Flowering in ryegrass and conservation of the photoperiodic response

Milan Gagic

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Biological Sciences, University of Auckland, August 2007

ACKNOWLEDGEMENTS

I would like to thank the following people for the time, help, and support they have given me during my PhD work.

Firstly and primarily I would like to thank my supervisors Assoc. Prof Jo Putterill for her encouragement, help, support, and expedience which were most appreciated. Also many thanks go to my second supervisor Dr Igor Kardailsky who was always a well of new ideas and discussion topics off all kinds. His support was invaluable.

I am most indebted to Natasha Forester for her kind and generous support every time I needed it. Her help in the field and glasshouse and impeccable filing and organisational skills were sometimes the only thing that kept me sane.

I would like to thank to my office mates Shalome and Steve whose tolerance, patience and proof reading skills were irreplaceable.

Thanks to everyone in the Plant Biotechnology and Plant Breeding groups within AgResearch, Grasslands for their additional advice and support. Thanks to them I looked forward to every coffee break to discuss new and interesting current issues. It was a pleasure and privilege working with such a great group of people.

A special thanks goes to my parents, sister, and my niece Ana for all their support and encouragement, without which I would never have made it this far.

Finally all this would not have been possible without amazing support from my wife Dragana and my daughter Mina whose presence and encouragement was always with me.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AFLP	Amplified fragment length polymorphism
CCT	CO, CO-like, TOC1 domain
cDNA	Complementary DNA
cv.	Cultivar
dCAPS	Derived cleaved amplified polymorphic sequences
DD	Continuous darkness
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
EST	Expressed sequence tag
GA	Gibberellin
gDNA	Genomic DNA
GFP	Green fluorescent protein
Kb	Kilobase
LD	Long day
LL	Continuous light
m	Minutes
Mb	Megabases
mRNA	Messenger RNA

NIL	Nearly isogenic line
NLS	Nuclear localization sequences
PCR	Polymerase chain reaction
QTL	Quantitative trait loci
Pfr	Far-red light absorbing form of the photochromic protein
Pr	Red light absorbing form of the photochromic protein
RACE	Rapid amplification of cDNA ends
RFLP	Restriction fragment length polymorphism
rpm	Revolution per minute
RT-PCR	Real time PCR
rt-PCR	Reverse transcription PCR
s	Seconds
SAM	Shoot apical Meristem
SD	Short day
SNP	Short nucleotide polymorphism
SSR	Simple sequence repeat
UTR	Untranslated region
wt	Wild type
ZT	Zeitgeber time

ABSTRACT

Grasslands account for almost one quarter of the world's cover of vegetation. Almost three quarter of the world's milk, beef and veal are produced from temperate grasslands. In New Zealand, ryegrass (*Lolium perenne*) is the main pasture constituent with more than half of the total export revenue coming from grass-related products. Much of ryegrass production and quality depends on the timing of flowering through seasonal progression. In many plants, day length is the critical environmental parameter that controls when plants begin to flower. In *Arabidopsis* the *CONSTANS (CO)* gene mediates day length response. Upstream of *CO* is the *GIGANTEA (GI)* gene which is associated with the circadian clock mechanism and is required to promote *CO* expression. The *FT* gene is the immediate downstream genetic target of *CO* and is a direct promoter of flowering.

In this study, cDNA libraries, sequence alignment and genome walking were used to sequence and describe three putative orthologues from the ryegrass photoperiod pathway: *LpGI, LpCOL1*, and *LpFT3*. All three behaved in a true photoperiod manner characterised by cycling expression patterns under continuous light conditions and differential expression patterns in LD and SD conditions. Different photoperiods brought about differences in expression of these genes measured either by the phase shift change (*LpGI* and *LpCOL1*) or by the change of the transcript level (*LpFT3*). Gene expression changes over a vernalisaton time course were also analysed and results indicated that *LpFT3* acts as the flowering integrator. The role of *LpGI*, *LpCOL1*, and *LpFT3* as putative photoperiod genes was further confirmed by genetic mapping, which placed them on linkage groups 3, 6, and 7, respectively. The syntenic positions in rice contain major heading date quantitative trait loci. The function of *LpFT3* was examined by over-expressing the gene in *Arabidopsis* under control of the cauliflower mosiac virus (CaMV) 35S promoter. Substantially higher expression of the endogenous *Arabidopsis AtFT* transcript was observed in the mutated *ft-1* line overexpressing *LpFT3,* suggesting a positive feedback loop either directly or through upstream intermediaries. Overexpression of the *LpGI* and *LpFT3* genes restored rapid flowering to the respective *gi-3* and *ft-1 Arabidopsis* mutants while overexpression of *LpCOL1* did not

accelerate flowering either in *co-2* or wild type *Arabidopsis* plants. However, overexpression of *LpCOL1* completely restored the late flowering phenotype of the *gi-3* mutant indicating the existence of another important link outside the well established hierarchy of *GI-CO-FT* in the photoperiod pathway.

This study revealed that the ryegrass photoperiod pathway genes show high similarity to their wheat, rice and *Arabidopsis* counterparts. Exploring ways to modulate flowering time in ryegrass could provide major benefits to the agricultural industry by increasing forage quality, controlling seed and pollen production, and addressing potential problems linked with climate change.

1 INTRODUCTION

1.1 General introduction

Flowering has long been a subject of intense study in plant biology due to its biological importance and agronomic impact. Plants have evolved multiple genetic pathways to regulate flowering time, which is influenced by both environmental factors such as day length and temperature and developmental factors associated with the age of the plant (Putterill et al, 2004; Corbesier and Coupland, 2006). Flower initiation is a key step of the reproductive process which precedes the development of fruit and seeds. The transition to flowering is defined as the conversion of the shoot meristem from vegetative to reproductive development. The shoot meristem is a group of cells formed during embryo development. It is located at the apex of the shoot where the identity of the shoot meristem is altered as the shoot develops. This becomes apparent when different organs start to be formed at distinct times. The genes and environmental conditions that control flowering time influence the change in identity of cells of the shoot meristem, which occurs during the switch from vegetative growth to flowering.

The genetics of flowering time has been studied extensively, often by identifying mutants or by following the phenotypes of the progeny of crosses between different varieties. Systematic attempts to understand the genetics and molecular biology of flowering time have been made in several plant species, with the most work done on *Arabidopsis thaliana* (*Arabidopsis*) and *Oriza sativa* (rice). Studies in *Arabidopsis* have led to the identification of components within individual signaling pathways that affect flowering, and to their positioning within regulatory hierarchies (Fig 1.1). To ensure reproductive success, flowering is controlled so that it occurs

in the optimal environmental conditions for seed production. In temperate species, environmental cues exhibiting seasonal variations such as photoperiod, temperature, and the developmental stage of the plant are important factors regulating the floral transition. All these factors are not perceived by the same plant organs. Daylength is perceived by the mature leaves and cold by both, young leaves and the shoot apex (Bernier, 1998; Searle et al., 2006).

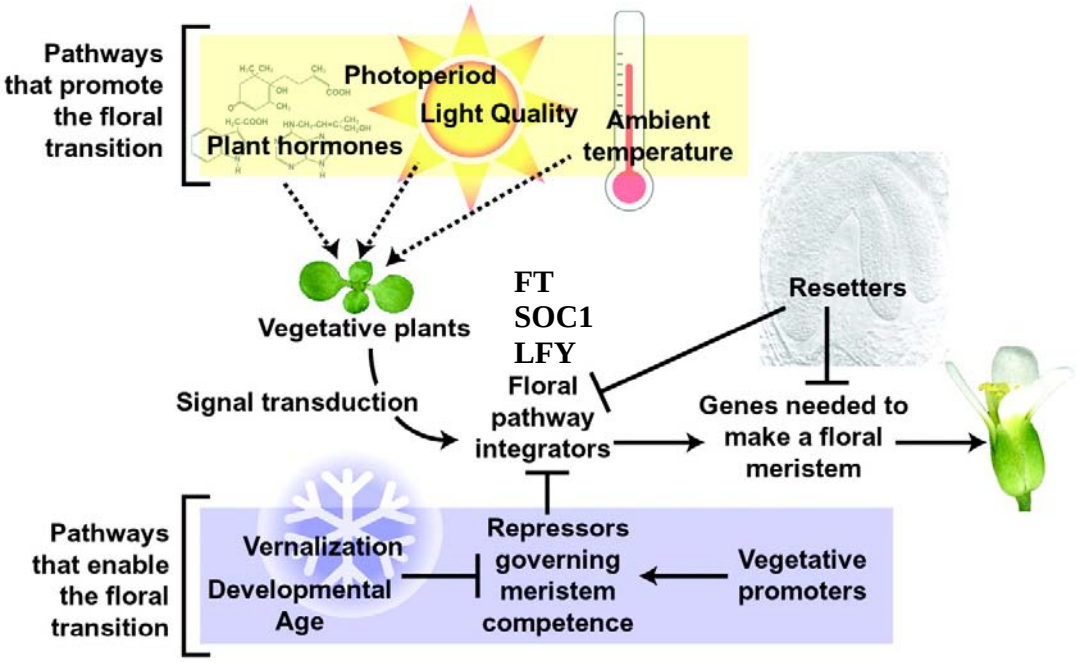


Figure 1.1 Pathways that enable or promote the floral transition (modified from Boss et al., 2004). Resetters, in this context, represent floral repressors whose mutant phenotypes are early flowering. Their alternative function might be to "reset" global expression states assuming that flowering is the default developmental pathway.

This thesis presents work carried out to isolate and characterize three crucial genes within the ryegrass photoperiod pathway, namely *GIGANTEA* (*GI*), *CONSTANS* (*CO*), and *FLOWERING TIME* (*FT*). In the introduction chapter several aspects of flowering in different model species will be discussed. To begin with environmental cues responsible for the initiation of flowering in *Arabidopsis* will be described followed by the analysis of genes from the photoperiod pathway and the molecular mechanism responsible for the flowering process. Next, conservation of the photoperiod pathway in rice and cereals will be discussed. In depth analysis of the ryegrass research so far will follow, focusing on issues such as

phylogeny and morphology, methodologies in ryegrass improvement, genetic resources, and molecular biology of flowering. This chapter will be concluded with the specific aims of this work.

1.2 Control of flowering by environmental cues in *Arabidopsis*

1.2.1 Vernalisation

Changes in daylength are a reliable indicator of seasonal progression, but daylength itself is not completely informative of the time of year. Some plants discriminate between equivalent photoperiods in the beginning of autumn and springs by flowering under long-day conditions only when these are preceded by a prolonged exposure to cold, a phenomenon that is known as vernalisation (reviewed in Amasino et al., 2005). The mechanism by which it is achieved involves suppressing the expression of genes that encode repressors of flowering. The suppression is an epigenetic event consisting of chromatin modification, rendering mitotically stable cells, even after the inducing signal (cold) is no longer present. Grafting vernalised and non-vernalised shoot apicies to non-vernalised and vernalised stocks, respectively, has demonstrated that the shoot apex is the site of cold perception. The vernalisation response can be facultative or obligate. Winter annuals have a facultative vernalisation response which means that cold exposure is not required for flowering but flowering will occur more rapidly after cold treatment. Biennials, in contrast, have an obligate requirement for cold treatment and thus cannot flower without prior cold exposure. These varieties differ at two loci, *FLC* and *FRI.* Dominant alleles at these loci in the winter annuals confer a vernalisation requirement (reviewed in Boss et al., 2004; Putterill et al., 2004). *FRI* is a key activator of *FLC* expression and the product of *FRI* gene increases *FLC* mRNA abundance. It represses flowering by upregulating *FLC* RNA levels (reviewed in Putterill et al., 2004; Searle and Coupland, 2004). Consistent with this, loss of *FLC* function eliminates the ability of *FRI* to delay flowering. *FLC* mRNA levels are downregulated by vernalisation. Most early-flowering varieties of *Arabidopsis* carry deletions or loss of function mutations that disrupt the open reading frame of *FRI* (Johanson et al., 2000; Shindo et al., 2005). It has been shown previously that early-flowering accessions have evolved on multiple independent occasions through loss of *FRI* function (Johanson et al., 2000; Gazzani et al., 2003). The *FLC* gene has also been shown to contribute to the natural variation of flowering time through a transposon

insertion in the first intron which modifies its function (Gazzani et al., 2003). In *Arabidopsis*, allelic variation at *FLC* and *FRI* contributes to natural variation in vernalisation requirement, with weak alleles leading to rapid-cycling habit seen in laboratory varieties of *Arabidopsis* such as Landsberg (L*er*) and Columbia (Col; Michaels et al., 2003i; Johanson et al., 2000).

The multiple pathways that regulate the floral pathway integrators in *Arabidopsis* are grouped into the promotion, enabling, and resetting pathways (Fig1.1; Boss et al., 2004). Enabling pathways determine the activity of the repressors of the floral integrators. Floral integrators such as *FT*, *SUPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), and *LEAFY* (*LFY*) cause floral transition when their expression level reaches certain threshold. The expression of the floral repressor *FLC* is regulated by several independent pathways. *FLC* is upregulated by a number of genes including *FRIGIDA* (*FRI*), and is downregulated by vernalization. *FLC* is a MADS-box transcriptional repressor, expressed predominantly in shoot and root apices and vasculature that quantitatively represses flowering by repressing floral pathway integrators (reviewed in Putterill et al., 2004; Amasino et al., 2005; Corbesier and Coupland, 2006). This mechanism involves repression of *FT* through direct binding, preventing the activation of *SOC1* at the shoot apical meristem (SAM; Searle et al., 2006). In addition *FLC* binds to the promoter regions of *SOC1* and *FD* and reducing their expression in meristem (Searle et al., 2006; Hepworth et al., 2002). This shows that *FLC* mediated repression involves both leaves and SAM as the repression sites. *FRI* is a key activator of *FLC* expression. It represses flowering by upregulating *FLC* RNA levels (reviewed in Putterill et al., 2004; Searle and Coupland, 2004). Consistent with this, loss of *FLC* function eliminates the ability of *FRI* to delay flowering. *FLC* mRNA levels are downregulated by vernalisation.

The vernalisation response is quantitative, with increasing durations of cold leading to progressively accelerated flowering once the plants return to ambient temperature. This process is also visible in *FLC* expression where longer periods of cold exposure lead to progressively lower expression of *FLC* mRNA (Michaels et al., 1999; Sheldon et al., 1999). Once acquired, the vernalised state is remembered by the plant during subsequent mitotic proliferation, but is reset in next generation which suggests an epigenetic basis. Insight into the molecular basis of the epigenetic repression came from studying vernalisation response in *Arabidopsis*.

Discovery of two other proteins implicated in the control of flowering time by vernalisation indicated that series of histone modifications are the main mechanism of *FLC* repression of flowering. *VERNALISATION1* (*VRN1*) and *VRN2* (Levy et al., 2002; Gendall et al., 2001) are required for the maintenance of low levels of *FLC* mRNAs established by cold treatment, once the plants are exposed to warmer conditions. *VRN1* encodes a plant specific DNA-binding protein, while *VRN2* is a transcriptional repressor, a member of the Polycomb group (PcG) with the gene members known to act as histone metyltransferases that catalyse covalent modifications of the amino terminal tails of histones (Kouzarides, 2002). It has been shown that dimethylation of K9 on histone H3 is crucial for the maintenance of stable *FLC* repression (Bastow et al., 2004) and it is possible that *VRN2*, given its properties, plays important role in this. Contrary to this, new work of Sheldon et al. (2006) suggest different role for *VRN1* and *VRN2*. They showed that *VRN1* and *VRN2* activity are needed for maximal reduction in *FLC* activity. With these proteins inactive the vernalisation response is reduced in magnitude with *VRN2* having a greater effect than *VRN1*. They also showed that neither VRN1 nor VRN2 are required for the maintenance of the repressed FLC state during growth and development of the plant subsequent to the cold exposure.

The molecular basis of the vernalisation response have been further characterised by the discovery of VERNALISATION INSENSITIVE 3 (VIN3) protein (Sung and Amasino, 2004) containing a PHD finger domain, which can act in association with the polycomb group protein homologues VRN2, FIE, SWINGER, and CURLY LEAF in a polycomb repression complex 2 (PRC-2) which is required for the repression of *FLC* by vernalisation (Wood et al., 2006; Sung et al., 2006). *VIN3* has been identified as the most upstream component of vernalisation process and most likely responsible for the establishment of the vernalisation-mediated repression of *FLC* in the shoot and root apex since the *FLC* repression does not occur until *VIN3* is induced. These results were further confirmed in wheat by characterising 3 VIN3-like (VIL) proteins (Fu et al., 2006). They all carry three conserved domains and transcription is up-regulated by vernalisation but they are also affected by photoperiod.

Parallel to the vernalisation there is the autonomous pathway which represses FLC expression and in the absence of the vernalisation input, this pathway is major regulator of *FLC* levels (reviewed in Simpson and Dean, 2002; Putterill et al., 2004; Boss et al., 2004; Simpson, 2004). Mutants in the autonomous pathway are late-flowering because of increased levels of

FLC mRNA, and this late-flowering is vernalisation responsive. All members of this pathway act to limit *FLC* expression but genetic analysis revealed that they have different functions. Expression of the genes from the autonomous pathway is not regulated by any other flowering time gene. There is also no evidence that photoperiod or low temperature regulate any of the genes. The input signals that regulate the autonomous pathway still remain unknown. Most of the autonomous genes encode proteins that contain plant specific RNA/DNA recognition motifs or encode factors that regulate *FLC* epigenetically by chromatin assembly, modification or remodeling. Some of the genes are required for the histone deacetylation of *FLC* chromatin which is the process that has already been associated with reduced levels of gene expression. Some features of this pathway may also be environmentally regulated. Recent findings indicate that that the control of flowering by ambient temperature may involve an *FLC* independent function of the autonomous pathway.

1.2.2 Gibberellin pathway

All currently known phytohormones (gibberellins, auxin, cytokinins, and brassinosteroids) have at some point been associated with flowering time control (Willson et al., 1992). Flowering of *Arabidopsis* is promoted by gibberellins (GAs), especially under short days (Chandler and Dean, 1994), and that is the case with many plant species (eg. tobacco, ryegrass). Mutants that disrupt either GA biosynthesis or signaling show alterations in flowering time. The mutant *ga1-3* fails to flower under SD and shows a slight delay in flowering under LD. The expression of *LFY* from a constitutive promoter can rescue the flowering defect of *ga1-3* mutant which suggest that the failure of *ga1-3* mutant to flower in SD is caused by the failure to upregulate *LFY* (Blazquez et al., 1998), through cis elements that are different from those that are sufficient for the daylength response (Blazquez and Weigel, 2000).

Recently Eriksson et al. (2006) have shown that gibberellin (GA_4) is the active component in the regulation of *LFY* transcription in *Arabidopsis*. This process is not mediated through PHYB, another negative regulator of the *LFY* promoter, since *phyB* mutation does not enhance the response of the *LFY* promoter to GAs (Blazquez and Weigel, 1999). Compared with SD, the flowering defect of *ga1-3* mutants is relatively minor in LD. Therefore GA is absolutely required for flowering under SD conditions in *Arabidopsis*. Also the dominant

gibberellic acid insensitive-1 (*gai-1*) mutant flowers extremely late under SD, and the flowering phenotype is not rescued by the exogenous treatment of GA (Wilson et al., 1992). On the other hand the mutant *spindly* (*spy*), which causes constitutively active GA signaling, flowers early under both, LD and SD days (Jacobsen and Olszewski, 1993).

It has also been shown that GA positively regulates *SOC1* expression under short days but does not have any effect on *FLC* and *FT* (Moon et al., 2003). According to this the GA pathway is the most important flower-promoting pathway under SD condition, and that removal of *FLC* is prerequisite but is not sufficient for flowering. Contrary to these studies, King et al. (2001) showed that in *Lolium temulentum* GA may serve as LD flowering signal because florally inductive LD leads to a large increase in GA_1 and GA_4. Another candidate gene relevant to the GA signal transduction pathway is the GAMYB transcription factor which acts as a transcriptional activator in the cereal aleurone (Gubler et al., 1995), and *L. temulentum* shoot apex (Gocal et al., 1999).

GA response has also been linked with *FLOWERING PROMOTING FACTOR 1* (*FPF1*, Kania et al., 1997). Overexpression of *FPF1* in *Arabidopsis* produced transgenic plants whose phenotype resembled that of wild type plants treated with GAs. The same plants treated with GA inhibitor paclobutrazol showed a typical dwarf phenotype of plants lacking GA indicating that *FPF1* in involved in GA-dependent signaling pathway by modulating GA response during the transition to flowering.

