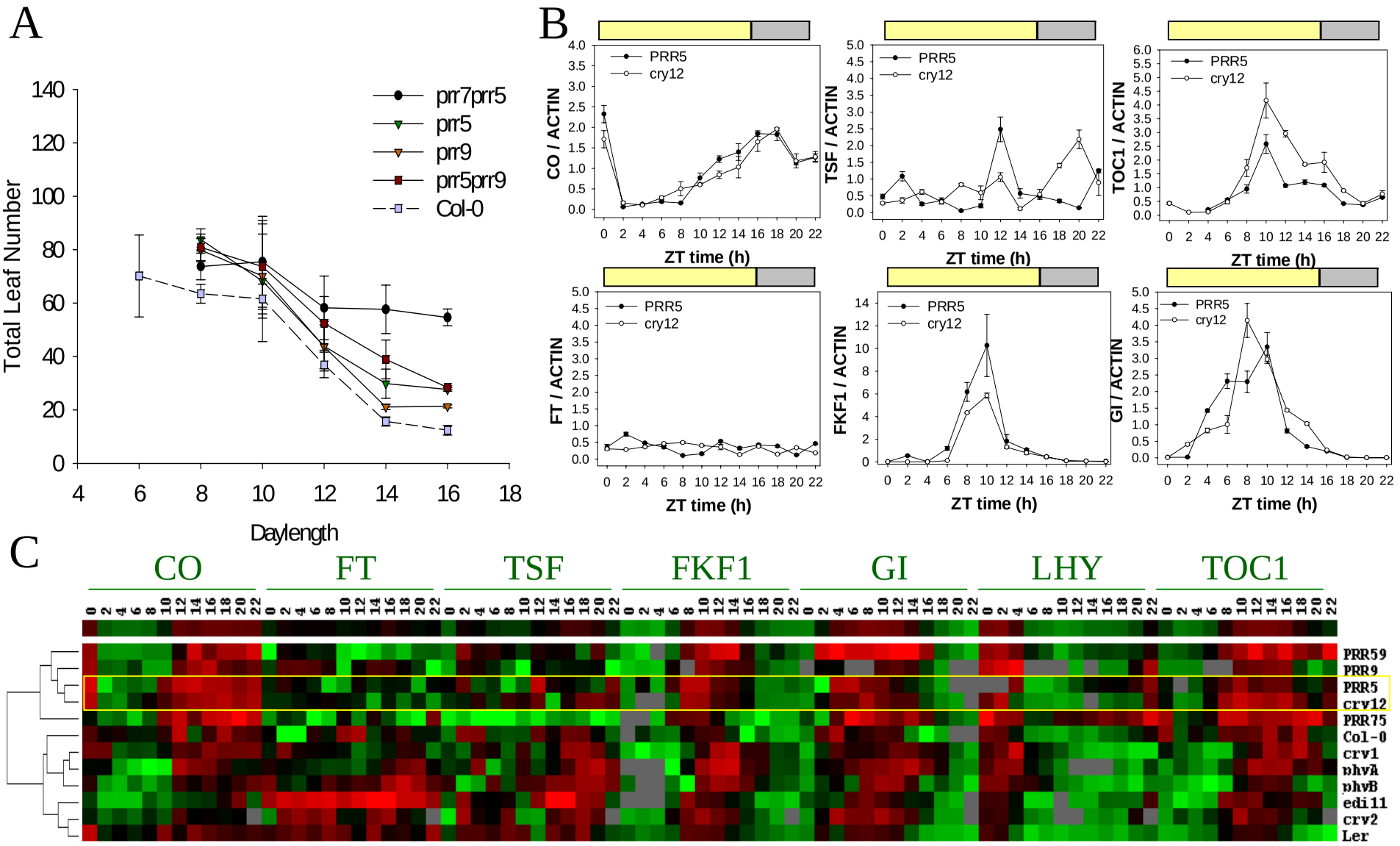


Figure 42 : Mutations in PRR genes cause similar effects to flowering time genes like *cry* mutants. A) Day length response curves of PRR mutant combinations and *CRY* combinations. B) Hierarchical clustering based on the mRNA expression under LD 16 h of 7 major flowering time genes between many photoreceptor and *prr* mutant combinations, places the *prr* mutations together with the double *cry1cry2* mutant. C) Comparison of the mRNA expression of flowering time genes between *prr5* and the double *cry1cry2* mutant, showing very similar effects.

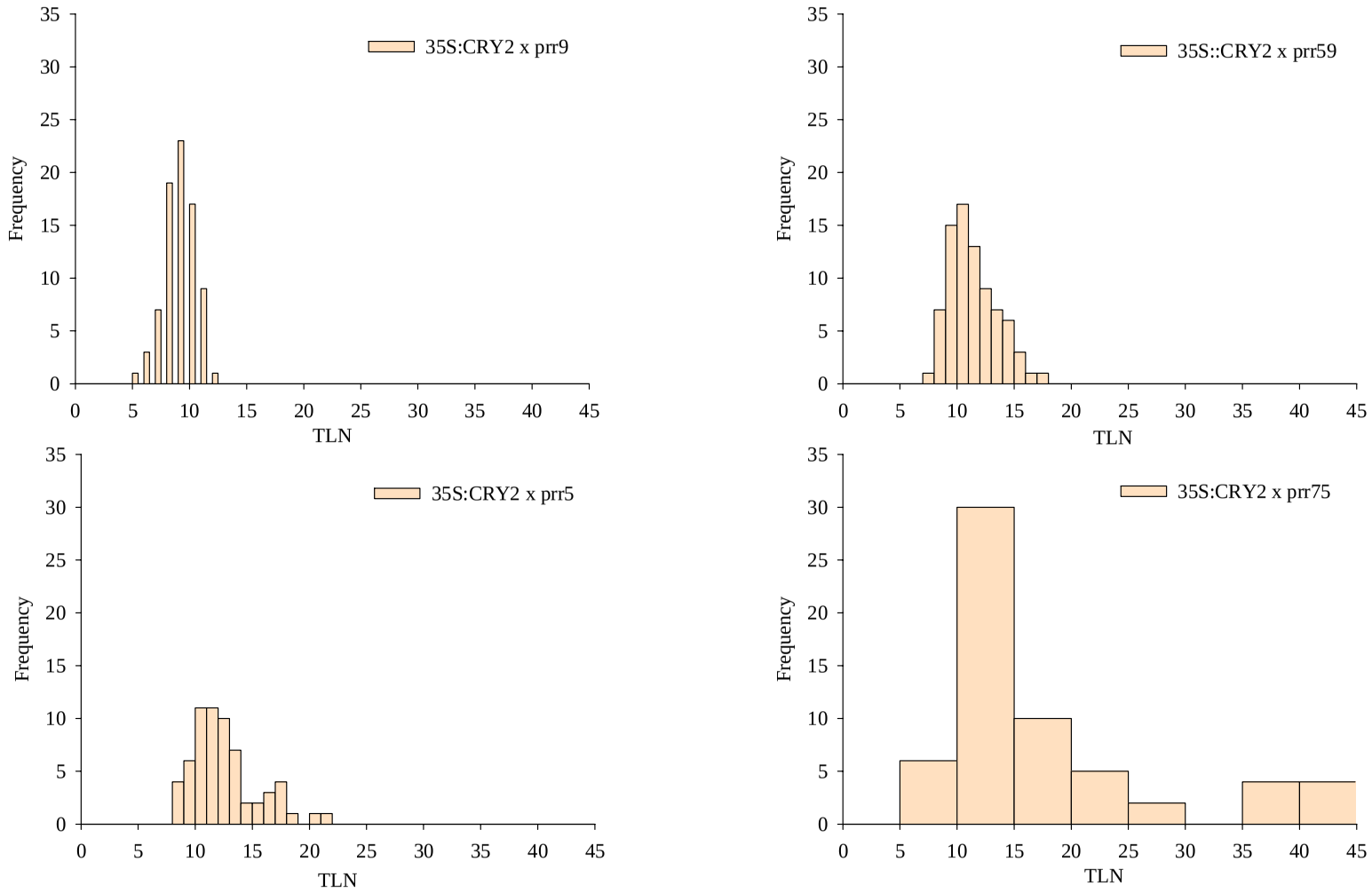


Figure 43 : Segregation of a cross between 35S::cry2 and several combinations of prr mutations. In all cases the extremely early flowering of the 35S::cry2 allele was totally masked by the prr mutations suggesting that the PRR genes act downstream of CRY2.

overexpressors was not affected by LFY mutations suggesting that the gene acts in parallel to LFY regarding the regulation of flowering inititation. However an attenuation of the homeotic transformations of the *lfy* mutations was observed in FPF1 overexpressors suggesting multiple roles for the gene. Additional results coupled *FPF1* to the function of AP1 and suggest a synergistic role between these genes in the regulation of flowering time (Melzer et al.,1999). Such a scenario suggests that *FPF1* is responsible for the competence of the SAM to the inductive signals, thus explaining why CO and GI mutants are partially rescued by FPF1 overexpressors.

An FPF1 overexpressor from *Arabidopsis* affected flowering time in several varieties of *Nicotiana* regardless of photoperiod requirements for flowering inititation, suggesting a general and conserved role for the gene between species (Smykal et al., 2004). Although FPF1 is a good candidate, the fact that the gene is expressed after floral transition in the SAM limits it role in only in modulation of meristem competence in the Bs-1 case. Additional limitations come from the fact that the gene is closely associated with gibberellins more than the photoperiod pathway.

Dijon-G x Sha population

Under short days, a QTL at the top of chromosome 4 was isolated in the Dijon-G x Sha population. This QTL most probably is FRI since Sha has an active allele and Dijon-G like Ler, a weak allele (Fig 12C, Johanson et al., 2000). The QTL at the bottom of chromosome 5 in that population is probably the *MAF* cluster, which confers late flowering specifically under SD in the absence of FLC strong alleles. *MAF1*, although it is on chromosome 1, it was been reported as a target of natural variation for flowering time in a cross with Sha as one parent (Werner et al., 2005). In the Dijon-G x Sha population no QTL around *FLC* is identified because both accessions carry weak FLC alleles (Michaels et al., 2003). Although vernalization response was not described yet in that population, a *FRI*/*MAF* interaction might be responsible in part for the transgressive variation observed under both SD of 10 and 8 h. Regarding day length discrimination the conditional QTL in the middle and bottom of chromosome 3 together with the conditional epiQTL at the top of the same chromosome have central roles. No obvious candidates are present in these regions, but study of published microarray data, suggest candidate genes which will be tested in RT-PCR (data not shown).

3.1.2.13 Summary and conclusions

Both Bs-1 and Cen-0 show a similar flowering response between the two nearly identical long day lengths of 14 and 16 h. However the lateness that is conferred by the Cen-0 genome seems to be controlled by a larger number of dominant loci, compared to Bs-1. Consistent with this observation, are the flowering times of the F1 hybrids. The vernalization treatment of the F2 populations suggest also that a vernalization response was restored in the Cen-0 x Ler F2 population and resulted in non-conditional transgressive variation between the two conditions. However this response was not enough to eliminate the differences between the two day lengths. Responsible for this response is the weak FLC^{Ler}, which responds to the strong FRI^{Cen-0} allele. Yet none of the two parental genotypes shows a response to vernalization since either *FRI* or *FLC* alleles are not active and therefore cannot explain the parental phenotypes. Good candidate for this is a strong *PHYC* allele in Cen-0, which could cause late flowering and overcomplement the naturally weak allele of Ler.

In the Bs-1 population the *PRR5* gene is a strong candidate, together with *HUA2* and *FPF1* genes. All of these genes have been associated with flowering time variation between different strains in a dosage-depended manner that fits to the semi-dominant nature of the Bs-1 QTL. *HUA2* and *PRR5* express their function in the vasculature by affecting the production of the *FT* through *FLC* and *CO* respectively, while *FPF1* changes the competence of the SAM to the florigen. *PRR5* is associated with crytochrome activity and acts together with other *PRR* genes downstream of *CO* mRNA and upstream of *FT*.

3.2 MOLECULAR ANALYSIS OF NATURAL VARIATION IN DAY LENGTH PERCEPTION FOR FLOWERING TIME

3.2.1 Introduction

Very often in the literature of natural variation in plants, humans, insects and yeast, transcriptional differences in key regulatory genes explain molecularly the observed variability between strains. For that reason, several different approaches aiming in the molecular characterisation of flowering time genes under specific day lengths and between selected accessions were selected. Initially an expression profiling of flowering time genes was conducted in interesting accessions aiming to identify cases of natural variation in flowering time genes expressed in the leaves, under steady-state day length conditions. The results of this approach were complemented with expression analysis of the same targets and conditions, in selected mutants of candidate genes from the QTL analysis.

Apart from the expression under standard continuous daylegths, these approaches involved an expression profiling of the flowering time genes such as the floral integrator *FT*, during a transient shift into long days. It is known that only under inductive long day conditions, *CONSTANS* mRNA peaks shortly before the beginning of the night (Suarez-Lopez et al., 2001, Imaizumi et al., 2003). Then, the stabilized by the cryptochromes and phytochrome A, CONSTANS protein accumulates and the expression of the downstream target *FT* is up regulated. What happens however, when *Arabidopsis* plants experience inductive long day conditions transiently and then return into non-inductive short days? Two possibilities exist. i) *FT* expression is up regulated only during long days. In other words, the expression of the gene is regulated in a transient manner, which is strictly dependent on CO and inductive photoperiods. ii) Once *FT* expression is up regulated, it is maintained and therefore not affected by photoperiod and CONSTANS. This could suggest the existence of a direct or indirect auto-regulatory loop, which once activated, ensures high levels of *FT* expression regardless of photoperiod.

Any of these two possibilities can be true and could lead to additional natural variation. In other terms accessions may differ in their ability of responding fast into inductive photoperiods, ensuring floral commitment. But also, within a given timeframe, different photoperiods may have differential strength in committing the same strain with a reproductive fate. And to expend this even further such natural variation, if identified for the first time, it can be attributed to flowering time genes expressed in the vasculature or the SAM or both, leading to a spatial aspect of natural variation in floral commitment.

3.2.2 Identifying candidate genes in the QTL regions using available microarray data.

As mentioned in the introduction, QTL mapping in the post-genomic era is often combined with genome-wide expression analysis. The fact that transcription is often responsible for natural genetic variation, was used in several mapping projects successfully in order to shortcut the steps of fine mapping and to identify candidate genes and in some cases the causal genes themselves. Microarray data between different accessions (including Cen-0 and Ler) are publicly available and were selected in order to be re-analysed with the aim of identifying candidate genes in the Cen-0 x Ler population. As mentioned in the previous chapter this population has many QTL and day length perception seems to relay in the regulation of many of these loci via complex interactions.

The first step in the analysis was to correlate the expression of all genes within each of the QTL regions in the Cen-0 x Ler population, between Ler and the Cen-0 accession (Fig 44 A to E). Transcripts with no differential expression between the two accessions are expected to align with the 45º-regression line in a scatter plot with balanced axis. Significant deviations from this trend indicate the existence of transcripts that are differentially expressed between the two accessions. Indeed such transcripts were identified in nearly all QTL regions with the QTL in the middle of chromosome 4 hosting most of them. If these transcripts indeed affect day length perception then they resemble *cis* type eQTLs under the notion that position and effect (transcriptional differences with physiological impacts) map in the same region.

The second step in the analysis was to search without bias in the whole genome for differentially expressed transcripts (Fig 44F). Many transcripts were found to be outside of the confidence limits of the regression analysis. A classification of such targets was made according to several criteria. First targets with obvious functions that do not fit flowering time and day length regulation were excluded. The second step was a classification of the remaining targets according to the differential regulation between Ler and Cen-0. Finally out of the remaining targets emphasis is given in those that are differentially expressed mainly or only between Cen-0 and Ler and not the rest of the accessions. Examples of such transcripts are shown in Fig 45. The microarray data also confirmed the differential up-regulation of FLC expression in Cen-0 vs Ler and indeed this up regulation was among the highest differences between the two accessions. These transcripts will be tested in the two accession under LD of 14 and 16 h with RT-PCR to search for targets which are differentially regulated also between the two day lengths. Also this analysis opens the path for further comparisons with a new

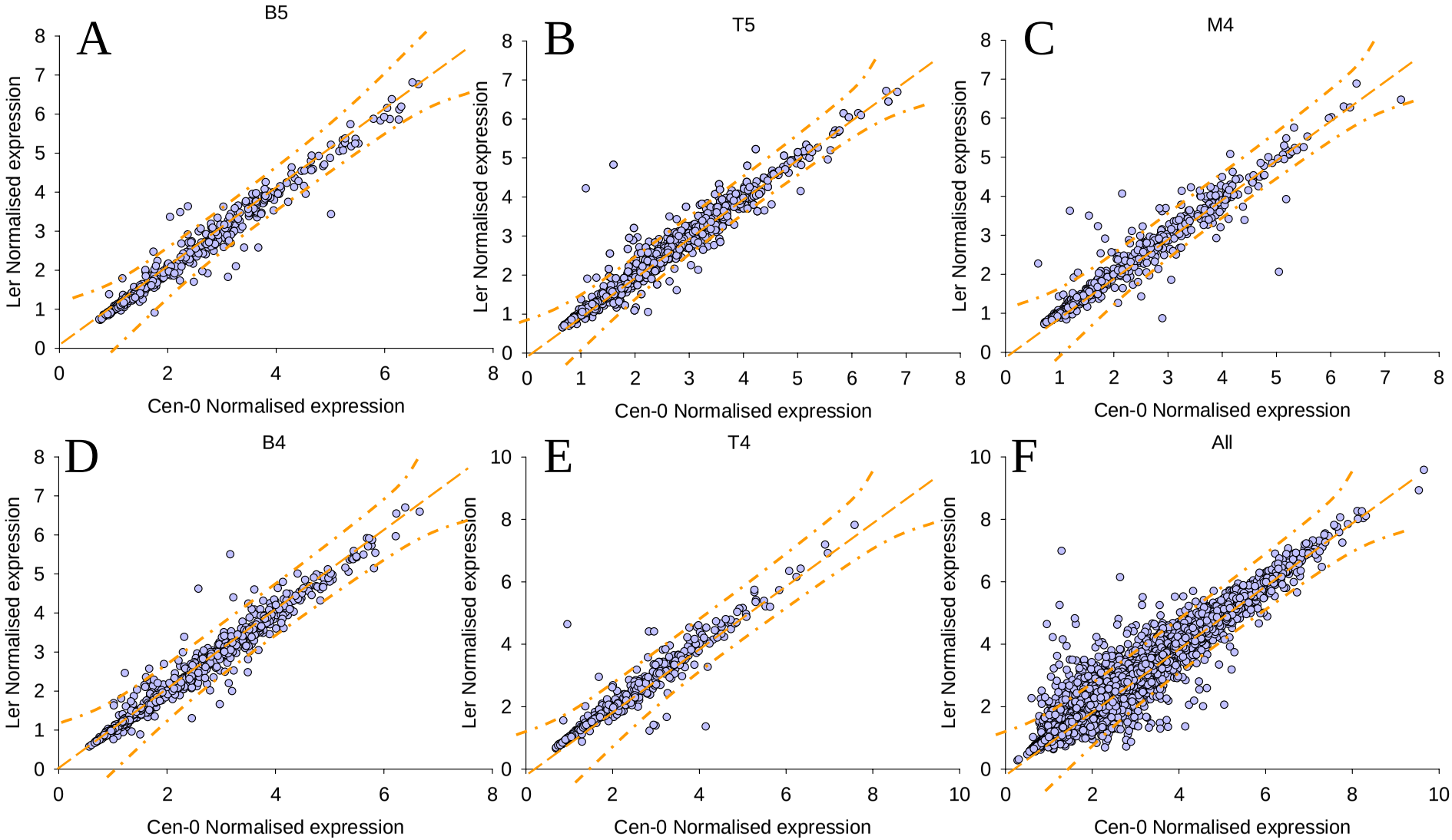


Figure 44 : Transcriptional scanning of the QTL regions in the Cen-0 x Ler. Published microarray data between accessions under LD 16 h were re-analysed. The QTL regions on Chromosome 5 Bottom (A) Top (B) and chromosome 4 middle (C), bottom (D) and top (E) were selected and scatter plots of the transcriptional activity of Ler vs Cen-0 are shown. The straight yellow line represents the expected equalibrium in the expression of genes between the two accessions. Th e curved yellow lines represent the confidence limits below and above which the expression of the corrrsponding targets differs between the two accessions. F) Genome-wide comparison.

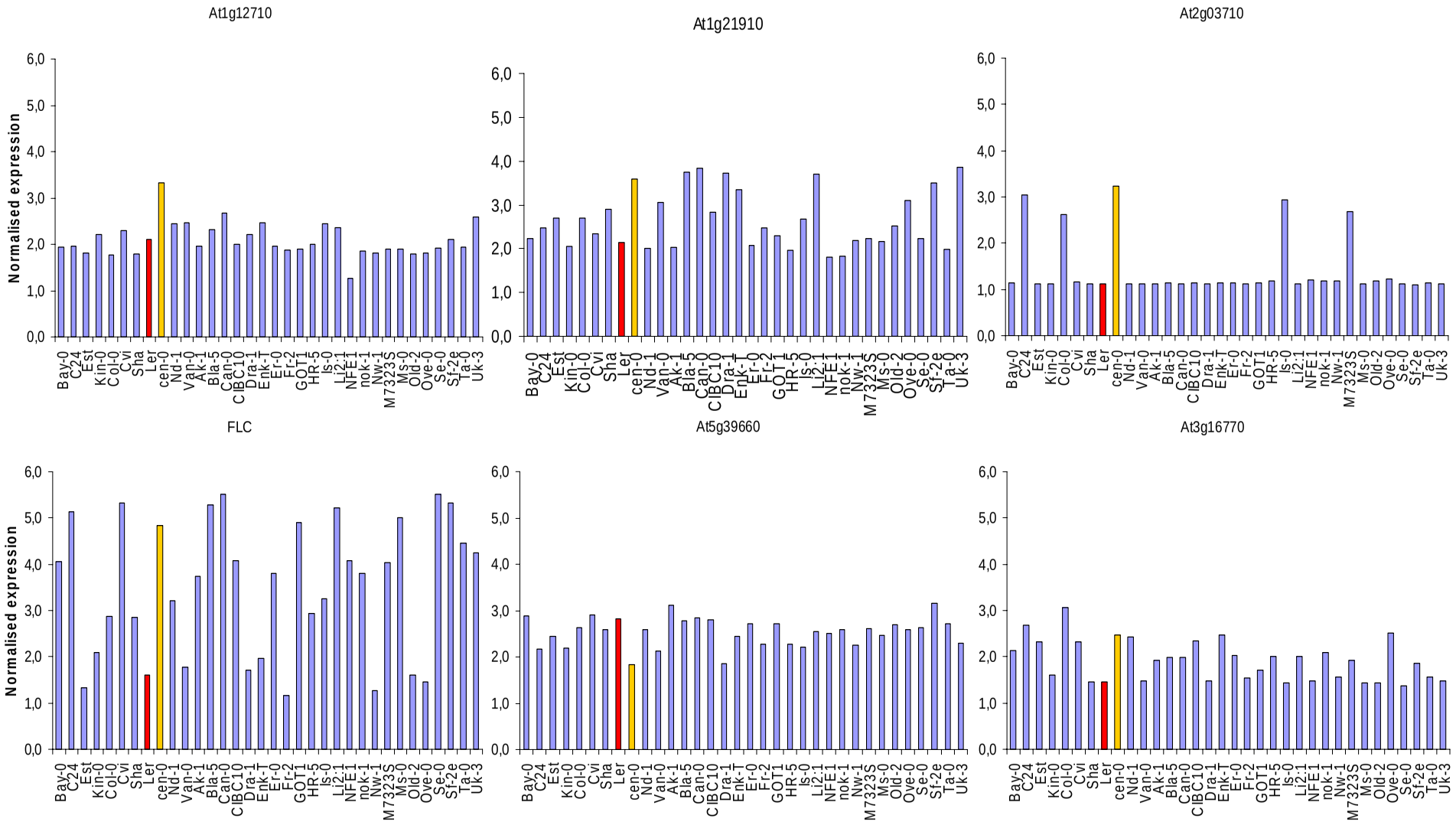


Figure 45 : Comparison of some of the most up- and down regulated targets between Cen-0 and Ler, with other accessions. Apart from the the differences between Ler and Cen-0, such targets are also selected on the basis of putative implications in flowering time regulation. Targets, which are downregulated in Cen-0 are avoided since in the absence of SFP signatures for Cen-0, miss-hybridised false positives cannot be excluded. Yellow bar : Cen-0. Red : Ler.

microarray analysis specifically designed under the conditions and accessions of this project, which is underway.

3.2.3 mRNA expression analysis of flowering time genes in selected accessions under constant day length conditions.

Based on the differences of flowering time (Fig 8D) Cen-0 and Bs-1 were selected and the expression of flowering time genes in these accessions was compared to Ler under both steady-state LD 16 and 14 h. As shown in Fig 46A, a two hours difference in phase was observed in *CO* mRNA expression in Cen-0 under LD of 14 h compared to LD 16 h. This difference was obvious for both the first evening peak as well as the second night peak. In contrast *CO* expression in Ler did not show any difference in phase or amplitude between the two day lengths (Fig. 46B). This phase difference in *CO* affected the expression of *FT* in Cen-0 (Fig. 46C). Under LD 16 h FT expression is significantly elevated at ZT 14, reaches its peak at ZT 16, and is followed by a small decay until ZT 18 before it returns to background levels of detection. However, under LD 14 h, a sharp peak is observed which now occurs two hours earlier, at ZT 14, compared to LD 16 h, and then decays much faster. Note that *FT* mRNA expression in Cen-0 shows no morning peak in both day lengths, but rather a gradual decay, which is more obvious under LD 16 h. In Ler, FT is expressed almost identically in both conditions, however the peak now occurs much earlier (ZT 12, Fig 46D). Consistent with the phenotypic difference, FT is expressed at much higher levels in Ler compared to Cen-0 in both day lengths. Note that the evening peak in *CO* mRNA expression in Ler occurs very early (ZT 10). This provides enough time for CO protein to overlap with light and to up-regulate *FT* mRNA expression in high levels under both conditions. It would be interesting to monitor *CO* and *FT* mRNA expression under LD 12 h (ZT 12), where Ler shows the first signs of a photoperiodic response.

Another difference is the occurrence of a morning peak under LD 16 h in Ler, which is reduced under LD 14 h (Fig 46D). Since Ler shows no difference in the flowering time between the two day lengths, this difference in *FT* mRNA expression in the morning may not be crucial for flowering time. However the high levels of *FT* expression in the afternoon in both day lengths can easily compensate for the morning difference, leading into photoperiodic insensitivity. Note also that the expression of *FT* in Ler never reaches zero levels. *SOC1* mRNA, like *FT*, was expressed at higher levels in Ler compared to Cen-0, but did not show any phase difference between the two day lengths in both accessions (Fig 46 E and F). Higher levels of *SOC1* expression were observed during the evening under LD 16 h compared to LD

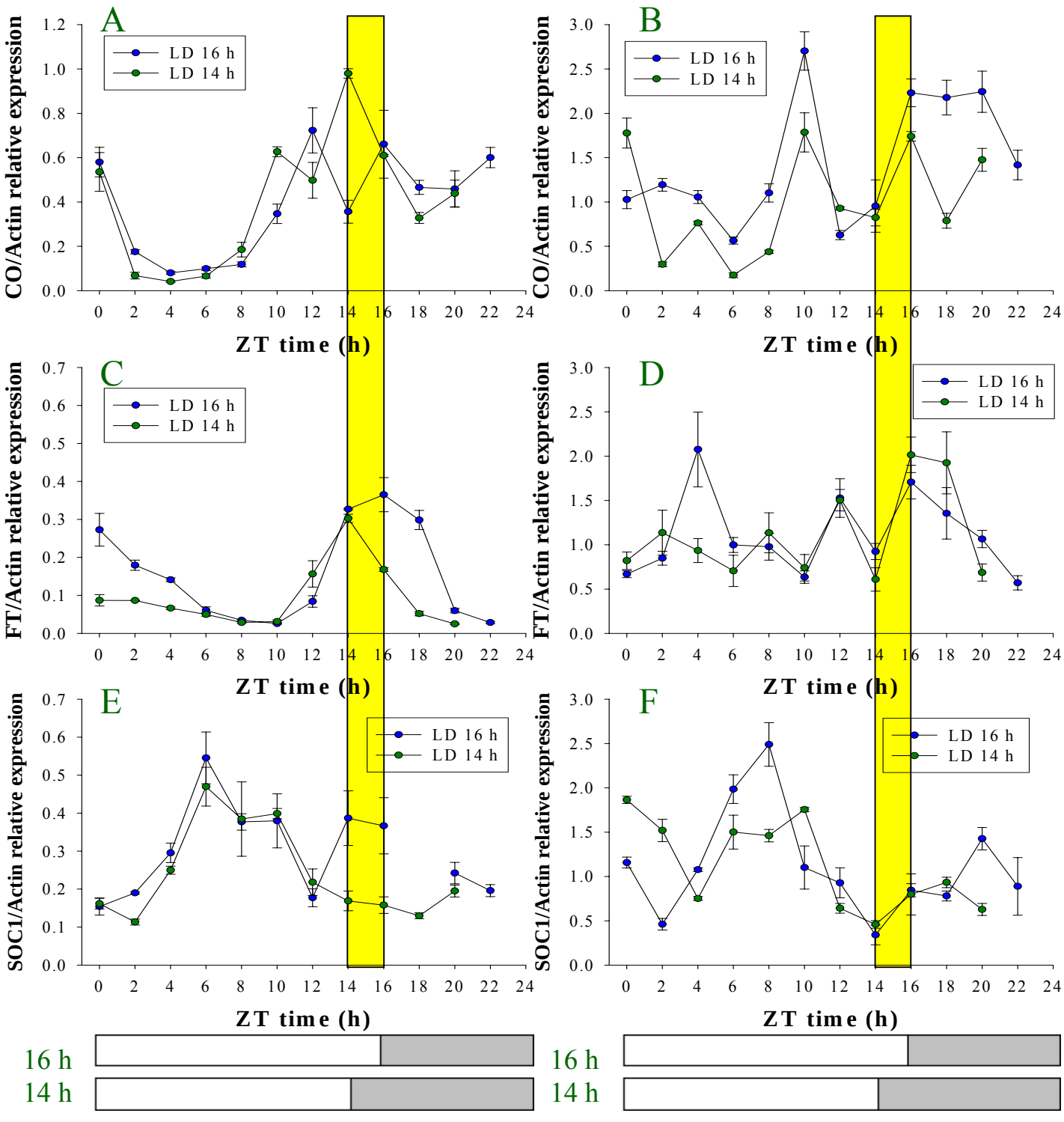


Figure 46 : Natural variation in the expression of known genes of the photoperiod pathway in selected accessions. A) C) E) Expression of *CO*, *FT* and *SOC1* in Cen-0 under LDs 16 h (blue curve) and LD 14 h (green curve). B) D) F) Expression of *CO*, *FT* and *SOC1* in Ler under LDs of 16 h (blue curve) and LD 14 h (green curve). The yellow-highlighted part corresponds to the critical window between the two day lengths. White bars : day, grey bars : night. Note that the LD 16 h specific morning peak in Ler is not present in LD 14 h in the same accession. In addition a gradual decrease instead a peak was observed during the morning in both conditions for Cen-0, resembling the expression pattern of Cvi (data not shown)

14 h. Last but not least *TSF* expression was very low (background levels) in both day lengths in Cen-0 (Fig 47A). In Ler the gene was expressed in similar high levels in both day lengths. Again a difference in the morning is observed, with high levels being observed under LD 16 h, compared to LD 14 h (Fig 47B). Like in Cen-0, *CO* mRNA expression in Bs-1 showed two hours of difference in phase between the two day lengths (Fig. 47C). *SOC1* expression was reduced in LD 14 h compared to LD 16 h, with more pronounced differences in the evening (Fig. 47D).

3.2.4 Expression profiling of the floral integrator FT during a transient exposure to inductive photoperiods

When plants experience constant conditions in day length, *FT* mRNA expression will rise in a developmentally dependent manner, since older plants are more sensitive to the inductive signal (unpublished data). However in nature, environmental conditions continuously change according to the season yet plants manage to commit their meristems with a floral fate. Which are the kinetics of this response and how this commitment is achieved in the molecular level? Is it possible that after an initial and environmentally dependent response phase, a mechanism is activated in order to maintain the inductive state regardless of the external conditions? In order to answer these questions, a shift assay from non-inductive SDs to inductive LDs was designed.

Initially the expression of *FT* was monitored during a transient exposure to long days, using semi-quantitative RT-PCR and Southern blotting. The results were also confirmed with real time PCR. 10 days old Col-0 plants grown under SD of 8 h, were shifted for two days into long days and then back to short days. Although, *FT* expression was down regulated before the induction, it increased transiently during the long day exposure (Fig 48). However once the plants experienced short days again, *FT* expression was immediately reduced to comparable levels to the ones before the induction. Regarding the expression of *CO*, a shift from the mono-phasic pattern of expression under short days, to the biphasic under long days was observed (Fig 49A). In addition, our short day control showed no expression of *FT* throughout the experiment and stable mono-phasic pattern for *CO*.

Landsberg plants that were grown for two weeks under SD 8 h were used in a second bigger experiment. These plants were exposed to three (instead of two) long days and then shifted back to short days, with leaves being harvested only. As shown in Fig 49B, *FT* expression in the leaves, like in whole seedlings, was transiently during the long day exposure (Corbesier et al., 2007). Interestingly, a morning peak at ZT 4 was also observed, with a gradual increase in

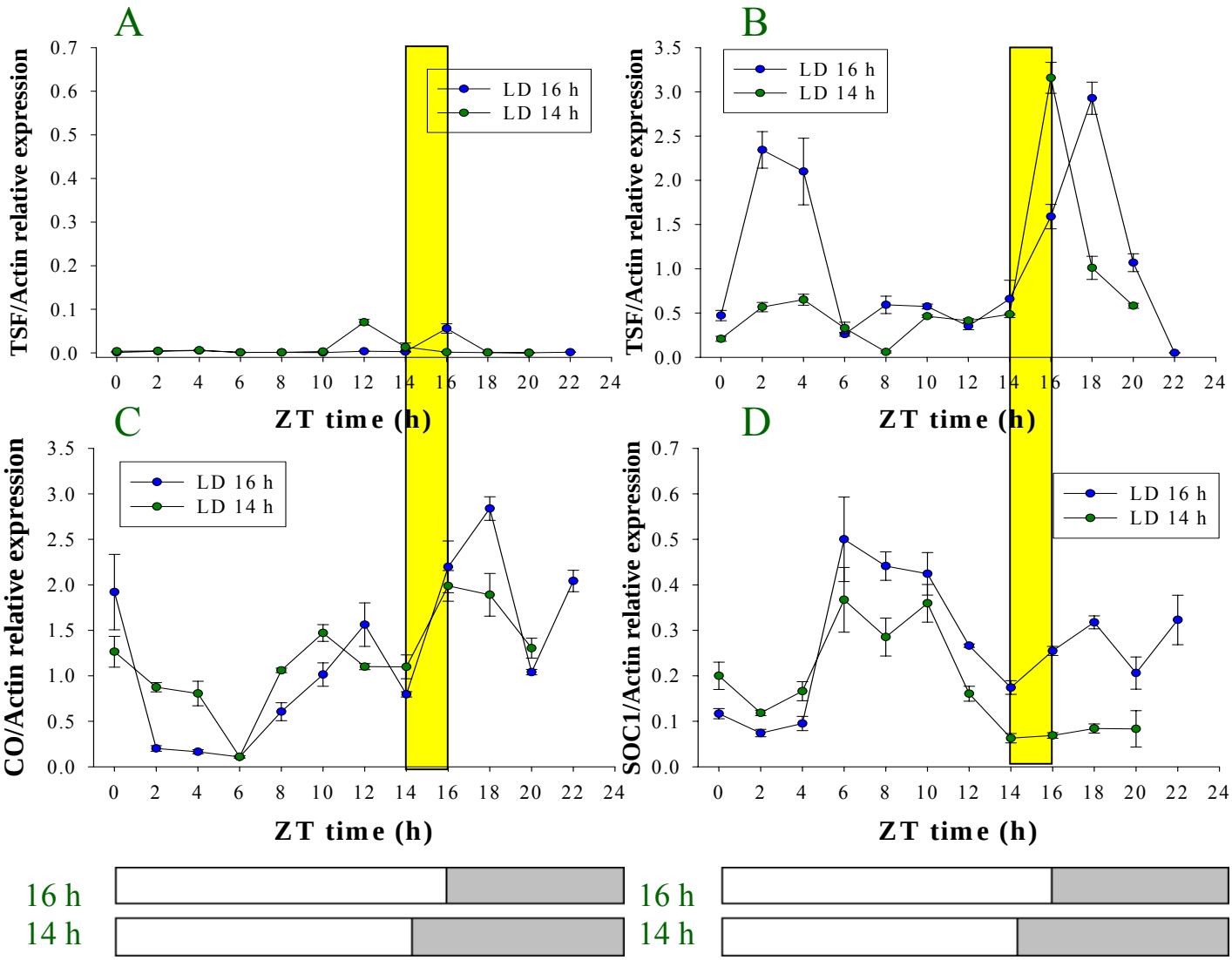


Figure 47 : Natural variation in the expression of known genes of the photoperiod pathway in selected accessions. A) Expression of TSF in Cen-0 under LD 16 h (blue curve) and LD 14 h (green curve). B) Expression of TSF in Ler under LD of 16 h (blue curve) and LD 14 h (green curve). C) Expression of CO in Bs-1 under LD of 16 h (blue curve) and LD 14 h (green curve). D) Expression of SOC1 in Bs-1 under LD 16 h (blue curve) and LD 14 h (green curve).

amplitude during the long day exposure. In the first long day of the induction, which followed the short days, only the normal evening peak and no morning peak were observed. In addition, in the first short day after the induction, which followed the three long days, *FT* expression showed a peak in the morning but not in the afternoon. This suggests that this morning peak requires at least one long day, and therefore the current model of external coincidence is not violated. During the remaining two short days the expression of the gene was similar to the levels before the induction or to the levels of the non-induced SD control (Fig 49B, Corbesier et al., 2007).

The expression of *TSF*, a gene similar to *FT*, showed similar patterns (Fig 49C). As an additional control the *co-2* mutation in Ler background was used. Both *FT* and *TSF* expression did not change during the induction, with similar levels to the non-induced control (data not shown). Flowering time data were consistent with the above results. *FKF1* and *GI* mRNA expression was also monitored but the resolution of the four hours was not sufficient to detect differences in the expression of the two genes (data not shown). The above data suggest that upon physiologically relevant photoperiodic induction conditions, *FT* and *TSF* mRNA levels increase transiently in the vasculature, regardless of the differences between the two accessions, the age, as well as the duration of the induction. This regulation occurs via a CO-dependent manner. However, the existence of a loop that is being triggered by extremely high *FT* levels, or that is present in even older plants and maintains *FT* expression, cannot be excluded.

3.2.5 Natural molecular variation in the kinetics of the response to day length upon transient exposure – Focusing in the vasculature

In nature, seasonal variation in climatological parameters that affect flowering time is known. Such parameters are overall levels of day length and temperature during the year. It is known that Northern sites reach higher and lower values in both parameters during the year and according to the season, compared to southern sites. This pattern sets a certain potential of earliness that each accession can reach in a specific niche, very much like the sigmoid trends in day length perception that where observed in this project. However another critical aspect of day length with adaptive significance, which has been neglected so far is the kinetics with which day length changes during the season. This is what the plant truly experiences in nature, not the steady conditions of the cabinet (or the *de facto* short days of winter or long days of summer) but a gradual increase in day length as the season moves to summer and most importantly a gradual decrease in the same parameter as the plant moves from autumn to

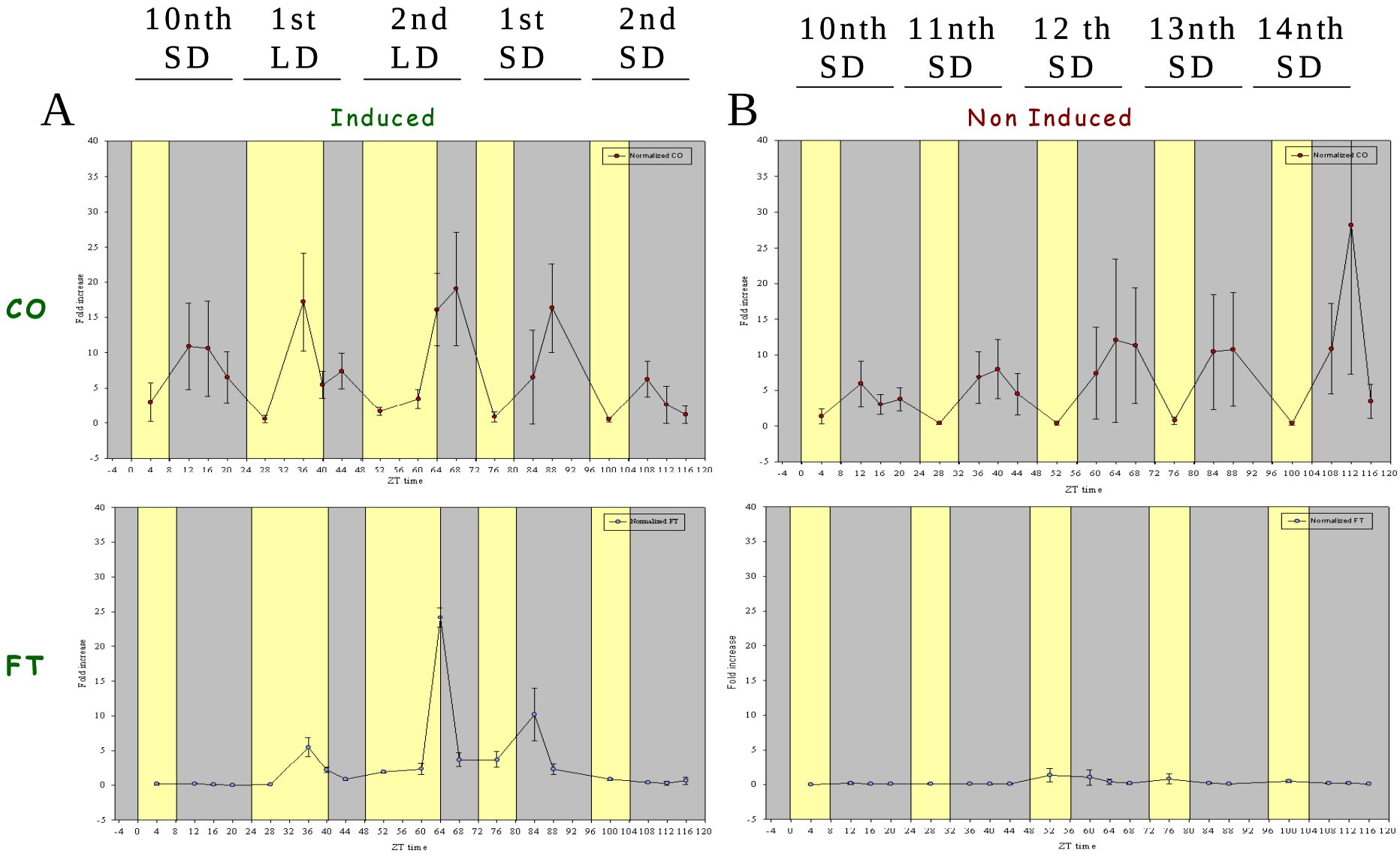


Figure 48 : Normalised results of the preliminary shift experiment. A) Upon exposure for two days under LD 16 h *CO* mRNA becomes byphasic and induces FT expression. B) However in the negative control plants which experienced constant SDs, *FT* expression never risen and *CO* mRNA pattern was continuously monophasic

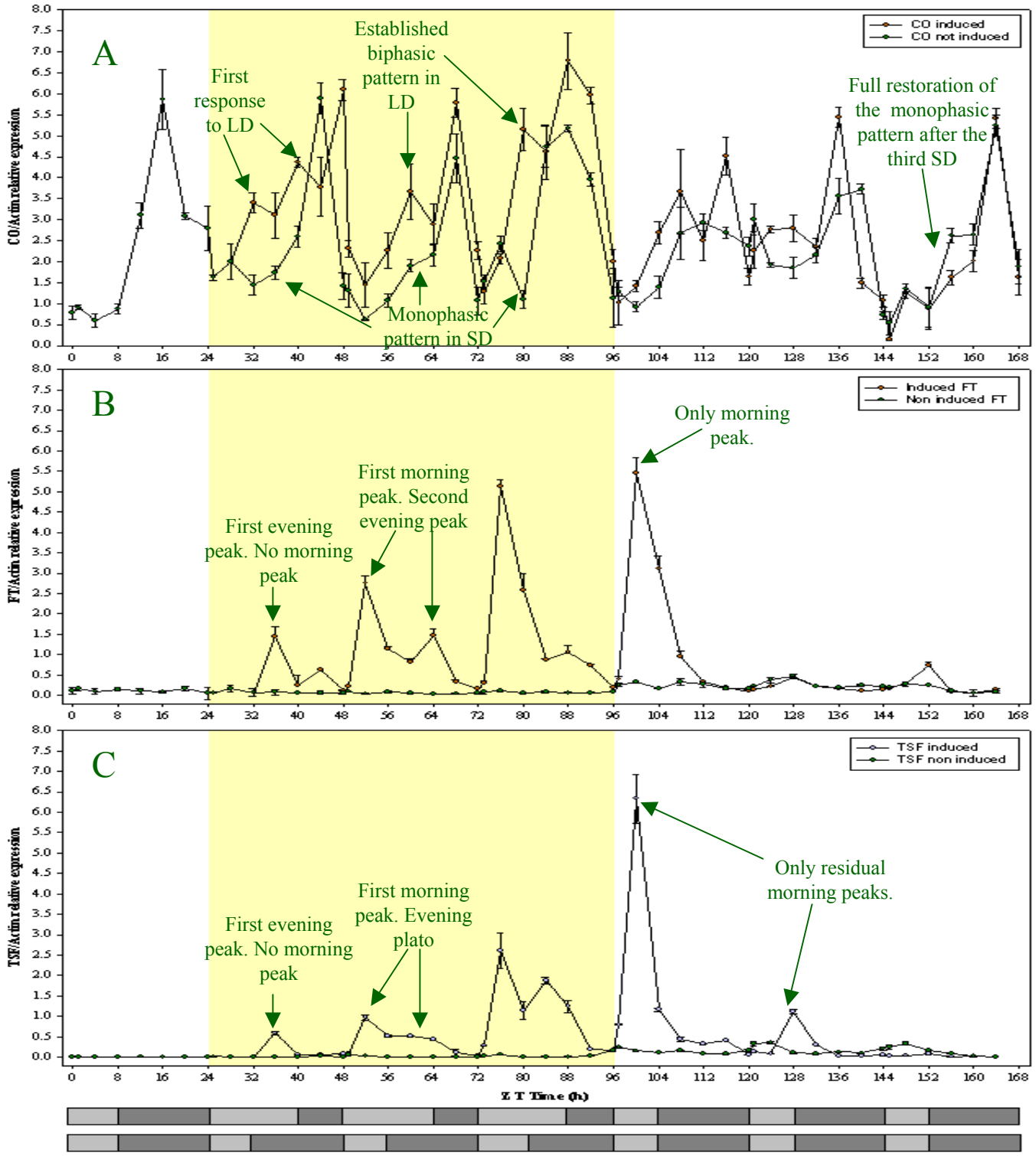


Figure 49 : Transcriptional profiling of flowering time genes in Ler background during a transient induction to LD 16 h. A) *CO* expression during a transient exposure to three long days. B) *FT* expression during a transient exposure to long days C) *TSF* expression during a transient exposure to long days. Arrows point critical changes compared to the SD control. Note that *CO* expression detects the changing day length from the first day, resulting in an evening peak in *FT* and *TSF*. However no morning peak was observed in the first day, in contrast to the second and third LDs and the first SD. This observation suggesting an absolute requirement of one previous LD for the morning specific peak of both *FT* and *TSF*. Therefore the external coincidence model is not violated. The yellow window highlights the critical three long days of the induction. Blue curve : induced plants, green curve : SD control. Light grey bars : day, dark grey bars : night.

winter. Climatological clines in day length kinetics are more than obvious and generally in the North photoperiod is altered much faster during the season compared to the south, leading to much sharper profiles (Fig 50, true data in collaboration with Prof. Annie Schmitt).

It was interesting therefore to ask whether such observations had an adaptive significance, leading to the formation of natural variation. In other terms, was it possible that accessions had differential properties of responding faster to the changing day length, in order to adapt to the different kinetics between the geographical sites? In order to test this hypothesis the same experimental setup, with the expression profiling of flowering time genes in Ler during a transient exposure to inductive long days, was selected and expanded. In the new setup, Ler plants together with Cen-0 and Bs-1 were initially grown for 2 weeks under non-inductive SDs and then were transiently challenged for a certain number of LD days of 16 and 14 h. After this transient induction the plants experienced again SD and where scored for flowering time. Harvesting of samples for RT-PCR was performed in selected days and this allowed the monitoring of the expression of *FT* mRNA in each accession and day length separately.

Natural molecular variation was identified both in the kinetics of the response between the accessions and within the same day length, but also between the day lengths and within the same accession (Fig 51). It was shown that in Ler three LD of 16 h are enough in order to induce flowering and commit the plants with a floral fate, even though FT mRNA expression is reduced rapidly after the return back to SD (Corbesier et al., 2007). It was interesting to ask first whether this rapid dampening of FT would occur also when the plant would experience longer exposure to inductive LDs. In other terms it was possible that three long days would be enough to commit the plant to flowering but five or seven would activate the postulated feedback mechanism, maintaining *FT* mRNA levels high after the shift back to SD. According to the obtained results even after seven long days, *FT* expression dampened quickly after the return to SDs. Seven long days in two weeks old plants are the saturation length of induction in Ler because at this point the plants bolt macroscopically. Interestingly LD of 16 and 14 h were in-disguisable in that accession, concerning the response of FT and the subsequent floral commitment (Fig 51A). These results confirm independently the results of the steady-state expression analysis and also the flowering time results, and prove that Ler responds equally in these two similar long days.

By contrast to Ler, Cen-0 plants required four times the exposure of Ler in long days of 16 h in order to up-regulate *FT* (3 LD of 16 h in Ler vs 12 in Cen-0, Fig 51B). Just like the steady-state conditions, the Cen-0 plants eventually bolted under LD 16 h with more leaves compared to Ler and of course like in Ler *FT* mRNA levels dampened rapidly after the return

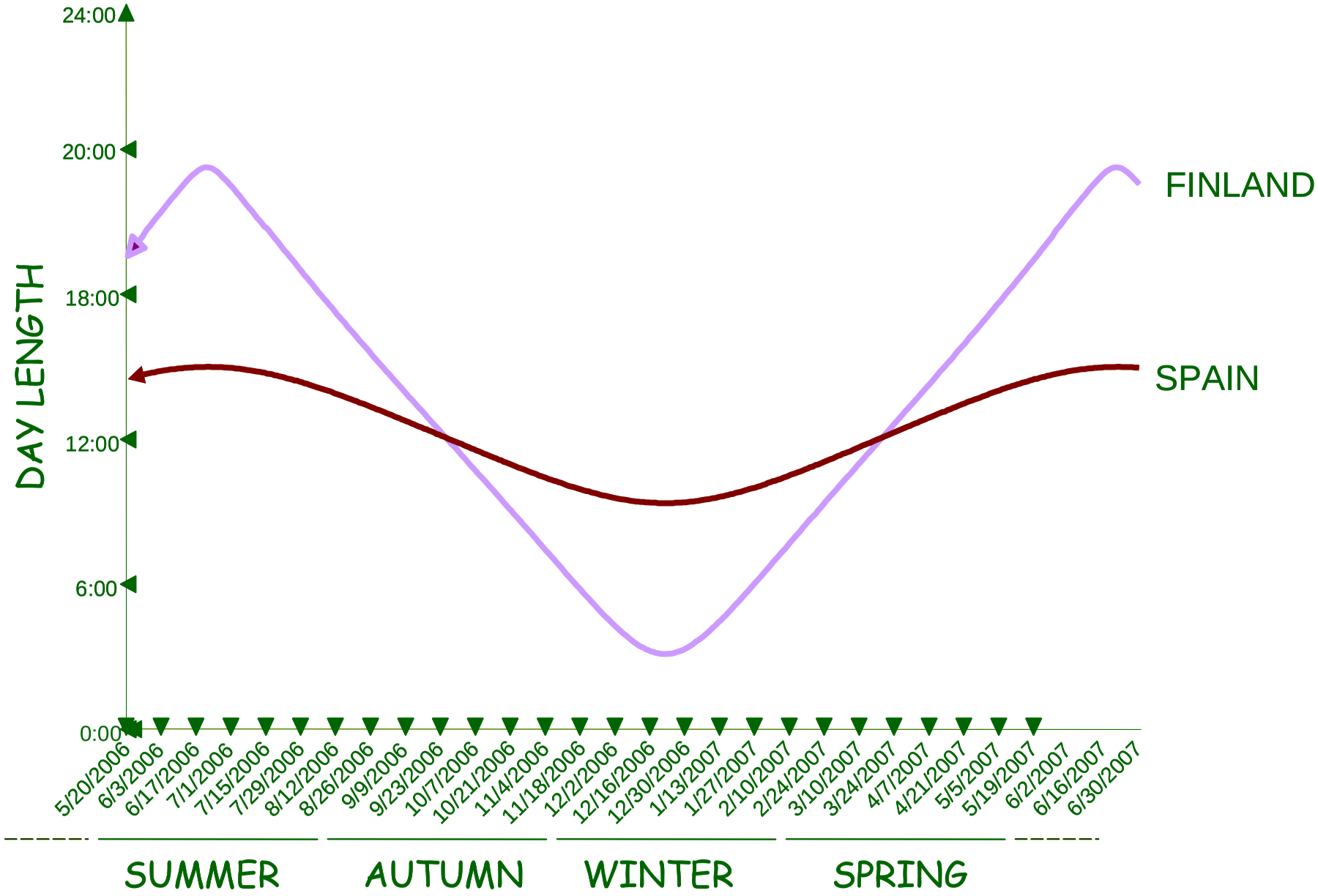


Figure 50 : Annual seasonal variation in photoperiod. Northern sites experience longer photoperiods which change faster compared to Southern sites, resulting into much sharper profiles.

to SDs. Bs-1 plants were even less responsive to the inductive conditions compared to Cen-0 and floral commitment in this accession requires at least 16 long days of 16 h (Fig 51C). Collectively these data suggest that natural variation indeed exists in the kinetics of the response to the same day length. Ler is very efficient not only in detecting 14 and 16 h as the same, but also in up-regulating *FT* very quickly, leading to floral commitment. Cen-0 requires approximately 4 times more exposure in LD of 16 h to ensure floral commitment while Bs-1 has the slowest kinetics of all, requiring more than 5 times (3 LDs of 16 h in Ler vs 16 in Bs-1) exposure to inductive LD of 16 h compared to Ler.

However the above results are only part of the story. As mentioned Ler responded equally in both LD of 16 and 14 h. This was not the case for Cen-0 and Bs-1. Both accessions responded much slower to LD of 16 h compared to Ler, but in both accessions LD of 14 h were not sufficient in up-regulating *FT* mRNA in the same levels as LD of 16 h (Fig 51 B and C). Both accessions failed to be committed to flowering even though induction took place with the same number of long days and in the same age as in the case of LD 16 h.

These results confirm once again the steady-state expression profiling and the physiological analysis, according to which Bs-1 and Cen-0 respond much weaker to LD of 14 h compared to LD of 16 h. In conclusion, natural molecular variation exists in both in the kinetics of the response to inductive conditions between the accessions, but also in the response between similar long days in the same accession. They also suggest that the observed natural molecular variation spatially is located in the vasculature since this is the site of expression of *FT*. Last but not least the mRNA profile of *FT* fits accurately to the flowering times of the induced plants, and therefore there is no need for speculation of post-transcriptional modifications at the level of FT that would complicate the phenotype. Nevertheless the production of FT in the vasculature is only the beginning in the process of floral commitment, a process that is completed in the SAM. Therefore the following question was raised : what can we learn by testing the consequences in the SAM of the observed natural variation in FT in the vasculature?

3.2.6 Natural molecular variation in the kinetics of the response to day length upon transient exposure – Molecular consequences in the SAM

We have shown that the accumulated FT protein in the vasculature of the leaves migrates in the apical meristem (SAM) and there initiates a molecular cascade ensuring floral commitment (Corbesier et al. 2007, see also introduction). One of the earliest targets of the FT protein in SAM is SOC1. *SOC1*, a MADs box transcription factor, is the earliest known

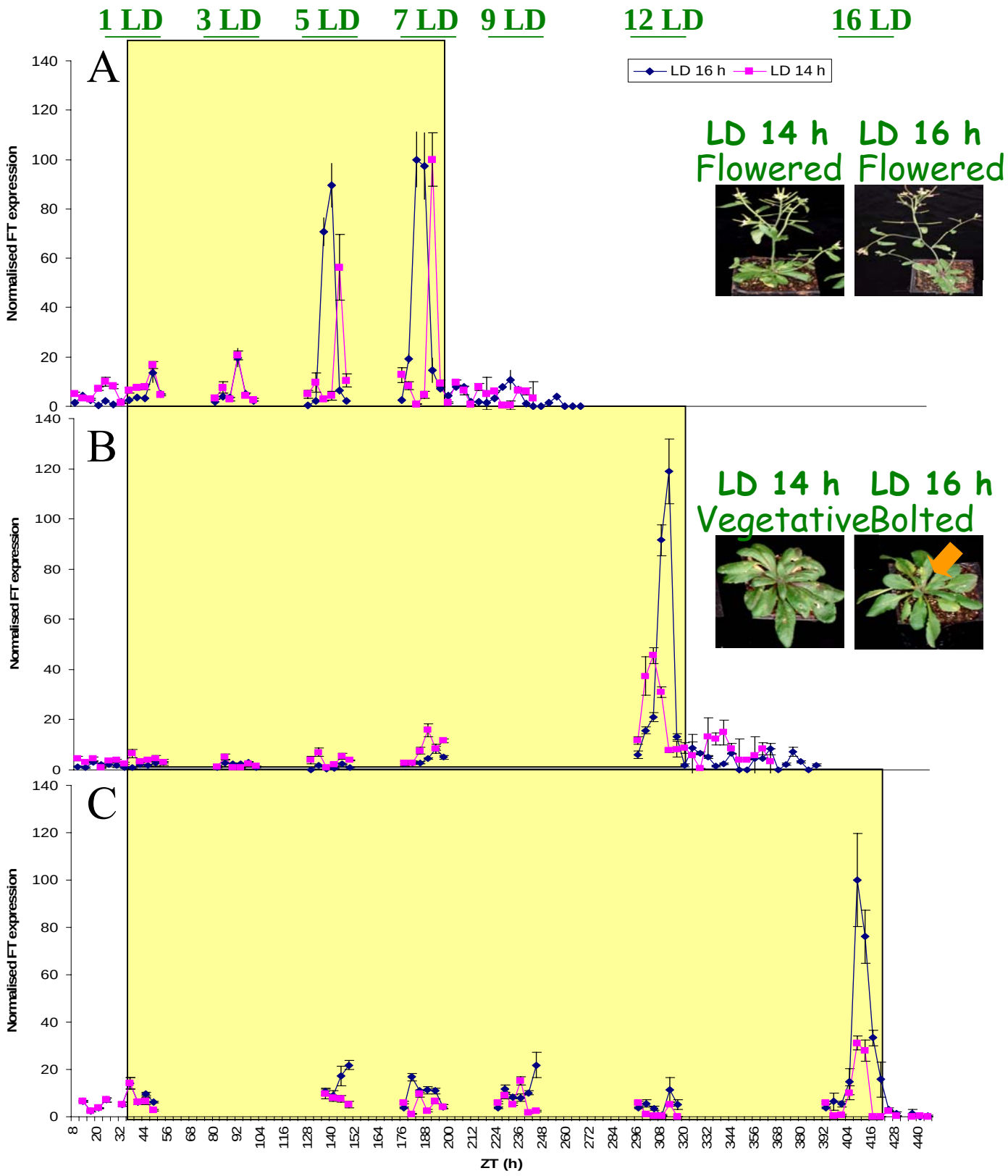


Figure 51 : Transcriptional profiling of FT in three accession backgrounds during a transient induction to LD 16 h and LD 14 h. A) Ler B) Cen-0 C) Bs-1. Blue curve : LD 16 h, red curve : LD 14 h. Ler responds equally strong to both LD 16 and 14 h and requires only three long days for floral commitment. Cen-0 and Bs-1 respond stronger to LD of 16 h compared to LD 14 h with Cen-0 requiring 12 and Bs-1 16 long days of exposure. After the shift back to SDs FT mRNA is decreases rapidly in all accessions. The flowering times of Ler and Cen-0 after the induction for 7 and 12 days respectively in the two daylengths is shown. Bs-1 flowering responses were similar to Cen-0 (data not shown)

marker of floral induction in the SAM and it was shown that the FT /FD protein complex upregulates *SOC1* mRNA in the SAM upon floral induction. Therefore *in situ* hybridisations, aiming to describe the expression of *SOC1* in the SAM according to the same experimental setup as the one described previously for *FT* in the vasculature, were performed.