1.2.3 Light quality pathway

The primary photosensory receptors of higher plants involved in flowering are the red/far-red light receptors called phytochromes and the blue/UV-A light receptors called cryptochromes (Lin, 2000). Blue light (400-500nm) and red light (600-700nm) are the two spectra of solar radiation that are most effectively absorbed and utilized by the photosynthetic system of plants. Light affects flowering time independently of photoperiod via a light quality pathway (Simpson and Dean, 2002). Light quality (the relative intensities of the various wavelengths that reach the plant) is important for monitoring local environment. A good example of light quality effect is the shade avoidance response. Light reflected from neighboring vegetation exhibits a reduced red/far-red ratio due to absorption of red light by chlorophyll. Far-red enriched light serves as a signal of competition and results in accelerated flowering in a

crowded environment. Phytochromes are photochromic proteins that exist in two photo-interconvertible isomeric forms: the red-light absorbing form (Pr) and the far red-light-absorbing form (Pfr; Hughes, 1999). In *Arabidopsis* the red light and far-red light region of the spectrum is perceived by a small family of chromoproteins, encoded by 5 genes; *PHYTOCHROME (PHY) A, B, C, D* and *E* (reviewed in Quail, 2002). Analysis of their mutants indicate that the light-stable PHYB, PHYD, and PHYE mediate responses to red light, whereas the light-labile PHYA is the photoreceptor that discriminates far-red light from darkness. The *Arabidopsis phyA* mutants flower later than wild-type plants in LD or quasi-LD conditions with either night-breaks or day extensions (reviewed in Neff and Chory, 1998). Consistent with this, transgenic *Arabidopsis* plants overexpressing *PHYA* flowered earlier than the wild type both in SD and quasi-LD conditions (Bagnall et al., 1995). The *phyA* mutant of pea, also showed a phenotype similar to that of the *Arabidopsis phyA* mutant. Interestingly, the pea *phyA* mutant accumulated a graft transmissible inhibitor that could delay the flowering of the grafted recipient plants, suggesting that PHYA signaling may influence the production of a floral suppressor (Weller et al., 1997). *PHYB* plays an inhibitory role in floral initiation by modulating *FT* expression through the PHYTOCHROME AND FLOWERING TIME 1 (PFT1) nuclear protein (Cerdan and Chory, 2003). The *Arabidopsis phyB* mutant flowered earlier than the wild type in both LD and SD conditions (Mockler et al., 1999). Remarkably, the early flowering phenotype of *phyB* was completely abolished when plants were grown at the slightly lower temperature of 16°C, suggesting that the early flowering of *phyB* is temperature dependent (Halliday et al., 2003). PHYB inhibits flowering in both LD and SD conditions but appears more apparent in the photoperiod that suppresses flowering in the respective plants. In *Arabidopsis* PHYC is functionally similar to the red light sensing PHYB and loss of function leads to early flowering in non-inductive condition (Monte et al., 2003). Recent findings also indicate that PHYC in *Arabidopsis* may be responsible for natural variation in flowering and growth response (Balasubramanian et al., 2006). In *Arabidopsis phyD* mutation is identified as a naturally occurring allele of the wild type Wasilewskija ecotype which encoded no functional PHYD protein (Aukerman et al., 1997). When introgressed into various genetic backgrounds, monogenic *phyD* mutant plants had no obvious phenotypic abnormality, whereas plants impaired in both the *phyB* and *phyD* genes flowered earlier than the *phyB* monogenic mutation in both LD and SD conditions

(Aukerman et al., 1997). This indicated that, like PHYB, PHYD inhibits flowering. Similarly the *phyE* mutant showed no phenotypic alteration unless it was in the *phyB* background. This indicated the function of PHYE is also similar to that of PHYB (Devlin et al., 1999). Another evidence of light-quality pathway comes from work on PFT1, a nuclear protein that acts in PHYB pathway and induces flowering in response to suboptimal light conditions (Cerdan and Chory, 2003). *pft1* was late flowering and completely suppressed the early–flowering phenotype of *phyB* in both LD and SD conditions, suggesting that PFT1 is essential for PHYB regulation of flowering time. In addition, the lack of significant correlation between *CO* and *SOC1* mRNA levels with *phyB* and *pft1* flowering times suggests that PHYB regulates *FT* mRNA levels, possibly by *PFT1* dependant mechanism that does not involve changes in *SOC1* or *CO* mRNA levels. It was also shown that the *FT* expression in *phyB* and *phyAphyBphyD* mutants was 20-fold, and 40-fold higher respectively when grown at 22ºC which correlated with *phyB* and *phyAphyBphyD* flowering phenotypes (Halliday et al., 2003).

Cryptochromes are another class of photosensory receptors responsible for perception of the blue/UV-A light. They are soluble flavoproteins identified originally for their role in blue-light-dependent inhibition of stem growth in *Arabidopsis*, in which they are encoded by two genes, *CRYPTOCHROME1* (*CRY1*) and *CRY2* (Ahmad and Cashmore, 1993; Lin et al., 1998). *cry2* mutants in *Arabidopsis* flower much later than wild-type plants during long but not short days which is an indication of the importance of *CRY2* in photoperiodic regulation of flowering time (Guo et al., 1998). The effect of *CRY2* under white light conditions requires the presence of the active *PHYB* (Mockler et al., 1999) with which it associates physically in the nucleus (Mas et al., 2000). In blue light, *CRY2* regulates flowering time in part redundantly with *CRY1* and *PHYA* (Mockler et al., 2003) indicating that *CRY2* and *PHYA* might be the principal photoperiodic photoreceptors, discriminating day from night in *Arabidopsis*. Recently another cryptochrome, CRY3 have been described and its crystal structure determined (Klar et al., 2007). It binds to DNA and carries an N-terminal sequence which mediates import into chloroplast and mitochondria.

Another group of blue-light photosensory receptors called phototropins regulate range of responses in order to optimize the photosynthetic efficiency of plants and promote growth (reviewed in Christie, 2007). *Arabidopsis* contains two phototropins (PHOT1 and PHOT2)

which show partially overlapping roles in regulating hypocotyl phototropism in response to high intensities to unilateral blue light.

Besides detection and transduction of the light signal, as described above, the ability to sense the changing photoperiod requires the second component, an endogenous timer and genes integrating these different signals. The next sections give an overview of the photoperiod pathway.

1.2.4 Photoperiod pathway

One of the most important factors controlling flowering time in temperate regions is the duration of the daily light period, or photoperiod. Photoperiodism is a biological response to the length and timing of day and night making it possible for an event to occur at a particular time of year (reviewed in Taiz and Zeiger, 2002). It was first described in detail by Garner and Allard through the demonstration that many plants flower in response to changes in daylength (Garner and Allard, 1920). They showed that some plant species flower when daylight falls below a critical daylength, whereas other plants show accelerated flowering in response to daylengths longer than a critical daylength. These plants are called short-day (SD), and long-day (LD) plants, respectively. Some of the SD plants are rice, and morning glory (*Pharbitis nil*), while examples of LD plants are *Arabidopsis*, ryegrass, and *Pisum sativum* (pea). Some species are day-neutral, i.e. day length has no influence on their flowering time (tomato and some *Nicotiana* species).

The photoperiodic control of flowering is brought about by the interactions of genes involved in the developmental control of floral initiation, circadian clock regulated flowering time genes, and the signal transduction of photoreceptors (Putterill et al., 2004; Blázqez, 2005). The alteration of flowering time in response to day length is mediated by complex interactions between environmental signals and the time keeping mechanism that is associated with the circadian clock. Erwin Bünning first proposed that the photoperiodic time-keeping mechanism is associated with the circadian clock (Bünning, 1936), an autonomous mechanism that generates biological rhythms with a period of approximately 24 h (reviewed in Hotta et al., 2007). The circadian system is often divided into three general parts (Dunlap, 1999): the central oscillator is the core of the system and generates the 24-hour rhythm; the oscillator is synchronized or entrained to daily cycles of night and day through light and

temperature signaling pathways; output pathways are controlled by the oscillator, which regulates a wide range of biochemical and developmental pathways including flowering. It has been known for some time that the total duration of light or darkness was not the critical factor determining the response to that environmental factor (reviewed in Roden et al., 2002). Instead photoperiodic time measurement may involve a circadian rhythm of responsiveness to light, known as the photoperiodic response rhythm. In order to perceive and respond to changes in photoperiod plants must detect light duration and couple this to the circadian clock (reviewed by Thomas and Vince-Prue, 1997). In *Arabidopsis*, light perceived by the red/far-red absorbing phytochromes and the UV/blue light absorbing cryptochromes act to reset the oscillator so that it remains in synchrony with its environment (Somers et al., 1999). Three of the proteins thought to be constituents of the central oscillator are CIRCADIAN CLOCK ASSOCIATED1 (CCA1), LATE ELONGATED HYPOCOTYL (LHY), and TIMING OF CAB EXPRESSION (TOC1) (Alabadi et al., 2001). Mutations in any of these gene shorten the circadian period, diminishing effect of daylength differences and in general the mutants are late flowering under LD conditions and early flowering in SD conditions (Schaffer et al., 1998; Somers et al., 1998; Strayer et al., 2000; Mizoguchi et al., 2002). A model in which these three proteins function to produce an autoregulatory transcriptional and translational negative feedback loop was proposed by Alabadi et al. (2001). According to this model *LHY* and *CCA1* bind evening elements within the promoter region of *TOC1* to negatively regulate its expression.

Members of *PSEUDO-RESPONSE REGULATOR* (*PRR*) gene family have also been implicated in the circadian clock associated functions and shown to redundantly regulate *CO* expression under LD (Nakamichi et al., 2007). Genetic studies suggested that *PRR5*, *PRR7*, and *PRR9* genes positively regulate *CO* expression through repression of *CYCLING DOF FACTOR 1* (*CDF1*) in late daytime. It was proposed that *LHY/CCA1* and *PPR5/PPR7/PPR9* act antagonistically by negatively regulating expression of *GI* and repressing transcription of *CDF1* respectively.

Two possible mechanisms have been proposed, by which a circadian clock might mediate perception of photoperiod (Pittendrigh and Minnis, 1964). According to the external coincidence model, a photoperiodic response may be induced when an external signal (light) coincides with a photoinducible phase of the circadian cycle (Fig. 1.2A). The photoinducible

phase may be determined by the diurnal oscillation of a key regulator. An alternative hypothesis, known as the internal coincidence model, suggests that inductive photoperiods may drive expression of two endogenous rhythms to a more favorable phase relationship (Fig. 1.2 B). As described below, the data fits the external coincidence model best so far.

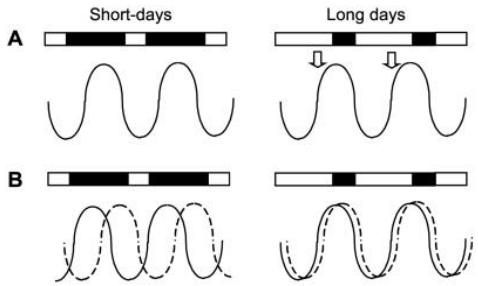


Figure 1.2 Coincidence models. **A**. External coincidence model proposes that light coincides with the light inducible phase which is regulated by an endogenous oscillator; **B**. Internal coincidence model suggests that inductive photoperiods may bring expression of the two endogenous rhythms to a more favorable relationship (from Roden et. al., 2002)

Arabidopsis is a facultative long day plant, which flowers earlier under long days but eventually flowers under short days. Under laboratory conditions, *Arabidopsis* will flower in response to a single long day (Corbesier et al., 1996). Molecular and genetic approaches have identified genes required for the daylength response, and some of these encode regulatory proteins specifically involved in the regulation of flowering, while others encode components of light signal transduction pathway or are involved in circadian clock function. A conserved pathway of regulatory proteins that induce flowering in response to day length has been described in *Arabidopsis*, and *GI*, *CO*, and its target gene, the floral integrator *FT* play central roles in this pathway.

These proteins have been defined by the mutations that modify day length-response by delaying flowering under photoinductive periods. Based on their mutant phenotypes (late-flowering in long days), transcriptional hierarchy, effects on CO protein, and the genetic interactions between the mutations they have been placed in the same genetic pathway (Searle and Coupland, 2004) forming a single epistatic group (Fig 1.3)

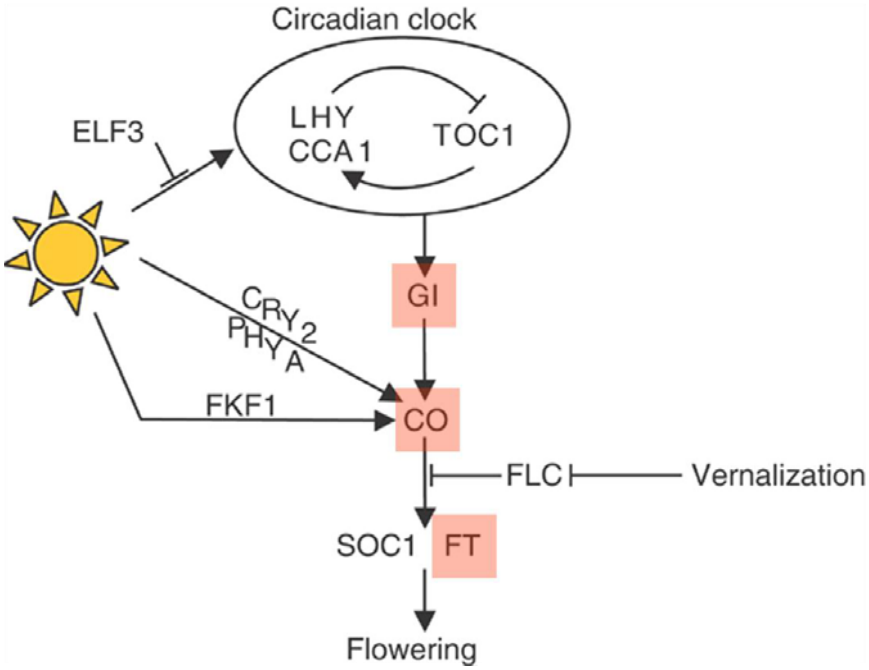


Figure 1.3 A genetic pathway that controls flowering in response to photoperiod in *Arabidopsis*. Arrows between genes represent promotive effects, whereas perpendicular lines represent repressive effects (from Searle and Coupland, 2004)

1.3 Genes in the photoperiod pathway

1.3.1 Different roles of the *GIGANTEA* gene

The *Arabidopsis GI* gene is important in regulating photoperiodic flowering and controlling circadian rhythms (Park et al., 1999; Fowler et al., 1999; Mizoguchi et al., 2005; Cao et al., 2005; Paltiel et al., 2006; Gould et al., 2006) and phytochrome signaling (Huq et al., 2000). It has also been implicated in resistance to paraquat (Kurepa et al., 1998), length of circadian rhythms (Park et al., 1999), and accumulation of starch in the leaves during photoperiod (Eimert et al., 1995). It encodes a protein which is nuclear localized (Huq et al., 2000; Mizoguchi et al., 2005) and does not show homology with genes of known function (Park et al., 1999; Fowler et al., 1999). It is plant specific, found as a single copy gene in angiosperms including monocots such as rice and in the gymnosperm loblolly pine but not in the moss *Physcomitrella* (Mittag et al., 2005). *GI* was shown to act as a promoter of flowering in *Arabidopsis* in LD but *gi* mutations had little or no effect on flowering under SD (Fowler et al., 1999). By contrast the *GI* ortholog in rice, *OsGI*, acts as a suppressor of flowering in LD

and promotes flowering in SD (Hayama et al., 2003). The *GI* transcript was detected at all stages of plant development tested, from seedlings at the two leaf stage to mature plants with developed siliques (Fowler et al., 1999). The *Arabidopsis GIGANTEA* gene appears to act upstream of *CO*. In the *gi* mutants the *CO* gene is expressed at very low levels and the late-flowering phenotype is corrected by ectopic overexpression of the *CO* gene (Suarez-Lopez et al., 2001). However, some *gi* mutants (*gi-1*, *gi-100*, and *gi-200*) caused additional defects including circadian rhythms to cycle faster under constant conditions, impairing red-light signaling from PHYB, and decreased photomorphogenesis in blue light respectively (Park et. al., 1999; Huq et. al., 2000; Martin-Tryon et al., 2007). A yeast two-hybrid system showed that *SPINDLY* (*SPY*), a negative regulator of gibberelin signaling, and *GI* can interact physically in yeast and that they may function in common pathways affecting hypocotyl elongation and photoperiodic response (Tseng et al., 2004).

The *GI* expression experiments in *Arabidopsis* grown in SD and LD showed that in SD, *GI* expression peaked ~8 h after dawn and reached trough levels ~3 h later, which was ~1 h after the transition to darkness. In LD, peak levels of *GI* expression were slightly lower and occurred 2 h later than in SD (~10 h after dawn) and reached trough levels later (~6 h later) before the transition to darkness (Fowler et. al., 1999).

Arabidopsis GI protein is post-transcriptionaly regulated accumulating during the day and declining at night (David et al., 2006). In addition, 35S::GI seems to be sensitive to day length. Exposing SD plants to an extended day resulted in stable GI accumulation rather than the decline usually observed at night in SD. Extending the light period of a SD with blue or red light produced the same result suggesting that accumulation of GI protein is not dependent solely on blue or red light photoreceptors (David et al., 2006). They also showed that the presence of proteasome inhibitors had major effect on GI accumulation in SD plants at ZT16 (ZT-zeitgeber time, number of hours after dawn) whereas in the absence of the proteasome inhibitors GI levels from the plants taken at the same time were significantly reduced (David et al., 2006). Taken together, these results strongly support the notion that GI protein levels respond directly to light/dark transitions, increasing in light and decreasing in dark via dark induced proteolysis by the 26S proteasome.

Of the three components of a circadian system (an input pathway(s), central oscillator, and output pathways; Dunlap, 1999), *GI* is unlikely to be a central oscillator component because

the putative null mutation does not abolish rhythmicity but alters period and reduces amplitude (Park et. al., 1999). In wild-type (wt) *Arabidopsis*, the free running period of a circadian clock lengthens with decreasing light intensity in continuous light (LL) (Somers et. al., 1998). The rate of period length increase with decreasing light is less in *gi* null mutant than in the wild type which implies that the *GI* increases sensitivity of the circadian system to the controlling effects of light on period length. Circadian expression pattern of the *GI* gene in continuous darkness (DD), after entrainment in long day, showed less severe effect on the amplitude and sustainability than under LL (Park et. al., 1999). This light dependent conditional effect of the *gi* mutations supports the notion that *GI* functions in a light input pathway. It has been suggested that this pathway represents a function of *GI* that is largely independent of its role in circadian clock function (Fowler et. al., 1999; Park et. al., 1999; Martin-Tryon et al., 2007). The *GI* transcript levels in *Arabidopsis* plants entrained in LD and transferred to either LL or DD, continued to cycle in a similar phase, indicating that they were controlled by the circadian clock.

In recent work Mizoguchi et al. (2005) showed that the effect of *GI* on the circadian rhythm is not only due to its role in light signaling since constitutive *GI* expression (*35S::GI*) and *gi-3* altered circadian rhythms under continuous light as well as under continuous darkness. Under diurnal cycles *35S::GI* delayed the phase of the expression of circadian clock-controlled genes *CCR2* and *LHY*, whereas *gi-3* delayed the phase of *CCR2* and reduced the amplitude of *LHY* expression. The effect of *35S::GI* and *gi-3* on the timing and amplitude of expression of the flowering time genes *CO* and *FT* is much more dramatic than on the expression of other clock-controlled genes (Mizoguchi et al., 2005). It is proposed that *GI* plays a significant role in controlling a subset of circadian rhythms in light and dark with an effect on phase in diurnal cycles but that its effect on flowering is distinct from its function in regulating these circadian rhythms. The epistasis of *gi* to the early flowering phenotype of *lhy-11 cca1-1* suggests that *GI* is part of the circadian clock and important for circadian clock-controlled flowering. The effect of *35S::GI* and *lhy-11 cca1-1* on flowering was partially suppressed by *co-2* and *ft-1* mutations, supporting the idea that the mechanism by which *GI* promotes early flowering includes *CO* and its target gene *FT*. *GI* also appears to promote flowering by the second mechanism that is independent of *CO* and *FT*. The delay in flowering of *lhy-11 cca1-1* caused by *co-2* and *ft-1* was weaker than that caused by *gi-3*, suggesting that as well as

promoting flowering by activating *CO* and *FT*, *GI* promotes flowering independently of these genes. Similarly, *co-2* and *ft-1* only partially suppressed the early flowering of *35S::GI* plants. It has been proposed that in wild type plants, *GI* regulates at least two circadian clock-controlled output pathways that promote flowering, one that includes *CO*, *FT*, and *SOC1*, and a second that promotes flowering independently of these genes (Mizoguchi et al., 2005). The hypothesis of distinctive *GI* roles was further supported by characterization of a missense allele *gi-200* which revealed that the *GI* action in the circadian clock is biochemically distinct from its regulation of the flowering pathway (Martin-Tryon et al., 2007). Although the expression pattern of *gi-200* mutants is similar to *gi-1* and *gi-3*, *gi-200* plants flower normally in photoinductive conditions due to the levels of *CO* transcript similar to wild type. At the same time *gi-200* exhibit a phase shift in *CO* expression since the pace of the clock is increased which in turn cause early flowering in SD. Recently, *GI* is also identified as a strong candidate for two of the QTLs associated with circadian leaf movement in *Arabidopsis* (Edwards et al., 2005).

Experiments by Cao et al. (2005) showed that the *Arabidopsis GI* gene can also be induced by cold stress which implicates involvement in the cold stress response. This hypothesis is further supported by analysis of freezing tolerance of the *gi* mutants. It was shown that the *gi-3* mutation decrease constitutive freezing tolerance and impair cold acclimation ability suggesting that *GI* gene is required for both constitutive freezing tolerance and cold acclimation in *Arabidopsis*. In another set of experiments *GI* transcript level increased 5-10 fold after 24 h exposure to low temperatures and remained increased in next 7 days (Fowler and Thomashow, 2002). Similar experiments done by Paltiel et al. (2006), showed that *GI* transcript accumulation responds differently when grown on different temperatures. They investigated *GI* expression in *Arabidopsis* and *Medicago truncatula* (*Medicago*) and found that dawn and evening trough levels were significantly higher in 7-day old *Arabidopsis* seedlings entrained at 30ºC compared to those entrained at 22ºC although the temperature did not significantly affect *AtGI* expression at peak levels. At the same time circadian regulated genes *AtCO* and *AtTOC1* displayed the same expression level and increased slightly, respectively. *GI* has also been implicated in the temperature compensation mechanism, responsible for maintaining a robust circadian rhythms over a broad range of physiological temperatures (Gould et al., 2006). In higher temperatures *GI* maintains rhythmicity and

accuracy by the regulation of *TOC1* through which CCA1/LHY maintain robust clock cycling.

1.3.2 *CONSTANS* a link between the circadian clock and floral integrators

The *CONSTANS* gene has an important role in the regulation of flowering by photoperiod and acts between the circadian clock and floral integrators (Samach et al., 2000; Suarez-Lopez et al., 2001). In *Arabidopsis*, *CO* belongs to a family of 17 putative transcription factors defined by two conserved domains (Putterill et al., 1995: Robson et al., 2001). The first is a zinc finger region near the amino terminus that resembles B-boxes, which regulate amino acid interactions in several animal transcription factors (Borden, 1998; Torok and Etkin, 2000). Usually it is of the type C-X^2-H-X^7-C-X^7-C-X^2-C-X^5-H-X^2-H, and was identified in a variety of animal proteins including several transcription factors, ribonucleoproteins, and proto-oncogene products (Borden, 1998). Further analysis showed that seven potential zinc-binding residues within the B-box consensus sequence are conserved in the CO and most of the CO-like (COL) B-boxes (Robson et al., 2001).

The second conserved domain is a region of 43 amino acids near the carboxy terminus named the CCT domain (CO, CO-like, TOC1), (Strayer et al., 2000; Robson et al., 2001). Although the conserved region of 43 amino acids is a longer stretch of contiguous homology than usually shown by nuclear localization sequences (NLS; Raikhel, 1992), the CCT region is thought to have an important function in nuclear localization based on its similarity to the consensus sequence for an NLS (Robert et al., 1998).

A transient expression assay in onion bulb epidermal cells showed that a translational fusion of GFP and a C-terminal segment of *AtCO* (amino acids 304-373; *GFP::CtermCO*), localised exclusively in the nucleus whereas GFP fused with N-terminal of *AtCO* (*GFP::NtermCO*) localized in both, the cytoplasm and nucleus suggesting that CCT region is sufficient to target GFP to the nucleus and that the only region of *AtCO* containing an NLS is between amino acids 304 and 373 (Robson et al., 2001). The CCT region is found in all 17 COL proteins in *Arabidopsis*.

In addition to nuclear localization, the CCT domain probably has other functions. The *co-7* mutant containing a mutation in C-terminal region had a severe effect on flowering time but

did not affect the nuclear localization (Robson et al., 2001) suggesting that this mutation was not enough to prevent nuclear import of the CO protein. A CCT domain is present in at least 18 proteins that do not contain B-boxes (Strayer et al., 2000). There are also 13 proteins that contain one or two B-boxes but do not have the CCT domain. Analysis of seven classical *co* mutant alleles demonstrated that all the mutations occur within either the zinc-finger region or the CCT domain, confirming that the two regions of homology are crucial for *CO* function (Robson et al., 2001).

The existence of proteins that contain only one of these domains, either B-box or CCT domain, suggests that these domains acts independently of one another. This is supported by the observation of Kurup et al. (2000), who showed that the CCT domains of *CO* and *TOC1* interact in yeast cells with the *Arabidopsis* transcription factor *ABI3*. This interaction was reduced approximately two-fold in both the *co-5* and *co-7* mutations. Therefore the terminal region probably has a role in protein-protein interaction as well as in nuclear localization. This has been further substantiated with the discovery that CCT region show similarities to yeast HEME ACTIVATOR PROTEIN2 (HAP2) and that CO interacts with AtHAP3 and AtHAP5 in yeast, *in vitro*, and in planta (Wenkel et al., 2006). They proposed a mechanism in which the CCT domain mediates the formation of the protein complex which is involved in the regulation of *FT* expression during the transition to flowering.

Arabidopsis CO mRNA shows a striking temporal pattern of expression that was proposed to provide a basis for the regulation of the pathway by day length. Expression analysis showed that *CO* mRNA levels oscillated under LD conditions, showing a broad peak between ZT12 and dawn (Suárez-López et al., 2001). The highest level of mRNA occurs at 16 h and dawn. Under SD, the peak of *CO* expression is narrower than under LD and occurs in the dark between 12 and 20 h. The higher abundance of *CO* mRNA under LD is most pronounced at dawn (Suárez-López et al., 2001). *CO* showed to be controlled by the circadian clock because in *Arabidopsis* plants entrained under LD and transferred to LL *CO* mRNA levels continue to oscillate with the period of ~24 h (Suárez-López et al., 2001). These results suggested that post-transcriptional regulation of CO by light under long days, might be responsible for the activation of CO, and thereby the response to long days.

Several *co* mutant alleles were shown to be semidominant, with the heterozygotes showing a phenotype intermediate between the homozygous mutants and wild-type (Koornneef et al.,

1991). Putterill et al. (1995) proposed that this was likely to be caused by haploinsufficiency in which the heterozygotes did not produce enough CO protein to promote early flowering, rather than the mutant allele encoding an altered gain of function protein. Recently Robson et al. (2001), showed that *co-3* mutant which carried an intact B-box domain and impaired CCT domain actively delayed flowering when overexpressed in wt plants.

Classical physiological experiments demonstrated that the initiation of flowering in response to day length involves a systemic signal formed in the leaves that induces floral development at the shoot apical meristem (SAM) (Zeevaart, 1976). Recently, Takada and Goto (2003), demonstrated CO involvement in *FT* mRNA expression in the vascular tissue of *TERMINAL FLOWER2* (*tfl2*) mutants which are known for exhibiting a pleiotropic phenotype (Larsson et al., 1998) by regulating chromatin structure to repress the expression of many genes including *FT* (Kotake et al., 2003). *tfl2* mutants are early flowering with upregulated *FT* mRNA. Early flowering of the *tfl2* mutant is dependant on *FT* as it is abolished in an *ft* mutant. Analysis of *CO::GUS* plants detected GUS expression throughout young leaf primordia, in the vascular tissue of mature leaves and cotyledons, as well as in the phloem and the protoxylem of stems with the weaker staining also detected in the SAM (An et al., 2004). The notion of *CO* acting in phloem to promote flowering is consistent with the grafting experiments which suggested that the perception of day length occurs in the leaf.

Transgenic *Arabidopsis* plants in which *CO* is overexpressed flower early and are insensitive to day length (Simon et al., 1996; Onouchi et al., 2000). CO directly activates the expression of another flowering time gene *FT* (Samach et al., 2000), which promotes flowering and whose expression is activated only under LDs (Kardailsky et al., 1999; Kobayashi et al., 1999).

1.3.3 *FLOWERING LOCUS T* acts as a floral integrator

FLOWERING LOCUS T (*FT*) gene acts partially downstream of *CO* and is among most potent activators of flowering, (Kardailsky et. al., 1999: Kobayashi et. al., 1999). It is similar to the sequence of *TERMINAL FLOWER1* (*TFL1*), an inhibitor of flowering (Bradley et al., 1997), Sequence structure also suggest that both of them belong to a family of possibly membrane associated proteins that includes phosphatidylethanolamine binding protein (PEBP) a membrane-associated mammalian protein (Kardailsky et al., 1999; Kobayashi et al.,

1999). It has been found that increased expression of the *FT* gene leads to accelerated floral development, and almost complete loss of a normal photoperiodic response. Early flowering was correlated with *FT* mRNA accumulation: plants with the highest levels flowered earlier then the others (Kardailsky et. al., 1999; Kobayashi et al., 1999). Experiments indicate that *FT* regulates flowering by acting in part downstream of *CO* since neither the SD photoperiod nor *co-1* mutation affect the early-flowering phenotype of transgenic plants. Consistent with this *ft-1* partially suppressed the precocious flowering phenotype of *35S::CO* (Kardailsky et. al., 1999; Kobayashi et al., 1999). In contrast, the semidominant *fwa-2* mutation, which did not affect *FT* expression partially suppressed the precocious flowering phenotype of *35S::FT* indicating that it might interfere with pathways downstream of *FT*. *FT/FWA* and *CO*, also interact differently with the meristem identity gene *LFY* because only *co* mutations affect transcriptional induction of *LFY* (Simon et al., 1996). Experiments also indicate that *FT* and *FWA* act redundantly with *LFY* to regulate *APETALA 1* (*AP1*; Ruiz-Garcia et al., 1997). Kardailsky et al. (1999), and Kobayashi et al. (1999) suggested that *FT* and *TFL* mediate signals for floral transition downstream of *CO* in an antagonistic manner. Recently it was shown that TFL2, which shows homology with HETEROCHROMATIN PROTEIN 1 (HTP1) of animals, functions as a negative regulator of flowering of *FT*. It represses the expression of the *FT* alone, without affecting other integrators such as *SOC1* or *LFY* with the mechanism possibly involving repression of specific euchromatin genes (Kotake et al., 2003). *TFL2* represses *FT* expression continuously throughout development, with *FT* eventually reaching a level sufficient for floral induction even in the presence of *TFL2*, suggesting that *TFL2* does not maintain *FT* in a silent state but counteract the effect of *CO* on *FT* activation (Takada and Goto, 2003). This close relationship between *CO*, *FT*, and *TFL2* has even been confirmed on the temporal level since all three of them are expressed together in leaf vascular tissues (Takada and Goto, 2003). A bZIP transcription factor *FD* is required for *FT* to promote flowering (Abe et al., 2005; Wigge et al., 2005). The interaction occurs in the shoot apical meristem where FT protein binds to the FD protein forming a dimmer which in turn binds the 130-bp response element within the *AP1* promoter (Wigge et al., 2005). Recent work of Teper-Banmolker and Samach (2005) suggests that *AtFT* promotes flowering through transcriptional activation of *FRUITFULL* (*FUL*), *SEPALLATA* (*SEP3*) and *AP1* within rosette leaves.

Recent genetic and physiological analysis revealed that flowering time correlated well with the expression levels of CRY2-GFP in vascular bundles, which are shown to be *CO* and *FT* expression sites in *Arabidopsis* (Endo et al., 2007) suggesting that *CRY2* regulates CO protein levels in a cell-autonomous manner. Expression of *FT* rises during the first LD after a shift from SD (Corbesier et al., 2007). In plants exposed to 3 LDs, *FT* mRNA abundance was increased in each of three days. After return to SD conditions the *FT* mRNA levels fell after one day to the SD characteristic levels. Although the expression of *FT* was transient these plants flowered much earlier than the plants grown exclusively in SD conditions.

It has been thought that pathways other than the *CO/GI* pathway are also involved in the activation of *FT*, because the expression of *FT* remains detectable in *co* mutants during late development (Kardailsky et al., 1999; Kobayashi et al., 1999). *Lily* MADS-box gene *LMADS3* up-regulated expression of *FT* and *SOC1* in the photoperiod pathway and floral meristem identity genes *AP1* and *LFY*, when overexpressed in *Arabidopsis* (Tzeng et al., 2003). 35S::*LMADS3* was unable to rescue the late-flowering phenotype of *ft-1* and *fwa-1* but it was able to compensate for the late-flowering phenotype in *gi-1* and *co-3*. The activation of mentioned photoperiod genes is indirect because their expression was unaffected in plants transformed with *LMADS3* fused with rat glucocorticoid receptor in the presence of both, dexamethasone and cyclohexamide. All this could indicate possible existence of parallel pathway alongside *CO*.

1.4 Molecular mechanism of control of flowering by photoperiod in *Arabidopsis*

Experiments in which the temporal pattern of *CO* expression was altered using mutants or by altering the length of the daily cycle showed that the exact timing of *CO* expression is important in distinguishing between long and short days (Roden et al., 2002; Yanovsky and Kay, 2002). The *toc-1* mutation causes circadian rhythms to cycle faster under constant light, and under SD causes *CO* mRNA abundance to peak earlier. This earlier peak in *CO* mRNA under SD occurs during the light rather then during the night and correlates with increased *FT* expression and early flowering under SD, and these effects require *CO* function since they are

largely abolished in a *co* mutant (Yanovsky and Kay, 2002). Altering the phase of *CO* expression involved changing the duration of the 24 h daily cycle. The timing of the expression of *CO* relative to the light-dark transitions could be altered by maintaining the ratio of light to dark within the daily cycle, but extending or shortening the cycle from 24 h to 21 or 30 h (Roden et al., 2002). This demonstrated a strong correlation between the expression of *CO* in the light, increased expression of the downstream gene *FT* and early flowering. Although these procedures are likely to affect the timing of expression of many clock-controlled genes, the correlation between *CO* expression during the photoperiod, upregulation of *FT* and early flowering suggest that post-transcriptional regulation by exposure to light is at least one mechanism by which flowering of *Arabidopsis* is activated in response to long days. The basis for this proposal can be found in two molecular mechanisms underlying the activation of *CO* by light. The stability of the CO protein was shown to be regulated by light, so that in plants exposed to blue or far-red light the protein accumulates in the nucleus, but in darkness or red light the protein is absent (Valverde et al., 2004; Fig 1.4). This correlates with blue and far-red light being the most effective in promoting flowering. Also genetic experiments demonstrate that the blue-light photoreceptors, *CRY1* and *CRY2*, as well as the far-red photoreceptor *PHYA* both promote flowering and stabilize the CO protein, whereas *PHYB*, which is activated by red light, delays flowering and promotes the degradation of CO protein (Johnson et al., 1994; Guo et al., 1998; Valverde et al., 2004).

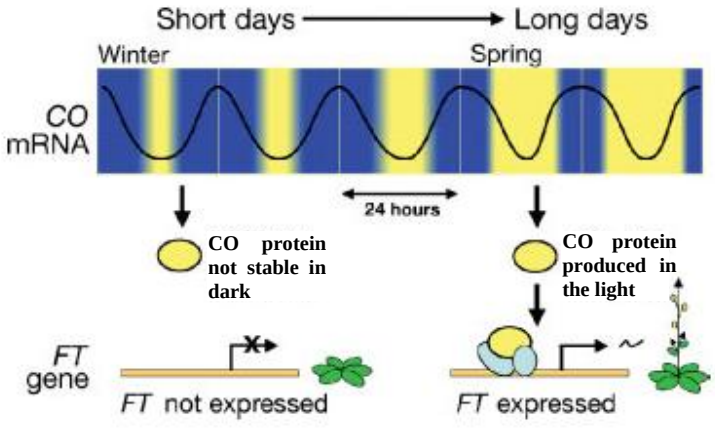


Figure 1.4 An LD model of flowering promotion in *Arabidopsis*.
During the winter (SD conditions) the peak of CO expression occurs in dark and is not stabilized by light. As the light period increases *CO* expression coincides with the light which stabilizes CO protein, subsequently inducing *FT* gene (from Putterill et al., 2004)

Another mechanism based on transcriptional regulation was recently shown to regulate *CO* in response to light. FLAVIN-BINDING, KELCH-REPEAT, F-BOX 1 (FKF1) is a protein whose mRNA is clock controlled which has an important role in generating the *CO* mRNA peak in the late afternoon. Mutations in *FKF1* both delay flowering and reduce *CO* expression at dusk, a crucial time regarding mRNA abundance which facilitates the coincidence between *CO* expression and light (Imaizumi et al., 2003). This regulation does not happen directly but through the regulation of transcription factor *CDF1* (Imaizumi et al., 2005). FKF1 directly controls the stability of the CDF1 by physically interacting with it rendering CDF1 protein less stable. Plants with elevated levels of CDF1 flower late and have reduced expression of *CO*.

The C-terminal, plant-specific CCT domain of *AtCO* and *Tomato COL1* (*TCOL1*) were shown to bind the trimeric CCAAT binding factor (CBF; Ben-Naim et al., 2006) which regulates transcription by directly binding to a common promoter motif in many eukaryotic genes (Maity and de Crombrugghe, 1998). The authors conclude that CBFs may be important in recruiting COL proteins to their DNA targets *in planta*. The responsiveness of *CO* activity to day length therefore depends on regulation at several levels. Circadian clock control of *CO* transcription underlies the system and restricts *CO* expression to the later part of the day/night cycle. The presence of light during the evening both enhances *CO* transcription and stabilizes the protein in the nucleus ensuring activation of the floral regulator *FT*. This requirement for light ensures that *CO* activation and flowering occur under long days.

Until recently spatial dimension of this cascade process was little known. It was shown that misexpression of *AtCO* from the phloem-specific promoters (*AtSUC2::CO; rolC::CO*), but not from meristem, epidermis, or root specific promoters (UFO, KNAT1, STM, and ML1), is sufficient to induce early flowering and complement *co* mutation (An et al., 2004). In these plants, in situ hybridization showed that *FT* was also upregulated specifically in the phloem and not in adjacent cells. *AtFT* also showed extreme early flowering when expressed in the phloem using *AtSUC2* or *rolC* promoters, but unlike *CO*, *FT* caused early flowering when fused to meristem, and epidermis specific promoters (An et al., 2004), suggesting that effectiveness of *FT* in promoting flowering is not restricted to these cells. In a similar experimental setup *gCO::GUS* and *pFT::GUS* expression was detected in the vascular tissues of cotyledons and leaves indicating that the phloem cells of leaves are the place of *FT*

activation by *CO* (Takada and Goto, 2003). Microarray experiments confirmed that the single gene, *FT*, is differentially affected by the *co* mutation and the shift from SD to LD which suggest that *FT* is the major target of *CO* in leaves (Wigge et al., 2005).

Using yeast two-hybrid screen (Wigge et al., 2005; Abe et al., 2005) isolated FD protein as a possible FT protein interactor, subsequently showing that FT and FD are sufficient to activate the expression of the floral marker genes *AP1* and *LFY*, possibly by mechanism which includes recruiting FT protein to the FD response element in the *AP1* promoter in an FD-dependent manner.

The concept of the mobile floral stimulus, generated in leaves upon exposure to inductive photoperiods and transported to the shoot apex where it acts to promote flowering has been present for a long time (Zeevaart, 1976). A paper indicating that *FT* mRNA is this mobile floral stimulus named florigen has been retracted since the published results could not be replicated (Böhlenius et al., 2007). Several other authors found no signs of *FT* mRNA in SAM in tomato, *Arabidopsis* and rice (Lifschitz and Eshed, 2006; Corbesier et al., 2007; Tamaki et al., 2007). Using heterologous promoters (*SUC2* and *KNAT1*), transient expression of *FT* in a single leaf was sufficient to induce flowering in *Arabidopsis* through the FT protein which is transported from the phloem companion cells, through the phloem, to the SAM (Corbesier et al., 2007). This result was supported by the similar study in rice where *Hd3a::GFP* fusion under the control of native HEADING DATE 3a (Hd3a, an *Arabidopsis* FT orthologue in rice), and phloem specific promoters (*rolC* and *RPP16*) produced early flowering plants with GFP signals detected in the vascular tissues of leaf blades and the stems of transgenic plants (Tamaki et al., 2007). This strongly suggest the role of FT/Hd3a protein as a florigen signal.