Like with *FT*, *SOC1* mRNA expression in Ler was up regulated rapidly in the SAM upon floral induction (Fig 52). A single long day was enough to detect expression of the gene, in contrast to the SD negative controls in which expression was below the detection level. This expression increased further 3 long days after the shift and floral primordial were visible in the flanks of the SAM at this stage. With 5 long days the plants started forming auxilary shoots. *SOC1* mRNA expression in the SAM of Ler plants was up regulated with exactly the same kinetics under LD of 16 and 14 h, reflecting accurately *FT* mRNA profile in the vasculature.

By contrast in the SAM of Cen-0 after three LD of 16 h, *SOC1* mRNA is nearly undetected while under LD of 14 h the expression is below the level of detection like in the short day negative control (Fig 53). Interestingly 9 long days of 16 h after the shift, *SOC1* mRNA is much stronger expressed and the expression increases further 12 long days of 16 h after the shift (Fig 53). Remember that this is the condition that *FT* mRNA is strongly up regulated in the vasculature and the plant is committed to flowering (Fig 51B). Nevertheless under LD of 14 h the expression is much lower compared to 16 h, yet the meristem is much more developed compared to the SD negative control. This proves that the failure of floral commitment is not due to a failure of the florigen to arrive in the meristem under LD 14 h. In that case the meristem would not have been developed. A more likely scenario is that in this accession, the photoperiod pathway is not able to accumulate FT protein in levels above a critical threshold that ensures floral commitment.

One can think that although the kinetics of *SOC1* up-regulation are much slower under 14 h compared to 16 h, eventually *SOC1* will reach the sufficient levels in order to allow commitment to flowering. However once the plants experience SDs again after the shift, *SOC1* expression is decreased very rapidly to levels comparable to the ones prior to the induction (Fig 54). This exciting result shows the very tight dependence of *SOC1* expression in the SAM from the arrived FT protein from the vasculature. In Bs-1, *SOC1* mRNA is detected for the first time 13 long days of 16 h after the shift while at the same time *SOC1* mRNA under LD of 14 h has much lower levels and the meristem is still undeveloped compared to the LD 16 h (Fig 55). Additional *in situs* need to be performed in order to check the expression 16 LDs after the shift together with the decrease after the return to SDs.

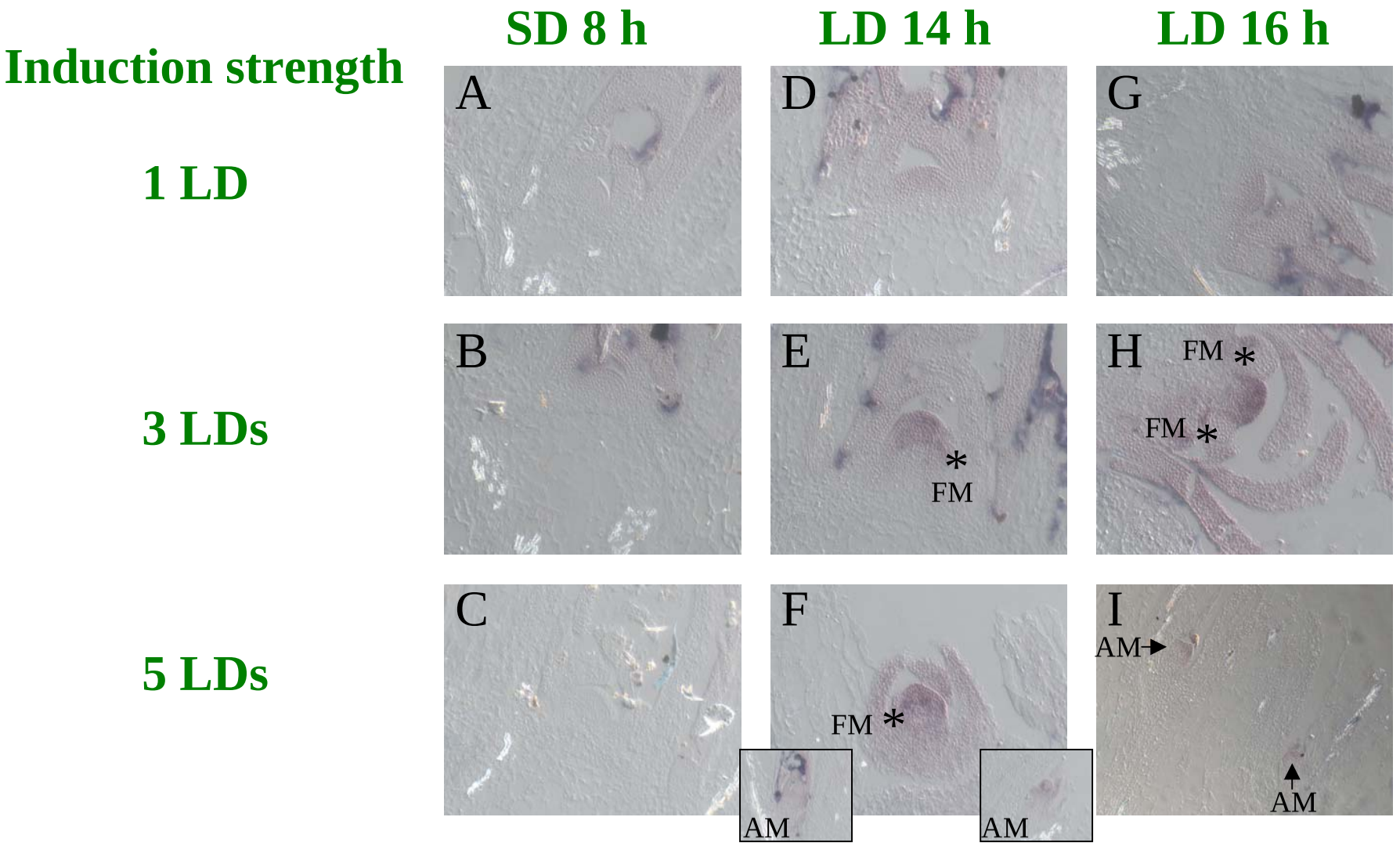


Figure 52 : *In situ* hybridizations detecting *SOC1* mRNA in the SAM of the Ler accession upon transient floral induction. A), B), C) SD 8 h D), E), F) LD 14 h, G) H) I) LD 16 h. A) D) and G) 1 LD of induction. B) E) H) 3 LD of induction C), F) and I) 5 LDs of induction. Both under LD of 16 and 14 h, equal levels of *SOC1* mRNA are detected in the SAM. This is in agreement with the *FT* mRNA levels in the leaves, proving molecularly that Ler responds equally strong in both day lengths. Three long days are enough to commit this accession to flowering since at this condition plants are already form floral primordia (FM) while 5 LDs after the shift, axillary primordia (AM) expressing *SOC1* are visible (inset in F, arrows in I). Seven days after the shift Ler plants bolt macroscopically (data not shown). Asterist indicates floral primordia, arrows axillary primordia.

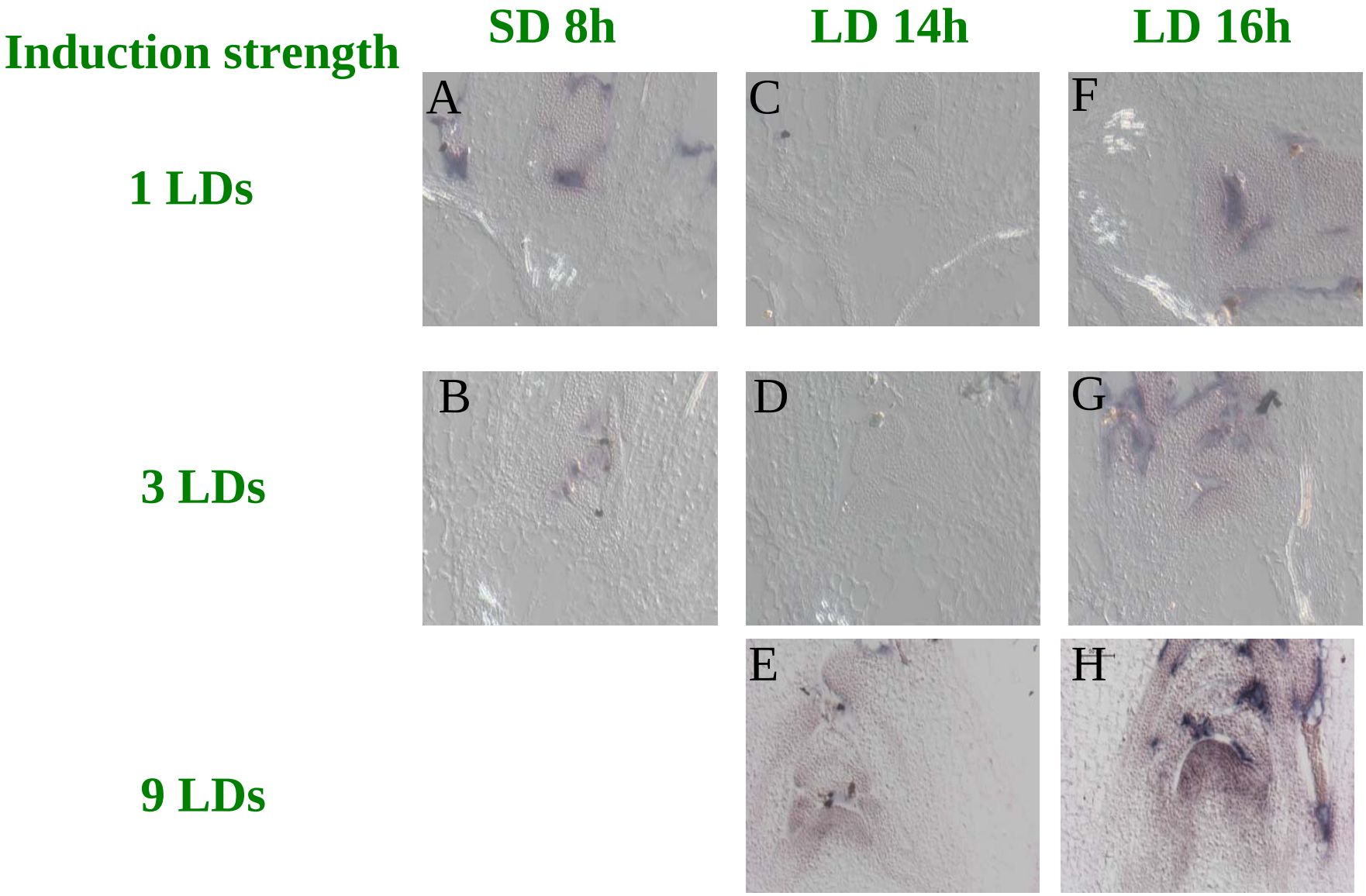


Figure 53 : *In situ* hybridizations detecting *SOC1* mRNA in the SAM of the Cen-0 accession upon transient floral induction. A), B) SD 8 h C), D), E) LD 14 h, F) G) H) LD 16 h. A) C) and F) 1 LD of induction. B) D) G) 3 LDs of induction E) and H) 9 LDs of induction. Under LD 14 h *SOC1* was undetected in the SAM until 9 LDs were weak signal was obtained. Under LD 16 h *SOC1* mRNA was weakly detected after 3 LDs while the expression was very strong after 9 LDs and the meristem was much more developed compared to LD 14 h.

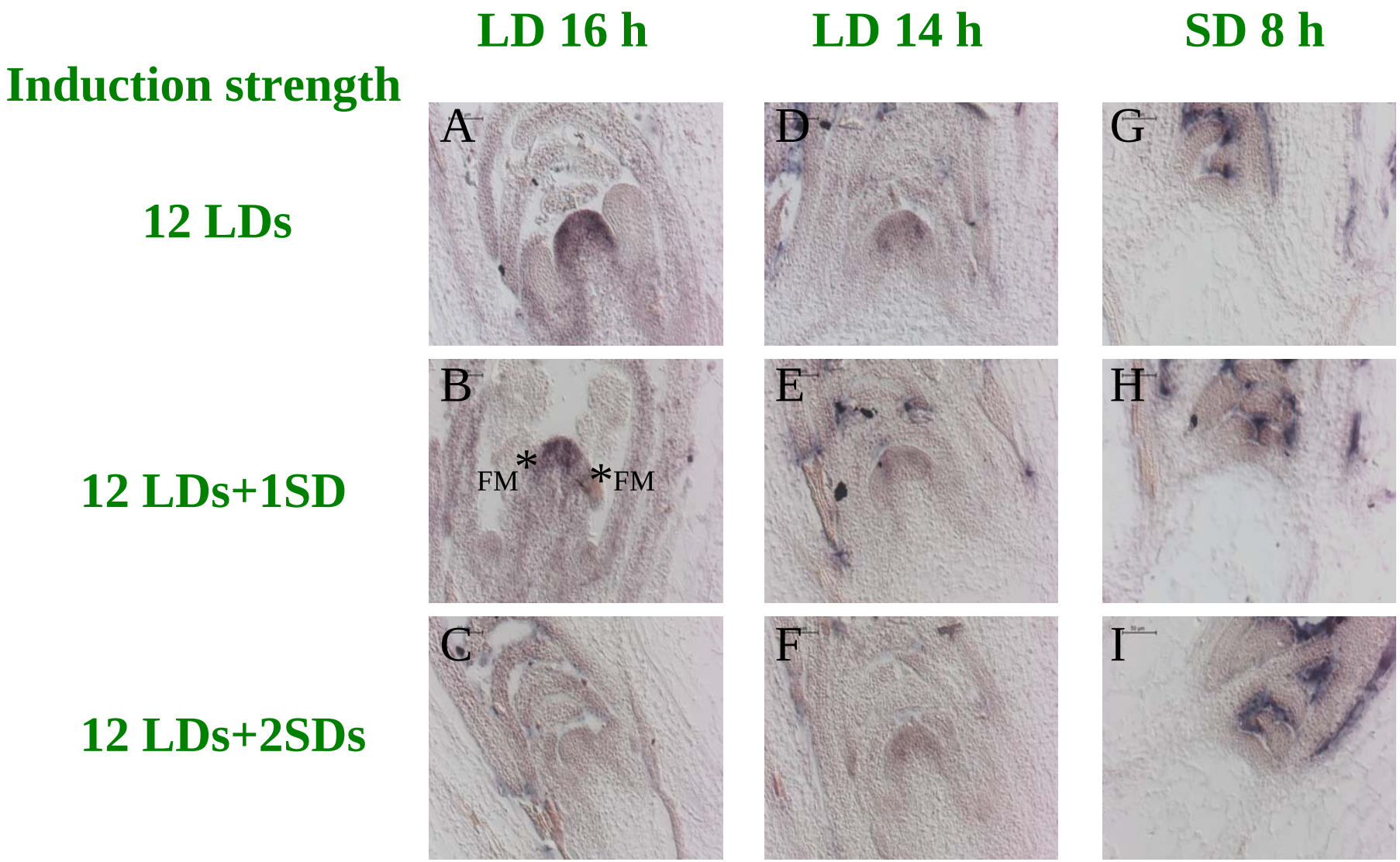


Figure 54 : *In situ* hybridizations detecting *SOC1* mRNA in the SAM of the Cen-0 accession upon transient floral induction. A), B), C) LD 16 h D), E), F) LD 14 h, G) H) I) SD 8 h. A) D) and G) 12 LDs of induction. B) E) H) 1 SD after 12 LDs of induction C), F) and I) 2 SDs after 12 LDs of induction. Under LD 14 h lower levels of *SOC1* mRNA are detected in the SAM compared to LD 16 h. However the meristem is more developed in comparison to the SD control. Interestingly after the shift back to SDs, *SOC1* expression fades away rapidly in a similar to *FT* mRNA in the leaves manner. Collectively these data indicate that *SOC1* mRNA upregulation is transient and under LD 16 h accumulation of *SOC1* mRNA to higher levels triggers floral commitment while under LD 14 h the accumulation is not sufficient retaining a late flowering phenotype. Asterist indicates floral primordia (FM).

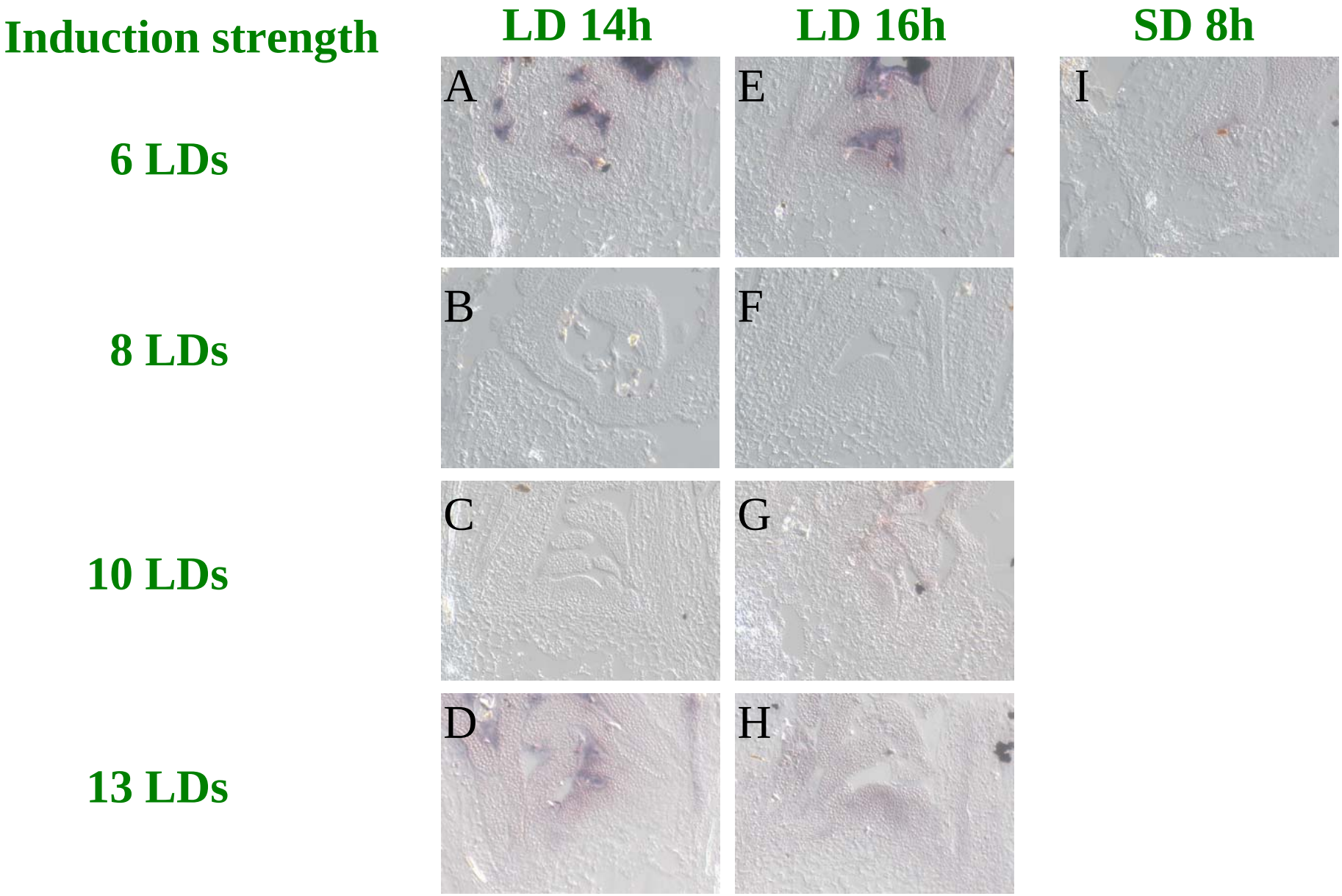


Figure 55 : *In situ* hybridizations detecting *SOC1* mRNA in the SAM of the Bs-1 accession upon transient floral induction. A) B) C) D) LD 14 h, E) F) G) H) LD 16 h, I) SD control before the induction. A) and E) 6 LDs of induction, B) and F) 8 LDs of induction, C) and G) 10 LDs of induction, D) and H) 13 LDs of induction. Under LD 14 h *SOC1* was undetected in the SAM until 13 LDs were weak signal was obtained. Under LD 16 h *SOC1* mRNA was weakly detected after 10 LDs while the expression was stronger after 13 LDs.

3.2.7 Discussion

Steady-state conditions

In the present study natural variation in the phase of CO expression was observed between different accessions (Fig 46 and 47). In poplar, the existence of a CO/FT regulatory network, similar to *Arabidopsis*, regulates the transition to flowering (Bohlenius et al., 2006). Natural variation in the expression of *PtCO2* in three different day lengths among four Poplar accessions originating from different latitudes resulted in differences in the up-regulation of the *PtFT1* gene. The peak in *PtCO2* occurred latter in the day in accessions originating from Northern latitudes, compared to southern ones. Therefore a requirement for a critical photoperiod exists, with northern accessions needing longer photoperiods compared to southern ones in order to overlap *PtCO2* expression with light, which will lead to up-regulated *PtFT2* expression and initiation of flowering. It is not clear whether a shift in the expression of *PtCO2* is observed among the different photoperiods for each accession and whether some of the Poplar accessions which come from very north latitudes, show a vernalization response. Nevertheless, taken together the above data confirm that natural variation in the phase of *CO* mRNA expression among different day lengths, not only is present but it is important for the regulation of photoperiodic responses in flowering time.

Transient induction of FT in Ler

The expression analysis for Ler under transient induction by photoperiod suggest that *FT* and *TSF* mRNA levels increase transiently in the vasculature (Fig 49). This regulation occurs via a CO-dependent manner. The existence of a loop that maintains *FT* expression, seems to be largely questioned. All data confirmed the first hypothesis, according to which, *FT* expression is regulated transiently by photoperiod, excluding the existence of an auto-regulatory loop. Published data initially supported the opposite theory (Huang et al., 2005). The authors of this publication used an alternative assay. A fusion between *FT* and a heat shock promoter (*HSp::FT*) was used in order to transform Ler plants. As a control, a GUS fusion (*HSp::GUS*) was used. Heat was then applied transiently into the leaves of 16-days old plants grown under SDs of 9 h. Although the expression of both *HSp::FT* and *HSp::GUS* was increased transiently, the expression of the endogenous *FT* in the leaves showed a gradual increase, which continued for three days after the initial heat shock. Based on these results, they proposed the existence of an auto regulatory loop that maintains the expression of *FT*.

This was in contrast to our results. In our experimental conditions, we used ten days old Columbia plants and two weeks old Ler plants. It is known that older plants are more sensitive

to inductive conditions. In addition Columbia generally is less responsive to long days and flowers later than Landsberg. Two more accessions were also tested. All four accessions (Col-0, Ler, Cen-0 and Bs-1), induction strengths (3, 5, 7, 12 and 16 long days) and day lengths (LD of 14 and 16 h) provided evidence against the existence of such a mechanism (Fig 48, 49 and 51). Additional evidence came from the *in situ* results (Fig 54). Eventually the publication that suggested such a mechanism was retracted recently (Huang et al., 2007).

Natural variation in the kinetics of the floral response – Evidence for a molecular thershold

Robert King back in 1973 was providing a quantitative basis regarding the translocation of the floral stimulus in *Perilla* that affects floral induction, confirming previous studies of Chailakhyan et al., in 1957. This floral stimulus moves together with photoassimilates through the phloem, but shares different kinetics compared to them. The authors also concluded that in the phloem sap there are traces of the floral stimulus, which in the absence of leaves are unable to induce flowering. This inability together with the quantitative nature of the response in these early physiological observations suggested the existence of a threshold that needs to be reached in order for flowering to be initiated.

In plants such as *Impatiens balsamina* floral reversion has been observed. In *Impatiens* and in contrast to *Arabidopsis*, axillary meristem primordia are formed prior to the floral induction and therefore their fate during induction can be analysed (Pouteau et al., 1998). That is a great advantage since the fate of these meristems will be different depending on the amount of inductive signal that the plant perceives. A clear correlation between the amount of signal and reversion of the floral meristem was observed, providing strong additional evidence about the existence of a threshold that needs to be reached in order to allow stable floral fate. A report of a continuous variation in the form of the inflorescence that reflects the amount and translocation rate of the produced floral stimulus from the leaves was clearly suggested. Last but not least the authors conclude that in many species, including *Arabidopsis*, a sufficient quantity of inductive signal is required to allow rapid initiation of flowering, a transition that otherwise will occur slowly and gradually as the plant ages.

In the present study, molecular evidence for the existence of this threshold both in the leaves and the meristem were provided. Interestingly natural variation shapes these responses since it affects the kinetics of the florigen production between accessions and day lengths. More specifically accessions vary in their ability to respond fast to the inductive conditions. Each accession has its own requirements, its own induction window and within that window it induces *FT* expression into levels that are sufficient for floral commitment. Accessions such

as Ler require only three LD of 16 h, while accessions such as Cen-0 and Bs-1 reach that threshold 4 and almost 6 times respectively slower in the same conditions (Fig 51). Even more interestingly natural variation shapes the response of the same accession within two very similar day lengths with Ler responding equally both under 14 and 16 h and Cen-0 and Bs-1 failing to respond equally fast between the two conditions (Fig 51). These observations indicate that in these accessions the photoperiod pathway is not activated to the same rate by the two conditions. This conclusion is further supported by preliminary experiments using the Bs-1 NILs, which indicate that the same component that causes late flowering under continuous LD of 14 h is responsible for the slower response under transient induction (Fig 41B). A two way ANOVA confirmed not only the identical behaviour of the NILs with the Bs-1 parent but also the statistically significant differences in flowering times between LD of 14 and 16 h (Fig 41C).

It was published that the FT protein reaches the SAM in order to initiate floral induction (Corbesier et al., 2007). In the SAM the FT protein forms a complex with FD that associates with the promoter of *AP1* in order to up-regulate its expression (Wigge et al., 2005; Abe et al., 2005). This complex regulates also the expression of *SOC1* (Searle et al., 2006). Collectively these studies predict a strong dependence of the SAM floral cascades on the leaves. This study provides solid proofs for this dependence in three different accessions with very diverse responses under two different day lengths. Exactly like *FT* mRNA up-regulation in the leaves, SOC1 expression in the SAM of Ler plants is induced equally strong within three long days of either 14 or 16 h (Fig 52). By contrast *SOC1* up-regulation is 4 and more than 6 times slower in Cen-0 and Bs-1 respectively in LD of 16 h while both accessions fail to induce *SOC1* mRNA to high levels under LD of 14 h (Fig 53, 54 and 55).

One of the most important observations is that *SOC1* mRNA expression in the SAM follows accurately the decay of *FT* mRNA expression in the vasculature once the plant experiences again non-inductive conditions after the induction (Fig 54). As a result the slowly accumulated *SOC1* levels under 14 h fail to exceed the threshold of floral induction within a given time frame, and the plant remains vegetative like the SD negative controls. However under LD of 16 h *SOC1* is accumulated much faster, it exceeds the threshold allowing stable floral commitment even if it decays to non induced levels after the shift back to SD. This indicates that other genes downstream of *SOC1* are stably activated and in turn mediate the threshold effect. Since *SOC1* regulates *LFY* expression through AGL24 (Yu et al., 2002), both *LFY* and *AGL24* are good candidates. *AP1* is regulated by FT and acts synergistically with *LFY* (Ferrandiz et al., 2000), therefore it is another very interesting candidate.

LFY expression was directly suggested to mediate the threshold of floral induction (Blazquez et al., 1997). Ferrandiz and colleagues have shown that the *LFY* regulators *AP1*, *FUL* and *CAL* are important in order for the plant to reach this threshold, since in their absence *LFY* never exceeds its initial low levels and floral transition does not occur (Ferrandiz et al., 2000). Hempel and colleagues have shown that transient promoter activity of *LFY* and *FUL* does not ensure stable floral induction and the plants remain vegetative after the return to non-inductive conditions (Hempel et al., 1997). These are additional proofs regarding the existence of a threshold in *LFY* expression that is required for stable floral commitment. Parcy and colleagues suggested that *LFY* express its role in the maintenance of the floral fate by negatively regulating directly or indirectly shoot meristem genes such as *TFL1* and positively regulating floral primordia genes such as *AG* (Parcy et al., 2002).

Both *LFY* and *AP1* expression is inhibited by TFL1, a floral repressor. TFL1 expression is increased during floral induction (Simon et al., 1996), the mRNA is confined to a central region yet the protein is diffused throughout the shoot meristem (Conti et al., 2007). The relative levels of *TFL1*, like *LFY*, control meristem behaviour and therefore a *LFY*:*TFL1* ratio was proposed to be the molecular mechanism that sets the long sought threshold of floral induction (Ferrandiz et al., 2000). Mutations in *AP1/CAL/FUL* genes, lower this ratio, increasing the threshold and leading to late flowering and homeotic transformations of the produced flowers due to ectopic *TFL1* expression (Ferrandiz et al., 2000). In addition *LFY* overexpression or *TFL1* mutations, increase the ratio, lower the threshold and lead to early flowering and homeotic transformation of the shoot into a terminal flower due to ectopic *LFY* activity in the SAM. The existence of a *TFL1-LFY* feedback loop was proposed through which a mutual exclusion of the two mediators of shoot and floral fate respectively controls the production of floral primordial (Conti et al., 2007). Yet this regulation is predicted to be indirect since the intra-cellular localization of *LFY* (nucleus) and *TFL1* (cytoplasm) differs.