1.4.1 Interaction of photoperiod pathway with vernalisation pathway

Photoperiod and vernalisation pathways act synergistically to regulate flowering time. The major role of *FLC* is to repress the expression of *FT* and *SOC1* which is opposite to the function of the long day floral promoter *CO* (Fig 1.5), whose role is to activate these genes (Putterill et al., 2004). Hepworth et al. (2002) showed that there is an antagonistic regulation of the transcription of *SOC1* by *FLC* and *CO*. *FLC* binds to a specific MADS-domain protein binding element (CArG) within the *SOC1* promoter which is thought to prevent *CO*

from binding to a separate *SOC1* promoter element. Vernalisation decreases the levels of *FLC* which in turn promotes flowering by making *SOC1* promoter element available to the *CO*. *FLC* is the major, but not the only target of the vernalisation pathway since *flc* null mutants still respond to vernalisation (Michaels and Amasino, 2001). Contrary to this a genetic analysis of the *ft-10* and *soc1-2* mutants and its expression patterns shows that *FT* is the major output for *CO* and that *CO* activates *SOC1* through *FT* to promote flowering in Arabidopsis (Yoo et al., 2005).

Several other transcriptional factors like *MADS AFFECTING FLOWERING 1-5* (*MAF1-5*), and *AGAMOUS-LIKE 24* (*AGL24*) have been implicated in the vernalisation response independently of *FLC* (Ratcliffe et al., 2001; Yu et al., 2002; Michaels et al., 2003a). Direct expression of *SOC1* in the SAM is able to promote flowering even in the absence of *CO* or *FT* while mutation of *FT* strongly delays the expression of *SOC1* in the SAM, indicating that *SOC1* acts downstream of *FT* in the SAM (Searle et al., 2006).

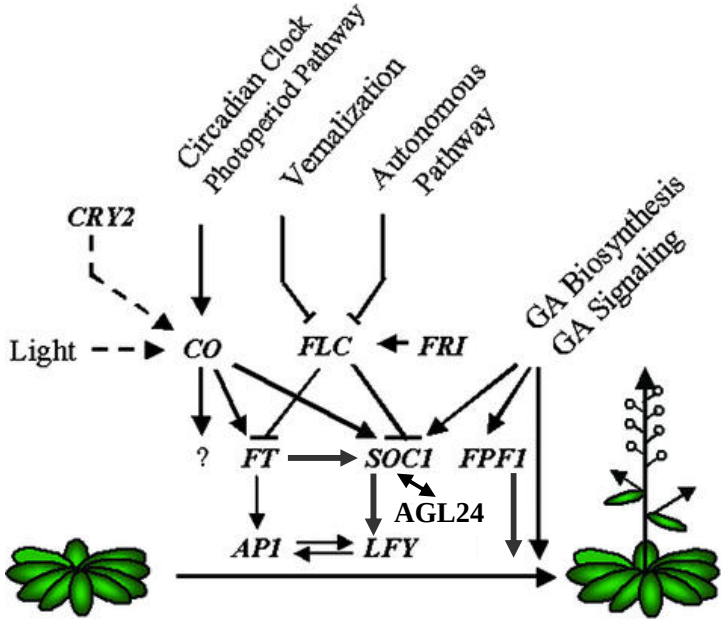


Figure 1.5 Overview of the relationships among Arabidopsis flowering pathways (modified from Mouradov et al., 2002)

1.5 Conservation and change to the photoperiod pathway

1.5.1 Rice

Data to support the *Arabidopsis* model of the genetic control of flowering time as applicable to other plant species was sparse until fairly recently. However, significant advances in the rice heading date (flowering) QTL mapping (Yano et. al., 2001) and identification of the genes underlying natural variation revealed the same genetic pathway, at least for the photoperiodic control of flowering (Hayama et. al., 2003; Kojima et. al., 2002) despite the fact that rice is very different to *Arabidopsis* in flowering behavior, being a tropical species that shows promotion of flowering in response to SD and lacks a vernalisation response (Figure 1.6; Table 1.1). It has been shown that three key regulatory genes for the photoperiodic control of flowering are conserved between *Arabidopsis*, a LD plant, and rice, a SD plant, but regulation of the *FT* gene by *CO* was reversed (Fig 1.5), resulting in the suppression of flowering in rice under LD conditions (Hayama et.al., 2003). This is also an example of the important developmental process which can be diversified by using the same set of regulatory genes but by regulating them differently.

Overexpression of *OsGI*, an orthologue of the *Arabidopsis GI* gene in transgenic rice, caused late flowering under both SD and LD conditions (Hayama et al, 2003) which suggested that *OsGI* acts as a suppressor of flowering in rice, which is a reversal of the role of *GI* in the photoperiodic control of flowering in *Arabidopsis* (Fowler et al, 1999). Contrary to this, transgenic rice in which *OsGI* expression was suppressed by RNA-mediated interference flowered early under LD but late under SD conditions indicating that *OsGI* is still required to promote flowering in SD conditions.

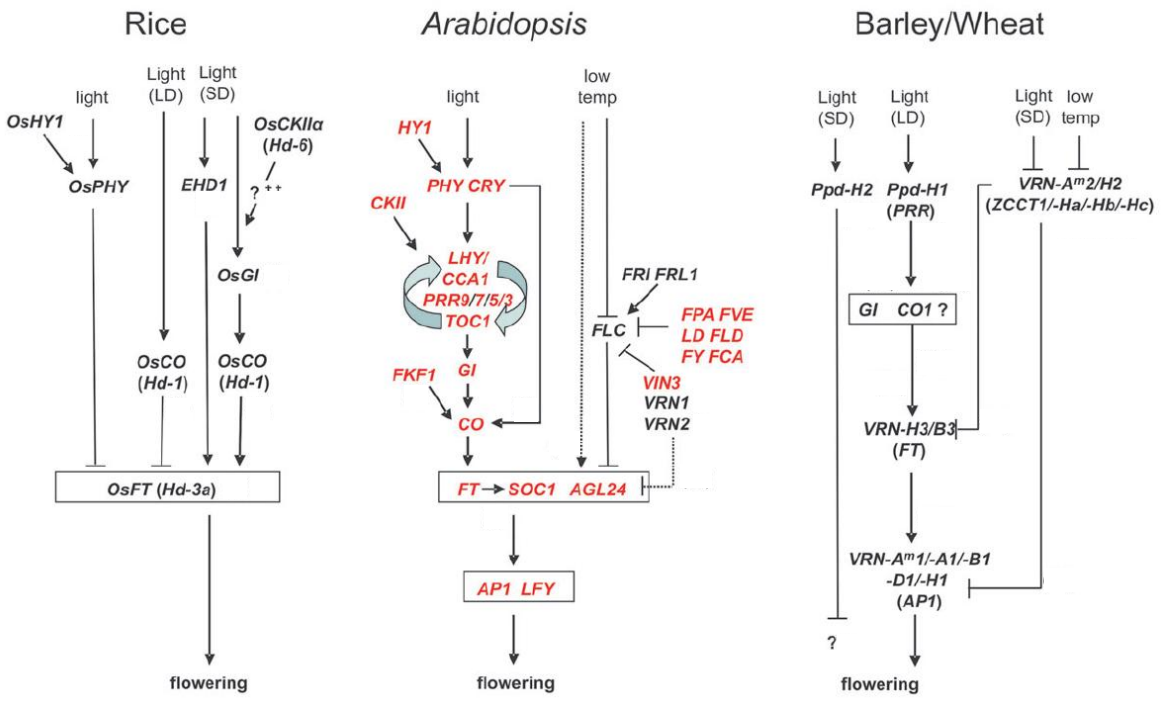


Figure 1.6 *Arabidopsis* flowering pathway genes and its counterparts in rice and wheat/barley Genes for which putative barley/wheat orthologues have been identified are in red (from Cockram et al., 2007)

In transgenic rice overexpressing *OsGI* mRNA levels of *Hd1* were positively correlated with *OsGI* with increased levels under both LD and SD conditions. In contrast, the *Hd3a* (*FT*) mRNA levels were negatively correlated with those of *OsGI*. These results suggests that *Hd1* acts as a suppressor of *Hd3a* gene expression as well as of flowering in transgenic rice. Expression analysis also showed that the quantities of *Hd1* mRNA did not change with the transition from LD to SD or when they were grown solely in SD or LD conditions (Yano et al, 2000). This results suggest that the expression of *Hd1* is not greatly affected by a change of photoperiod.

High resolution linkage mapping coupled with the transformation analysis showed that *Hd1* might be bi-functional promoting heading under SD conditions and inhibiting it under LD conditions when compared with lines that displayed loss of function of *Hd1* (Yano et al, 2000). In *Arabidopsis*, the CO protein promotes flowering under LD conditions but has no phenotypic effect on flowering time under SD conditions (Putterill et al, 1995). This suggests that rice *Hd1* contains specific structure(s) responsible for the specific transfactors activated

under LD conditions, and/or rice produces certain transfactors to regulate CO activity under LD conditions.

Table 1.1 List of flowering genes from *Arabidopsis*, *H.vulgare, T. monococcum*, *T. aestivum*, and *S.cereale*. Modified from Cockram et al., 2007

Locus/gene name	Predicted protein	Pathway	Function
Arabidopsis			
AGL24	MADS-box	Pathway integrator	Activate floral organ identity genes
AP1	MADS-box	Meristem identity	Activate floral organ identity genes
CCA1	Myb-related transcription factor	Photoperiod	Component of the central oscillator
CO	B-box, CCT domain	Photoperiod	Promotes flowering
CRY1-2	FAD-binding domain	Light quality/Photoperiod	Blue light perception
FCA	RNA-binding	Autonomous	FLC repression
FKF1	Flavin-binding, kelch repeat	Photoperiod	Promote peak CO transcription
FLC	MADS-box	Vernalisation	Central repressor of flowering
FLD	HDAC-associated protein	Autonomous	FLC repression
FPA	RNA-binding protein	Autonomous	FLC repression
FRI	Coiled-coil	Vernalisation	Up-regulate FLC
FRL1	Related to FRI	Vernalisation	Up-regulate FLC
FT	Putative kinase inhibitor	Pathway integrator	Activate floral organ identity genes
FVE	MSI4	Autonomous	FLC repression
FY	Polyadenylation factor	Autonomous	FLC repression
GI	Nuclear protein	Photoperiod	Output of central oscillator
HY1	Haemoxygenase	Photoperiod	Chromophore synthesis
LD	Homeodomain protein	Autonomous	FLC repression
LFY	Plant-specific transcription factor	Pathway integrator	Activate floral organ identity genes
LHY	Myb-related transcription factor	Photoperiod	Components of central oscillator
PHYA-E	Phytochrome	Light quality	Light sensors
PRR9/7/3/5	CCT domain	Photoperiod	Components of central oscillator
SOC1	MADS-box	Pathway integrator	Activate floral organ identity genes
TOC1	CCT domain	Photoperiod	Component of central oscillator
VIN3	PHD, VID-domain	Vernalisation	Cold-mediated FLC repression
VRN1	B3-domain, DNa-binding	Vernalisation	FLC repression-post vernalisation
VRN2	Su(z)12-like polycomb protein	Vernalisation	FLC repression-post vernalisation
Rice			
EHD1	B-type response regulator	Photoperiod	Promote flowering under SD

Hd1(Se1)/CO	B-box, CCT domain	Photoperiod	Promote flowering under SD
Hd3a/FT	Putative kinase inhibitor	Photoperiod	Promote flowering under SD
Hd6/CKX2α	Protein kinase	Photoperiod	Promote flowering under SD
Se5/HY1	Haem oxygenase	Photoperiod	Chromophore synthesis
H.vulgare			
VRN-H1/BM5A^{b}	MADS-box, AP1 like	Vernalisation	Recessive alleles promote flowering after vernalisation
VRN-H2/ZCCT	B-box, CCT domain	Vernalisation/photoperiod	Dominant alleles promote flowering after vernalisation
VRN-H3/HvFT	Putative kinase inhibitor	Vernalisation/photoperiod	Recessive alleles promote flowering after vernalisation, up-regulated in LD
Ppd-H1/PRR	Pseudo-receiver and CCT domain	Photoperiod	Light sensitive alleles promote flowering under LD
Ppd-H2	not cloned	Photoperiod	Light sensitive alleles delays flowering under SD
T. monococcum			
VRN-A^{m}1(VRN-1)	MADS-box, AP1-like	Vernalisation	Promote flowering after vernalisation
VRN-A^{m}2(VRN-2) /ZCCT1	B-box, CCT domain	Vernalisation	Promote flowering after vernalisation, up-regulated under SD
T. aestivum			
VRN-A1(Vrn1)	MADS-box, AP1-like	Vernalisation	Promote flowering after vernalisation
VRN-B1(Vrn2)	MADS-box, AP1-like	Vernalisation	Promote flowering after vernalisation
VRN-D1(Vrn3)	MADS-box, AP1-like	Vernalisation	Promote flowering after vernalisation
VRN-B3(Vrn-B4) /TaFT	Putative kinase inhibitor	Vernalisation	Recessive alleles promote flowering after vernalisation, up-regulated in LD
Ppd-A1(Ppd1)	not cloned	Photoperiod	Promote flowering in SD and LD
Ppd-B1(Ppd2)	not cloned	Photoperiod	Promote flowering in SD and LD
Ppd-D1(Ppd3)	not cloned	Photoperiod	Promote flowering in SD and LD
S.cereale			
Vrn-R1		Vernalisation	Recessive alleles promote flowering after vernalisation

These characteristics of rice as an SD plant raise the question whether or not the *Hd1/CO* genes from LD plants can influence heading in the background of SD plants. Recent experiments showed that introduction of the genomic complement of one of the wheat (a LD plant) *CO-like* genes, *TaHd1*, into nearly isogenic line (NIL) of Nipponbare type rice with functional alleles *Hd1* and *Hd2* replaced with Kasalath (Indica type) non-functional alleles, produced earlier heading phenotype under SD conditions in comparison to those with only the

vector sequence and NIL (*Hd1/Hd2*) (Nemoto et al, 2003), suggesting that the regulatory sequences in *CO/Hd1* genes are conserved to the level they can function in heterologous plant species. Conservation between rice and *Arabidopsis* also suggests that *CO-like* genes are likely to be involved in flowering time control in other cereals and grasses which are quantitative LD plants such as ryegrass.

Recently Doi et al. (2004), have described a B-type response regulator *Ehd1* which plays an important role in photoperiodic control of flowering in rice and whose counterpart could not be found in *Arabidopsis*. *Ehd1* can induce expression of *FT-like* genes in *Hd1* deficient background which reveals the second signaling pathway integrated with the existing photoperiod pathway. *Ehd1* mRNA is induced only under SD conditions which is different from the *Hd1* mRNA expression patterns that stay similar irrespective of the light regime. *Ehd1* is thought to function as a transcription factor by directly binding DNA and regulating the expression of the downstream *FT-like* genes. Further proof to that could be the expression patterns of *FT-like* genes that closely follow expression of *Ehd1* under SD condition.

So far vernalisation has not been reported in rice so it is not surprising that no apparent *FLC* homologue has been found in rice. *FRI* the activator of *FLC* in *Arabidopsis* has not been found in rice also (Johanson et al, 2000). Furthermore no orthologues of the *VRN1* and *VRN2*, required for vernalisation in *Arabidopsis* have been found in rice which suggest that the vernalisation related genes have been lost from the rice genome during *the* evolution.

1.5.2 Temperate cereals

In barley (LD plant), two major loci regulating photoperiod response have been identified by the QTL and genetic analyses. In vernalised plants flowering time under short days was largely controlled by the Photoperiod-H2 (*Ppd-H2*) locus on the long arm of chromosome 1H while under LD conditions flowering was largely controlled by *Ppd-H1* locus of the short arm of chromosome 2H (Laurie et al., 1995). A pseudo response regulator (*PRR*) underlying *Ppd-H1* locus has been cloned recently using fine mapping procedure (Turner et al., 2005). *PRR* genes in *Arabidopsis* are circadian clock associated and include *TOC1* a member of the central oscillator.

Analysis of cereal ESTs and rice genomic sequence revealed 16 barley genes that fell into the COL group of genes (Griffiths et al., 2003). In barley *HvCO1* and *HvCO2* were most similar

to the rice *Hd1* and *Arabidopsis CO*, *COL1*, and *COL2* genes. Subsequent mapping analysis placed *HvCO1* on barley chromosome 7H, previously shown to be syntenic with the region of rice LG6 containing *Hd1*. Although *HvCO1* was the most similar to rice *Hd1*, the B-box was not well conserved since *HvCO1* peptides lacked three highly conserved C residues that would be predicted to abolish B-box2 function. B-box2 from *HvCO2* was more conserved but still had two non-consensus amino acids in the position of *co-1* deletion, which suggest that in cereals the B-box2 region is less important to the function of most *CO-like* genes in cereals (Griffiths et al., 2003). *HvCO2* had no counterpart in rice but is found in wheat (Nemoto et al., 2003). Although *CO/Hd1* homologues exist in barley and are likely to have roles in the control of flowering they do not correspond to the major photoperiod loci which are located on chromosomes 1H and 2H. These results could be an indication that temperate cereals may have evolved a new mechanism for controlling photoperiod response.

Using a similar approach three *CO-like* genes (*TaHd1-1*, *TaHd1-2*, and *TaHd1-3*) were isolated from hexaploid wheat (Nemoto et al., 2003). Transcripts of *TaHd1-2* could not be detected at any stage presumably because of a 63 bp deletion at the promoter region. The location of *TaHd1* gene was assigned to the long arm of chromosome 6 which is highly re-arranged region with some duplicated restriction length fragment polymorphism (RFLP) markers mapped between chromosomes 2 and 7, and between 2 and 5 of Triticae. Diurnal oscillation patterns of *TaHd1* mRNA were similar to those in rice and *Arabidopsis* with *TaHd1-1* gene complementing rice *Hd1* and function normally in the rice background (Nemoto et al., 2003).

The barley photoperiod gene *HvGI* was amplified and sequenced from a barley BAC library and comparison of its genomic and cDNA sequences revealed 14 exons and 13 introns with intron/exon structure identical to rice *GI* and similar to *AtGI* (Dunford et al., 2005). It was mapped on the short arm of barley chromosome 3H which is collinear with the rice chromosome 1S and shows overall conserved pattern of transcriptional regulation to the *Arabidopsis GI* (Dunford et al., 2005). The same experiments on *Zea mays* (maize) identified two copies of *ZmGI*, on chromosomes 3 and 8, both partially syntenic to the rice chromosome 1S (Dunford et al., 2005). The wheat *GI* gene (*TaGI*) was isolated through sequence homology and it showed high sequence similarity with *OsGI* and *AtGI* (Zhao et al., 2005). Similarly to *AtGI* and *OsGI*, *TaGI* expression is under the control of circadian clock and

photoperiod; overexpression of *TaGI* functionally complements *gi* mutant in *Arabidopsis*. *In situ* hybridization showed that *TaGI* is expressed in both SAM and leaves with interesting observation that *TaGI* expression cycled only in leaves since the expression there was stronger at ZT10 than ZT0.

Although vernalisation-dependent flowering has been studied extensively in *Arabidopsis*, the respective pathway in grasses has not been completely described. Two vernalisation related (*VRN*) genes (unrelated to the Arabidopsis *VRN* genes) have been positionally cloned in einkorn wheat (*Triticum monococcum*, Yan et al., 2003; Yan et al., 2004), with their alleles associated with the major vernalisation effects in wheat and barley. They have also been shown to co-segregate with major quantitative trait loci (QTL) for vernalisation in cereals and forage grasses (Laurie, 1997; Jensen et al., 2005). Detailed genetic study showed that the variation in vernalisation habit in wheat was completely linked to the *VRN1* gene, which is similar to the *Arabidopsis AP1* gene in sequence and expression similarity (Yan et al., 2003). Similar results were obtained in barley (*Hordeum vulgare*) where it was found that *HvVRN1* is regulated by vernalisation and development (Trevaskis et al., 2006).

So far *AP1* has not been associated with vernalisation requirements in *Arabidopsis* in which natural variation for vernalisation requirement has arisen through the nonfunctional or weak *FRI* and *FLC* alleles (Gazzani et al., 2003). In wheat, VRN2 locus was determined by positional cloning and it was linked with two genes *ZCCT1* and *ZCCT2* with similarities to *CO* and *CO-like* genes within the 43 aa CCT region (Yan et al., 2004). Experiments in barley showed that *HvZCCTa* and *HvZCCTb* genes found at the VRN2 locus were regulated by daylength in vernalisation responsive winter varieties (Trevaskis et al., 2006). Recently *VRN-H3* gene in barley and its wheat orthologue *VRN-B3* have been mapped and characterized (Yan et al., 2006). Their map position to the homologous chromosome 7HS and expression patterns strongly suggests the identity between *FT* and *VRN-H3*/*VRN-B3* in barley and wheat respectively.

Five barley FT-like genes were isolated recently using BAC library screening (Faure et al., 2007). Sequence and expression analysis coupled with gene mapping showed that only one, *HvFT1* (*VRN-H3*), could fulfill the role of the floral integrator in barley.

Relationship between VRN1 and VRN2 genes and between VRN1/VRN2 and other genes within vernalisation/photoperiod pathway is very complex and still poorly understood. In

cereals VRN2 was found to modulate the quantitative levels of *FT* providing a link between the vernalisation pathway and *FT*. In the light of these results it is possible to conclude that *Arabidopsis* and temperate grasses developed different vernalisation pathways using different genes down-regulated by vernalisation (*ZCCT1*, *FLC*), and similar genes with different regulatory profiles (*AP1*). Comparative studies on vernalisation between dicots and monocots are incomplete at the moment but from the mentioned studies direct comparisons can be made within monocot species.

1.6 Grass - staple food for grazing animals

Crop domestication has been viewed as the single most important event in human cultural development over the last 10000 years with grass in the broad sense being most prominent (Buckler et al., 2001) and providing the vast majority of the world's food. Grass is a very important world pastoral crop. In some countries including the UK, Australia and New Zealand, animal products from grassland make a greater contribution to the value of agricultural production than any other crop. In total grasslands account for almost one quarter of the world's cover of vegetation (Jones and Lazenby, 1988). Eighty per cent of the world's cow milk and 70% of the world's beef and veal are produced from temperate grasslands. In New Zealand there are nearly 10 million hectares of sown pasture (mostly perennial ryegrass) compared with 1.3 million for the combined areas of crop, horticulture, and commercial forestry resulting in the export revenue from animal products of over 50% of the total export revenue (http://www.nzte.govt.nz/section/11751.aspx).

1.6.1 Ryegrass – importance and phylogeny

The ryegrasses have several important performance characteristics which account for their widespread use and popularity. Among them are high herbage yield, a long growing season, tolerance to a wide range of environmental conditions and grazing practices, excellent persistence under close grazing, compatibility with white clover, and high forage quality and palatability. Perennial ryegrass is a member of family Poaceae and is, therefore, related to more extensively studied grass species, such as the cereals, wheat, barley, maize, and rice.

Two species in the genus *Lolium* are among most important pasture and turf grasses in the world; *L. perenne*, perennial ryegrass or English ryegrass; and *L. multiflorum*, Italian ryegrass

or annual ryegrass, both of them self-incompatible and cross-pollinated with a gametophytic self-incompatibility system controlled by two genetic loci designated S and Z (Cornish et al. 1979). Despite its name, some cultivars of Italian ryegrass are annuals while others are perennials with stand persistence of four or more years. There are diploid and autotetraploid cultivars of both species. Diploids are used mainly for permanent pastures. They tiller (bunch) more, making very solid and durable, long lasting pastures. They have higher dry matter content which makes them good for grazing. Tetraploids have larger cells and are thus lower in dry matter making them lusher, somewhat sweeter, therefore a bit more palatable. They are generally a bit more productive than diploids and have better disease resistance. On the negative side, tetraploids do not tiller as well, so the stands are more open than a diploid pasture (Moser et al., 1996). This factor, together with having a higher crown, makes them less hardy under a grazing regime. Also, tetraploids do require more nutrients and moisture than diploids making them more susceptible to stress if these components are deficient, and they are generally shorter lived. In New Zealand ryegrass species and cultivars are a predominant component in nearly all pasture mixtures with perennial ryegrass and white clover forming the basis for permanent pastures. Italian ryegrass may be included in permanent pasture seed mixtures to provide rapid early feed production. *Lolium* species are interfertile and will cross with *Festuca* species as well, including meadow fescue and tall fescue (Wratt and Smith, 1983).

1.6.2 Objectives and methodologies in ryegrass improvement

The most important traits affecting the feeding value of herbage are *in vitro* dry matter digestibility (DMD), the ratios of crude protein, water-soluble carbohydrate (WSC) and fibre and the concentration of alkaloid toxins. Untill the 1980s, the main aim of grass breeding was to improve persistency and dry matter yield (DMY) but in recent years quality aspects, traits for efficient use of inputs and a long growing season are also being bred into modern grass varieties (Wilkins et al., 2000). The fact that *Lolium* species and varieties vary in their fatty acid and linoleic acid content (Dewhurst et al., 2003) open up opportunities for breeding varieties with positive effect on human health. Dry matter digestibility of forage plants decreases markedly (>10%) as plants flower and senesce (Radojevic et al., 1994). The changes in dry matter digestibility greatly contribute to the lowering of the nutritive value of

forage during summer. Increasing dry matter digestibility has been ranked as the most important goal in genetic improvement of nutritive value of forage grasses for dairy pastures (Smith et al., 1997). However, since heritability of dry matter digestibility is low and a large number of genes control it, the potential for rapid genetic improvement by traditional methods is low (Barnes, 1990). Lignification of plant cell walls has been identified as a major factor responsible for lowering digestibility of forage tissues as they mature (Buxton and Russel, 1988). As a consequence during late spring and summer, the digestibility of grass pasture declines. A proportion of the decline is associated with maturation and flowering and may be offset in practice by intensive grazing to reduce development of reproductive stems (Johnson and Parsons, 1985). Plant phenology, especially flowering time and extent, has probably more impact than all other traits on grassland production, forage quality and persistency. Studying flowering time genes in ryegrass and finding the mechanism through which they control floral development could eventually contribute developing varieties with extended vegetative growth, effective pollination and forage production or/and varieties in which flowering could be controlled by regulatable transgene systems.

1.6.3 Ryegrass genetic and genomic resources

Progress is also being made in breeding techniques, with molecular markers used to follow the introgression of genes. Genetic linkage maps and quantitative trait locus (QTL) analysis including linkage disequilibrium (LD) can locate genes and linked markers associated with important agronomic traits to facilitate QTL introgression and selection in crop breeding programmes. Linkage mapping of the major human food crops where most of the important species are self-fertile is well advanced. They have been highly inbred, and consequently have a relatively narrow genetic base meaning that well defined F2, or recombinant inbred lines are easy to produce (Humphreys et al., 2006). In contrast, most temperate pasture grasses are outbreeders which make production of suitable mapping families more difficult. For this reason mapping and QTL analysis in forage grasses has lagged behind work in the related cereals but considerable progress has been made in recent years.

The first-generation molecular marker-based genetic maps for agronomically important plant species have been largely based on anonymous genetic markers. In perennial ryegrass the reference genetic map contains large number of amplified fragment length polymorphism

(AFLP) and genomic DNA-derived simple sequence repeat (SSR) markers (Jones et al., 2002a) as well as heterologous RFLP markers, some of which correspond to unanotated cDNA sequences from other Poaceae species (Jones et al., 2002). These types of markers can anchor maps across different pedigrees but they are not in general closely associated with variations in genes controlling phenotypic traits. Molecular variation based on functionally defined genes underlying specific biochemical or physiological functions will provide the next generation of molecular markers for forage species (Faville et al, 2004). Such markers often called gene-derived or candidate gene-based markers are becoming available in *Lolium* and promise to have very close association with loci controlling variation for the trait in question, allowing development of the perfect markers and the direct selection of genotypes with desirable allele content.

Lolium genetic maps are now well aligned with maps of other species in the Poaceae (Jones et al., 2000) and are being used in comparative analysis to identify gene orthologues in other crops such as rice (Armstead et al., 2004; Jensen et al., 2005). The alignment of the ryegrass map with those of the Triticae cereals confirms the relatively close taxonomic relationship of genus *Lolium* to the Triticae genera *Triticum, Hordeum,* and *Secale* (Fig 1.7, Jones et al., 2002b). Perennial ryegrass seems to be more closely related to the Triticeae that to oat in terms of genome structure, despite the closer taxonomic affinity that exists between perennial ryegrass and oat. This observation was supported by the general agreement of the syntenic relationship between perennial ryegrass, oat, and rice, and those between the Triticeae and these species.

Marker assisted selection (MAS) will particularly improve the efficiency of conventional plant breeding in situations where the traits to be selected are difficult or expensive to evaluate, where the traits are expressed late in the growth cycle, where numbers of traits are needed to be improved simultaneously or when the traits are controlled by recessive alleles. It is clear that advances in genomics will lead to improved efficiency of forage breeding.

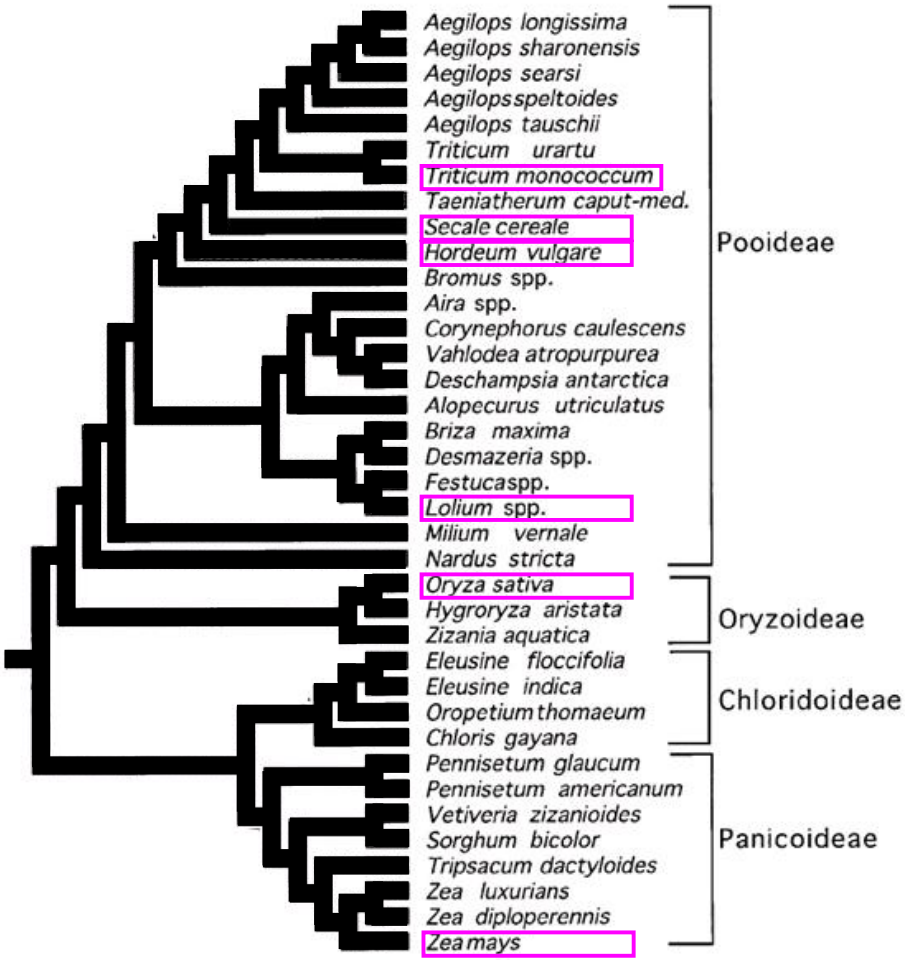


Figure 1.7 Cladogram showing the relationship between different grass species Relevant species are highlighted (modified from Kellogg, 1998)

Numerous ryegrass EST libraries from different tissues have been developed by different companies and research facilities around the world. This has greatly facilitated ryegrass gene discovery and development of the next generation of molecular markers from expressed sequences. AgResearch, New Zealand in collaboration with AgVic, Australia has developed ryegrass EST library which generated 44636 sequences providing the resource for the development of a large number of candidate gene-associated markers.

Today genetic engineering is complementing traditional breeding in the development of improved germplasm. A range of gene transfer protocol have been described for the production of the transgenic perennial ryegrass plants (Ge et al., 2007; van der Maas et al., 1994) including *Agrobacterium tumefaciens* mediated transformation, biolistic gene transfer using DNA coated microprojectiles, silicon carbide fibre-mediated gene transfer, and direct

gene transfer into protoplast. All of them require a long tissue-culture periods with the very low transformation rate which made them not so popular among molecular breeders. Recently improved and accelerated biolistic transformation and selection protocols have been developed alongside *Agrobacterium*-mediated perennial ryegrass transformation protocol which should make generation of transgenic lines much easier.

1.6.4 Morphology of ryegrass

Although the perenniality of grasslands is dependent on vegetative regeneration, the seasonal cycle of the crop is dominated by flowering. In spring, following a precise sequence of environmental conditions, the shoot apex of many mature established tillers switches from the production of leaf initials to that of floral parts, and the stem begins to elongate. The transition from vegetative to reproductive growth is closely associated with a number of changes in the basic physiology of the plant which, in turn, play a major role in bringing about high and sustained rates of dry matter production characteristic of the spring crop.

In terms of plant development, the aerial parts of ryegrass are produced by the apex positioned on the base crown a few millimeters above the ground and surrounded by the developing leaves (Fig 1.8A). During vegetative growth the apical meristem generates lateral meristems initially recognized as semicircular ridges along the main axis which become the leaf primordia. This morphological pattern does not change until the apex has been induced to flower by elevated temperatures and increasing day length.

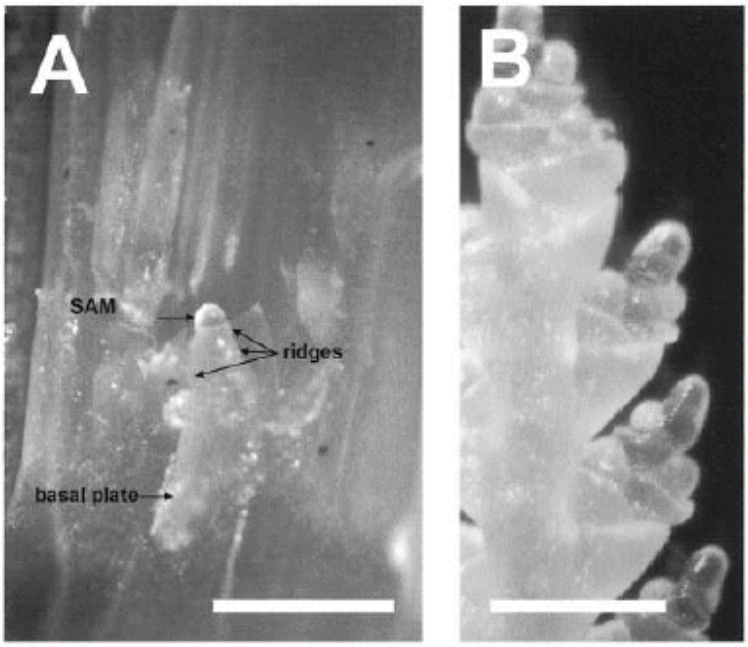


Fig 1.8 Shoot apical meristem (SAM) morphology of *Lolium perenne*.
A. Vegetative apex is compact with the SAM and the semicircular ridges that will give rise to leaves and tillers. **B**. The ryegrass inflorescence consists of spikelets alternately attached to the main axes. Each spikelet consists of 3-10 flowers. (from Jensen et al., 2001).

Upon transition to reproductive growth, the apical meristem and later the lateral meristems start to expand and eventually turn into groups of inflorescences (spikelets) each containing 3-10 floral meristems, attached alternately and directly to the central axis (Fig 1.8B). Each floret consists of four whorls of organs (Fig 1.9A). The outermost whorl consists of the palea and the lemma surrounding the lodicules (whorl 2), the three stamens (whorl 3), and the gynoecium (whorl 4) which is syncarpous, consisting of two or three carpels forming the ovary.

The flowers of the ryegrass inflorescence are arranged in a cymose pattern, always terminating apical growth with the production of the apical flower which represents determinate plant architecture (Fig 1.9B). In contrast, plants such as *Arabidopsis* have an indeterminate (racemose) inflorescence (Fig 1.9C).

The minimal requirement for flower induction in cold-responsive perennial ryegrass is a vernalisation period of 10-12 weeks below 5ºC (primary induction-PI), followed by secondary induction with LD photoperiod (16 h of light, 8 h of darkness) and temperatures above 20ºC.

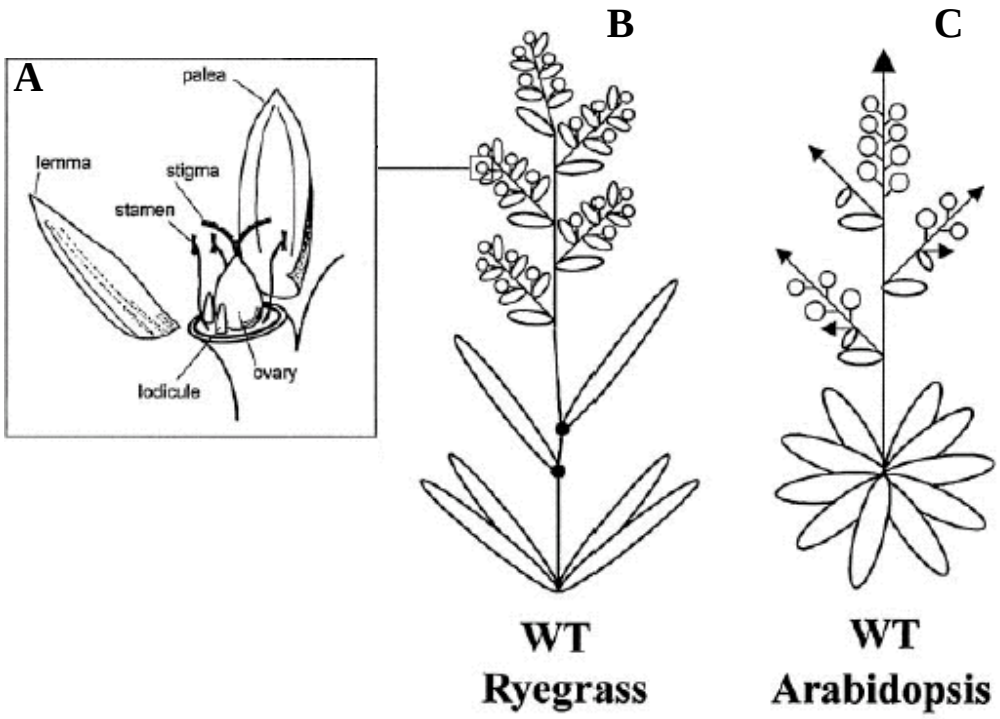


Fig 1.9 Comparative morphology of perennial ryegrass and *Arabidopsis*.
A. Each floret consists of 4 whorls of organs; pallea and lemma, surrounding lodicules, the three stamens, and the ovary. **B**. In ryegrass flowers develop in cymose (determinate) patterns. **C**. In *Arabidopsis* the flowers are arranged in racemose (indeterminate) pattern (from Jensen et al., 2001)

The requirement for PI varies greatly within perennial ryegrass. In general, the PI requirement increases with increasing latitude of origin of the germplasm; some plants of Mediterranean origin require only LD in order to flower (Aamlid et al. 2000). Furthermore, day length has no effect at low temperatures (>5ºC) during vernalisation. After vernalisation the critical number of LDs for secondary induction depends on the ecotype. Some Mediterranean ecotypes (Veyo) require only 1-3 days while the Scandinavian ecotypes (Falster and Kleppe) need 14-16 long days (Aamlid et al. 2000).

1.6.5 Molecular biology of flowering in perennial ryegrass

Last several years have seen substantial progress in understanding flowering time regulation in temperate grasses. Initial research of heading-date QTL in perennial ryegrass revealed major QTL on LG7 with high degree of synteny with the *Hd3a* region of rice chromosome 6 but due to the insufficient resolution it was not possible to determine if that was single QTL or number of tightly linked QTLs (Armstead et al., 2004). Further work identified *LpCO* (*LpHd1*) as one of the genes within this QTL and demonstrated that the physical order of genes can also be conserved at the microsyntenic level (Armstead et al., 2005). Comparative mapping showed that *L. perenne* and *Festuca pratensis* (meadow fescue) chromosome 7 has high degree of genetic synteny with rice chromosome 6 which is consistent with previous work where it was shown that barley *HvCO1* (the sequence with which *LpHd1* and *FpHd1* has the greatest homology), is collinear with the region of rice chromosome 6 that contains *Hd1* (Griffiths et al., 2003). Five QTL for the vernalisation response in perennial ryegrass was identified and mapped to 4 linkage groups which individually explained between 5.4 and 28% of the total phenotypic variation (Jensen et al., 2005). In the same study a major QTL on LG4 was found to co-localise with the marker *vrn-1*. *VRN1* and *VRN2* are the most important loci regarding the vernalisation response in diploid wheat (Tranquilli and Dubcovsky, 2000) mapping to the distal end of the long arm of chromosome $5a^{m}L$. By means of comparative mapping it was found that several regions of wheat chromosome 5 and oat chromosome 6 are syntenic to perennial ryegrass LG4 (Alm et al., 2003) supporting the hypothesis that the QTL for vernalisation response on LG4 could be an orthologue of the wheat *VRN1*.