3.2.8 Conclusions

The above data provide molecular evidence about the existence of a threshold that regulates floral induction and prove that natural molecular variation exists the kinetics to reach it between accessions and day lengths. This natural variation is placed in the *FT* expression in the vasculature and shapes variability in *SOC1* expression in SAM, perhaps the earliest known target of floral induction in this tissue. Such molecular variation is not only between accessions in the same day length but also between day lengths in the same accession, with Ler being hyper-responsive both under LD of 16 and 14 h and Cen-0 and Bs-1 showing much

slower kinetics of *SOC1* mRNA up-regulation under LD of 16 h compared to Ler and much weaker expression at the same developmental stage under LD of 14 h compared to LD of 16 h. The observed pattern of *SOC1* mRNA in the meristem fits tightly to the *FT* mRNA profile in the vasculature, proving once more the tight dependence of the SAM on the florigen during the floral induction. It is important to distinguish between two different phases of floral transition : the induction and the maintenance phase. Within the induction phase, FT accumulation from the leaves, its association with FD in the SAM and the subsequent up-regulation of *SOC1* and *AP1* are included. Also the initial up-regulation of *LFY* by *AGL24* and *SOC1* as well as *TFL1* mRNA up-regulation can be considered part of the induction. However for the maintenance key role has the subsequent, stable and stronger expression of *LFY* by the combined action of *AP1/CAL/FUL* and the determined ratio of *LFY/TFL1* expression in the SAM that will set the threshold and will allow the formation of floral primordial in the flanks of SAM. A proposed model for floral induction is shown in Fig 56.

Several questions still remain without answer. How the Cen-0 and Bs-1 plants are eventually committed to flowering under LD of 16 h even though *SOC1* up-regulation is transient? Downstream factors must have a different profile ensuring maintenance of the floral commitment process once the critical threshold is achieved. Is it indeed only *LFY* and *TFL1* the genes that create this threshold and what are the regulatory steps between them and FT? Under LD of 14 h is *TFL1* expression affected or is it only *LFY* expression lower, reducing in turn the *LFY/TFL1* ratio and preventing flowering? Finally what is the genetic infrastructure behind this molecular variability in the kinetics of the response and can we connect it to the already identified QTL for flowering time responses under steady-state conditions in these accessions? In Bs-1 we can. Transient induction of selected NILs from the Bs-1 x Ler population revealed that the introgressed QTL region at the top of chromosome 5 was enough to ensure the same behaviour of these genotypes with parental homozygous Bs-1 plants during transient induction both under LD of 16 and 14 h. *In situs* using the constructed NILs, an experiment already under way, will help describing with accuracy this connection.

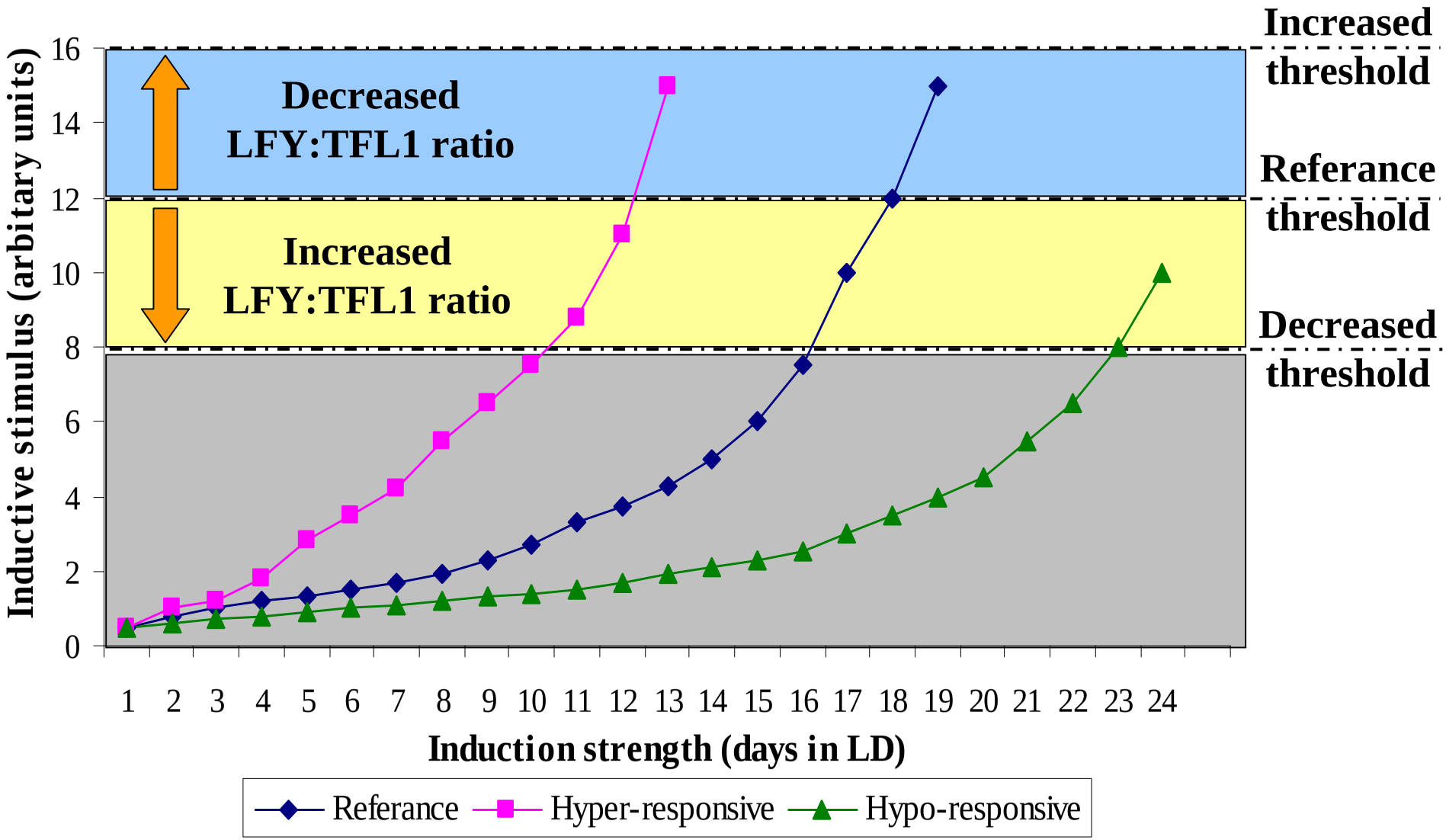


Figure 56 : Proposed model for floral induction. Hyper-responsive conditions reach the threshold for floral initiation much faster compared to reference conditions. The opposite occurs in hypo-responsive conditions. Such conditions in general can be environmental (LDs vs SDs) or genetic (mutant vs WT or accession vs laboratory strain). The floral threshold in turn can be modulated. Mutations that lower LHY expression such as the AP1/CAL/FUL genes, increase the threshold while mutations that up-regulate LFY, decrease the threshold allowing rapid floral commitment. The curves indicate the accumulation of inductive stimulus, such as FT, with inductive strength represented as days under inductive conditions.

3.3 NATURAL VARIATION IN *GIGANTEA* CIRCADIAN EXPRESSION

3.3.1 IDENTIFYING NATURAL VARIATION IN GIGANTEA EXPRESSION

3.3.1.1 Introduction

Several reasons justified our selection to also search for natural-genetic variation in the expression of the *GIGANTEA* gene. Being the link between the circadian clock and CONSTANS (Suarez-Lopez et al., 2001), GIGANTEA can help evaluating the contribution of the circadian clock to natural-genetic variation in photoperiod discrimination. Apart from being a strict clock output, GIGANTEA also feeds back to the central oscillator. The expression of the gene is robust in a wide range of photoperiods including constant darkness or light, in contrast to CONSTANS. Therefore high through-output systems such as top count imaging, can be used to screen a large numbers of plants. This together with the fact that the expression of the gene is regulated both by light and by the circadian clock increases the chances of identifying natural variation. Last but not least GIGANTEA have been recently connected with the ability of ambient temperature perception by the clock (Gould et al., 2006). Ambient temperature also influences flowering time. Therefore interesting accessions can be used in future studies to further elucidate the mechanisms of temperature perception by the clock.

In the attempt to describe natural variation in the clock, several circadian parameters of *GI* expression were selected (Fig 57). The primary interest is *GI* phase. Circadian phase refers to the time duration between two previously defined points or stages of the rhythm. In chronobiology, usually phase describes the duration between dawn and the peak of the expression. Phase in this project was critical in understanding how the gene is regulated in certain photoperiods. However in genes such as GI, expression is regulated by the circadian clock and by light independently. Therefore phase alone cannot separate these two parameters and another property of the rhythm, strictly circadian and without the influence of light must be studied as well. Since GI rhythm is robust under constant dark (DD) conditions, the free running period under DD was also monitored. Period in general is the time duration of a complete oscillation and free running period under DD means the time duration of a complete oscillation (one circadian cycle) under constant dark conditions. Other parameters that were tested were the amplitude of the rhytm and RAE, which is an estimate of the robustness of the

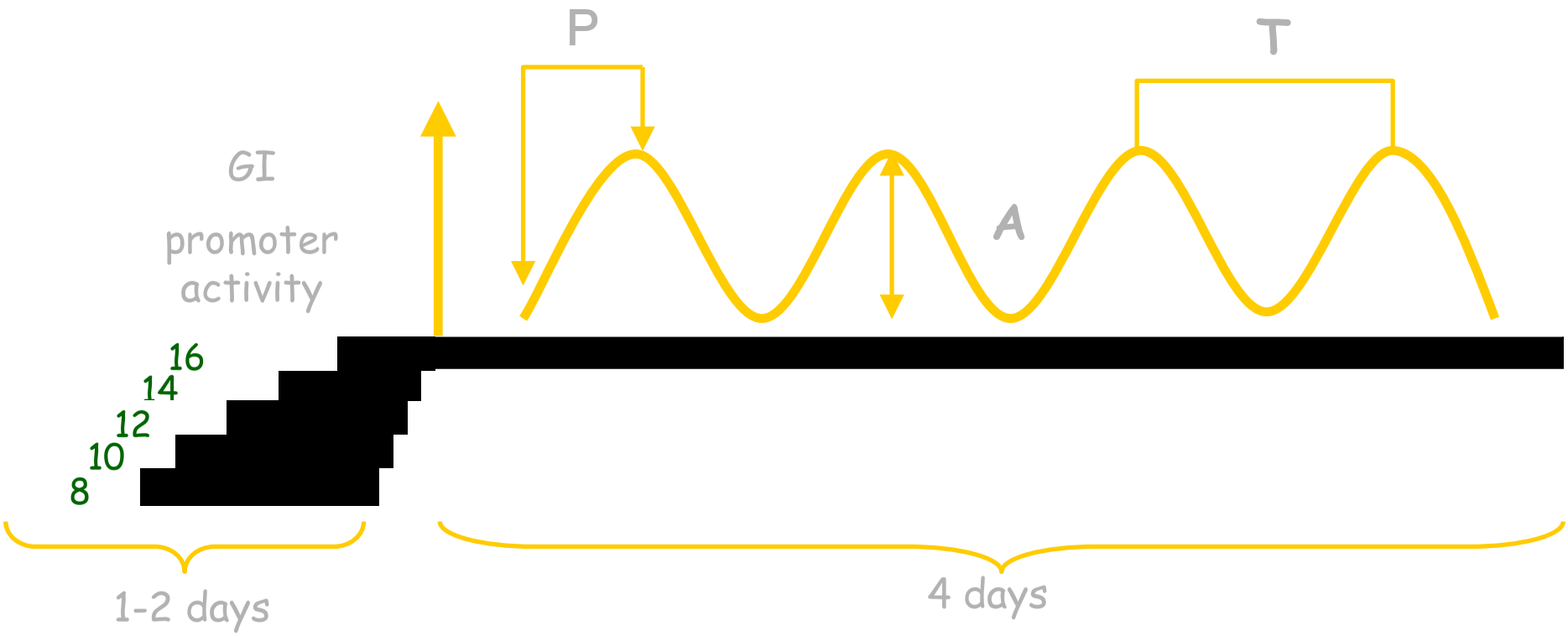


Figure 57 : Circadian parameters and imaging protocols used in this project. Phase (P) is the time distance between the peak in the expression and a predefined time point, such as dawn. Period (T) is the duration of time that is need under constant dark or light for the completion of a single circadian cycle. Amplitude (A) is the peak of the expression. In these experiments, plants were grown for 10 days in a specific daylength. On days 11 and 12 the expression of GI was monitored under this particular daylength conditions. At the end of these measurements, plants were released continuously into constant dark (DD) and the rhythm was monitored for 4days continuously

rhythm under constant conditions. Here focus mainly in the results from phase and free running period under constant dark will be made.

3.3.1.2 Natural variation in *GIGANTEA* phase

The expression of *GI::LUC* was monitored under five different day lengths in eighty accessions. A strong influence of day length, on the phase of *GI::LUC* expression was observed. Like with flowering time analysis, we could detect a general pattern in the day length responses. As the day became longer, the peak in *GI* expression was shifted later compared to shorter photoperiods (Fig 58). A strong response was observed among all accessions, between SD of 8 and 10 h. However, under LDs milder responses and higher variability were frequently the case.

In order to analyse the data and to detect natural-genetic variation in the circadian responses, a similar approach to the flowering time analysis was selected. Based on the hierarchical clustering results, two major clusters and three main types of response were identified (Fig 59A). In a large group of accessions, *GI::LUC* expression revealed a phase delay under all day lengths, accompanied with a strong response to the changing photoperiod (Fig 59C and F). Three representatives of this group are Dijon-G, S96 and Lip-0 (Fig 60 and 61). The differences of these accessions with Col under SDs are less than half an hour, however under LDs this difference reaches the level of 1,5 hours suggesting a more pronounced response to long photoperiods compared to Col-0.

In contrast, another group of accessions revealed a phase advance under all day lengths (Fig 59 D and G). Although under SDs of 8 and 10 h the values were similar to the average, *GI::LUC* expression shifted slightly later under long days, resulting into a weak photoperiodic response. Accessions like Dijon-M, Ei and C24 showed this type of response (Fig 62 and 63). Like Dijon-G, S96 and Lip-0, the differences of these accessions with Col under SDs are less than half an hour. Under long days however, the phase of *GI::LUC* expression demonstrates moderate changes. Therefore the difference with Col reaches the levels of 1.5 hours, with the latter being more responsive to day length. The differences between hyper- and hypo-responsive accessions were maximum one hour under short days and approximately three hours under long days (Fig 64B). These results for Dijon-G and Lip-0 were confirmed with at least four independent transformants (Fig 64C and D). In order to further verify the results the expression was also tested in two constitutive days of measurement under each photoperiod, confirming the robustness of the assay (Fig 64E)

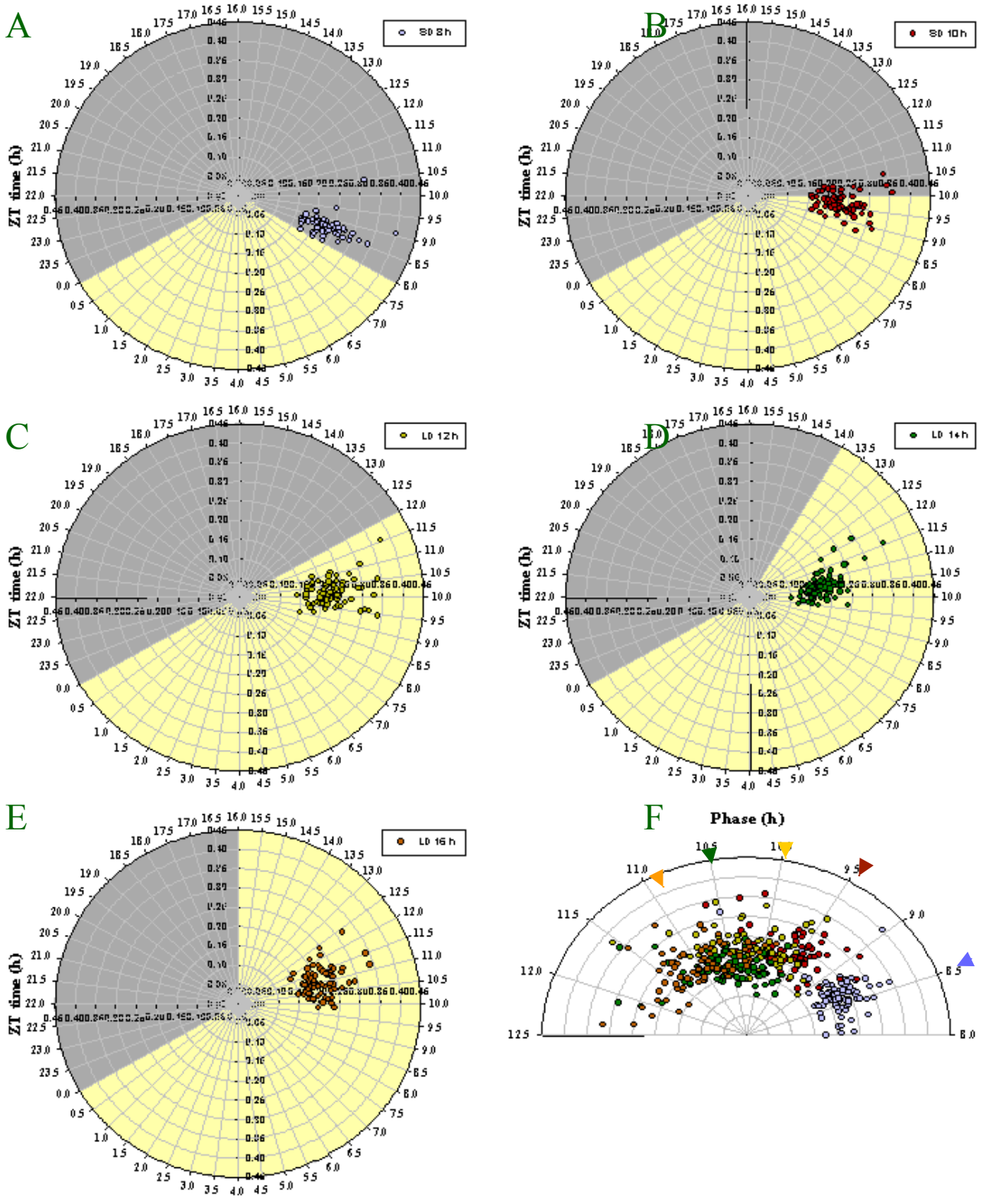


Figure 58 : Polarized scatter plots with the phase of GI expression vs RAE in five day lengths. A) SD 8 h B) SD 10 h C) LD 12 h D) LD 14 h E) LD 16 h F) Merged day length plot. The green arrows indicate the average of the peak times of all accessions in the corresponding daylength. Note the partial overlap of the phase values and the latter shifting with the lengthening of the day, reflecting the effect of the changing daylength on the expression of *GI*. Angular axis corresponds to the ZT transformed times of the first 24 h of the experiment. Radial axis corresponds to the RAE of each accession. RAE values approaching zero correspond to very robust rhythm. RAE was calculated based on 4 days in constant dark. Yellow : day, grey : night.

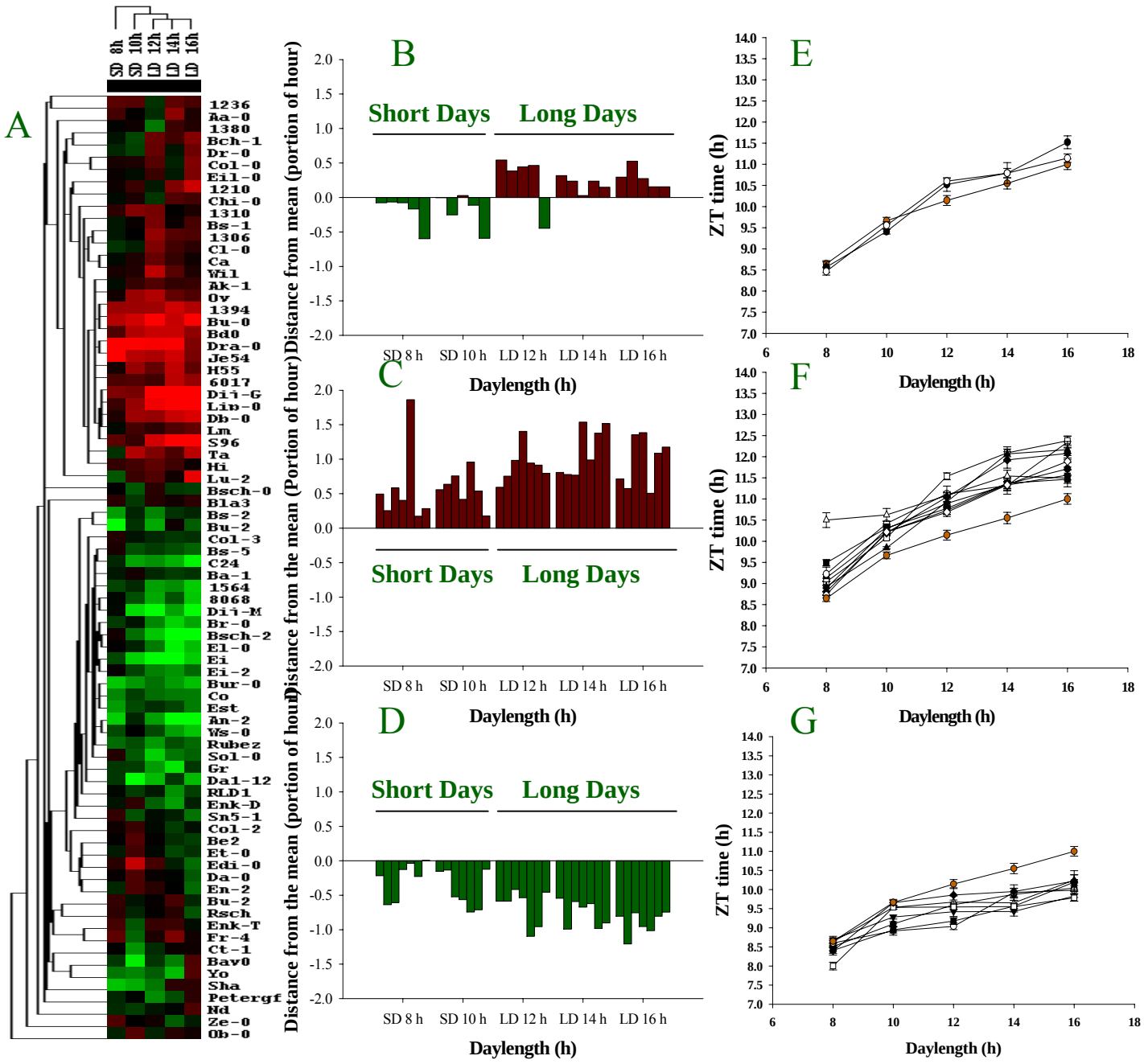


Figure 59 : The hierarchical clustering identifies natural variation in the response of *GIGANTEA* expression under different photoperiods. A) Following the same principles as for flowering time, the experimental values of *GI* peak time from 80 accessions and under five different day lengths, were clustered according to their general response pattern. Three major groups were identified. The first contains accessions in which the expression of GI is earlier than the wild type under short days and latter under long days. These differences result in a moderate type of response. The second cluster consists of accessions with a gradually increasing with day length, phase delay compared to the average. The corresponds to a strong type of response. Finally the third cluster consists of accessions with a phase advance especially under long days. This results into a reduced response under long days. B) C) D) Distance from the mean of representative accessions from the three different types of responses. E) F) G) Day length response of *GI* for representative accessions from the three types of responses. The average response is shown in orange.

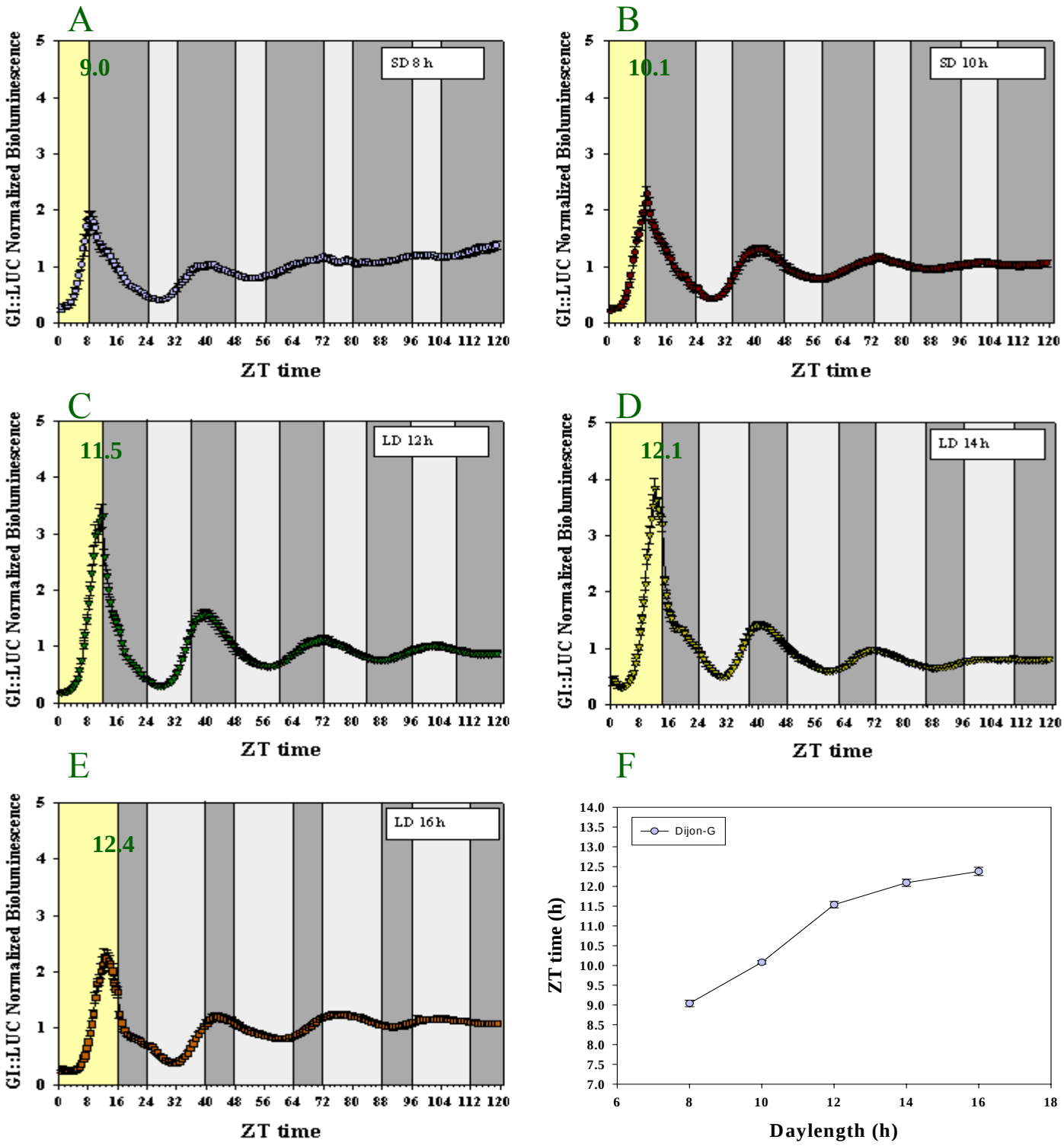


Figure 60 : Real time monitoring of GI::LUC bioluminescence in Dijon-G in different photoperiods and in constant dark. A) SD 8 h B) SD 10 h C) LD 12 h D) LD 14 h E) LD 16 h F) Daylength response curve. The numbers indicate the time of the peak. The data in A) B) C) D) and E) correspond to average values of 24 seedlings, after normalization to the average of the expression. The data in F) represent the average from four independent transformants.

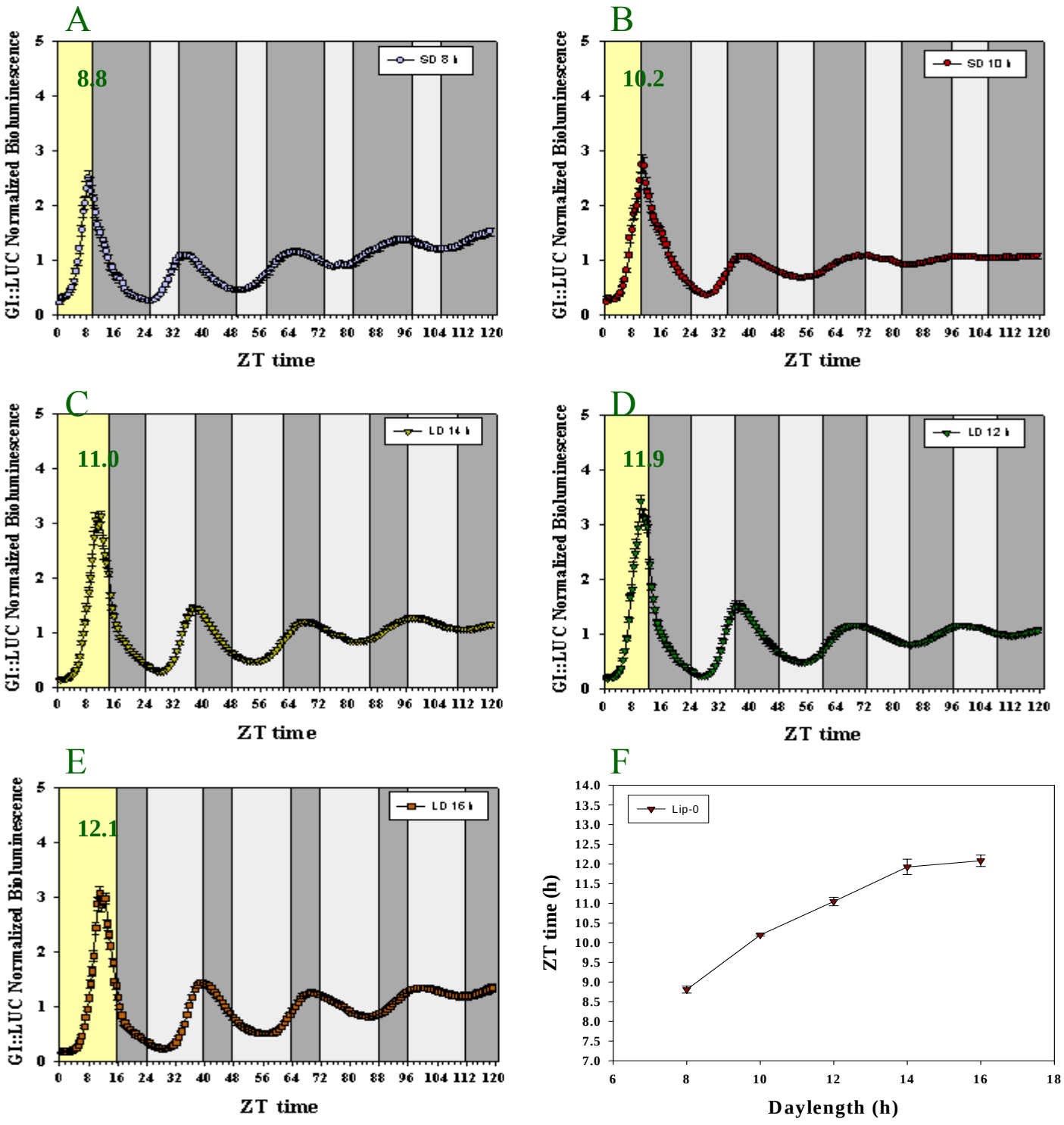


Figure 61 : Real time monitoring of GI::LUC bioluminescence in Lip-0 in different photoperiods and in constant dark. A) SD 8 h B) SD 10 h C) LD 14 h D) LD 12 h E) LD 16 h F) Daylength response curve. The numbers indicate the time of the peak. The data in A) B) C) D) and E) correspond to average values of 24 seedlings, after normalization to the average of the expression. The data in F) represent the average from three independent transformants.

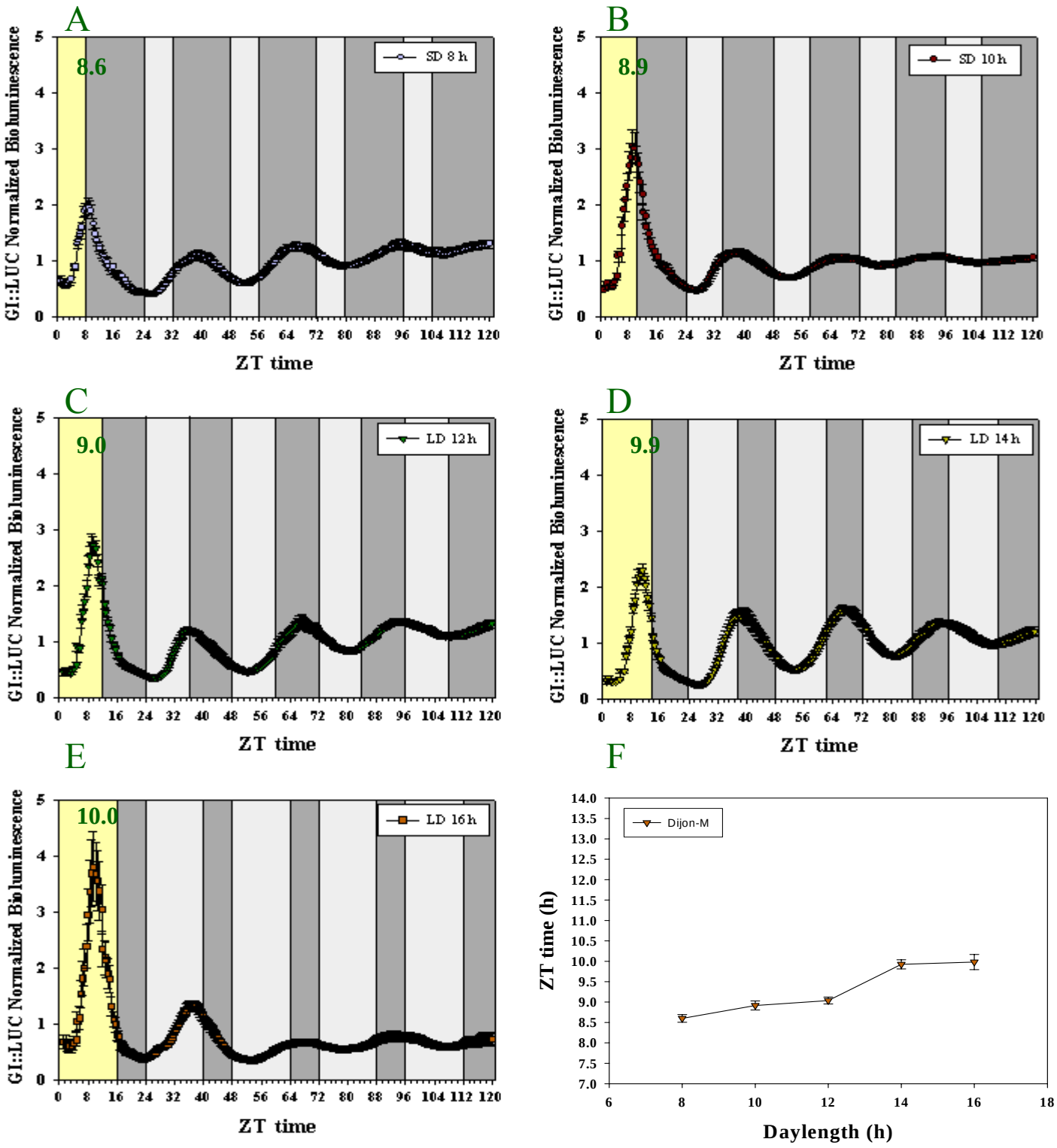


Figure 62 : Real time monitoring of GI::LUC bioluminescence in Dijon-M in different photoperiods and in constant dark. A) SD 8 h B) SD 10 h C) LD 12 h D) LD 14 h E) LD 16 h F) Daylength response curve. The numbers indicate the time of the peak. The data in A) B) C) D) and E) correspond to average values of 24 seedlings, after normalization to the average of the expression. The data in F) represent the average from one transformant.

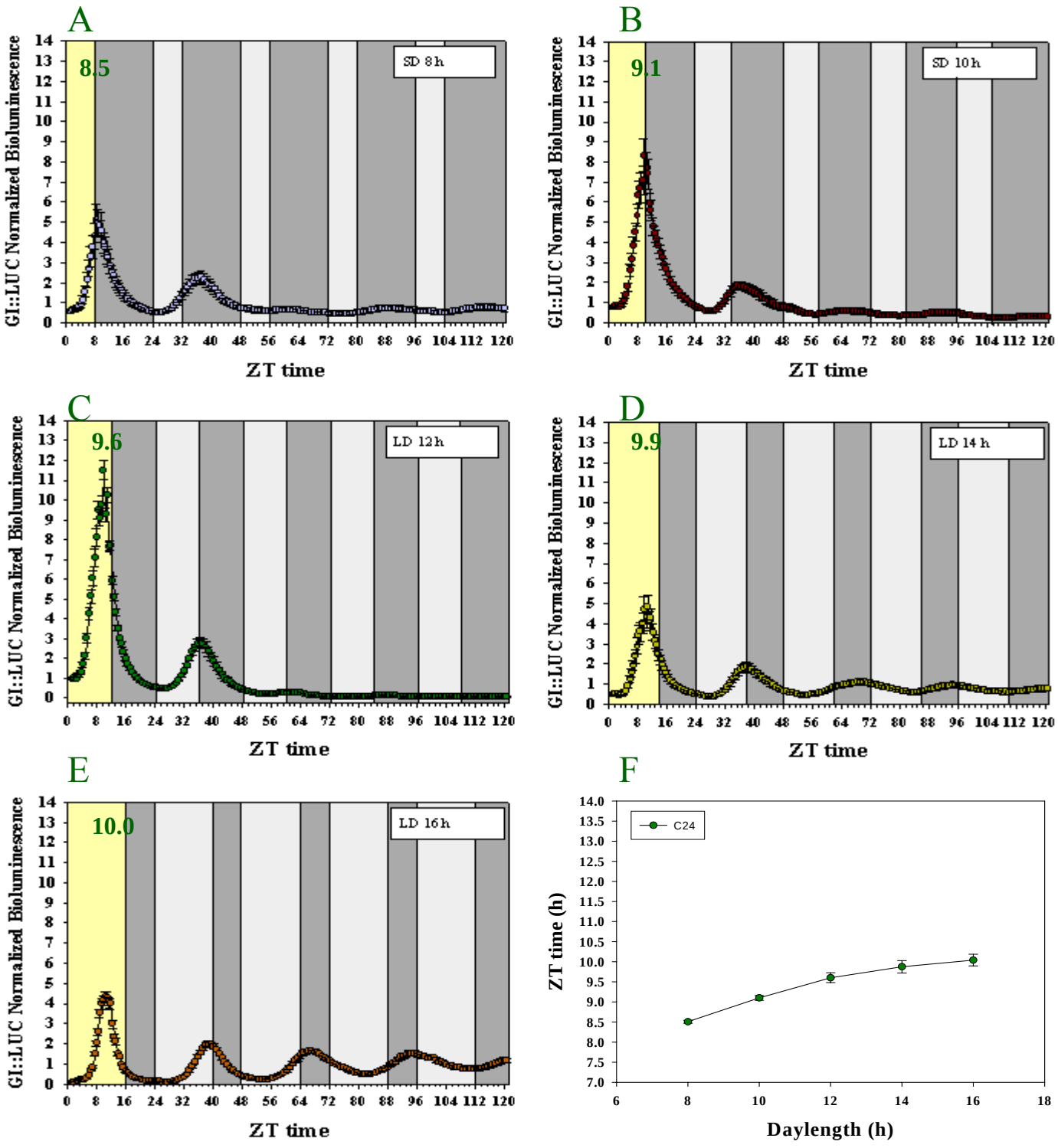


Figure 63 : Real time monitoring of GI::LUC bioluminescence in C24 in different photoperiods and in constant dark. A) SD 8 h B) SD 10 h C) LD 12 h D) LD 14 h E) LD 16 h F) Daylength response curve. The numbers indicate the time of the peak. The data in A) B) C) D) and E) correspond to average values of 24 seedlings, after normalization to the average of the expression. The data in F) represent the average from one transformant.

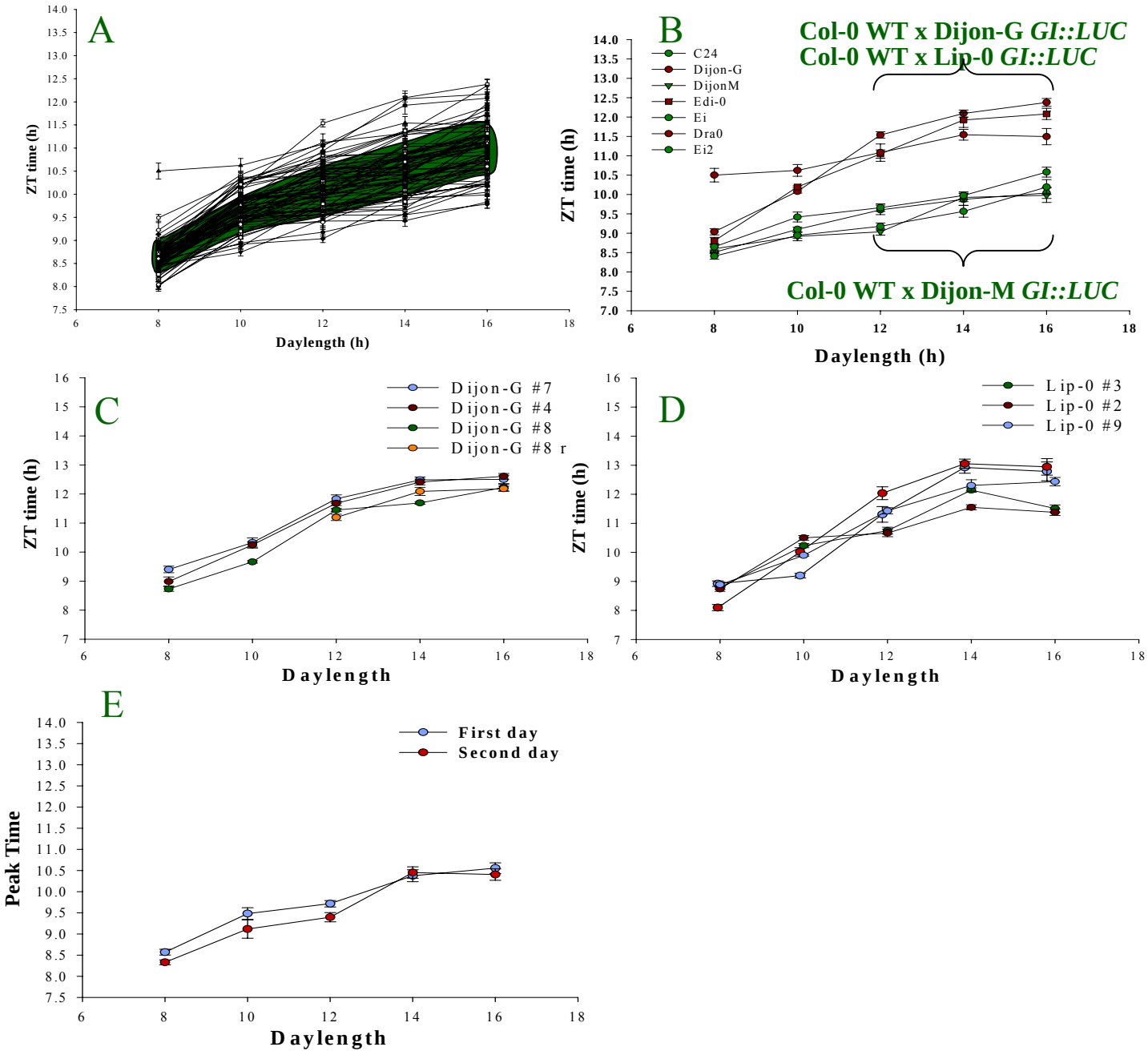


Figure 64 : Identified natural variation under long days. A) The general pattern of the response of *GIGANTEA* under five day lengths for 80 accessions. A strong response is observed for almost all accessions between SD 8 and 10 h. *GI* tracks the lengthening of the day length more efficiently in some accessions compared to others, resulting in a broad pattern of responses under LDs (green shade). B) Accessions with different response to day length. In the red accessions, *GI* show a phase delay under LDs. In the green accessions, the ability of *GI* to track the changing day length is arrested. Note that both groups of accessions share similar peak times under SD of 8 h and respond equally strong between SD 8 and 10 h. Dijon-G and Lip-0, two hyper responsive accessions were crosses to Col-0 WT in order to identify QTLs which control the observed phase delay under LDs. In addition Dijon-M will be crossed to Col-0 WT in order to detect QTLs which reduce the response under LDs. C) and D) Confirmation of the response curve for Dijon-G and Lip-0 with additional transformants. E) Comparison of the shape of the daylength response curve based on one and two constitutive days of imaging. No differences were observed.

Although the expression of *GIGANTEA* is shifted later within the day by the increasing day length, gradually a bigger portion of the expression is under light (Fig 60 to 63). Since hyper-responsive accessions such as Dijon-G or Lip-0 and hypo-responsive ones like Dijon-M and C24, show different timing of the peak in *GI::LUC* expression, differences also occur in the portion of the rhythm, which is under light. As shown in Fig. 60 and 61, in the hypo-responsive accessions Dijon-M and C24 the portion of the rhythm under light greatly increases as the day becomes longer, in contrast to the hyper-responsive Dijon-G and Lip-0 (Fig 62 and 63).

3.3.1.3 Natural variation in *GIGANTEA* free running period

In contrast to the strong influence that photoperiod had on the phase of *GI::LUC*, as expected the free running period under constant dark was not affected by the entraining conditions (Fig 65). However, differences up to four hours in the free running period between the accessions were detected in any given day length, suggesting natural variation in the general circadian regulation. According to the hierarchical clustering two major clusters where observed and accessions with long or short free running periods were detected (Fig 66). Interestingly, accessions with a phase delay like Dijon-G, Lip-0 and S96 had longer free running periods than the average while accessions with a phase advance like Ei and C24 had shorter to the average free running periods.

However that was not always the case. Dijon-M and Co are two accessions with a phase advance and a short free running period under constant dark. This suggests that targets of natural variation on phase not always control free running period under constant dark. Since GI is controlled both by light and by the circadian clock, these phenotypic differences might genetically correspond to these different types of regulation. Last but not least natural variation was also detected in RAE, suggesting that the accessions differ in their ability to maintain *GI* rhythm in the absence of light.

3.3.1.4 Discussion and conclusions

We could observe phase shifts in *GIGANTEA* expression that were caused by the increasing day length. Higher plasticity in these responses was observed under long days compared to short days. The difference between hyper-responsive and hypo-responsive accessions regarding phase, was maximum one hour under SDs and approximately 3 hours under LDs. This resulted into differences also in the portion of the expression of GI that is under light in each day length. This observation could suggest differences also in the efficiency of the

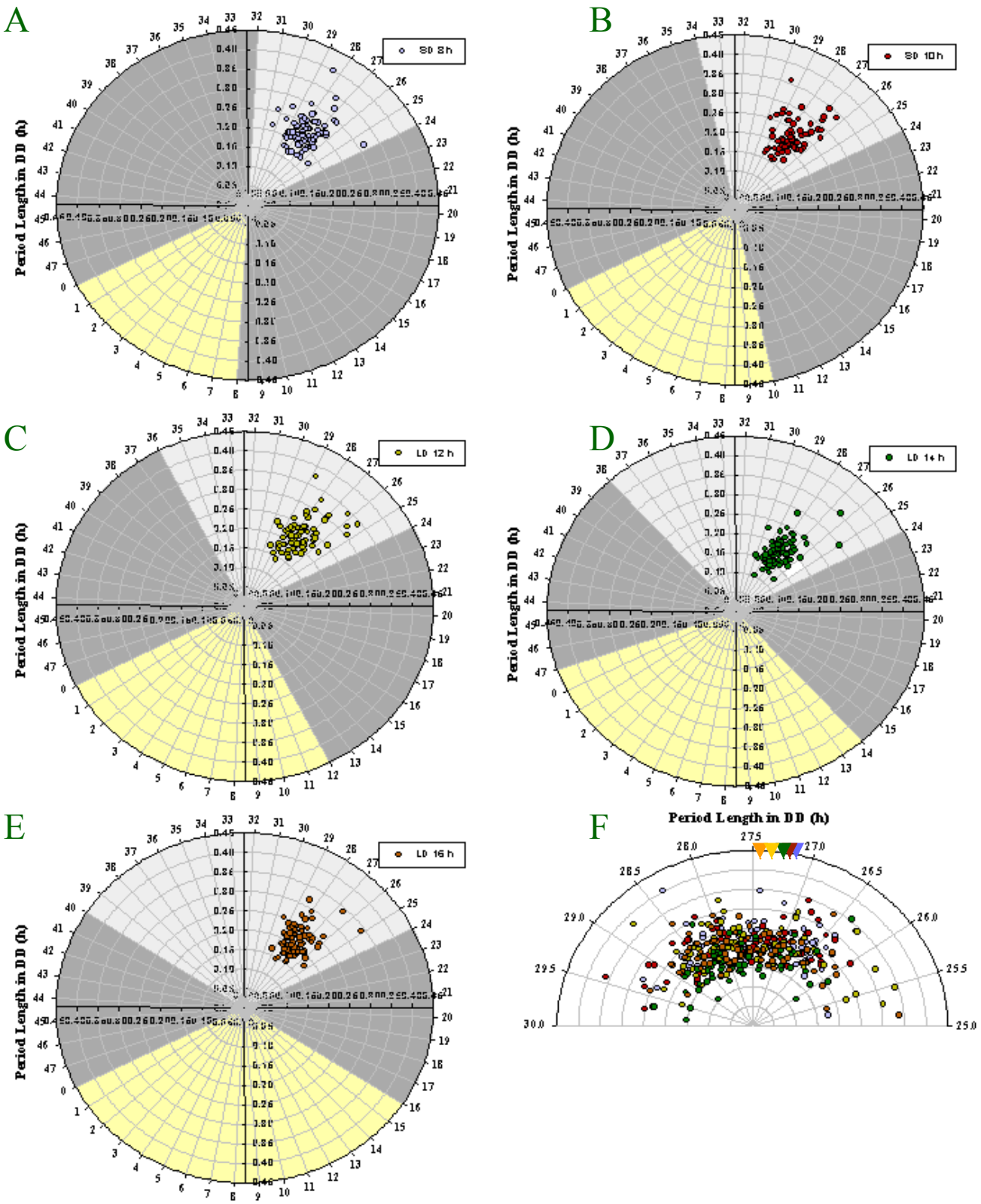


Figure 65 : Polarized scatter plots with the free running period of *GI* expression vs RAE after entrainment in five day lengths. A) SD 8 h B) SD 10 h C) LD 12 h D) LD 14 h E) LD 16 h F) Merged day length plot. Arrows indicate the average of the period length of all accessions in the corresponding day length. Not the overlapping values, reflecting the absence of a general effect of the changing day length on the free running period. Angular axis corresponds to the ZT transformed times of the first 48 h of the experiment. Radial axis corresponds to the RAE of each accession. Yellow : day, light grey : subjective day, dark grey : subjective night.

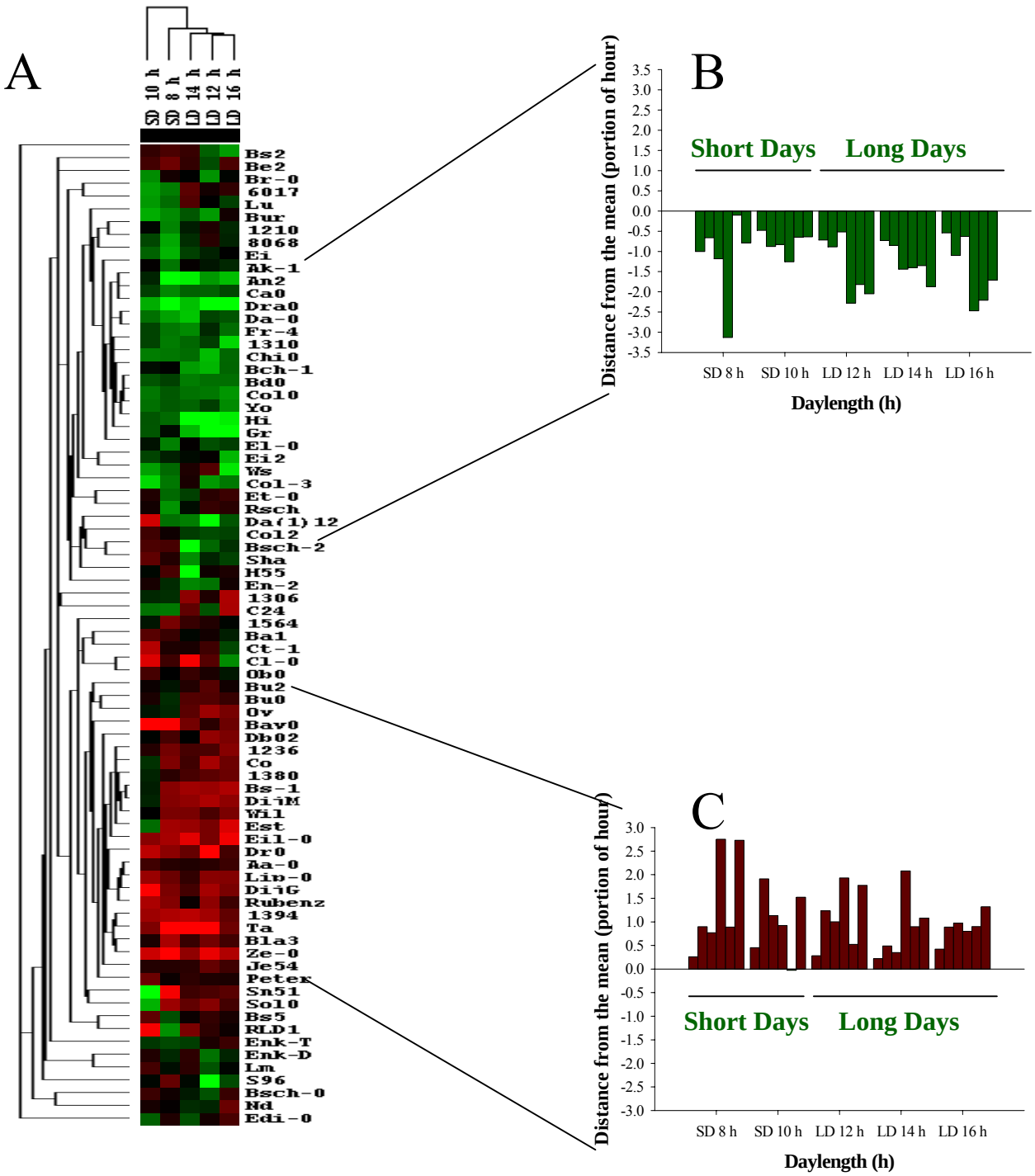


Figure 66 : The hierarchical clustering approach identifies natural variation in the free running period of *GIGANTEA* under constant dark. A) Output tree based on the period lengths of GI in 80 accessions under constant dark, after entrainment to five different day lengths. Two classes were observed. B) Distances from the mean of accessions with short period lengths under DD. C) Distances from the mean of accessions with long period lengths in DD. No effect of the entraining conditions could be observed in the free running period.

mechanism that involves GI and resets the clock by dusk among the accessions. In contrast to the influence of photoperiod on the phase of *GIGANTEA* expression, no changes in the free running period of the gene that were caused by the changing day length were detected. However, natural genetic variation was present between the accessions and resulted in differences of maximum 3-4 hours.

Recently it was shown that the GI protein is more stable in the light (David et al., 2006). Collectively the observations of phase regulation support a scheme of an increasingly abundant pool of GI protein as the day becomes longer, since bigger portion of the thyrhm coincides with light. Since GI positively regulates flowering time at least through CO, these results can at least in part explain the gradually early flowering with increasing photoperiods. However, epistasis with downstream elements of the photoperiod pathway, or from the vernalization pathway can revert the predicted flowering time responses in certain accessions. C24 and Yo for example are two hypo-responsive accessions regarding *GI::LUC* expression, but since both respond to vernalization, they exhibit a late flowering phenotype. Therefore what is more likely is that GI sets the base upon which other genes act in order to shape the final flowering response to the changing photoperiod.

According to the classical model, dawn resets the clock through *LHY* and *CCA1*, two morning genes with an acute response to light that is more pronounced in the morning. *GI* expression shows an acute response to light, which is gated (Frederic Cremer and Reka Toth, in preparation). Under short days this acute response is similarly strong throughout the oscillation of the expression, in contrast to long days where it becomes more evident during the time of the peak. The interlocked model of the circadian clock predicts the existence of a mechanism through which, dusk also resets the clock (Locke et al., 2005). Additional results suggest that GI participates in this mechanism and one of the roles of the gene is to reset the clock by dusk (Frederic Cremer and Reka Toth, in preparation).

A hyperactive acute response to light under short days could make GI a more sensitive marker in detecting changes in the occurrence of dusk and consequently, in day length. Under long days the restricted light response would act in order to prolong the peak of *GI* expression. It is therefore possible that in hyper-responsive accessions like Dijon-G and Lip-0, the acute response to light of GI, is gated differently compared to hypo-responsive accessions such as Dijon-M and C24, under long days. As a result, the mechanism of clock resetting by dusk would operate more efficiently in these accessions, allowing a better track of the changing day length.

3.3.2 GENETIC AND MOLECULAR ANALYSIS OF NATURAL VARIATION IN GI EXPRESSION

3.3.2.1 Introduction

Extensive variation was identified in the circadian parameters of GI expression under a range of different photoperiods. Various statistical approaches identified, in an unbiased way, accessions that differ significantly in their responses. Examples of such accessions are Dijon-G and Lip-0. Both accessions have a phase delay of 2 h under LDs, compared to Col-0, suggesting that the daylegth depended GI phase shifts, are hyperactive in these two genetic backgrounds. Interestingly natural variation was observed also in the freed running period of GI under constant dark, a strictly circadian parameter.

In order to elucidate the genetic basis that underlies this variability both accessions, carring the GI::LUC transgene, were crossed independently to Col-0 WT plants. The resulting two mapping populations went through a typical QTL mapping project, tailored to the needs of the in vivo real time luciferase imaging in order to identify for the first time QTL controlling the phase of an evening gene. QTL, which control the circadian period where also identified and comparisons of the two parameters were made. Last but not least like for flowering time, the resulting mapping populations went through a genome wide search for genetic interactions aiming to identify additional regulatory loci of the circadian traits of GI.

3.3.2.2 Phenotypic analysis of the GI expression in two mapping populations.

Both the Col-0 (WT) x Dijon-G (GI:LUC) and the Col-0 (WT) x Lip-0 (GI:LUC) populations were phenotyped for the circadian phase of GI under five day lengths for two constitutive days. After these two days the plants were released for four constitutive days in the dark and the free running period and the RAE circadian parameters were also monitored after entrainment in different day lengths. In both populations a strong influence of day length on phase was observed (Fig 58). By contrast the free running period remained unaffected by the previous conditions of entrainment. These results reflect accurately what has been observed at the level of the accessions for both parameters. Since the number of conditions is large, directly comparative histograms would have very low clarity due to overlaps of the segregation range between the conditions. For period since there are almost no differences between the day lengths, normal histograms would be pointless. To overcome these difficulties, all histograms were converted to cumulative frequencies.

For phase, under SD of 8 h both populations have a very sharp cumulative histogram, suggesting a monophasic segregation pattern (Fig 67). The transition point in the cumulative curve (or the peak of the histogram) fits the peak times of the two parents under this condition, indicating absence of transgression. As the day becomes longer the cumulative curve shifts towards higher ZT values, indicating the day length depended phase advance of GI expression. This shift is gradient in both populations like in the accessions. Interestingly the segregation pattern is also changed since the sharpness of the curve is gradually reduced. Thus comparing the SD of 8 h with the LD of 16 h one can observe that as the day becomes longer, a tail towards higher ZT values is formed or in other words the segregation range is less concentrated compared to SD. Since this range is far from being characteristic of the two parents, it suggests the presence of conditional trangression, indicating the activation of genes causing later phase specifically under LDs.

By contrast to phase, the cumulative segregation of period is not affected by the previous conditions of entrainment, exactly as it has been observed for the accessions (Fig 67). Under all day lengths, the cumulative histograms are almost identical for both populations with a slightly higher similarity in the Lip-0 population over the Dijon-G. Nevertheless the sharpness of the cumulative frequencies is not extreme and the saturation of the segregation occurs rather gradual in all cases. The segregation range extends beyond the period values of the two parents and suggests again the presence of transgressive variation, which in contrast to phase is not conditional. Collectively these data confirm the strong effect of day length on phase but not period, which was observed in the accessions and indicate the presence of balancing effect genes regulating phase in the parental genotypes. These effects are uncoupled in the F2 population leading to strong conditional trangression.