The transcription of *LpVRN1* is induced during vernalisation (Andersen et al., 2006) with the expression patterns similar to diploid wheat (Yan et al., 2003). In wheat, *VRN2* maps closely to *VRN1* but only one QTL was found in this region in ryegrass raising question of whether the second QTL is present or simply not detected. Another two MADS-box genes (*LpMADS1* and *LpMADS2*), regulated by vernalisation, co-localise to the *VRN1* locus in ryegrass (Petersen et al., 2006). *LpMADS1* was found to be increasingly induced by cold exposure with putative MADS-box protein-binding site (CArG-box) possibly responsible for the vernalisation-regulated expression. Functionally it acts more like SOC1 from *Arabidopsis* forming dimers with the same partners as SOC1 as shown on yeast two-hybrid screen (Ciannamea et al., 2006). On the other side *LpMADS10* (*SVP-like* MADS-box gene) was

found to be increasingly repressed by cold exposure. In yeast two-hybrid assay two of them interacted together but when ectopically expressed neither affected flowering time significantly (Petersen et al., 2004). In search for *LpVRN2* Andersen et al. (2006) mapped two *LpVRN2* candidates (*LpVRN2_2* and *LpVRN2_3*) in close proximity to the QTL for vernalisation response on LG7 with the deduced amino sequences similar to *TmVRN2* of diploid wheat.

Contrary to *Arabidopsis* photoperiodic response in ryegrass has not generated lots of interest with only small number of papers published in this area. In perennial ryegrass two *CO* and *COL* genes have been described (Martin et al., 2004; Armstead et al., 2005; Ciannamea et al., 2006). *Lolium perenne CO* gene showed the highest identity with *LtCO* from *Lolium temulentum* and phylogenetically grouped in Ia *CO-like* group together with *LtCO*, rice *Hd1*, barley *HvCO1*, and wheat *TaHd1-1/3*. *LpCO* was shown to promote flowering and restore *Arabidopsis co-2* mutant to the *wt* flowering phenotype with expression patterns similar to *AtCO*, suggesting role in photoperiod signal mediation (Martin et al., 2004). *LpCO* was mapped on LG7 which has degree of conserved genetic synteny with rice chromosome 6 (Armstead et al., 2005). Using microarray approach another *COL* gene was described (Ciannamea et al., 2006). This gene designated *LpCOL1* is homologous to the *CO-like* gene family but appears to be different from *LpCO* by belonging to group III of these transcription factors. Strong up regulation towards the end of the cold period for *LpCOL1* suggest a role in vernalisation-induced flowering.

Lolium temulentum LtFT gene expression was observed in relation to the expression of GA in floral signaling where it was found that GA induced flowering without increasing *LtFT* expression indicating separate roles for GA and *LtFT* in floral signaling (King et al., 2006). In addition to this it was noticed that the response of *LtFT* to the LD exposure was almost immediate and within the first day of the exposure.

1.7 Aims of this thesis

The mechanism of the photoperiod response in plants is still not sufficiently understood. Although genetic variation in photoperiod and vernalisation response is well documented in crops, the identities of the majority of underlying genes remain unknown. Therefore the aim of this PhD thesis will be to identify, characterize, map, and analyse the function of the three genes of the photoperiod pathway (*GI*, *CO*, and *FT*) in ryegrass which should provide a better understanding of how plant development is controlled by environmental cues such as day length and temperature. Work on this segment of plant development should also help to reveal aspects of *Arabidopsis* model which are functionally conserved in ryegrass, and the changes accumulated due to evolutionary divergence.

Material and Methods are presented in Chapter II.

The analysis of all three genes are presented in the separate sections of the Chapter III starting with *LpGI* gene, followed by *LpCOL1*, and *LpFT3*. Detailed study included the amplification and characterization of the genes, expression analysis under various environmental conditions and gene mapping, described in consecutive sections. The final sections for analysis of each of the genes are dedicated to the functional analysis of the overexpressed ryegrass genes in *Arabidopsis* and expression analysis of the endogenous *Arabidopsis* genes in transgenic plants. Chapter III is dedicated only for the experimental setups and results obtained without extensive analysis.

Detailed discussion and integration of the results is presented in Chapter IV. Here results are compared with the other plant species, differences and similarities deduced, and a regulatory cascade of the ryegrass photoperiod pathway proposed. Experiments that would complement present study and further contribute to understanding of a photoperiod initiation of flowering are proposed at the end of this chapter within the concluding remarks.

2 MATERIAL AND METHODS

2.1 Materials

2.1.1 Chemicals, enzymes, oligonucleotides, and cloning vectors

Chemicals used for this work were purchased from Sigma-Aldrich (USA), Merck (Germany), Eppendorf (Germany), Gibco BRL (USA), and Invitrogen (Germany). Modifying enzymes and kits were purchased from Invitrogen (Germany), Dynal Biotech (Norway), and Clontech (USA). Oligonucleotides were synthesized at Invitrogen and Sigma-Aldrich. Cloning vectors used were pCR2.1, pENTR/D-TOPO® (Invitrogen), and pRSh1 (unpublished). *E. coli* DH5α and TOP10 competent cells were purchased from Invitrogen (Germany). *A.tumefaciens* was a kind gift from Dr Igor Kardailsky.

2.1.2 Buffers, solutions, and media

Standard buffers, solutions, and media were prepared as described in Sambrook and Russell (2001).

2.1.3 Bacterial strains

E. coli DH5α	*sup*E44 Δ*lac*U169 (*Φ*80*lacZ*ΔM15) *hsd*R17 *rec*A1 *end*A1*gyr*A96 *thi-1 rel*A1
E. coli TOP10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

A. tumefaciens GV3101:pMP90

2.1.4 Plant materials

Lolium perenne	cv. Impact A1662-1
Lolium perenne	cv. Samson GA66
Arabidopsis thaliana	Columbia (*Col*)
Arabidopsis thaliana	Landsberg (L*er*)
Arabidopsis thaliana	*gi-3* mutant in L*er* background (kindly provided by Assoc. Prof Joanna Putterill)
Arabidopsis thaliana	*co-2* mutant in L*er* background (kindly provided by Assoc. Prof Joanna Putterill)
Arabidopsis thaliana	*ft-1* mutant in L*er* background (kindly provided by Dr. Igor Kardailsky)

2.1.5 Database accessions

Accession numbers of protein sequences: *TaGI* (AF543844), *AtGI* (AJ133786), *MtGI* (ABE81212), *HvGI* (AY740524), *OsGI* (AP003047), *AtCOL3* (AC006585), *AtCOL6* (Al035679.1), *AtCOL9* (Z97338.2), *HvCO1* (AF490468), *HvCO2* (AF490469), *LpCO* (AAT42130), *TaHd1* (AB094490), *OsHd1* (AB001882), *AtCO* (X94937), *HvCO3* (AF490472), *OsB* (AB001887), *HvCO4* (AF490475), *OsC* (AAAA01001728), *HvCO7* (AY082963), *HvCO9* (AY082965), *OsF* (AAAA01022688), *HvCO8* (AY082964), *OsG* (AAAA01008321), *OsJ* (AAAA01000385), *OsN* (AB001888), *OsH* (AAAA01000838), *OsI* (AAAA01011539), *HvZCCT2* (AY485977), *TdZCCT1* (AY485979), *TmZCCT1* (AY485969), *HvZCCT2* (AY485978), *TdZCCT2* (AY485980), *TmZCCT2* (AY485975), *TaHd1-1* (AB094490), *OsHd3a* (ABO52943), *TaFT* (AAW23034) *HvFT* (DQ100327), *AtFT* (BAA77838)

Accession numbers of nucleotide sequences: *TaGI* (AF543884), *AtGI* (AJ133786), *OsGI* (AJ133787), *HvGI* (AF411229), *AtCO* (X94937), *OsHd1* (AB001882), *TaHd1-1* (AB094490),

Accession number of the sequences submitted from this study: *LpGI* (DQ534010), *LpCOL1* (DQ534011), and *LpFT3* (DQ309592).

2.1.6 Plasmids

Plasmid	Source	Resistance
pENTR/D-TOPO	Invitrogen	Kanamycin
pCR® 2.1-TOPO®	Invitrogen	Kanamycin
pRSh1	Unpublished	Spectinomycin

2.1.7 Antibiotics and herbicides

On LB plates the final concentration of antibiotics was 50 µg/mL for Kanamycin, and 100 µg/mL for Spectinomycin. BASTA selection on plants was performed with gluphosinate ammonium, with the final concentration of 240 µg/mL.

2.1.8 Primers

Table 2.1 Primers used for PCR reactions. In degenerate primers: S (G or C), M (A or C), R (G or A), K (G or T), V (G, C, or A).

No	Name	Sequence	Degeneracy
1	MG003	CCATGGGCSMGGCCRTGCGA	8 fold
2	MG005	GAACCATGGARCAGTAYYATARCTAC	16 fold
3	MG006	TGGATYGATGGKCTYGAGTTCTC	8 fold
4	MG008	TGATCCAWGGAGAYARWGGCCT	8 fold
5	MG010	CTCCATTRGCAACAGCCCATCT	2 fold
6	MG012	GCAATKCCTTGBGCAGTGGC	6 fold
7	MG017	CGCAAGGCTTCAACTGCTATTGT	N/A
8	MG018	ACAGATGGGATGCTTGTTGATGG	N/A
9	MG019	CAGACTGGGAAGTCATCCTTGAGC	N/A
10	MG020	GAGATGCCAGGGTCCTCAGGTACA	N/A
11	MG021	CTGAATCTGCATGGCATTCAAGCA	N/A
12	MG022	GTGCGTCACAGGTAGTACACAGTGCTGC	N/A
13	MG023	CGATGCCAAACGGTTAGCAGA	N/A
14	MG024	GATGCCAGTGCATTCCAGCTCCATAG	N/A
15	MG025	GAGCAATTCCCGGAGGATGTA	N/A
16	MG026	AGGCCCTTAACGAGTGACAGTCTACATTTGA	N/A
17	MG032	TCTCACAGGTTACACAGACACGCACTCGCTCAT	N/A
18	MG034	TGGCAAGTGAGCAGCAAGAGAGCACTTATGGAA	N/A
19	MG036	TGCACCTTTGGGCATTAGAAG	N/A
20	MG037	CTGCTGATACGGCCGCTGC	N/A
21	MG038	ATCATGTGCCTCAATGAACCTT	N/A

22	MG039	ACTGTGTCATCCATCGTGAACA	N/A
23	MG041	GGTATGGCTGGTGTTGTCAACATC	N/A
24	MG042	GCTCCAAGCTTCTCTCTCCTCG	N/A
25	MG045	TTCTCATCTCATTGCATCAGATT	N/A
26	MG046	CACCATGTCTGTCTCAAATGGGAAGTG	N/A
27	MG047	CACCATGTTAATGAATTGTGATTTCAAT	N/A
28	MG048	CTCTGCTAACCCAATTGCTCA	N/A
29	MG052	CTCCCGAAACACAGATGTCT	N/A
30	MG053	CTCTCCACATCTTCTGACCTT	N/A
31	MG054	CCTCCAACAACTGCTTACAG	N/A
32	MG056	CGTTGAATGGGGAGATCC	N/A
33	MG057	GGATCTCCCCATTCAACG	N/A
34	MG058	CACCATGGCCGGGAGGGATA	N/A
35	MG060	AGGAATACAGACTGCAAGAACAGG	N/A
36	MG061	TTTCCCAGAGTGAACACCTT	N/A
37	MG062	CCAGAAGAAGCTAACTGTGCT	N/A
38	MG063	CAGCTTCAAATATGGCTGTG	N/A
39	MG064	GGCATCGGATGGTAGGTGAT	N/A
40	MG065	ATCTCGAGCAATCCAGGGCT	N/A
41	MG066	GCTTCCGCCTGTAAAGTTGC	N/A
42	MG067	CAGCATCCCACTTACATCGC	N/A
43	MG068	GTCTTCGTAACATTCTTGCAAAGTA	N/A
44	MG069	TCAGCAACAGCAAGGCTCTT	N/A
45	MG070	TCAACATCAGGACCATGAGC	N/A
46	MG071	CCAGTAGCATAATCTTCGGCT	N/A
47	MG072	ACTCCAGCTCCATTGGCAAC	N/A
48	MG073	GTGTCGTTACATTCTGCAACTTT	N/A
49	MG074	ATCCTCCGGGAATTGCTCAC	N/A
50	MG086	GTGGTGCTAAGAAGAGGAAGA	N/A
51	MG087	TCAAGCTTCAACTCCTTCTTT	N/A
52	GIK16	CATGAAATGTGCAGAACAGG	N/A
53	GIK17	TGGTTGACTCTCCACATCTTC	N/A
54	GIK18	GACGTGCTGGACSCVTTCGTSMG	32 fold
55	GIK19	CGCCGGAGCCKGCCTCGCG	2 fold
56	GIK20	GACAACTGGTGCTTCCTTCGGGCA	N/A
57	GIK31	GACAACTGGTGCTTCCTTCG	N/A
58	GIK32	AAAAACTCATCAGCATCATCATTC	N/A
59	GIK40	TCATCGATCTTGTGTAGGTCTG	N/A
60	GIK55	CTCAAGGGCATTTTGGGTTA	N/A
61	GIK56	GCTGTATCCCCACTCGTTGT	N/A
62	GIK57	CAGTTCAACCAAAGCTGCTG	N/A
63	GIK58	AAGGAGTCCACCCTCCACTT	N/A
64	5'RACE	CGACUGGAGCACGAGGACACUGA	N/A
65	AP1	GTAATACGACTCACTATAGGGC	N/A
66	AP2	ACTATAGGGCACGCGTGGT	N/A

2.2 Methods

Ryegrass *Lolium perenne* cv. Impact was used in all experiments unless otherwise stated. For each time point three tillers (leaves) were collected. *Arabidopsis thaliana* ecotype Columbia (Col) and Landsberg (L*er*) seed, alongside *gi-3*, *co-2*, and *ft-1* mutants deficient in *GI*, *CO*, and *FT* expression respectively, were used in gene expression experiments in transgenic *Arabidopsis* plants.

2.2.1 Plant growing conditions

2.2.1.1 Diurnal expression experiment (CL & CS)

For the diurnal expression in long days (CL) and short days (CS) genetically identical (clonal) ryegrass plants were grown in standard soil in both glasshouse and growth cabinets. For the primary induction (vernalisation) plants were grown in a SD (8 h light : 16 h dark) growth chamber at 5ºC for 10 weeks. Following vernalisation plants were grown in the glasshouse in LD (16 h light : 8 h dark) or in SD at ~ 22°C. When grown in LD conditions the supplemental light was turned on between 1 am and 9 am so that the light period was between 1 am till the natural dusk (~5pm) with the dark period between 5 pm - 1 am. Collection was done after 7 days in respective light regimes, every two or three hours for 24 h. The samples were stored at -80°C before the RNA extraction. Total RNA extraction was done using the Trizol® protocol (Chapter 2.2.2.1.3). Reverse transcription or cDNA synthesis was performed with the SuperScript® (Invitrogen) reverse transcription system (Chapter 2.2.2.5). Real time PCR (RT-PCR) was performed in triplicate with SYBR-green used as a fluorescent reporter on the i-cycler (Chapter 2.2.3).

2.2.1.2 Free running experiment (FR)

Ryegrass plants were grown outside in standard soil, in the glasshouse, and in a growth chamber. Initially plants were vernalised outside for 8 weeks at an average temperature of ~6-8°C. Following vernalisation plants were transferred to the glasshouse in LD with the same conditions as described in Chapter 2.2.1.1. After 7 days they were transferred into the 22°C growth chamber with continuous light (LL). Collection was done every two hours for four

days. The samples were stored at -80°C before the RNA extraction. Expression analysis of the *LpGI* gene was done in triplicate as described in chapter 2.2.1.1.

2.2.1.3 Vernalisation and LD dependent gene expression (GV & LD)

In order to check the expression of the *LpGI* and *LpFT3* genes during and after vernalisation, plants were differentially vernalised between 0 and 10 weeks at 4°C. Upon vernalisation they were transferred into the LD glasshouse at 22°C. After 7 days under LD, samples were collected at ~ZT16 when the expression of these genes was thought to be the highest. Expression analysis of the *LpGI* gene was done in duplicate.

To check the expression pattern of the *LpFT3* gene regarding LD exposure, the plants were fully vernalised for 10 weeks at 4°C and then transferred into LD conditions. The sampling was done every day at ~ ZT16 for next 26 days. For both experiments the samples were stored in the -80°C freezer before being processed. Expression analysis was done in duplicate as described in Chapter 2.2.1.1.

2.2.1.4 Gene expression in transgenic Arabidopsis plants

Seeds of *Arabidopsis thaliana* (Col and L*er*), *gi-3*, *co-2*, and *ft-1* lines as well as transgenic *Arabidopsis* lines were soaked in water and stratified at 4°C for two days. Plants were grown in standard soil in LD growth chambers at 22°C. On average five plantlets from each line were used for DNA extraction and expression analysis. Plants were collected at two time points (ZT0, and ZT16) and stored at -80°C before RNA isolation.

2.2.2 Nucleic acid techniques

2.2.2.1 DNA and RNA isolation

2.2.2.1.1 Plasmid isolation

Plasmid DNA was routinely isolated by an alkaline lysis method. Briefly, 1.5 mL of over night (o/n) culture was pelleted by centrifugation at 12000 *xg* and the supernatant discarded. 100 mL of Solution 1 (50 mM glucose, 10 mM EDTA, 25 mM Tris pH 8.0) was added and briefly vortexed. 200 mL of Solution 2 (1% SDS, 0.2M NaOH) was added afterwards and incubated at room temperature for 5 m followed by 150 mL of Solution 3 (1.3M CH_3COOK

pH 4.8-5.2) and additional incubation for 5 m. Samples were then centrifuged for 10 m at 12000 *xg*, supernatant isolated, 0.6 V of isopropanol added and centrifuged again for 5 m. The precipitate was washed with 70% ethanol, dried and dissolved in 50 µL of TE+RNAse. *A. tumefaciens* plasmid isolation was performed using the following protocol; a single colony was grown o/n in 5 mL of LB with 50 µg/mL of kanamycin. 3 mL of cells were centrifuged on 12000 g for 45 s and the supernatant removed. 100 mL of solution 1 (50 mM glucose, 10 mM EDTA, 25 mM Tris pH 8.0) was added, and the sample vortexed and incubated for 5 m at room temperature. 20 µL of a 20 mg/mL lysozyme solution was added, vortexed and incubated an additional 15 m at 37°C. After adding 200 mL of solution 2 (0.2 N NaCl, 1% SDS,) reactions were incubated on ice for 5 m. In the next step, 150 mL of solution 3 was added (5 M CH_3COOK, 0.1 M CH_3COOH), vortexed and incubated 5 m on ice. The reactions were centrifuged for 5 m (12000 *xg*), the supernatant removed to a new tube and 400 mL of phenol/chloroform/isoamyl alcohol (25:24:1) added, followed by the centrifugation for 5 m. The supernatant was again transferred into new tube and 300 mL of isopropanol added, followed by centrifugation (5 m at 12000 *xg*). The pellet was washed with 70% ethanol, dried and dissolved in 50 µL of TE+RNAse

2.2.2.1.2 Plant DNA isolation

The plant DNA protocol was based on the CTAB protocol for plant DNA extraction. Leaf samples were ground with a mortar and pestle in the presence of liquid nitrogen. Approximately 100 mg of tissue was transferred into an eppendorf tube and 400 µL of extraction buffer (250 mM NaCl, 200 mM Tris pH8.0, 25 mM EDTA, 0.5%SDS) prewarmed to 65°C was added and the sample incubated for 10 m at 37°C. Samples were centrifuged for 2 m, the upper phase discarded and 0.6 V of isopropanol added and incubated at room temperature for 5 m. Samples were centrifuged for 10 m, and the pellet washed with 70% ethanol. The resulting DNA pellet was dissolved in 100 µL TE+RNAse.

2.2.2.1.3 Plant RNA isolation

For ryegrass expression analysis the three biggest tillers were collected at appropriate times and ground in mortars with liquid nitrogen using Retch (Germany) MM301 mixer mill. Approximately 100 mg of ground material was used for RNA extraction. For *Arabidopsis* expression plants were ground in eppendorf tubes with liquid nitrogen using plastic pestles.