3.3.2.3 QTL mapping identifies regulatory loci controlling phase and period of GI expression.

Using the same approaches as for flowering time, the resulting two mapping populations were genotyped with the aim of identifying QTL regulating primarily phase under LD of 16 h and free running period under constant dark. For the Dijon-G population, a single QTL located between 14 and 16.9 Mbps has been detected for phase (Fig 68). By contrast three QTL have been detected for period. The first is located between 0,2 and 0,8 Mbps on top of Chromosome 4, the second between 8 and 9 Mbps while the third is located between 11,5 and 12,1 Mbps on the same chromosome (Fig 68). In the same population several QTL were also observed for RAE, reflecting the robustness of the rhythm under constant conditions. For this

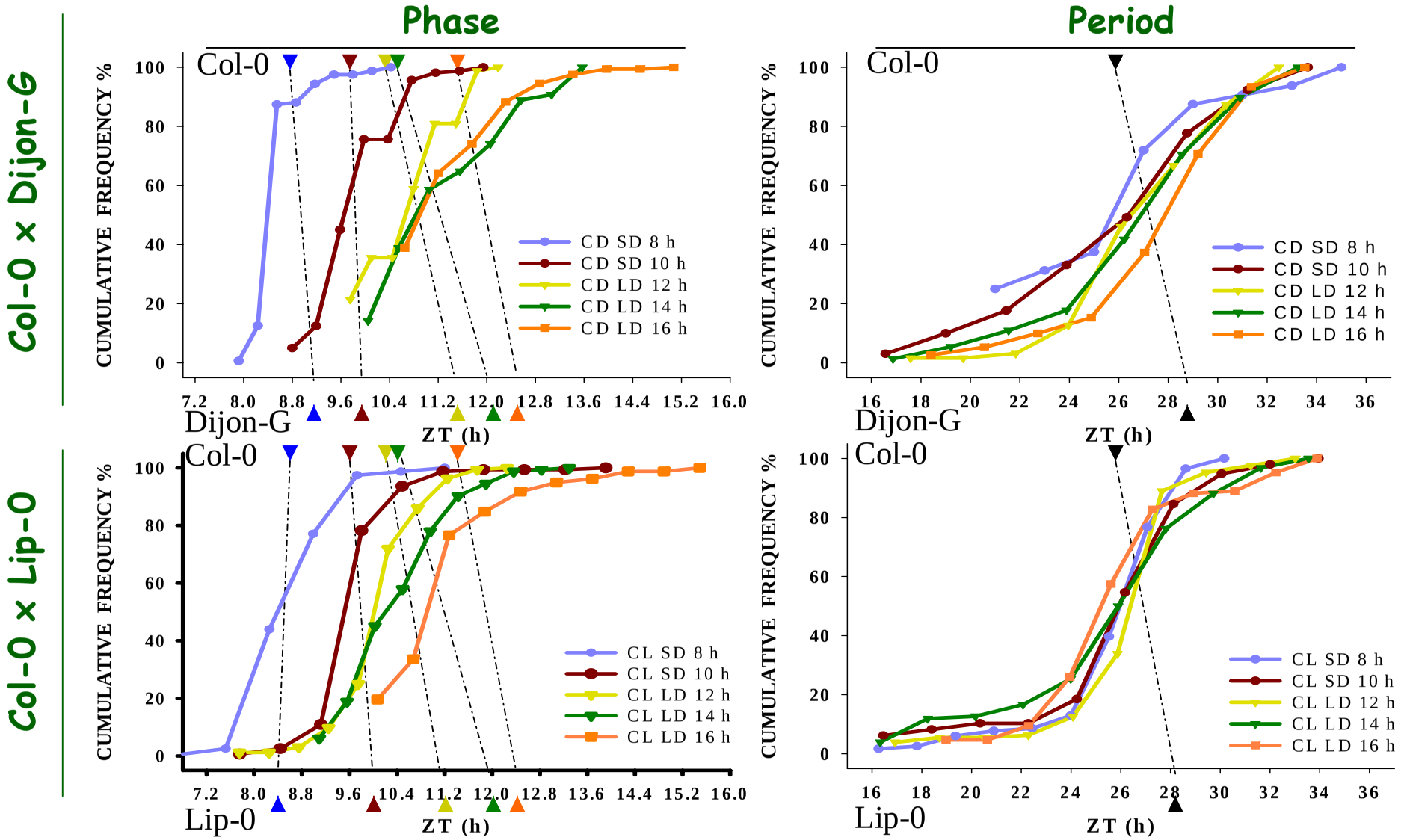


Figure 67 : Cumulative frequency distributions of the segregating Col-0 x Dijon-G GI::LUC and Col-0 x Lip-0 GI::LUC populations under five day lengths, for phase and free running period in DD. As expected, free running period is not affected from the entraining conditions. In contrast there was a strong effect of day length on phase. The sharp rise under SD for both populations reflects the strong kurtosis of the segregation. The gradual cumulative increase as the day becomes longer, indicates that the segregation range was increased due to the presence of conditional LD transgressive variation. The colored arrowheads represented the parental values, while the dashed lines the difference between the parental phenotypes. Upper arrowheads : Col-0, lower arrowheads : Dijon-G or Lip-0 parents. For period no difference was observed between the conditions, therefore the parental phenotypes are represented as a single arrowhead.

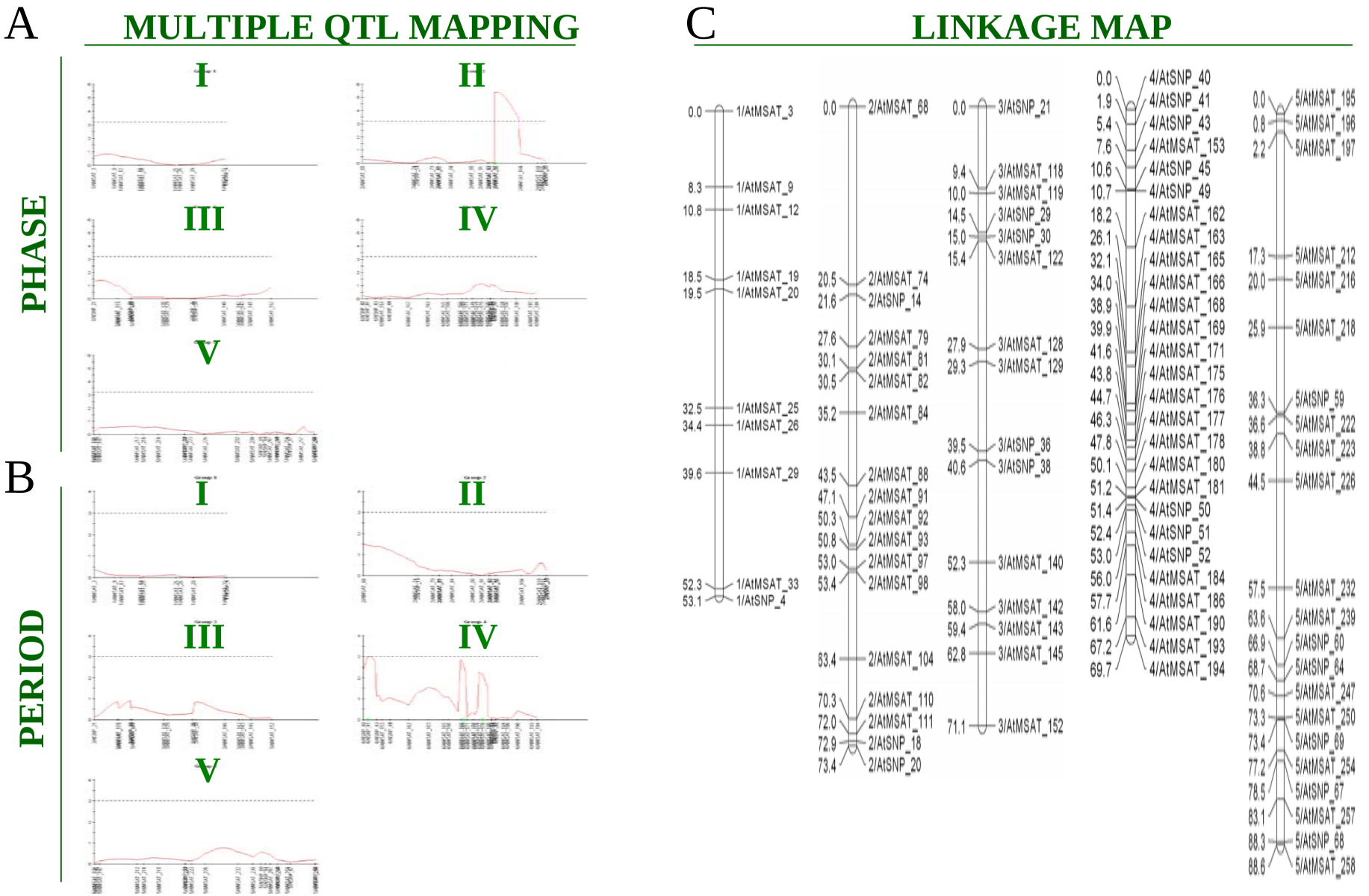


Figure 68 : QTL mapping and linkage map for the Col-0 x Dijon-G GI::LUC population. A) QTL map for phase under LD 16 h. B) QTL mapping for period under DD after entrainment in LD 16 h. C) Linkage map of the cross.

trait one QTL at the top of chromosome 2 between 0,3 and 2,4 Mbps and another QTL at the top of chromosome 3 between 0,6 and 5,3 Mbps were detected. A third QTL is located on top of chromosome 4 between 0,3 and 0,8 Mbps (data not shown). Interestingly the QTL for period and phase do not overlap in the Dijon-G population, suggesting that the genetic basis of phase and period for GI in different.

For the Lip-0 population three QTL for phase were detected (Fig 69). The first is located on chromosome 1 between 17,5 and 20 Mbps. The second QTL is located on the top of chromosome 3 between 9,3 and 2,2 Mbps. A third QTL was located on chromosome 4 between 12,1 and 14,7 Mbps. A fourth putative QTL was located on chromosome 2 between 7,6 and 8,7 Mbps, however the LOD score of this locus was below the significance threshold. For period four QTL were located (Fig 69). The first two were located on chromosome 2 between 6,5 and 8,4 Mbps and 9,4 and 12 Mbps. The third QTL was located at the top of chromosome 3 between 2 and 3,6 Mbps while the last QTL was located on chromosome 4 between 11 and 12 Mbps. In contrast to the Dijon-G population, in the Lip-0 population most of the QTL between phase and period either overlap or are located very close to each other. Conclusively the above suggest that phase in the Dijon-G population is not strictly circadian regulated since genetically is detached from a plain circadian parameter, period. However in the Lip-0 population QTL from the two parameters co-localise indicating common regulatory factors. The fact that most QTL for phase and period identified between the two populations, do not match suggests also that there are multiple ways of creating the same phenotype, an observation that was made also for flowering time.

3.3.2.4 Identifying genetic interactions for circadian parameters on a genome wide scale.

The observed transgressive variation both for phase and free running period in the Dijon-G and Lip-0 populations suggest the existence of balancing effects between multiple regulatory loci that control both traits in the parental genomes. Such QTL interact and therefore it is highly likely that epistasis can be detected in both populations on a genome wide scale. Like for flowering time, the saturated linkage maps of both populations allowed the search of genetic interactions directly in the F2 level on an unbiased, global scale.

Indeed several significant cases of interactions were identified for both populations and traits (Fig 25). For Dijon-G phase, three epiQTL were identified. The first is located on chromosome 1 at approximately 13 Mbps and interacts with a second locus on chromosome 3 between 9,3 and 9,7 Mbps. The same locus on chromosome 3 interacts with another locus on

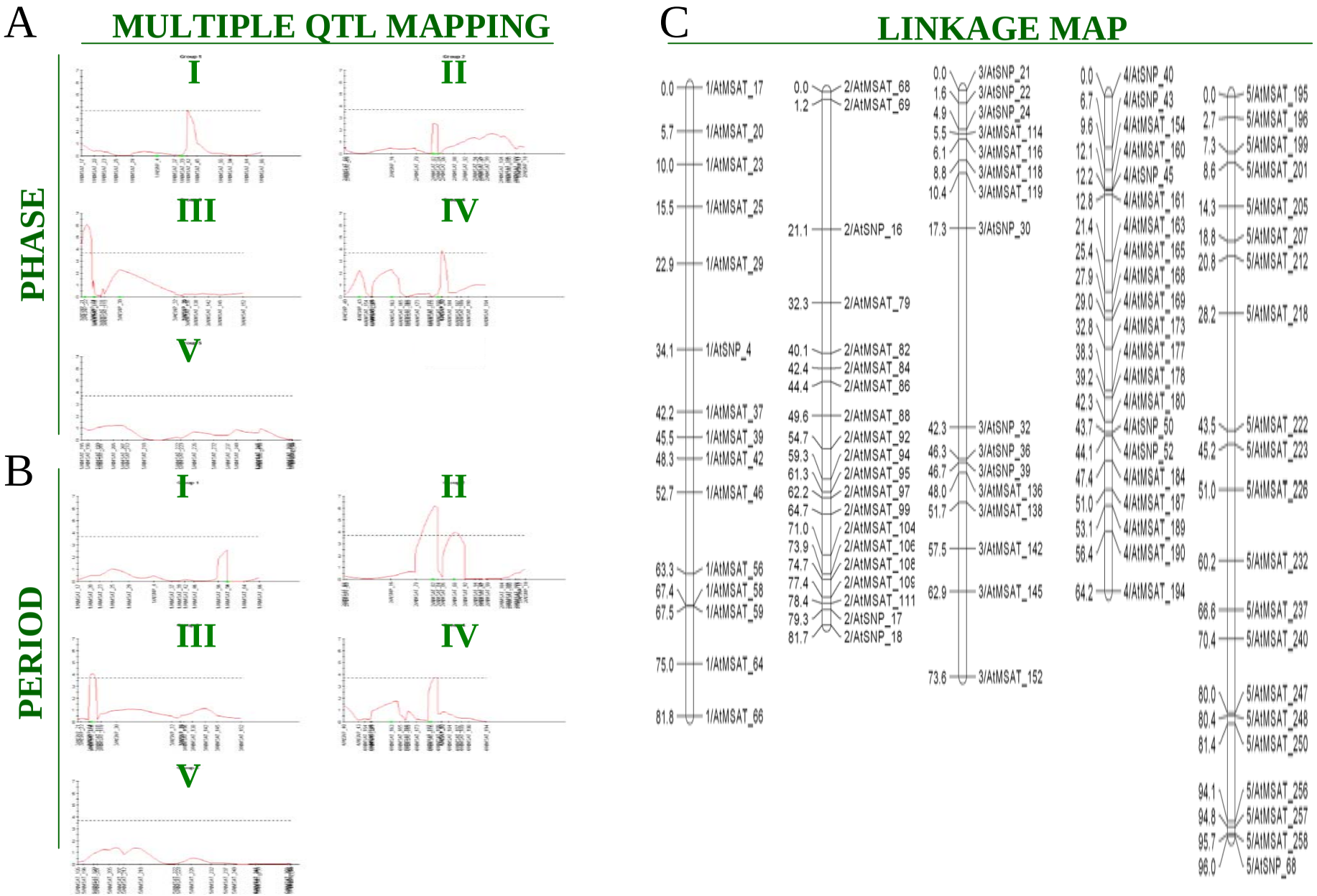


Figure 69 : QTL mapping and linkage map for the Col-0 x Lip-0 GI::LUC population. A) QTL map for phase under LD 16 h. B) QTL mapping for period under DD after entrainment in LD 16 h. C) Linkage map of the cross.

chromosome 4 between 9 and 11,1 Mbps, forming the major epiQTL of this population. Last but not least one epiQTL is located at the top of chromosome 4 (0 and 2,8 Mbps) and interacts with another locus on chromosome 5 between 20 and 23,8 Mbps.

For period in the same population two major epiQTL have been identified. The first is located on chromosome 2 between 1,2 and 1,4 Mbps and interacts strongly with another locus on chromosome 5 between 1 and 8 Mbps. A second very strong epiQTL controlling period in this population is located on chromosome 3 between 13,5 and 23,4 Mbps and interacts significantly with a second locus on chromosome 5 between 13,6 and 17,1 Mbps.

Several epiQTL for phase and period have been identified also in the Lip-0 population. For phase, three epiQTL were significant (Fig 25). The first was located on chromosome 1 between 9,3 and 11,1 Mbps and interacts with another locus at the bottom of the same chromosome at position 25,4 Mbps approximately. The major epiQTL of this population for phase is located at the top of chromosome 4 between 0,8 and 4,9 Mbps and interacts strongly with another locus on chromosome 3 between 0,6 and 0,9 Mbps. This epiQTL is of particular interest since at this locus on top of chromosome 4 one QTL for period has been identified. The last major epiQTL in that population for phase is located on chromosome 3 at position 10,7 and 13,5 Mbps and interacts with another locus on chromosome 5 between 13,6 and 16,3 Mbps.

For period three major epiQTL have been identified. The first is located on chromosome 1 at position 25 Mbps approximately and interacts with a second locus on chromosome 2 between 6,5 and 8,7 Mbps. The other two epiQTL are both located on chromosome 2 and interact with loci on chromosome 4. The first epiQTL of these two is located on chromosome 2 between 5 and 6,5 Mbps and interacts with the middle of chromosome 4, between 8 and 9,5 Mbps. The last epiQTL is the most significant of this population for period and is located in the middle of chromosome 2 at position 6,5 and 8,7 Mbps and interacts significantly with a locus at the bottom of the same chromosome in position between 17,2 and 18,5 Mbps. This locus on chromosome two is the same that interacts also with the locus on chromosome 1 of the first epiQTL for period in this population. Collectively the above results show that a significant yield of epiQTL both for phase and period has been made for both populations. In the majority of the cases these epiQTL are novel, they do not fit the positions of the classical QTL and are not shared between traits or populations. This indicates that complex genetic interactions are responsible for a large part of the variability observed in the circadian parameters of GI expression. Such epistatic interactions underlie also the strong transgressive segregation observed in the F2 populations for all traits and under various conditions. Fig 36

summarises the identified QTL together with the complex genetic interactions for phase and period.

3.3.2.5 Light response of GI is differentially gated between the accessions

It was shown that the phase delay in GI expression occurs strongly under LDs creating the difference of 2 h in the peak times of Dijon-G and Lip-0 compared to Col-0. By contrast both under SD of 8 and 10 h the peak times of the accessions where very similar. This suggested that a mechanism is activated upon LDs creating the observed differences. Interestingly based on the QTL mapping results, period from phase were genetically detached in the Dijon-G population since the identified QTL are not shared between the two parameters, in contrast to the Lip-0 population. Collectively the above indicate that in Dijon-G the mechanism that creates the phase difference might be associated with the way that light affects the rhythm of GI. It is already mentioned that GI expression is regulated both by light and by the circadian clock. Free running period under constant dark is a strictly circadian parameter since it reflects the ability of a rhythm to be maintained in the absence of light. Common feature in the aspects of light regulation is the fact that light is not equally efficient to affect the rhythm throughout the day. In other terms, there are only certain times of the day that the effect of light on the rhythm is maximised and during the rest of the day the effect is buffered. This phenomenon is called gating.

It would be therefore interesting to test whether natural variation exists in the gating of light in the rhythm of GI. In such a scenario the gating profile of Dijon-G should be different from the rest of the accessions specifically under LDs but not under SDs. This would ensure a direct connection with light regulation and at the same time would fit to the LD effect of the phase advance. For that reason the gating profile of two hyper-responsive accessions (Dijon-G and Lip-0, Fig 60 and 61) and one hypo-responsive one (Dijon-M, Fig 62) was examined after entrainment in SD of 8 h and LD of 16 h. Under SDs all three accessions showed a similarly low response to light, which was equal throughout the day (Fig 70). Sole exception were two time points in the Lip-0 population at ZT 34 and 36, at which the response was significantly lower compared to the rest time-points.

However under LDs the effect was much different. Both the Lip-0 and the Dijon-M profiles were almost identical with the common feature of increased response to light especially around the peak time of GI and low response in the middle of the day. By contrast the light response of Dijon-G was equally low throughout the day and resembled strikingly the SD

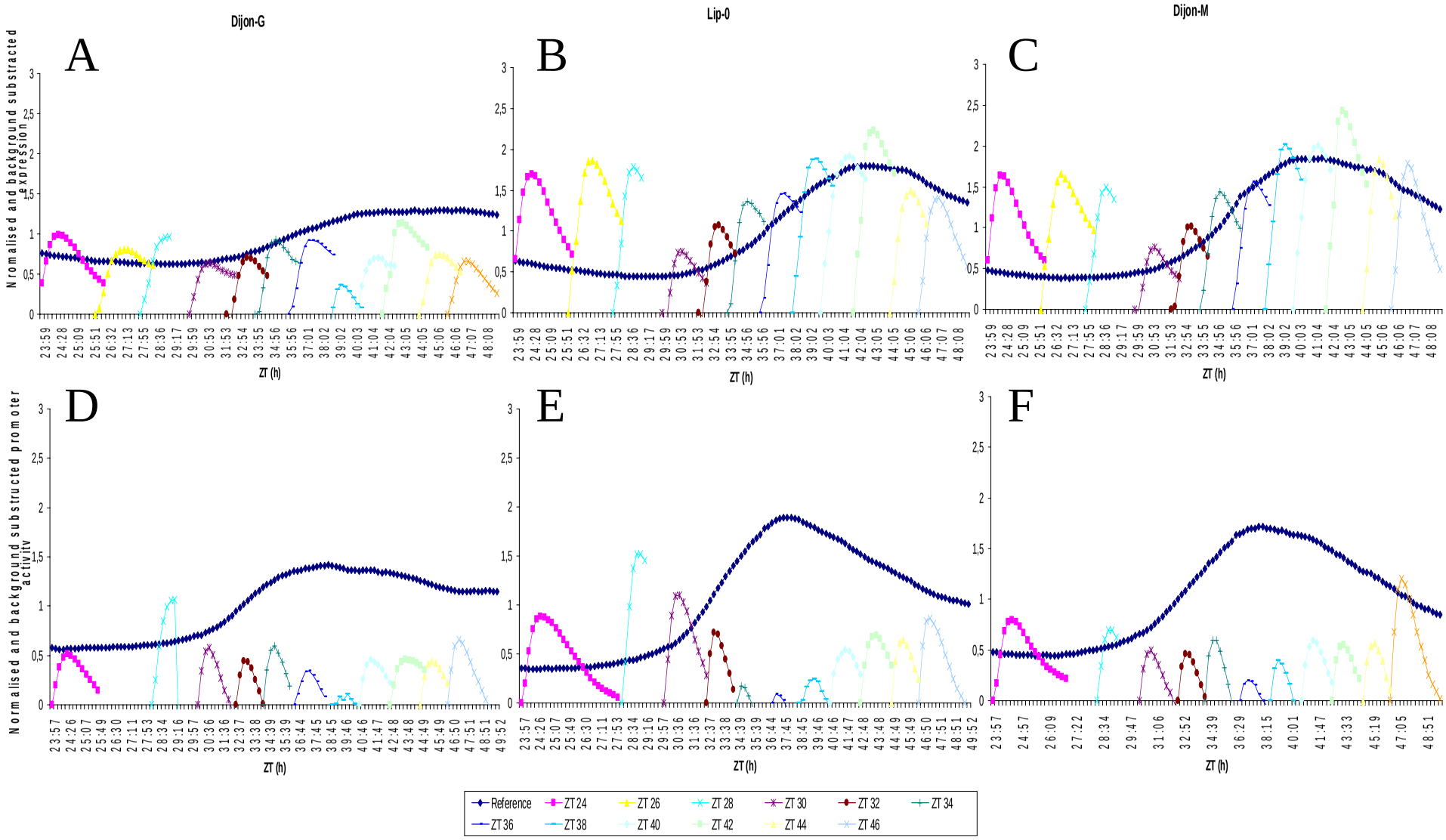


Figure 70 : Identified natural variation in the gating of the response of *GIGANTEA* to light. Two hypersensitive accessions Dijon-G (A and D) and Lip-0 (B and E) have been assayed together with the hyposensitive Dijon-M (C and F) in two daylegths LD 16 h (A, B and C) and SD 8 h (D, E and F). Under SD all accessions show a similar and equal response to light during the day with the exception of Lip-0. Under LD 16 h Lip-0 and Dijon-M show higher response to light at at around the circadian peak of GI. In Dijon-G however a pattern similar to SDs is observed with equally low response during the day. Since GI expression is regulated also by light, this lack of gating under LDs could be responsible for the phase delay of GI under LD in Dijon-G.

profile (Fig 70). Collectively these data confirm that light indeed has a different effect in the rhythm of GI in Dijon-G population specifically under LDs.

3.3.2.6 Comparing the promoter activity and mRNA profile of clock genes in different accessions and photoreceptor mutants.

If light is responsible for the phase shifts in the rhythm of GI in Dijon-G, then affecting the function of photoreceptors in a reference accession should phenocopy the LD specific phase delay. In order to test this hypothesis two different routes were selected. Initially the effect of the photoreceptor mutants was evaluated using real time imagine of bioluminescence, in the same manner as for the accessions. This allowed thorough characterization of the effect in a variety of conditions however it was based strictly on promoter activity. In a second step the effect under LD of 16 h was compared at the mRNA level using RT-PCR.

According to the results of the LUC imaging, certain but not all photoreceptor mutants indeed have an effect in the phase of GI, which phenocopies remarkably Dijon-G. More specifically the *hy1* mutation that corresponds to the CRY1 loss of function does not have any effect in the phase of GI under the examined five day lengths in the Ler background (Fig 71). But that was not the case for the *cry2* and *phyA* single mutants as well as the *cry1cry2* double. Both the *cry2* and the *cry1cry2* mutants had a significant phase advance of less that an hour specifically under SDs compared to the reference Ler rhythm. The effect of phyA was insignificant and matched quite accurately the Ler rhythm. Under LDs, a significant phase delay was observed in all three mutants, starting from LD of 12 h. Interestingly this phase delay in both the single *phyA* and the double *cry1cry2* became maximum at LD of 14 h and remained the same in LD of 16 h. The difference with Ler was approximately 1 h at this point with *phyA* and approximately 1,5 h with *cry1cry2*. *cry2* was similar with the exception that the effect of phase delay was not maximised at LD of 14 h but continued to increase in LD of 16 h, fitting accurately to the pattern of Dijon-G. Overally the pattern of *cry2* was the most hyper-responsive across all day lengths with phase advance under SDs and phase delay under LDs.

These results confirm that the effect of photoreceptor mutants mimics the pattern of Dijon-G, and further strengthen the scenario of a light and not circadian dependent effect on GI expression in this accession. However if indeed photoreceptors are involved, their effect should be more general and should affect more circadian clock genes. In order to test this hypothesis and to confirm the promoter activity results at the mRNA level, RT-PCR under LD of 16 h (where the effect on the rhythm is maximum) were performed (Fig 72). The

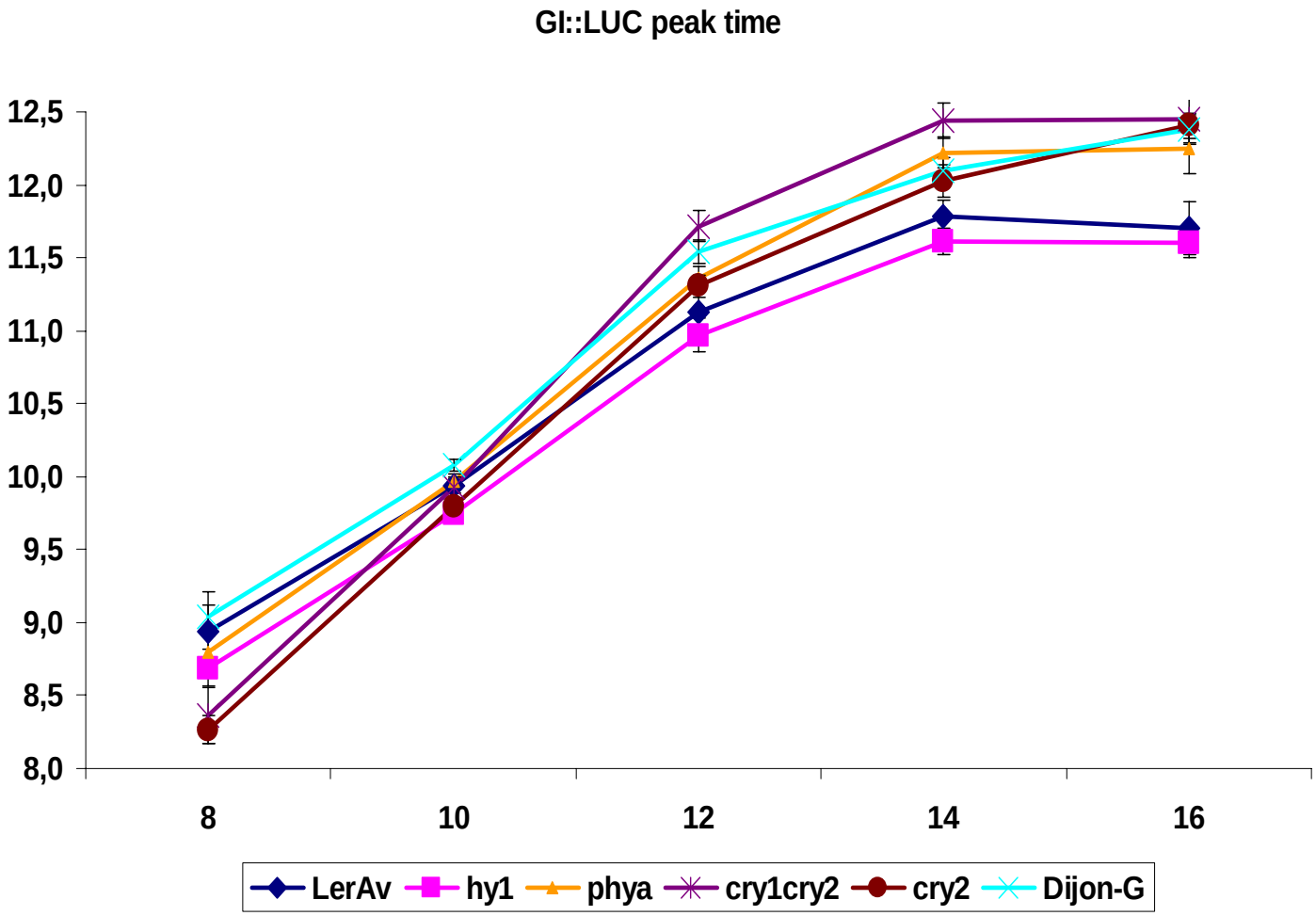


Figure 71 : Effect of photoreceptor mutations on *GIGANTEA* phase over a range of different daylengths**.** The hy1 mutant is similar to WT. By contrast the double cry1cry2 and the single cry2 and phyA mutants cause a phase delay, which becomes more pronounced as the day becomes longer.

expression of three circadian regulated genes was tested. Two evening genes (GI and TOC1) and one morning gene (LHY) were selected in order to identify more general effects in the clock.

GI mRNA expression under LD of 16 h in Dijon-G showed a two hours shift towards dusk compared to the peak times of Col-0 and Cen-0 which were used as reference, confirming the results of the promoter activity-based luminescence imaging (Fig 72). Interestingly analogous phase shifts were observed in the case of *cry* and *phyA* mutants. But was this effect specific for GI? The answer is no, since LHY expression was also affected. LHY shows an acute response to light in the morning. This acute light response was dramatically enhanced in Dijon-G compared to Cen-0, Ler and Col-0 (Fig 72B). Exactly the same effect was observed again in *cry* mutations (Fig 72E). Last but not least similarities were observed also in the case of *TOC1* rhythm, the effect of which was phase advance in Dijon-G compared to Ler (Fig 72 C and F). It is important to remind that Dijon-G and Ler are very similar accessions genetically yet significant differences are observed in the rhythm of *GI*. Collectively all the data indicate that the effect in the rhythm of *GI* is light regulated, it can be related with a failure to gate the response to light under LDs and it seems to be associated to cryptochromes and PHYA.

3.3.2.7 Discussion

In this study QTL for the circadian phase of an evening gene were isolated for the first time. Only one previous study used a similar approach in order to isolate QTL that affect circadian phase of CAB, a morning gene (Darrah et al., 2006). In this study circadian phase was determined as the circadian peak in the first cycle in constant dark after entrainment into different photoperiods. In the present study circadian phase was evaluated as the peak during exposure in a certain photoperiod and was verified in a constitutive measurement in the same conditions for selected accessions before the plants were shifted to constant dark. By contrast to phase, several studies are reported for the isolation of QTL that control circadian period, mainly under continuous light (Swarup et al., 1999).

Phase and not free running period is a better trait in order to link the circadian clock to the environment because in nature plants do not proliferate in areas where constant dark or light occur (e.g the poles). Therefore natural variation in phase has the potential of being adaptive. So far all of the studies that identified genes that affect the circadian clock, showed a tight connection between phase and period. Mutations that affected one trait, affected the other as well in genes such as TOC1 and ZTL (Somers et al., 2000 ; Strayer et al., 2000). However in

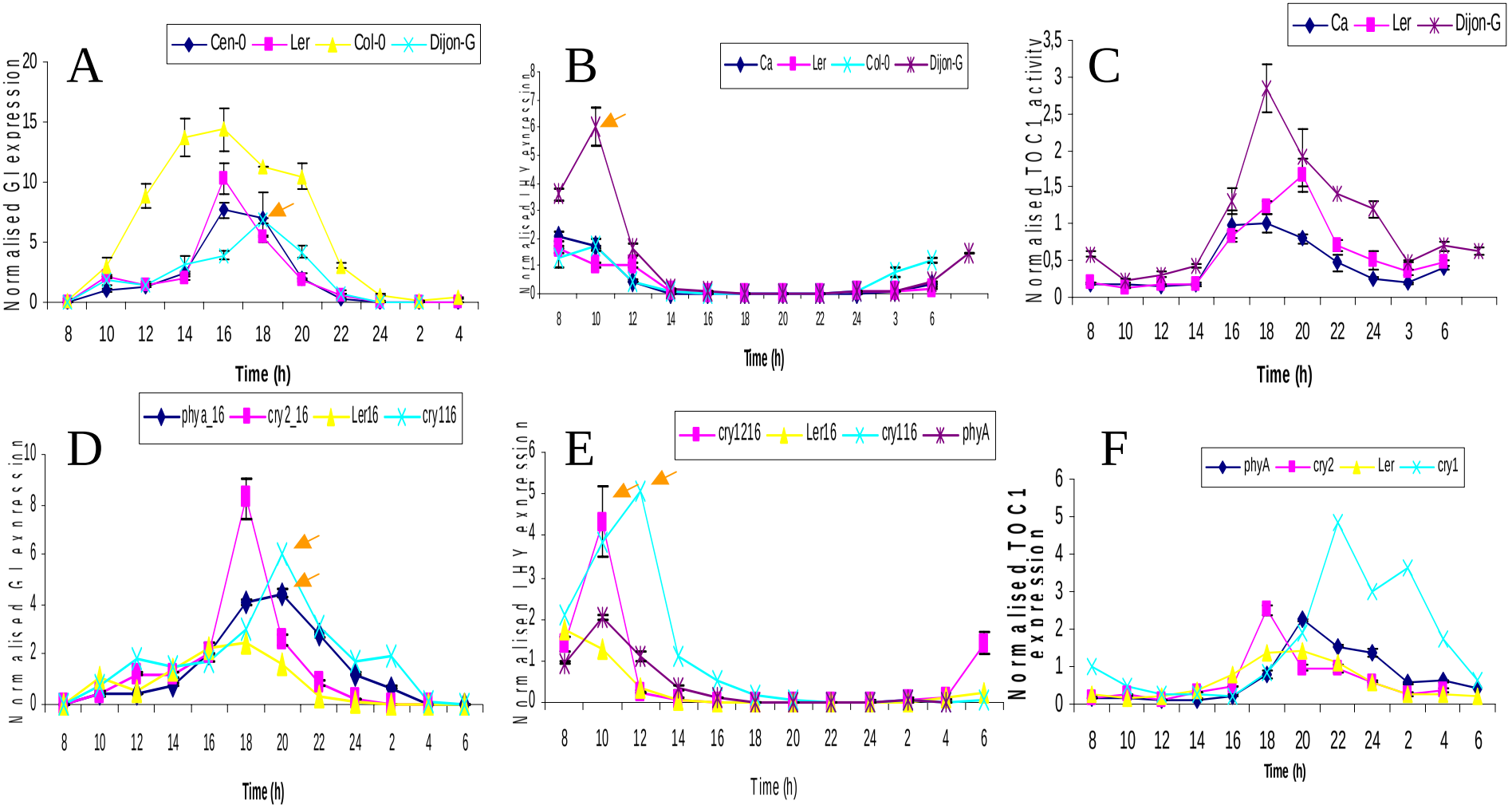


Figure 72 : Expression analysis of clock genes under LD 16 h in selected accessions and photoreceptor mutants. A) *GIGANTEA* mRNA expression in selected accessions. Orange arrow indicates the phase delay of Dijon-G compared to the rest accessions, confirming the luciferase data. B) *LHY* expression in selected accessions. Note the increased acute response of the rhythm in Dijon-G in the rest of the accessions. (orange arrows) C) *TOC1* expression in selected accessions. D) *GIGANTEA* expression in photoreceptor mutants. Note the phase delay in *cry* mutations (orange arrows). E) *LHY* expression in photoreceptor mutants. Note the increased acute response in *cry* mutants (orange arrows). F) *TOC1* expression in photoreceptor mutants.

studies that utilised accessions including this one, a correlation between phase and period has been lacking (Michael et al., 2003). These data suggested that in nature the two traits are not connected and instead are regulated by different sets of genes. Such dissociation was observed also in the *Drosophila* circadian clock (Hall et al., 1990). Like for GI (Fig 58F) a strong effect for day length was observed for CAB phase (Darrah et al., 2006). By comparing the segregation of phase for the F2 populations with the range of phase values among the accessions for each day length, transgressive variation was observed for GI (Fig 67). This transgressive variation was also observed for CAB phase in a similar manner (Darrah et al., 2000), suggesting that both for a morning and an evening gene loci with balancing effects regulate the trait in wild strains of *Arabidopsis* suggesting the existence of stabilizing selection operating in the plant circadian clock. Yet like in previous studies (Darrah et al., 2000) this work failed to detect robust associations of circadian parameters with latitude, although such associations have been reported for the *Drosophila* clock (see introduction). The reason like for flowering time may lie in the complex population structure and the admixture within populations that has been observed for *Arabidopsis*.

A very strong trend for conditional regulation was identified for CAB phase since no QTL was common between the different conditions (Darrah et al., 2006). However it must be noted that the experimental conditions of this work not only included day lengths which were truly extreme (3 h of light or 21 h of light compared with 12 h of light) but also the difference between the photoperiods was really big in order to expect any type of correlation genetically and physiologically. In such extreme conditions like 3 h of light or dark only, clock function must be expected to deviate from the normal pattern. At least for flowering time this is true, since clock mutants converge already in SD of 6 h even though their flowering phenotypes are very diverse from LD of 16 h until SD of 8 h (Fig 8F). For that reason in the present study more physiologically relevant conditions were used and the day length range was from SD of 8 h until LD of 16 h, while even the SD of 6 h were excluded from the analysis.

In the previous studies no epistatic interactions were identified among the discovered QTL and no attempt to search for such interactions in a genome-wide scale was made. In this study several epiQTL were identified for phase both in the Dijon-G and the Lip-0 population indicating that genetic interactions are largely responsible for the control of the circadian characters. This study also shows for the first time that natural variation exists in the gating of the light responses of any circadian gene and that this is associated with different photoperiods (Fig 70). The connection of the Dijon-G phase phenotype with response to light, requires involvement of the photoreceptors. Indeed certain photoreceptor mutants have an

effect which phenocopies the Dijon-G phenotype of many clock genes with diverse roles and function (Fig 71 and 72).

Photoreceptors have been postulated to regulate *CAB* phase as well since genes that affect the function of phytochromes have been assigned as candidates for identified QTL and were confirmed using NILs (Darrah et al., 2006). Another candidate in that study was *PRR3*, which belongs to the same family as *PRR5* and *PRR7*. Here it was shown that these two genes act as positive elements in cryptochrome signalling and are therefore closely associated with normal photoreceptor function. As mentioned in the discussion for flowering time many photoreceptors such as *PHYA*, *B* and *C* together with *CRY2* have been identified as targets of natural variation influencing light responses and flowering time. Therefore increasing evidence exist that in natural populations, in which phase is detached genetically from period, light has a very important regulatory role. Plants with attenuated function of photoreceptors do not have any period phenotypes either in constant light of dark, thus such genes are suitable candidates in order to explain the dissociation between the two traits (Somers et al., 1998). Under continuous white light period effects are abolished due to functional redundancy between the photoreceptors while under dark of course not effect should be expected.

In this study several QTL controlling circadian period under constant dark were also isolated for the first time in *GI* (Fig 68 and 69). Unlike phase, no association was observed between period and the entraining conditions highlighting the strict circadian nature of the trait (Fig 65). Hierarchical clustering identified two main clusters with early and late period values among the accessions (Fig 66). Like for phase, transgressive variation was observed in the segregation for period, not only compared to the values of the F2 population parents but also in comparisons with all the accessions (Fig 65). This was also observed in published reports of leaf movement under constant light (Swarup et al., 1999), and together these two diverse sets of data indicate that stabilizing selection acts in nature in order to maintain circadian period at approximately 24 h. Interestingly genetic interaction in a global wide scale were identified in the present study, yet the previous studies failed in detecting such interactions between the identified QTL and did not proceed in interactions across the genome.

One of the roles of this study was to determine whether natural variation in the clock has any effects on flowering time. Indeed natural variation has been identified for phase of GI and had the potential of changing day length perception for flowering time. However a robust association between circadian phase of GI and flowering time was not observed (data not shown). Several reasons can explain this observation. In the Dijon-G population it seems that

light through the plant photoreceptors, has a major role in the determination of phase. However light has additional roles downstream of GI in the photoperiod pathway, since it regulates the GI protein abundance, the activity of the FKF1-GI protein complex and most importantly CO protein activity and accumulation (see introduction). Therefore any effect of light in the circadian phase of GI, might be masked by downstream effects mediated by the very same components of this response, the plant photoreceptors. In addition flowering time regulation is very complicated and does not rely solely on the photoperiod pathway. Genes such as FLC are epistatic to the long day pathway genes, yet they show a circadian clock phenotype both for phase and period verified in many studies (Swarup et al., 1999; Darrah et al., 2006). In order to unquestionably search for such interactions, NILs with verified roles in the clock, must be examined for flowering phenotypes.

3.3.2.8 Conclusions

In this study novel QTL for period and phase of an evening gene were detected using a high-throughput method of real time imaging of bioluminescence. In both Col-0 x Dijon-G and Col-0 x Lip-0 genetic interactions were observed in a genome wide scale, which complement the genetic architecture of both circadian traits. Transgressive variation was evident in all cases, reflecting the presence of stabilising selection for the circadian clock in the wild. Interestingly a detachment of period from phase was observed in the genetic level especially for the Dijon-G population. Natural variation was observed in the gating of light response, with that accession showing an equally low response to light throughout the day under LD, a pattern reminiscent of the SD conditions. Other accessions gated light normally, in order to increase response at the time of the circadian peak of the gene. Such observation suggested that the normal function of photoreceptors in this accession is compromised and indeed photoreceptor mutants phenocopy the pattern of expression of many clock genes in Dijon-G. Fine mapping together with sequencing of candidate genes will help elucidating the genetic nature of the detected phenotypes.

4. PERSPECTIVES

4.1 NATURAL VARIATION IN FLOWERING TIME

As a first step to characterise the $PHYC^{Cen-0}$, complementation crosses between Cen-0 and *phyc* in Col-0 will be made and F1 flowering time will be scored. In addition the allele from Cen-0 will be first genotyped with markers against known polymorphisms and eventually sequenced. Last but not least transgenic complementation of *phyC* Col-0 by $PHYC^{Cen-0}$-OX plants is already planned. In an attempt to split the effect of the different QTL, selected individuals have been backcrossed in order to create NILs for this population.

In order to further challenge the possible implication of HUA2 in the Bs-1 phenotype, the produced NILs were vernalised and their flowering times are being scored while this manuscript was written. In addition, samples during vernalization were taken from the NILs and the Bs-1 parent, in order to test the levels of FLC in these individuals. Remember that these NILs have FLC^{Ler} homozygous alleles. Differences in the FLC levels and vernalization response under 14 would be a strong argument for the HUA2 involvement. In such case the $HUA2^{Bs-1}$ allele will be sequenced and used in transgenic complementation tests.

In order to further establish the role of the PRR genes in the regulation of flowering time, CO protein levels will be tested in the *prr* mutants in collaboration with Ryosuke Hayama and Bhavna Agashe, two members of the Coupland group. The effect of the genes on hypocotyls lengths under certain spectra of light will help defining further the role of these genes in light regulated processes of plants. The $PRR5^{Bs-1}$ will be sequenced and subsequently used in complementation studies of the *prr5* mutant.

The produced NILs from the Bs-1 population will be used in order to define molecularly the pattern of *FT* and *SOC1* mRNA up-regulation in the vasculature and the SAM respectively and to show that the QTL that causes lateness under 14 h directly affects the kinetics of the floral induction. In an attempt to understand the reasons that prevent floral commitment under LD of 14 h, the mRNA expression of *TFL1* and *LFY* will be tested and the *LFY:TFL1* ratio will be determined. In order to fine map even further and to eventually clone the responsible gene, a BC2 F2 population of 3600 individuals sown under LD of 14 h, is growing as this manuscript is written. This population will be phenotyped and subsequently genotyped with the aid of new polymorphic markers designed every 100 kbp within the introgressed region. Finally in order to further understand the molecular consequences of the observed phenotypes for Bs-1 and Cen-0, and to help identifying the causal genes, microarray experiments during a transient exposure to inductive LD of 14 and 16 h are underway. These

microarray experiments will be compared with the published microarrays under continuous LD of 16 h between Cen-0 and Ler in order to detected similarities, while SFP signatures for these accessions be a valuable aid both as a control against mis-hybridizations and as molecular sources of sequence diversity.

Under SD, selected F2 individuals will be tested in the F3 generation in order to confirm the role of the conditional QTL on chromosome 3 in day length perception. The F2 population will be vernalised in order to prove that the QTL on top of chromosome 4 is indeed FRI and confers a vernalization response. Samples from Dijon-G and Sha have been collected under SD of 8 h and the expression of flowering time genes will be tested with RT-PCR.

4.2 NATURAL VARIATION IN GIGANTEA EXPRESSION

For GI, F3 famillies have been already selected from both the Col-0 x Dijon-G and the Col-0 x Lip-0 population, in collaboration with Amaury de Montaigu in order to validate the observed QTL. Selected F3 individuals will be used in order to create NILs that will help splitting the effect of the QTL. Eventually fine mapping with markers designed in the QTL regions will be performed. Once the molecular nature of the Dijon-G and Lip-0 phase phenotypes is characterised it will be compared with the effects of *cis* variation between accessions in the promoter of GI, a project that is already running for comparisons between different *Arabidopsis* relatives in the lab. This will help characterising biochemicaly the interactions between transcriptional activators of GI and its own promoter and will help understanding how natural variation affects them. Significant amounts of *cis* natural variation in GI are reported in poplar (Ove Nilsson, personal communication).

LITERATURE

1. Abe, M., Y. Kobayashi, et al. (2005). "FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex." Science 309(5737): 1052-1056.

2. Adam, E., L. Kozma-Bognar, et al. (1997). "Tobacco phytochromes: Genes, structure and expression." Plant, Cell & Environment 20(6): 678-684.

3. Ahmad M, Cashmore AR., 1993. HY4 gene of A. thaliana encodes a protein with characteristics of a blue light photoreceptor. Nature 366: 162-166.

4. Alabadi David, Oyama Tokitaka, Yanovsky J. Marcelo, Harmon G. Franklin, Mas Paloma and Kay A. Steve, 2001. Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. Science 293: 880-883.

5. Alonso-Blanco, C. and M. Koornneef (2000). "Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics." Trends in Plant Science 5(1): 22-29.

6. An, H. L., C. Roussot, et al. (2004). "CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis." Development 131(15): 3615-3626.

7. Ausin I., Alonso-Blanco C., and Martinez-Zapater J. M., (2005) Environmental regulation of flowering. Int. J. Dev. Biol. 49: 689-705.

8. Balasubramanian Sureshkumar, Sureshkumar Sridevi et al. (2006) The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. Nature Genetics 38:711.

9. Bäurle, I. and C. Dean (2006). "The timing of developmental transitions in plants." Cell 125: 655-664

10. Bentsink L. Jowett j et al., (2006) Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis 103(45):17042

11. Berger J., Suzuki T. (2001) Genetic mapping with SNP markers in Drosophila. Nat Genet. 29(4):475

12. Bernier G and Perilleux C. (2005) A physiological overview of the genetics of flowering time control. Plant Biotech. J. 31: 3-16.

13. Blazquez, M. A., L. N. Soowal, et al. (1997). "LEAFY expression and flower initiation in Arabidopsis." Development 124(19): 3835-3844.

14. Bohlenius, H., T. Huang, et al. (2006). "CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees." Science 312(5776): 1040-3.

15. Borevitz J., Hazen R., et al., (2007) Genome-wide patterns of single-feature polymorphism in Arabidopsis thaliana PNAS 104(29): 12057

16. Borthwick HA., Hendricks SB, Parker MW., Toole EH., Toole VK., 1952. A reversible photoreaction controlling seed germination. Proc. Natl. Acad. Sci. USA, 38: 662-666.

17. Bradley, D., O. Ratcliffe, et al. (1997). "Inflorescence commitment and architecture in Arabidopsis." Science 275(5296): 80-83.

18. Brand U., Fletcher J., (2000) Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. Science 289(5479): 617

19. Brand, U., M. Hobe, et al. (2001). "Functional domains in plant shoot meristems." Bioessays 23(2): 134-141.

20. Brem R., Yvert G., et al., 2002 Genetic dissection of transcriptional regulation in budding yeast. Science 296(5568): 752.

21. Cashmore AR, Jarillo JA, Wu YJ, Liu D., 1999. Cryptochromes : blue light receptors for plants and animals. Science, 284: 760-765.

22. Cerdan, P. D. and J. Chory (2003). "Regulation of flowering time by light quality." Nature 423(6942): 881-885.

23. Chae E., Tan Q., et al., (2008) An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. Development 135(7): 1235.

24. Chailakhyan M and Butenko R (1957) Translocation of assimilates from leaves to shoots during different photoperiodic regimes of plants. Fiziol. Rast (trans) 4: 426.

25. Chen X. and Meyerowitz E. (1999). HUA1 and HUA2 are two members of the floral homeotic AGAMOUS pathway. Mol. Cell 3: 349.

26. Clark R., Schweikert G., et al. (2007) Common sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana Science. 317(5836): 338

27. Clark, S. E., R. W. Williams, et al. (1997). "The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis." Cell 89(4): 575-585.

28. Conti Lucio and Bradley Desmond (2007) TERMINAL FLOWER 1 is a mobile signal controlling Arabidopsis architecture. The Plant Cell 19: 767.

29. Coupland, G. (1995). "Genetic and environmental control of flowering time in Arabidopsis." Trends in Genetics 11(10): 393-397.

30. Cress W. D. and Triezenberg S. J., 1991. Science 251, 87.

31. Cutler SR., Ehrhardt DW., Griffitts JS., Somerville CR., 2000. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. Proc. Natl. Acad. Sci. USA, 97: 3718-3723.

32. Darrah Chiarina Taylor Bethan et al. (2006) Analysis of phase of LUCIFERASE expression reveals novel circadian quantitastive trait loci in Arabidopsis. The Plant Physiology 140: 1464.

33. David M. Karine, Armbruster Ute, Tama Nga and Joanna Putterill (2006). Arabidopsis GIGANTEA protein is post-transcriptionally regulatd by light and dark. FEBS letters

34. de Folter S. Immink R., et al., (2005) Comprehensive interaction map of the Arabidopsis MADS Box transcription factors Plant Cell. 17(5): 1424

35. Dodd N. A., Salathia N., Hall A., Kevei E., Toth R., Nagy F., Hibberd M., J., Millar J. A. and Webb R. A. A. (2005). Plant circadian clocks increase photosynthesis, growth, survival and competitive advantage. Science 309: 630-633.

36. Doyle M. Bizzell Colleen et al., (2005). HUA2 is required for the expression of floral repressors in Arabidopsis thaliana. The Plant Journal 41: 376.

37. El-Din El-Assal Salah, Alonso-Blanco Carlos, Peeters Anton, Raz Vered and Koornneef Maarten (2001) A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2.Nature Genetics 29: 435.

38. Eriksson Sven, Henrik Böhlenius, Thomas Moritz and Ove Nilsson. (2006). GA_4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. The Plant Cell 18:2172-2181.

39. Fankhauser, C., K. C. Yeh, et al. (1999). "PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in Arabidopsis." Science 284(5419): 1539-1541.

40. Ferrandiz, C., Q. Gu, et al. (2000). "Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER." Development 127(4): 725-734.

41. Filiault Daniele, Wessinger Carolyn et al. (2008). Amino acid polymorphisms in Arabidipsis phytochrome B cause differential responses to light. PNAS 105:3157.

42. Foster T. Lough T. et al., (2002) A surveillance system regulates selective entry of RNA into the shoot apex Plant Cell. (7):1497

43. Fowler, S., K. Lee, et al. (1999). "GIGANTEA: A circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains." Embo Journal 18(17): 4679-4688.

44. Gazzani Silvia, Gendall Anthony, Lister Clare and Dean Caroline (2003). Analysis of the molecular basis of flowering time variation in Arabidopsis accessions. Plant physiology 132 : 1107-1114.

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

46. Gisel, A., S. Barella, et al. (1999). "Temporal and spatial regulation of symplastic trafficking during development in Arabidopsis thaliana apices." Development 126(9): 1879-1889.

47. Glazier A., Nadeau J and Aitman T. (2002) Finding genes that underlie complex traits. Science. 298(5602): 2345

48. Goto, K. and E. M. Meyerowitz (1994). "Function and regulation of the Arabidopsis floral homeotic gene Pistillata." Genes & Development 8(13): 1548-1560.

49. Goto, K., J. Kyozuka, et al. (2001). "Turning floral organs into leaves, leaves into floral organs." Current Opinion in Genetics & Development 11(4): 449-456.

50. Gould D. Peter, Locke C.W. James, Larue Camille, Southern M. Megan, Davis Seth, Hanano Shigaru, Moyle Richard, Milich Raechel , Putterill Joanna, Millar J. Andrew and Anthony Hall. The molecular basis of temperature compensation in the Arabidopsis circadian clock. (2006). The Plant Cell. 18: 1177-1187.

51. Gregis V., Sessa A., et al., (2006) AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis Plant Cell 18(6): 1373

52. Griffiths, S., R. P. Dunford, et al. (2003). "The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis." Plant Physiology 131: 1855-1877.

53. Guo HW, Duong H., Ma N., Lin CT., 1999. The Arabidopsis blue light receptor cryptochrome 2 is a nuclear protein regylated by a blue light-dependent post-trascriptional mechanism. Plant J. 19: 279-287.

54. Gustafson-Brown, C., B. Savidge, et al. (1994). "Regulation of the Arabidopsis floral homeotic gene APETALA1." Cell 76(1): 131-143.

55. Hagenblad Jenny, Tang Chunlao et al., (2004) Haplotype structure and phenotypic associations in the chromosomal regions surrounding two Arabidopsis thaliana flowering time loci. Genetics 168: 1627-1638.

56. Hall J. and Kyriacou C (1990). Genetics of biological rhythms in Drosophila. Adv. Insect Physiology 22: 221.

57. Hansen B., Halkier B. and Kliebenstein D. (2008) Identifying the molecular basis of QTLs: eQTLs add a new dimension Trends Plant Sci. 13(2): 72

58. Harmer L. Stacey, Hogenesch B. John, Straume Marty, Chang Hur-Song, Bin Han, Zhu Tong, Wang Xun, Kreps A. Joel and Kay A. Steve, 2000. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science, 290: 2110-2113.

59. Hartmann, U., S. Hoehmann, et al. (2000). "Molecular cloning of SVP: A negative regulator of the floral transition in Arabidopsis." Plant Journal 21(4): 351-360.

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

61. Hempel, F. D., D. Weigel, et al. (1997). "Floral determination and expression of floral regulatory genes in Arabidopsis." Development 124(19): 3845-3853.

62. Hepworth, S. R., F. Valverde, et al. (2002). "Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs." Embo Journal 21(16): 4327-37.

63. Hicks KA, Albertson TM, Wagner DR (2001) EARLY FLOWERING3 encodes a novel protein that regulates circadian clock function and flowering in Arabidopsis. Plant Cell 13: 1281–1292.

64. Hoffman PD, Batschauer A., Hays JB., 1996. PHH1, a novel gene from Arabidopsis thaliana that encodes a protein similar to plant blue-light photoreceptors and microbial photolyases. Mol. Gen. Genet., 253: 259-265.

65. Hoffmann M (2002) Biogeography of Arabidopsis thaliana (L) Heynh (Brassicaceae) J. Biogeogr. 29: 125.

66. Huang, T., H. Bohlenius, et al. (2005). "The mRNA of the Arabidopsis gene FT moves from leaf to shoot apex and induces flowering." Science 309(5741): 1694-6.

67. Huang, T., H. Bohlenius, et al. (2007). "The mRNA of the Arabidopsis gene FT moves from leaf to shoot apex and induces flowering." Science retraction.

68. Huq, E., J. M. Tepperman, et al. (2000). "GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis." Proceedings of the National Academy of Sciences of the United States of America 97: 9789-9794.

69. Imaizumi, T., H. G. Tran, et al. (2003). "FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis." Nature 426(6964): 302-306.

70. Imaizumi, T., T. F. Schultz, et al. (2005). "FKF1 F-BOX protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis." Science 309(5732): 293-297.

71. Immink G. H. Richard, Gadella W.J. Theodorus, Jr., Ferrario Silvia, Busscher Marco and Gerco C. Angenent (2002). Analysis of MADs box protein-protein interactions in living plant cells. PNAS 99: 2416-22421.

72. Introduction to Plant Physiology, William G. Hopkins, Second edition

73. Jack, T., G. L. Fox, et al. (1994). "Arabidopsis homeotic gene APETALA3 ectopic expression: Transcriptional and posttranscriptional regulation determine floral organ identity." Cell 76(4): 703-716.

74. Jack, T., L. L. Brockman, et al. (1992). "The Homeotic Gene Apetala3 of Arabidopsis-Thaliana Encodes a Mads Box and Is Expressed in Petals and Stamens." Cell 68(4): 683-697.

75. Jander G., Norris S., et al., (2002) Arabidopsis map-based cloning in the post-genome era Plant Physiology 129(2): 440.

76. Jang, I. C., J. Y. Yang, et al. (2005). "HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling." Genes & Development 19(5): 593-602.

77. Jansen R. (2001) mRNA localization: message on the move Nat Rev Mol Cell Biol. (4):247

78. Johanson, U., J. West, et al. (2000). "Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time." Science 290(5490): 344-347.

79. Kania Thomas, Russenberger Doris et al. (1997) FPF1 Promotes flowering in Arabidopsis. The Plant Cell 9: 1327.

80. Keurentjes J., Fu J., et al. (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. PNAS 104(5): 1708.

81. Kim I., Cho E. et al., (2005) Cell-to-cell movement of GFP during embryogenesis and early seedling development in Arabidopsis PNAS 102(6):2227

82. Kim Jae-Yean, Song Hae-Ryong, Taylor L. Bethan and Carre A. Isabelle, 2003. Light-regulated translation mediates gated induction of the Arabidopsis clock protein LHY. The EMBO Journal, 22: 935-944.

83. Kim W., Fujiwara S et al. (2007) ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature 449 (7169): 356.

84. King, R. W. and J. A. D. Zeevaart (1973). "Floral stimulus movement in Perilla and flower inhibition caused by noninduced leaves." Plant Physiology 51: 727-738.