Plant RNA was isolated using Trizol® reagent (Invitrogen) according to the manufacturer's instructions. Purification of poly A RNA from total RNA was done using the Dynabeads mRNA purification kit (Dynal Biotech, Norway), according to the manufacturer's instructions.

2.2.2.2 DNA digestion and ligation

Digestion and ligation of DNA fragments with restriction enzymes and ligases, respectively, were carried out according to the manufacturers instructions and in the buffers provided. For the cloning of ryegrass transforming constructs, digestions were performed using only the minimally required enzyme amounts and incubation times for complete digestion, as indicated by the enzyme manufacturer instructions. The fragments were separated on a 1% low-melting agarose gel (Invitrogen) and excised from the gel. Low melting agarose was melted by incubation at 65°C for 10 m, and ~2 µL aliquots of each fragment was used in a ligation reaction. For TOPO® TA and pENTR Directional TOPO® cloning (Invitrogen) the respective manufacturer's protocols were followed. For site specific recombination of the entry clone (pENTR/D) with a destination vector (pRSh1), Gateway™ LR Clonase™ Plus Enzyme Mix (Invitrogen) was used according to the manufacturer's protocol.

2.2.2.3 Polymerase chain reaction (PCR) amplification

Polymerase Chain Reaction (PCR) is a rapid and versatile *in vitro* method for amplifying defined target DNA sequences present within a source of DNA. Usually the method is designed to permit selective amplification of a specific target DNA sequence, or sequences within a heterogeneous collection of DNA sequences.

In this study PCR amplifications were carried out with either Platinum Taq polymerase (Invitrogen) or the TripleMaster™ PCR system (Eppendorf) on a Bio-Rad i-Cycler Thermal Cycler (Bio-Rad, USA).

Table 2.2 Reagents used in PCR reactions

Components	Volume	Final Concentration
10X PCR buffer, no Mg^{2+}	2 µL	1X
50 mM $MgCl_2$	0.6 µL	1.5 mM
25 mM dNTP mixture	0.16 µL	0.2 mM each
10 µM sense primer	0.4 µL	0.2 mM
10 µM anti-sense primer	0.4 µL	0.2 mM
Template DNA	1 µL	As required
Platinum Taq DNA Pol.	0.1 µL	1 unit
H_2O	15.34	n/a

Table 2.3 The thermal profile used for PCR reactions. The annealing temperature was set according to the primers used and extension was adjusted according to the length of the amplicon. The number of cycles were estimated after the PCR optimization procedure.

Thermal profile		
1	95°C	2 m
2	95°C	30 s
3	Annealing	30 s
4	72°C	Extension
5	Repeat steps 2-4	Cycles
6	72°C	5 m
7	12°C	∞

Primers for PCR were obtained commercially from Invitrogen and Sigma-Aldrich (Table 2.1). If not stated otherwise, the PCR reaction volume was 20 µL with the reaction components as presented in Table 2.2. Thermal profiles of the reactions are presented in Table 2.3.

2.2.2.4 Touchdown PCR

Touchdown is a simple method to optimize yields of amplified DNA when the melting temperature (Tm) of hybrids between oligonucleotide primers and their target sequences is not known with certainty. In touchdown PCR a range of annealing temperatures are utilized in a single PCR reaction. During the first two cycles the Tm is set 3-10ºC above the calculated

melting temperature. The annealing temperature is then reduced by one degree for each subsequent cycle until the temperature for specific priming is reached. Eventually the Tm will be reached where non specific priming can occur but by that time the specific amplification product will have such an abundance that it will actively suppress the accumulation of non-specific products. For *LpGI* 5'RACE-PCR, a Tm of 63ºC was used as a starting point for 3 cycles, followed by 15 cycles in which the Tm dropped 1ºC every cycle. After that, an additional 30 cycles were performed with a Tm of 51ºC. The extension time was 70 s throughout the PCR reaction. For *LpCOL1* 5' gene walking the PCR starting Tm was 69ºC for 5 cycles in which the Tm dropped 0.5ºC every cycle, followed by an additional 30 cycles with a Tm of 67ºC. Extension time was 3 m throughout the reaction. For *LpCOL1* 3' gene walking the PCR reaction starting Tm was 69ºC for 10 cycles in which the Tm dropped 0.5ºC every cycle, followed by an additional 30 cycles with a Tm of 67ºC. The extension time was 2 m 30 s.

2.2.2.5 cDNA synthesis

For cDNA synthesis the Thermoscript™ rt-PCR two-step system (Invitrogen) was used according to the manufacturer's instructions. RNA used for the cDNA synthesis was obtained as described in Chapter 2.2.2.1.3 and equalized to similar concetrations by adding DHPC-treated water. Random hexamer primers were used for the synthesis. After synthesis the cDNA was diluted with 50 µL of water.

2.2.2.6 RACE-PCR

Rapid amplification of cDNA ends (RACE-PCR) is the method used to obtain full length 5' and 3' ends of cDNA using known cDNA sequence. In this work the GeneRacer™ kit (Invitrogen) which is based on RNA ligase mediated, and oligo-capping rapid amplification of cDNA ends, and results in the selective ligation of an RNA oligonucleotide to the 5' ends of decapped mRNA using T4 ligase was used. The reactions were performed according to the recommended protocol.

2.2.2.7 Genome walking

Genome walking is a method for deducing upstream or downstream sequence in DNA from a known sequence such as a cDNA. This method utilizes "libraries" of uncloned, adaptor

ligated, genomic DNA fragments. The libraries are DNA families digested with different restriction enzymes that recognize a 6-base site leaving blunt ends ready to be ligated to the genome walker adaptor. In this work the GenomeWalker™ kit (Clonetech) was used. Enzymes used for the preparation of the DNA libraries were *DraI, EcoRV, PvuII, StuI*, and *SmaI/HpaI/HincII*. For the 5' genome walking experiment on *LpCOL1,* AP1, and nested AP2 primers were used, in combination with the gene specific MG021 and MG023 primers, respectively. For the 3' end the AP1/MG020 primer pair was used according to the recommended protocol.

2.2.2.8 Nucleic acid sequencing

DNA sequencing is the determination of the deoxynucleotide order of a specific segment of DNA. In this study DNA sequences were determined using the Bio-Rad i-Cycler Thermal Cycler (Bio-Rad), and the BigDye® Terminator Cycle Sequencing Kit (ABI), according to the manufacturer's instructions. The fragments were separated with the ABI Prism® 3100 genetic analyzer (ABI), using either the 22 cm, or 36cm array depending on the length of the template. The results were analysed with the ABI Prism sequencing analysis software V 3.1 (ABI).

2.2.2.9 Sequence analysis

Standard sequence analysis was performed using components of the Vector NTI Suite 9 (Invitrogen). Database searches were routinely carried out using the BLAST algorithm at GeneBank (http://www.ncbi.nlm.nih.gov/BLAST/) Sequences were aligned using ClustalW in AlignX, a component of Vector NTI suite (Invitrogen). Separate sequences were aligned into contigs using Contig Express, another component of the Vector NTI suite. Phylogenetic trees were calculated from the alignments using the Neighbor Joining method; construction and bootstrapping of the trees was performed using programs from the Vector NTI suite. Transmembrane domains in the *LpGI* gene were determined through the TM prediction programs TMpred (http://www.ch.embnet.org/software/TMPRED_form.html), and TopPred (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html). Localisation of the LpGI protein was done using PSORT (http://psort.nibb.ac.jp), a computer program for the prediction of protein localization sites in cells. The *LpCOL1* nucleotide sequence was analysed for

transcription factor binding sites using an in-house version of the TRANSFAC® professional database version 9.2.

2.2.2.10 dCAPS analysis

dCAPS (Neff et al., 1998) is a PCR-based detection of single nucleotide polymorphisms where a cleaved amplified polymorphic sequence is used for the detection of single nucleotide polymorphisms. However, this technique is not limited to mutations or SNPs which create or disrupt a restriction enzyme recognition site. Modification of this technique was designed where mismatches in a PCR primer are used to create a polymorphism based on the target mutation. This technique is useful for following known mutations in segregating populations and genetic mapping of isolated DNAs used for positional based cloning of new genes. For the *LpCOL1* mapping the MG048/041 primers which amplify genomic segment of 1358 bp were used. One of the SNPs was within the *MboI* restriction site at position 201 of the amplicon. The complete amplicon contained 5 *MboI* sites producing restriction digest segments ranging from 394 to 23 bp. Absence of an *MboI* site at position 201 (second allelic variation) produced an additional band of 499 bp used for allelic discrimination.

2.2.3 Quantitative RT-PCR

The real-time PCR system is based on the detection and quantitation of a fluorescent reporter. The signal increases in direct proportion to the amount of PCR product in a reaction. By recording the amount of fluorescence emission at each cycle it is possible to monitor the PCR reaction during the exponential phase where the first significant increase in the amount of PCR product correlates to the initial amount of target template. In this study SYBR Green as a fluorescent reporter in a final dilution of 1:75000 was used. In addition fluorescein (BioRad) was used in all wells in a final concentration of 10nM in order to generate the well factors (these are internal calibration values that the iCycler uses for each well for each reaction). The fluorescein generates sufficient signal for the iCycler to measure fluorescence from the well but not enough signal to interfere with the reaction. The PCR reaction volume used was 20 μL with the PCR components as presented in Table 2.4. Thermal profiles of the reactions are presented in Table 2.5.

Table 2.4 Reagents used in RT-PCR reactions

Components	**Volume**	**Final Concentration**
10X PCR buffer, no Mg	2 µL	1X
50mM $MgCl_2$	0.8 µL	2.0 mM
25mM dNTP mixture	0.16 µL	0.2 mM each
10 µM sense primer	0.4 µL	0.2 mM
10 µM anti-sense primer	0.4 µL	0.2 mM
CYBR Green	0.2 µL	1:75000
Fluorescein	0.2 µL	10 nM
Template DNA	1 µL	as required
Platinum Taq DNA Pol.	0.1 µL	1 unit
H_2O	14.74	n/a

Table 2.5 Thermal profile of the RT-PCR reactions. The annealing temperature was set according to the primers used. Extension was adjusted according to the length of the amplicon. In step 5 the number of cycles were estimated after the PCR optimization procedure. In step 7 the temperature dropped 0.5°C every cycle for 82 cycles in order to obtain the melting profile.

Thermal profile		
1	95°C	2 m
2	95°C	30 s
3	Annealing	30 s
4	72°C	Extension
5	Repeat steps 2-4	Cycles
6	95°C	1 m
7	95°C ↓0.5°C	10 s
8	Repeat step 7	X 82
9	12°C	∞

The results were analysed using an in house excel spreadsheet. This spreadsheet analysed raw, background subtracted data determining the threshold (Ct) value as the interpolated cycle number when the rate of increase of fluorescence is maximum (infliction point), which should correspond to the end of the exponential phase. Steps 7 and 8 were included in order to obtain the melting profile of the amplicon.

The software package Quantity One (BioRad) that comes with the system uses a maximum curvature approach where the fluorescence has to be above a certain arbitrary threshold in order to determine the Ct value. Standard curves were generated from a dilution series constructed from a reference sample using the same primers as those used in the respective RT-PCR. The results were normalized using one of the house keeping genes; ubiquitin (GIK57/58, Tm 60°C, extension time 20 s, 45 cycles) and GAPDH (GIK55/56, Tm 60°C, extension time 20 s, 45 cycles) for ryegrass and ubiquitin (MG086/087, Tm 55°C, extension time 20 s, 45 cycles) for *Arabidopsis* expression analysis. The normalization was done by dividing the copy number of the sample by the copy number of the house keeping gene from the respective sample. Every sample was run in triplicate, if not stated otherwise, from which the mean value was derived. House keeping genes were run in duplicate. Only samples which showed the normal sigmoid amplification curve were considered. In all other cases they were excluded from the calculations. Generally, results showed a high level of reproducibility. Occasionally some samples would differ from the other two by an order of magnitude or more, most likely due to the inaccurate pipetting. In these cases such results would be discarded and only two replicates used to calculate the mean value. Negative controls containing no DNA were run alongside the samples. The standard deviation (SD) was obtained using the same set of samples.

2.2.4 Transformation and selection procedures

2.2.4.1 Bacterial transformation and selection

Chemically competent DH5α and TOP10 *E. coli* cells were purchased from Invitrogen and transformed according to the manufacturer's protocol. Chemically competent *Agrobacterium tumefaciens* cells GV3101:MP90 (gifted by Dr. Igor Kardailsky) were transformed as follows; competent *A. tumefaciens* bacteria were kept at -80°C and thawed on ice (100 mL per reaction) immediately prior to use. Approximately 1 µg of DNA was added, incubated on ice for 5 m and transferred into liquid nitrogen for 5 m. Following this, the cells were incubated at 37°C for 5 m, 1 mL of LB added, and the tubes rolled for 4 h. A 100 µL of cells were plated onto LB plates containing the appropriate antibiotic.

2.2.4.2 Arabidopsis transformation and selection

The plasmids pRSh1-*LpGI*, pRSh1-*LpCOL1*, and pRSh1-*LpFT3* were transformed into *A. tumefaciens* strain GV3101 : pMP90 using a modified floral dip procedure (Clough and Bent, 1998). Briefly, a single colony resistant to 50 µg/mL kanamycin was inoculated for preculture in liquid LB medium (Sambrook and Russell, 2001) supplemented with the same antibiotic. The presence of the *LpGI*, *LpCOL1*, and *LpFT3* inserts within the pRSh1 plasmid was confirmed by PCR amplification using respective primers. 1 mL of the preculture aliquot was used to inoculate 300 mL of YEB medium (Sambrook and Russell, 2001) supplemented with 50 µg/mL of kanamycin. Cultures were grown o/n in the shaking incubator (200 rpm) at 28°C. When the cultures reached an OD_{900nm} they were transferred into 50 mL plastic tubes, centrifuged at 4°C for 10 m and pellet dissolved in 25 mL of 2% sucrose solution. Prior to transformation, 0.001% of Silwet copolymer L-77 (GE silicones), acting as a silicone surfactant was added. No less then 10 *Arabidopsis* plants were transformed using following protocol: *Arabidopsis* plants were grown in LD conditions until flowering when the first bolts were clipped to encourage proliferation of many secondary bolts. Eventually transformed plants had many immature flower clusters and not many fertilized siliques. Above ground parts of plants were dipped in the *Agrobacterium* solution for 5-10 s with gentle agitation. The dipped plants were placed under the cover on their side so that the solution did not reach the soil and allowed to remain there for 48 h. Afterwards they were watered and grown normally and the dry seed collected. The collected seed was sown again and transformants selected on the basis of BASTA resistance. On average, transformation showed a success rate of 1%. Transformed plants were self pollinated to generate T2 populations that segregated for the transgenes. T2 lines were again selected for BASTA resistance. Lines with only one insert were characterized by 3:1 (live/dead) segregation ratio and only those lines were carried into the next generation by self pollination. In T3 generation, BASTA selection was again used to isolate non-segregating, homozygous lines. These were further characterised for the leaf count. The characterized lines were also checked for the inserts using PCR.

2.2.5 Phenotypical analysis of transgenic *Arabidopsis* plants

Prior to sowing, *Arabidopsis* seeds were kept at 4°C for two days in order to break seed dormancy. Germination was induced by incubation for 5 days with slightly elevated humidity

under the LD conditions (16 L:8 D). After germination the pots were weeded out so that ~ 10 plants per line was left to be analysed. The total number of leaves (rosette and cauline) was counted until appearance of the inflorescence bud. Average flowering time and standard deviation were calculated for each line. The growth of the plants was documented photographically.

2.2.6 Mapping procedure

SNPs comprise the largest set of sequence variants in most organisms and mapping of the ryegrass photoperiod genes was done based on SNPs. In order to facilitate SNP search within the mapping population a segment of the gene was amplified from gDNA of parents (Impact/Samson) of the mapping population using GSP primers. The amplified segments were cloned into the pCR2.1 TOPO vector and used to transform DH5α cells. Transformants were selected and on plates containing kanamycin antibiotic. Up to ten colonies from each genotype were each grown overnight, followed by DNA miniprep and insert sequencing. Alignment of the amplified regions revealed candidate SNPs which could be suitable for the mapping experiment (heterozygous in one or both parents). After determining SNP positions within the parent amplicons, the same primers were used to amplify genes of interest within 188 mapping population genotypes. The allelic status of the genes was determined by direct sequencing of amplification products with respective oligonucleotides. Sequences were analysed on an ABI 3100 Genetic Analyser (Chapter 2.2.2.8) using a 22 cm capillary array with POP-7™ polymer (Applied Biosystems). Electropherograms were analysed using ABI Prism GeneScan (v 3.7, Applied Biosystems), and genotype data was scored using Mutation Surveyor™ (Softgenetics). In order to fully utilize the mapping population and to check the obtained results from the direct sequencing, the dCAPS method was devised for *LpCOL1* mapping (Chapter 2.2.2.10).

2.2.7 Ryegrass genetic map (previous work)

Through the work of a separate group within AgResearch (Faville et al., 2004), a genetic linkage map was constructed for perennial ryegrass (*Lolium perenne* L.) using 165 EST-SSR markers (simple sequence repeat markers derived from expressed sequence tags) in population RM2, a full-sib F1 population (n=188) developed from a pair cross between

individual heterozygous genotypes from cv. 'Grasslands Impact' and cv. 'Grasslands Samson'. This population structure is sometimes referred to as a two-way pseudo-testcross. Genetic linkage analysis was conducted with JoinMap® 3.0 software (www.kyazma.nl), following procedures for cross pollinator (CP) population type. Initially separate parental maps were constructed using dominant marker types. These maps were aligned to verify map positions of markers common to both parents, following which the parental datasets were combined to construct a single bi-parental consensus map using a mixture of dominant and co-dominant marker types. This map served as the framework for mapping the *LpGI, LpCOL1,* and *LpFT3* genes, as a single nucleotide polymorphism (SNP) marker.

2.2.8 Comparative mapping

Rice orthologues of perennial ryegrass ESTs mapped as SSR and STS markers were identified as tentative consensus (TC) sequences by BLASTn analysis against the TIGR rice gene index (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=rice). Criteria used were: E value $<e^{-20}$; sequence identity >85% over >100 bp alignment length. Rice genome positions for the rice TCs were identified on the TIGR rice pseudomolecule assembly (http://www.gramene.org/Oryza_sativa/index.html).

3 RESULTS AND DISCUSSION

The aim of this thesis was to isolate and characterise three ryegrass orthologues of *Arabidopsis GI*, *CO*, and *FT* genes. Each gene was separately analysed. The first sections describe the search through the ryegrass EST database and multiple alignments of respective orthologues genes from other species followed by the sequence analysis and genetic mapping of the isolated genes. After that expression of each gene was analysed under various environmental conditions and their regulation by the circadian clock was determined. Subsequent functional analysis was performed by overexpressing ryegrass genes in relevant *Arabidopsis* mutant and wt lines. Finally expression of relevant *Arabidopsis* endogenous genes was checked in the transgenic lines overexpressing ryegrass genes to establish a mode of action within the ryegrass photoperiod pathway.

3.1 Plant varieties used and their phenotypic characterization

Two main cultivars of ryegrass (*Lolium perenne*), cv.”Grassland Impact” and cv.”Grassland Samson”, were used in this work. All the plants analysed in this thesis were clones of the A1662-1 Impact and GA66 Samson cultivars (Chapter 2.1.4). Phenotypic characterization of the ryegrass plants confirmed their need for vernalisation and long day induction as a prerequisite for the subsequent transition to flowering (joint effort of myself and the whole flowering time group at AgResearch; Fig 3.1). It was established that 8 weeks of vernalisation followed by 40 long days will cause >75% of *Lolium perenne* cv. Impact plants to flower. It was also shown that transition to flowering is a stochastic process to some degree with the majority of plants flowering after 8 weeks of vernalisation period and 40 long days but at the same time some of the clonal plants under the same conditions flower much earlier and some

of them do not flower at all. This observation will surface again in some of the expression experiments with the three photoperiod genes in the later parts of this chapter.