85. King, R. W., L. T. Evans, et al. (1968). "Translocation of the floral stimulus in Pharbitis nil in relation to that of assimilates." Z. Pflanzenphysiol. 59: 377-388.

86. Kleiner O., Kircher S., Harter K., Batschauer A., 1999. Nuclear localization of the Arabidopsis blue light receptor cryptochrome 2. Plant J., 19: 289-296.

87. Koornneef, M., C. Alonso-Blanco, et al. (2004). Naturally occurring genetic variation in Arabidopsis thaliana. Annual Review of Plant Biology 55(1): 141-172.

88. Koornneef, M., C. J. Hanhart, et al. (1991). "A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana." Molecular and General Genetics 229(1): 57-66.

89. Korstanje R and Paigen B. (2002). From QTL to gene: the harvest begins. Nat Genet. (3): 235.

90. Krizek, B. A. and E. M. Meyerowitz (1996). "The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function." Development 122(1): 11-22.

91. Kwiatkowska D (2008) Flowering and apical meristem growth dynamics J Exp Bot. 59(2): 187

92. Kyriacou C., Peixoto A., et al. (2007) Clines in clock genes: fine-tuning circadian rhythms to the environment Trends in Genetics 24 (3) : 124.

93. Lamb R., Hill T., et al., (2002) Regulation of APETALA3 floral homeotic gene expression by meristem identity genes Development 129(9): 2079

94. Lang, A. (1965). Physiology of flower initiation. Encyclopaedia of Plant Physiology. Vol 15 : W. Rhuland, ed (Spinger-Verlag, Berlin), 1380-1536.

95. Laurent Corbesier, Coral Vincent, Seonghoe Jang, Fabio Fornara, Qingzhi Fan, Iain Searle, Antonis Giakountis, Sara Farrona, Lionel Gissot, Colin Turnbull, George Coupland (2007). FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. Science 316: 1030.

96. Laux, T., K. F. X. Mayer, et al. (1996). "The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis." Development 122(1): 87-96.

97. Lee, H., S.-S. Suh, et al. (2000). "The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis." Genes & Development 14(18): 2366-2376.

98. Lee, H., S.-S. Suh, et al. (2000). "The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis." Genes & Development 14(18): 2366-2376.

99. Lempe Janne, Balasubramanian Sureshkumar, Sureshkumar Sridevi, Singh Anandia, Schmid Markus and Weigel Detlef (2005). Diversity in flowering time responses in wild Arabidopsis thaliana strains. PLOS Genetics 1: e6.

100. Levy Y.Y. Mesnage S., Mylne J.S., Gendall A.R. and Dean C., (2002). Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control. Science 297: 243-246.

101. Li Yan, Roycewicz Peter, Smith Evadne and Borevitz Justin (2006). Genetics of local adaptation in the laboratory : Flowering time quantitative trait loci under geographic ans seasonal conditions in Arabidopsis. PLOS One 1: e105.

102. Lifschitz, E., T. Eviatar, et al. (2006). "The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli." Proc Natl Acad Sci U S A 103(16): 6398-403.

103. Lin Chentao and Shalitin Dror, 2003. Cryptochrome structure and signal transduction. Annu. Rev. Plant Biol. 469-496.

104. Lin, Q., J. Li, et al. (1998). "Molecular cloning and chromosomal mapping of type one serine/threonine protein phosphatases in Arabidopsis thaliana." Plant Molecular Biology 37(3): 471-481.

105. Lin, Q., W. D. O. Hamilton, et al. (1996). "Cloning and initial characterization of 14 myb-related cDNAs from tomato (Lycopersicon esculentum cv Ailsa Craig)." Plant Molecular Biology 30(5): 1009-1020.

106. Liu C., Zhou J., et al., (2007) Specification of Arabidopsis floral meristem identity by repression of flowering time genes. Development. 134(10):1901.

107. Liu J, Zhang Y, et al., (2008). COP1-Mediated Ubiquitination of CONSTANS Is Implicated in Cryptochrome Regulation of Flowering in Arabidopsis. Plant Cell. 20(2):292.

108. Liu, L.-J., Y.-C. Zhang, et al. (2008). "COP1-Mediated Ubiquitination of CONSTANS Is Implicated in Cryptochrome Regulation of Flowering in Arabidopsis." Plant Cell: tpc.107.057281.

109. Locke, J. C. W., A. J. Millar, et al. (2005). "Modelling genetic networks with noisy and varied experimental data: the circadian clock in Arabidopsis thaliana." Journal of Theoretical Biology 234(3): 383-393.

110. Lohmann, J. U., R. L. Hong, et al. (2001). "A molecular link between stem cell regulation and floral patterning in Arabidopsis." Cell 105(6): 793-803.

111. Lough T and Lucas W (2006) Integrative plant biology: role of phloem long-distance macromolecular trafficking Annu Rev Plant Biol. 57:203

112. Lucas, W. J., B. C. Yoo, et al. (2001). "RNA as a long-distance information macromolecule in plants." Nature Reviews Molecular Cell Biology 2(11): 849-857.

113. Mackay T (2001) Quantitative trait loci in Drosophila Nat Rev Genet. 2(1): 11.

114. Macknight, R., I. Bancroft, et al. (1997). "FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains." Cell 89(5): 737-745.

115. Maity S. and de Crombrugghe B. (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription Trends Biochem Sci. 1998 23(5):174.

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

117. Mandel, M. A., J. L. Bowman, et al. (1992). "Manipulation of Flower Structure in Transgenic Tobacco." Cell 71(1): 133-143.

118. Mas P., Devlin PF., Panda S., Kay SA., 2000, Functional interaction of phytochrome B and cryptochrome 2. Nature, 408: 207-211.

119. Mas Paloma, Kim Woe-Yeon, Somers E. David and Kay A. Steve, 2003. Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana. Nature 426: 567-570.

120. Matsushika Akinori, Makino Seiya, Kojima Masaya, Yamashino Takafumi and Mizuno Takeshi, 2002. The APRR1/TOC1 quintet iplicated in circadian rhythms of Arabidopsis thaliana :II. Characterazation with CCA1-overexpressing plants. Plant Cell Physiol., 43: 118-122.

121. Mayer, U. and G. Jurgens (1998). "Pattern formation in plant embryogenesis: A reassessment." Seminars in Cell & Developmental Biology 9(2): 187-193.

122. Melzer Siegbert et al., (1999) FPF1 modulates the competence to flowering in Arabidopsis. The Plant Journal 18: 395.

123. Michael, T. P. and C. R. McClung (2003). "Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis." Plant Physiology 132(2): 629-39.

124. Michael, T. P., P. A. Salome, et al. (2003). "Enhanced fitness conferred by naturally occurring variation in the circadian clock." Science 302(5647): 1049-1053.

125. Michaels Scott, and Amasino Richard (2001) Loss of FLOWERING LOCUS C activity eliminates the late flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. The Plant Cell 13: 935-941.

126. Michaels, S. D. and R. M. Amasino (1999). "FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering." Plant Cell 11(5): 949-956.

127. Michaels, S. D., Y. H. He, et al. (2003). "Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis." Proceedings of the National Academy of Sciences of the United States of America 100(17): 10102-10107.

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

129. Moon, J., S. S. Suh, et al. (2003). "The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis." Plant Journal 35(5): 613-623.

130. Morley M., Molony C., et al. (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430(12) : 743.

131. Mueller R., Borghi L. et al., (2006) Dynamic and compensatory responses of Arabidopsis shoot and floral meristems to CLV3 signaling Plant Cell. 18(5): 1188

132. Nakamichi et al. (2007) Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. Plant Cell Physiol 48: 822-832.

133. Nilsson, O., I. Lee, et al. (1998). "Flowering-time genes modulate the response of LEAFY activity." Genetics 150(1): 403-410.

134. Nordborg M., Hu T., et al., (2005) The pattern of polymorphism in Arabidopsis thaliana. PLoS Biol. 3(7):e196

135. O'Neil C, Morgan C. et al (2008) Six new recombinant inbred populations for the study of quantitative traits in Arabidopsis thaliana TAG 116(5):623

136. Osterlund, M. T., C. S. Hardtke, et al. (2000). "Targeted destabilization of HY5 during light-regulated development of Arabidopsis." Nature 405(6785): 462-6.

137. Parcy F. (2005). Flowering: a time for intergration. Int. J. Dev. Biol. 49: 585-593.

138. Parcy Francois, Bomblies and Weigel Detlef (2002) Interaction of LEAFY, AGAMOUS and TERMINAL FLOWER1 in maintaining floral meristem identity in Arabidopsis. Development 129: 2519.

139. Parcy, F., O. Nilsson, et al. (1998). "A genetic framework for floral patterning." Nature 395(6702): 561-566.

140. Park D.H., Somers D.E., Kim Y.S., Choy Y. H., Lim H. K., Soh M.S., Kim H.J., Kay S.A. and Nam H.G., (1999). Control of circadian rhythms and photoperiodic flowering by the Arabidopsis GIGANTEA gene. Science 285: 1579-1582.

141. Pelaz, S., C. Gustafson-Brown, et al. (2001). "APETALA1 and SEPALLATA3 interact to promote flower development." Plant Journal 26(4): 385-394.

142. Pergota G., Nimu L., Flamma F., Weller JL., Kendrick RE., et al., 2000. Tomato contains homologues of Arabidopsis cryptochrome 1 and 2. Plant Mol. Biol., 42: 765-773.

143. Pouteau, S., D. Nicholls, et al. (1997). "The induction and maintenance of flowering in Impatiens." Development 124(17): 3343-3351.

144. Pouteau, S., F. Tooke, et al. (1998). "Quantitative control of inflorescence formation in Impatiens balsamina." Plant Physiology 118(4): 1191-1201.

145. Putterill, J., F. Robson, et al. (1995). "The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors." Cell 80(6): 847-857.

146. Quail PH., 1991. Phytochrome : a light-activated molecular switch that regulates plant gene expression. Annnual Review of Genetics 25, 389-409.

147. Quail, P., 1997. An emerging molecular map of the phytochromes. Plant Cell Environ. 20, 657–666.

148. Quail, P.H. (1994). Phytochrome genes and their expression, In Photomorphogenesis in Plants, 2nd ed.(eds. R.E Kendrick and G.H.M. Kronenberg) pp. 71-104. Kluwer, Dordrecht, Netherlands.

149. Ratcliffe, O. J., G. C. Nadzan, et al. (2001). "Regulation of flowering in Arabidopsis by an FLC homologue." Plant Physiology 126(1): 122-132.

150. Riechmann, J. L. and E. M. Meyerowitz (1997). "Determination of floral organ identity by Arabidopsis MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity." Molecular Biology of the Cell 8(7): 1243-1259.

151. Samach, A., J. E. Klenz, et al. (1999). "The UNUSUAL FLORAL ORGANS gene of Arabidopsis thaliana is an F-box protein required for normal patterning and growth in the floral meristem." Plant Journal 20(4): 433-445.

152. Sawa, M., D. A. Nusinow, et al. (2007). "FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis." Science 318(5848): 261-265.

153. Schaffer, R., N. Ramsay, et al. (1998). "The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering." Cell 93(7): 1219-1229.

154. Schmid K., Soerensen Rosleff T., et al. (2003) Large-scale identification and analysis of Genome-Wide single-nucleotide polymorphisms for mapping in Arabidopsis thaliana. Genome Res. 13: 1250.

155. Schmid K., Toerjek O., et al. (2006) Evidence for a large-scale population structure of Arabidopsis thaliana from genome-wide single nucleotide polymorphism markers TAG 122: 1104.

156. Schmid, M., N. H. Uhlenhaut, et al. (2003). "Dissection of floral induction pathways using global expression analysis." Development 130(24): 6001-6012.

157. Scortecci Katia, Scott D. Michaels and Ricard Amasino (2003). Genetic interactions between FLM and other flowering-time genes in Arabidopsis thaliana. Plant Molecular biology 52:915-922.

158. Scortecci, K. C., S. D. Michaels, et al. (2001). "Identification of a MADS-box gene, FLOWERING LOCUS M, that represses flowering." Plant Journal 26(2): 229-236.

159. Searle I, Men A et al., (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. Science 299(5603):109

160. Searle, I. and G. Coupland (2004). "Induction of flowering by seasonal changes in photoperiod." EMBO Journal 23: 1217-1222.

161. Searle, I., Y. He, et al. (2006). "The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis." Genes & Development 20(7): 898-912.

162. Senger H., 1980. The blue light syndrome, Berlin: Springer-Verlag.

163. Seo, H. S., J. Y. Yang, et al. (2003). "LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1." Nature 424(6943): 995-9.

164. Sessions, A., M. F. Yanofsky, et al. (2000). "Cell-cell signaling and movement by the floral transcription factors LEAFY and APETALA1." Science 289(5480): 779-781.

165. Shalitin, D., H. Y. Yang, et al. (2002). "Regulation of Arabidopsis cryptochrome 2 by blue-light-dependent phosphorylation." Nature 417(6890): 763-767.

166. Sharbel T., Haubold B., et al., (2000) Genetic isolation by distance in Arabidopsis thaliana: biogeography and postglacial colonization of Europe Mol Ecol. 9(12): 2109

167. Shaw R., Byers D, Darmo E. (2000) Spontaneous mutational effects on reproductive traits of Arabidopsis thaliana Genetics 155(1): 369

168. Shindo Chicado, Aranzana Maria Jose et al., (2005). Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. Plant physiology 138 : 1163-1173.

169. Shinozaki K., Yamaguchi-Shinozaki K.,2000. Current Opinion of Plant Biology, 3: 217.

170. Simon, R. and G. Coupland (1996). "Arabidopsis genes that regulate flowering time in response to day-length." Seminars in Cell & Developmental Biology 7(3): 419-425.

171. Simon, R., M. I. Igeno, et al. (1996). "Activation of floral meristem identity genes in Arabidopsis." Nature 384(6604): 59-62.

172. Simpson, G. G., A. R. Gendall, et al. (1999). "When to switch to flowering." Annual Review of Cell & Developmental Biology 15: 519-550.

173. Simpson, G. G., P. P. Dijkwel, et al. (2003). "FY Is an RNA 3' End-Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition." Cell 113: 777-787.

174. Smykal Petr, Gleissner Roland et al. (2004). Modulation of flowering time responses in different Nicotiana varieties. Plant Molecular Biology 55: 253.

175. Sohn J., Rojas-Pierce M., et al., (2007) The shoot meristem identity gene TFL1 is involved in flower development and trafficking to the protein storage vacuole PNAS 104(47): 18801

176. Somers D and Kay S (1998). Genetic approaches to the analysis of circadian rhythms in plants. Oxford: BIOS Scientific.

177. Somers E. David, Schultz F. Thomas, Milnamow Maureen and Kay A. Steve, 2000. Zeitlupe encodes a novel Clock-associated PAS protein from Arabidopsis. Cell, 101: 319-329.

178. Somers, D. E., T. F. Schultz, et al. (2000). "ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis." Cell 101(3): 319-329.

179. Sridhar V., Surendrarao A., et al., (2004) Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development PNAS 101(31):11494

180. Steinmetz L., Midrinos M. and Oefner P. (2000) Combining genome sequences and new technologies for dissecting the genetics of complex phenotypes. Trends in Plant science Trends Plant Sci. 5(9): 397

181. Stenoien Hans Fenster Charles Kuittinen Helmi and Sovolainene Outi (2002) Quantifying latitudinal clines to light responses in natural populations of Arabidopsis thaliana (BRASSICACEAE). American Journal of Botany 89: 1604-1608.

182. Stinchcombe, J. R., C. Weinig, et al. (2004). "A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA." Proceedings of the National Academy of Sciences of the United States of America 101(13): 4712-4717.

183. Stock J. B., Stock A. M., Mottonen J. M., 1991. Nature 344, 395.

184. Strayer Carl, Oyama Tokitaka, Schultz F. Thomas, Raman Ramanujam, Somers E. David, Mas Paloma, Panda Satchidananda, Kreps A. Joel and Kay A. Steve, 2000. Cloning of the Arabidopsis clock gene TOC1, a autoregulatory response regulator homolog. Science, 289: 768-771.

185. Strayer, C., T. Oyama, et al. (2000). "Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog." Science 289(5480): 768-771.

186. Suarez-Lopez, P., K. Wheatley, et al. (2001). "CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis." Nature 410(6832): 1116-1120.

187. Sung S., Amasino R., Vernalization in Arabidopsis thaliana is mediated by the PHD-finger protein VIN3 (2003). Nature 427: 159-164.

188. Sung S., He Y. et al., (2006) Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nature Genetics 38(6):706.

189. Sung, S. B. and R. M. Amasino (2004). "Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3." Nature 427(6970): 159-164.

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

191. Swiezewski S., Crevillen P et al. (2007) Small RNA-mediated chromatin silencing directed to the 3' region of the Arabidopsis gene encoding the developmental regulator, FLC PNAS 104(9):3633

192. Tamaki, S., S. Matsuo, et al. (2007). "Hd3a protein is a mobile flowering signal in rice." Science 316(5827): 1033.

193. Teper-Bamnolker, P. and A. Samach (2005). "The Flowering Integrator FT Regulates SEPALLATA3 and FRUITFULL Accumulation in Arabidopsis Leaves." Plant Cell 17(10): 2661-2675.

194. Tonsor S.J, Alonso-Blanco A. and Koorneef M. (2004). Gene function beyond the single trait: natural variation, gene effects and evolutionary ecology in Arabidopsis thaliana. Plant, Cell and Environment 28, 2-20.

195. Turck F., Fornara F and Coupland G (2008) Regulation and identity of florigen: Flowering locus T moves center stage Annual Reviews of Plant Biology Vol. 59.

196. Turner, A., J. Beales, et al. (2005). "The Pseudo-Response Regulator Ppd-H1 Provides Adaptation to Photoperiod in Barley." Science 310(5750): 1031-1034.

197. 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.

198. van de Pol, P. A. 1972. Meded. Landbouwhogesch. Wageningen 72(9):1-89

199. Wang H., Ma LG., Li JM., Zhao HY., Deng XW., 2001. Direct interaction of Arabidopsis cryptochromes with COP1 in light control development. Science, 294: 154-158.

200. Wang Qing, Sjja Uday et al. (2007). HUA2 caused natural variation in shoot morphology of Arabidopsis thaliana. Current Biology 17: 1513.

201. Wang, Z. –Y. and Tobin, E.M, 1998. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell 93: 1207-1217.

202. Wang, Z. Y., D. Kenigsbuch, et al. (1997). "A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene." Plant Cell 9(4): 491-507.

203. Weigel D., Nordborg M. (2005). Natural variation in Arabidopsis. How do we find the causal genes? Plant Physiology, 138:567-568.

204. Weigel, D. (1995). "The genetics of flower development: From floral induction to ovule morphogenesis." Annual Review of Genetics 29: 19-39.

205. Weigel, D. and E. M. Meyerowitz (1994). "The ABCs of Floral Homeotic Genes." Cell 78(2): 203-209.

206. Weigel, D., J. Alvarez, et al. (1992). "Leafy Controls Floral Meristem Identity in Arabidopsis." Cell 69(5): 843-859.

207. Weites R. and Simon R., (2000) Signaling cell fate in plant meristems. Three clubs on one tousle Cell. 103(6): 835

208. Wellmer F., Alves-Ferreira M., et al., (2006) Genome-wide analysis of gene expression during early Arabidopsis flower development PLoS Genet. 2(7):e117

209. Wenkel, S., F. Turck, et al. (2006). "CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis." Plant Cell 18(11): 2971-2984.

210. Werner Jonathan, Borevitz Justin et al. (2005) Quantitative trait locus mapping and DNA array hybridisation identify an FLM deletion as a cause for natural flowering-time variation. PNAS 102: 7.

211. West M., Kim K. (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics 175(3): 1441.

212. Wigge, P. A., M. C. Kim, et al. (2005). "Integration of spatial and temporal information during floral induction in Arabidopsis." Science 309(5737): 1056-1059.

213. Wu, X. L., D. Weigel, et al. (2002). "Signaling in plants by intercellular RNA and protein movement." Genes & Development 16(2): 151-158.

214. Yamaguchi Rumi, Nakamura Masanobu, Mochizuki Nobuyoshi, Kay A Steve and Nagatani Akira, 1999. Light-Dependent translocation of a Phytochrome B-GSPfusion protein to the nucleus in Transgenic Arabidopsis. The journal of Cell Biology, 145: 437-445.

215. Yang H-Q, Wu Y-J, Tang R-H, Liu D., Liu Y. et al., 2000. The C termini of Arabidopsis Cryptochromes mediate a constitutive light response. Cell, 103: 815-827.

216. Yanovsky, M. J. and S. A. Kay (2002). "Molecular basis of seasonal time measurement in Arabidopsis." Nature 419(6904): 308-12.

217. Yanovsky, M. J. and S. A. Kay (2003). "Living by the calendar: how plants know when to flower." Nature Reviews Molecular Cell Biology 4(4): 265-75.

218. Yoo, B. C., F. Kragler, et al. (2004). "A systemic small RNA signaling system in plants." Plant Cell 16(8): 1979-2000.

219. Yu H., Xu Y., et. al., (2002) AGAMOUS-LIKE 24, a dosage-dependent mediator of the flowering signals. PNAS 99(25): 16336

220. Zeevaart, J. A. D. (1976). "Physiology of flower formation." Ann. Rev. Plant Physiol. 27: 321-348.

APPENDIX

Genetic markers

www.arabidopsis.org/cereon/

www.naturalvariation.org/snp

www.naturalvariation.org/sfp

www.walnut.usc.edu/210.html

www.mpiz-koeln.mpg.de/masc/

Genetic tools

www.insertion.standford.edu/melt.html

www.patho.mgh.harvard.edu/ausebel-web

http://helix.wustl.edu/dcaps/dcaps.html

Microarray data

https://www.genevestigator.ethz.ch/

http://ipmb.sinica.edu.tw/affy/link.html

http://jsp.weigelworld.org/expviz/expviz.jsp

http://www.weigelworld.org/resources/microarray/microarray/

RIL collections

http://www.arabidopsis.org/abrc/catalog/recombinant_inbred_set_1.html

http://urgi.versailles.inra.fr/vnat/Fichier_collection/Rech_rils_pop.php

http://www.inra.fr/internet/Produits/vast/RILs.htm

STOCK SOLUTIONS

• **DEPC treated H_2O**
Add 1 ml fresh diethylpyrocarbonate (Sigma D-5758) to 1 liter distilled H_2O in fumehood. Shake well, leave overnight at room temperature and then autoclave for 15 min. This water is now RNAase free.

Probe labeling

• **Template DNA**
After restriction digestion, phenol/chloroform extract, ethanol precipitate and resuspend in DEPC treated H_2O at a concentration of 1 mg/ml. Store at -20°C.

• **5mM ATP, GTP, CTP**
Dilute 100 mM stocks (AP Biotech # 27-2025-01) with DEPC-treated H_2O. Store at -20°C.

• **1mM DIG-UTP**
Dilute Roche stock (#1-209-256) with DEPC-treated H_2O and store at -20°C. (With 10 mM stock, dilute 1:10 in a final volume of 250 µl)

• **10 x RNA polymerase buffer**
Roche #1-465-384. Store at -20°C.

• **RNA polymerases (T7, T3 & Sp6)**
Roche. Store at -20°C.

• **RNAse inhibitor**
RNAguard - AP Biotech # 27-0815-01. Store at -20°C.

• **MS X 1**

	20 ml
10mM Tris-HCl pH 7.5	100 µl 2M
10mM $MgCl_2$	200 µl 1M
50mM NaCI	666 µl 8.5%

Make 2M Tris by preparing 2M Tris-HCI and 2M Tris-base in DEPC-treated H_2O, and then mixing at a ratio of 8 acid : 2 base to get correct pH. Make 1M $MgCI_2$ arid 8.5% NaCI in distilled H_2O and then treat with DEPC as above.

• **tRNA**
Make tRNA (Sigma #R4251) at 100 mg/ml in DEPC treated H_2O. Store at -20°C.

• **DNAase 1**
Gibco-BRL #18068-015. Store at -20°C.

• **2 x carbonate buffer**
80 mM $NaHCO_3$ 0.672 g/100 ml
120 mM Na_2HCO_3 1.277 g/100 ml
It should be pH 10.2. Add DEPC to 0.1% and treat as above.

• **deionised formamide**
Sigma # F-9037, aliquot and store at -20°C.

• **DIG labelled control RNA**
Roche #1585.738. Store at -20°C.

Tissue pretreatment

• **Histoclear**
Fisher #H/0468/17

• **8.5% NaCI**

NaCI	85 g

in 1L of dH_2O. Treat with 0.1% DEPC, as above, before use.

• **10 X PBS**
(1.5M NaCl, 0.07M Na_2HPO_4, 0.03M NaH_2PO_4 pH 7)

NaCI	85 g
Na_2HPO_4	9.94 g
NaH_2PO_4	4.14 g

in 1L of dH_2O. Treat with 0.1% DEPC, as above, before use.

• **Paraformaldehyde**
Sigma #P6148.

• **10% GIycine**
10 g glycine in 100 ml of dH_2O. Treat with 0.1% DEPC as above and then store at 4°C.

• **Triethanolomine**
Sigma #T1377
Then throw into organic waste

• **Acetic anhydride**
BDH #100022M.

Hybridization

• **10 X salts**
(3M NaCI, 0.1M Tris-HCl pH 6.8, 0.1M $NaPO_4$ buffer, 50mM EDTA)

NaCl	175.3 g
2M Tris-CI pH 6.8	50 ml
NaH_2PO_4 . H_2O	6.9 g
Na_2HPO_4	7.1 g
0.5M EDTA	100 ml

Combine NaCl and phosphate salts and make up to 850 ml distilled H_2O. DEPC treat as normal and then after autoclaving add 100 ml DEPC-treated 0.5 M EDTA pH 8 and 50 ml 2M Tris pH 6.8 made up with DEPC treated H_2O (9.5 acid : 0.5 base).

• **tRNA**
Make tRNA (Sigma #R4251) to 100 mg/ml in DEPC treated H_2O

• **Deionised formamide**
Sigma # F-9037. Aliquot and store at -20°C.

• **50 x denhardt's**
Sigma #D-2532. Store at -20°C.

• **50% Dextran Sulphate**
AP Biotech #17-0340-02. Make up with DEPC-treated H_2O and heat in a boiling water bath to dissolve. Store at -20°C in aliquots.

Washes

• **20 X SSC**

(3M NaCI, 0.3M Na_3Citrate)

NaCl	175.3 g
Na Citrate	88.2 g

up to 1 L of dH_2O. Treat with 0.1% DEPC as above.

• **0.1 %SSC**

20 x SSC	5 ml
DEPC treated H_2O	995 ml

Antibody staining

• **10 x DIG Buffer 1**

1 M Tris	500 ml 2 M pH7.5
1.5 M NaCI	87.7 g

in 1 L of d H_2O

• **DIG Buffer 2** (make fresh each day)
0.5% (w/v) blocking reagent (Roche #109 6176) in DIG Buffer 1. Dissolve for ~1 hour at 60-70°C

• **DIG Buffer 3** (make fresh each day)
1% BSA (Roche 735 086 - Sigma A7906)
0.3% Triton-X-100 (Sigma T8787)
Dissolve in DIG Buffer 1

• **DIG Buffer 4** (make fresh each day)
1 in 3000 dilution of anti-digoxigenin-AP FAB fragments (Roche #1093-274) in DIG Buffer 3.

• **10 x DIG Buffer 5A**

1 M Tris pH 9.5	250 ml 2 M
1 M NaCl	100 ml 5 M

up top 500 ml with autoclaved distilled H_2O

• **10 x DIG Buffer 5B**
0.5 M $MgCl_2$ in autoclaved distilled H_2O.

• **4-nitro-blue-tetrazolium chloride (NBT) 75 mg/ml**
Roche #1087-479 made up in 50% dimethylformamide (Sigma # D4254), aliquoted and stored in the dark at -20°C.

• **5-bromo-4-chloro-indolylphosphate 4 toluidine salt (BCIP) 50 mg/ml**
Roche #1383-221 Store at -20°C.

• **Levamisole**
Sigma #L9756. Make up to 24 mg/ml in distilled H_2O and store at -20°C in aliquots.

Erklaerung

"Ich versichere, dass ich die von mir vorgelegte Dissertation selbstsändig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschliesslich Tabellen, Karten und Abbildungen-, die anderen Werken Im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keener anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden."

Köln, den Mai 2008

Teilpublikationen

Laurent Corbesier, Coral Vincent, Seonghoe Jang,Fabio Fornara, Qingzhi Fan, Iain Searle, Antonis Giakountis, Sara Farrona, Lionel Gissot, Colin Turnbull, George Coupland (2007). *"FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis"*. Science, Vol. 316. no. 5827, pp. 1030 – 1033.

Giakountis A., Mavromatis A., Arvanitogianis I., (2004). *"Allergens in food : What are they and how can be encountered with Biotechnology"*. Agricultural and Cattle Breeding, Vol. 5 pp : 58-64.

Mavromatis A., Giakountis A., Korkovelos A., Goulas C., (2004)**.** *"Analysis of the genetic relationship between local varieties of Phaseolus vulgaris with the use of RAPD molecular markers"*. Bio, Vol. 11 pp 42.

Sakellariou M., Grammatikaki E., Mavromatis C., Doxastaki M., Avgelis A., Korkovelos A., Giakountis A., Goulas A., (2004)**.** *"In vitro regeneration in banana and study of soma-clonal variation with the use of cytological and molecular methods"*. Bio, Vol. 11 pp 43.

Lebenslauf

Persönliche Informationen

Name	Giakountis
Vorname	Antonis
Geburtsort	Athen
Geburtstag	17. Juni 1980
Nationalität	Griechisch
Familienstand	Ledig

Ausbildung

1998	Abitur, 2nd Gymnasium und Lyceum Immitous, Athen.
1999-2004	Studium der Agriculture, Universität von Thessaly, Volos
2003-2004	Diplomarbeit "Molecular analysis of local varieties of bean with the use of RAPD markers and preparation of an antibody against the PvLHY protein that controls photoperiodism" Universität von Thessaly – National Zentrum für wissenschaftliche Forschung "Demokritos". Betreuer: Prof. Dr. Christos Goulas
September 2004	Beginn der Doktorarbeit am Max-Planck-Institut für Züchtungsforschung, Köln in Der Arbeitsgruppe von Prof. Dr. George Coupland.

Köln, den 29 April 2008