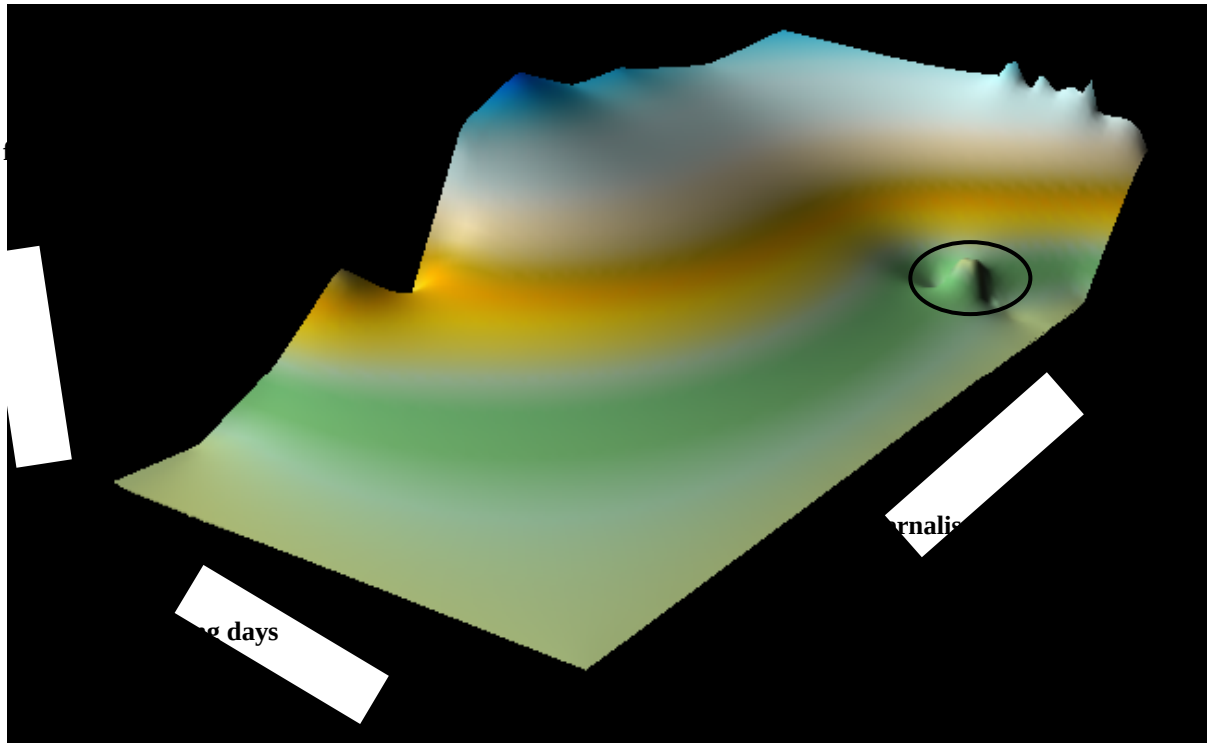


Fig. 3.1 Effect of vernalisation and LDs on the transition to flowering in *Lolium perenne* cv.Impact Plants were differentially vernalised (1-10 weeks), and transferred into the LD conditions for different number of days (0-40). After that the plants were transferred back into the SD conditions and heading date (flowering) was measured. All three sets of data (number of LD, duration of vernalisation, number of floral tillers) were plotted on the chart. Different shades of yellow orange and blue represent points with the higher floral tiller percentage. Circled area in the chart represents the deviation from the expected heading date (chart produced by Dr Igor Kardailsky, AgResearch, Palmerston North, NZ).

For the expression analysis experiments, cultivar Impact was used. Table 3.1 lists the growth conditions and collection times of the plants used for the analysis (a detailed description can be found in chapter 2.2.1). Overall, the expression of 27 flowering related genes was checked under the different environmental conditions. Over 180 time points were analysed with more than 2000 copy measurements produced using real time PCR (RT-PCR) method as described in chapter 2.2.3. This was a joint effort of myself and the flowering time group at AgResearch, Palmerston North.

Table. 3.1 Growth conditions and the harvesting regime for the ryegrass plants used for expression analysis. **CL** Diurnal expression-long days, collection done on LDs 1 and 8. **CS** Diurnal expression-short days, collection done on SDs 1 and 8. **CV** Circadian expression-vernalisation series, collection done on the first and last day of vernalisation. **FR** Free running, collection started after transfer to the LL conditions. **GV** Vernalisation series-general, collection done on the day of transfer from SD to LD and on days 3 and 7 of the LD treatment. **LD** Long day series, collection was done from the day the plants were transferred from SD to LD condition and during the next 28 days.

Conditions	Vernalisation(weeks)	Short days	Long days	Collection
CL	10	7	1 & 8	Every 3 h for 24 h
CS	10	1 & 8	0	Every 3 h for 24 h
CV	0 & 8	0	0	Every 3 h for 24 h
FR	12	5	7	Every 2 h for 72 h
GV	0-10	7	0, 3, 7	Just before dusk
LD	*8*	*7*	*0-28*	*Just before dusk*

In the following sections, some of the expression results from the Table 3.1 linked with and relevant to the three genes from the photoperiod pathway in ryegrass, namely *LpGI*, *LpCOL1*, and *LpFT3* are described and analysed. Observed results were subsequently linked to the results from gene mapping and functional analysis in order to produce detailed picture of the photoperiod pathway in *Lolium perenne*.

3.2 *GIGANTEA* (*LpGI*)

3.2.1 Sequencing and characterization of the *LpGI* gene

The search through the EST database was a joint effort between Dr Igor Kardailsky and Milan Gagic. The reminder of the work was done solely by Milan Gagic. The initial search of the AgResearch *Lolium perenne* EST database (Sawbridge et al., 2003) revealed one clone (GI-1, EST#14RGB5271HX2), with a DNA sequence resembling the *Arabidopsis GI* gene. The GI-1 clone had an 800 bp insert which showed significant similarity with two different segments from the 3' UTR regions of the *Arabidopsis GI* gene as well as wheat, rice and barley *GI* genes from Genbank (Appendix, Fig 5.1). It is likely that this is a fusion of two separate segments of *LpGI* cDNA, and occured during the cloning of the ryegrass EST library. Since the *GI* gene itself is very long, trying to recover the rest of the gene by RACE-PCR from ryegrass cDNA would be very difficult.

Therefore, in order to obtain the 5' segment of the gene, after aligning *GI* genes from the different species (*Arabidopsis*, wheat, rice, and barley), several degenerate primers were designed with low degeneracy (MG006, MG008, and MG010, Figure 3.2, Table 2.1) as well as gene specific primer (GSP) from the 3' end of the *LpGI* gene (MG019, Fig. 3.2, Table 2.1). Out of several different primer combinations two of them, MG006/MG008, and MG006/MG010 (Fig 3.2, Table 2.1) with Tm of 52ºC and 57ºC, respectively, and extension time of 90 s, over 40 cycles, produced the most distinctive bands (Figure 3.3). The selected bands were gel purified and TOPO-TA cloned into the pCR2.1 vector for further characterization (Chapter 2.2.2.2)

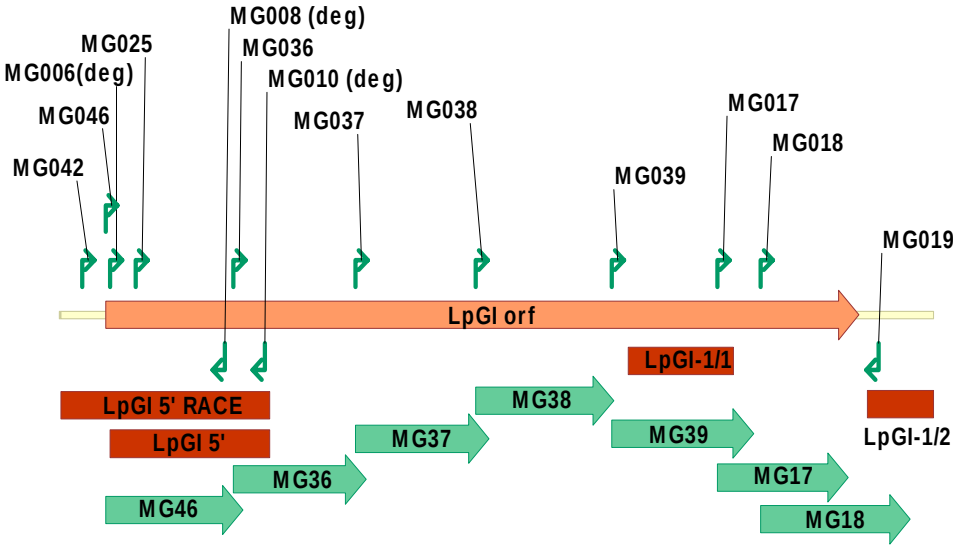
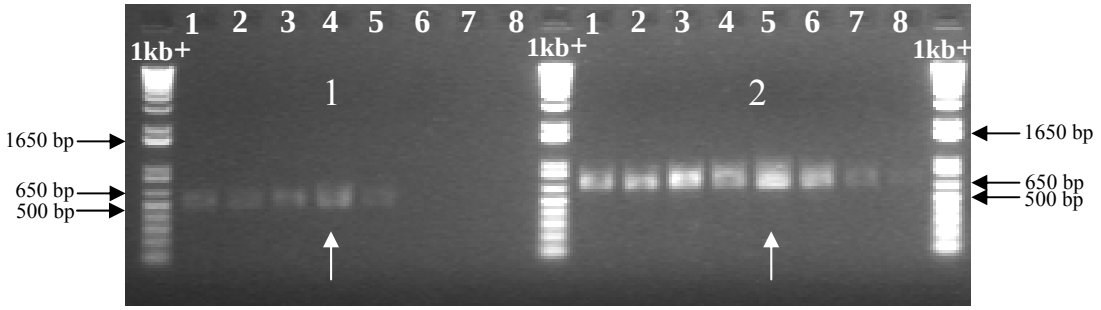


Figure 3.2 Diagram of the *LpGI* gene
Position of the EST mapped segments (LpGI1/1and LpGI1/2), 5'GI segment from the degenerate primers (LpGI 5'), and LpGI 5'RACE segment from the RACE-PCR reactions are labeled red. Full sequence was obtained using primers MG046, MG036, MG037, MG038, MG039, MG017, and MG018 (long green arrows). Short green arrows indicate position of the used primers.

Sequencing was performed according to the recommended protocol (Chapter 2.2.2.8) and as expected, the insert (*LpGI*5') was found to be similar to the 5'end of the orthologous *GI* genes from other species (data not shown).



. **Figure 3.3** Degenerate PCR reaction on *LpGI L. perenne* cDNA
Primer pairs MG006/MG008 (1), and MG006/MG010 (2), were used with a temperature gradient of 45-65ºC (lanes 1-8), and 90 s extension time, over 40 cycles, run alongside 1kb+ DNA standard. White arrows show the bands used for subsequent cloning and sequencing. The samples were run alongside 1kb+ DNA marker

Since ~30 bp was predicted to be missing from the 5' end of the gene and around 2 kb between the 5' and 3' ends, 5' RACE-PCR was used to recover the missing 5' segment and to design additional primers to amplify and sequence the missing region between the 5' and 3' ends. Ryegrass total RNA was isolated from the ryegrass cv. Impact using Trizol reagent as described in Chapter 2.2.2.1.3. The GeneRacer Kit was used to obtain a full-length *LpGI*

cDNA (Chapter 2.2.2.6). Two types of cDNA were produced with regards to the primers used to obtain cDNA: oligodT cDNA, and random hexamer cDNA. 5'RACE-PCR was performed using several GSPs and degenerate oligos (antisense primers) from the known 5' region of *LpGI* as well as the 5'RACE primer supplied by the manufacturer (sense primer) using random hexamer cDNA as a template. Touchdown PCR was performed with the four primer pairs using the conditions as described in the Chapter 2.2.4. Only primer pair MG010/5'RACE (Table 2.1) produced a single distinctive band (Fig 3.4) which was used for subsequent TOPO-TA cloning (data not shown).

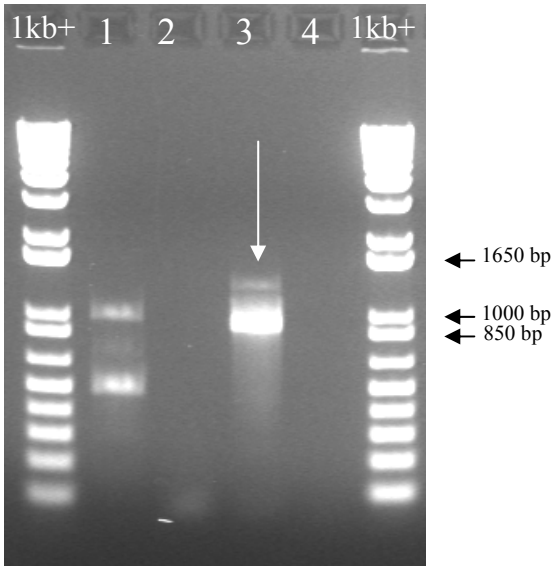


Figure 3.4 *LpGI* 5' touchdown 5'RACE-PCR to obtain the 5' end of the gene and corresponding UTR.
Four sets of primers used: 1. MG026/5'RACE; 2. MG008/5'RACE; 3. MG010/5'RACE; 4. MG012/5'RACE, and only primer set 3 produced a distinctive band (white arrow). The band was gel purified and cloned into the pCR2.1 vector. 1kb+ DNA size standard.

The 5'RACE insert was sequenced and upon aligning produced a contig with the previously sequenced *LpGI* 5' extending into the 5'UTR region of the *GI* gene (*LpGI* 5' RACE, Fig 3.2). The region between 5' and 3' end of the gene was amplified using GSP MG025/MG019 (Figures 3.2 and 3.5A; Table 2.1) which produced ~ 3.5 kb band (Tm 67ºC, extension time of 4 m 30 s, 40 cycles).The band was cloned and sequenced using 14 different sequencing primers (MG019, MG037, MG063, MG064, MG065, MG066, MG067, MG068, MG069, MG070, MG071. MG072, MG073, and MG074; Table 2.1), and the contig assembled (Appendix, Fig 5.2). A protein sequence of 1149 aa was deduced from the nucleotide sequence (Appendix, Fig 5.3).

After the successful sequencing of the *LpGI* cDNA an effort was made to determine the introns and their positions within the *GI* gene. *LpGI* gene was amplified from the Impact

ryegrass gDNA using primers MG042/019 (Fig 3.5B) and the product of ~5.7 kb TOPO-TA cloned (data not shown). MG primers 046, 025, 036, 054, 061, 037, 062, 038, 039, 052, 053, 018, 057, 056, and 019 (Table 2.1) were used to sequence the genomic fragment.

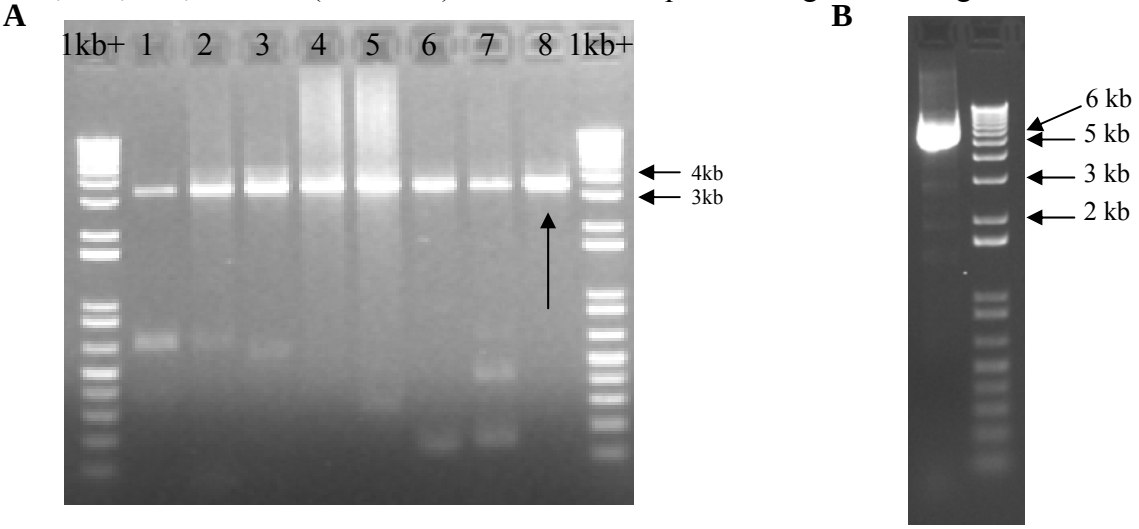


Figure 3.5 PCR amplification of the *LpGI* cDNA and gDNA
A. For the cDNa amplification, primers MG025/MG019 were used with temperature gradient (Tm 52-67ºC, lanes 1-8, Extension time 4 m 30 s, 40 cycles). The temperature of 67ºC (lane 8) produced single distinctive band (black arrow) subsequently used for the cloning experiment **B**. Primers MG042/MG019 were used on the ryegrass gDNA template under the following conditions: Tm 68ºC, extension 10 m, 35 cycles. 1kb+ DNA size standard.

3.2.1.1 LpGI sequence analysis

The *LpGI* cDNA has an ORF of 3447 bp, encoding 1149 amino acids (GeneBank-DQ534010). At the protein level it shows 89.6% identity with the GI of wheat, 87.6% identity with rice, and 65.8% identity with *Arabidopsis*. At the nucleotide level the identity is 91%, 86%, and 67% respectively.

Comparison of the genomic and cDNA *LpGI* sequences revealed 12 introns and 13 exons spanning the region of 5497 bp (Fig 3.6). At the genomic level the overall structure of the *LpGI* gene is similar to the *OsGI*, and *AtGI* genes in rice and *Arabidopsis* except for the fact that the rice *GI* has 13 introns contrary to the *Arabidopsis* and ryegrass *GI* which have 12 (Fig 3.5).

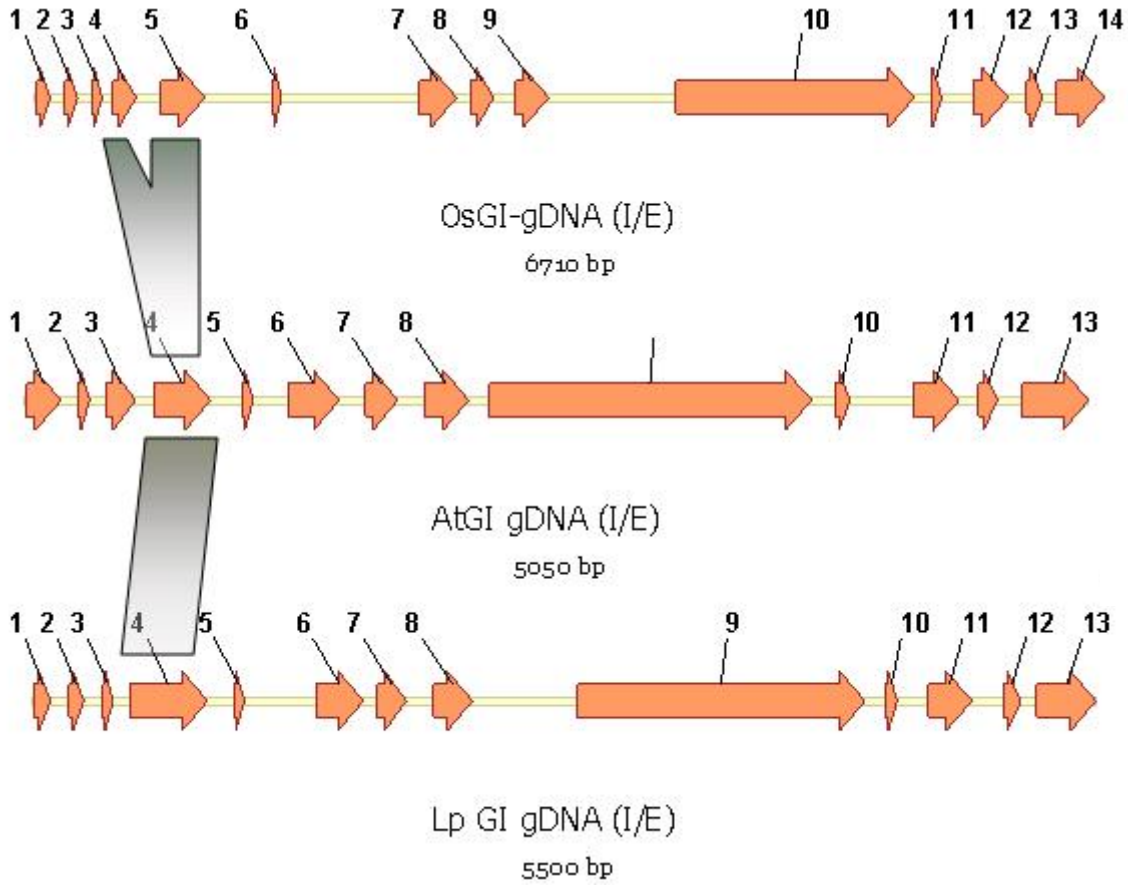


Figure 3.6 Comparison of the genomic DNA between rice *Arabidopsis* and ryegrass. The genomic segment of the *LpGI* gene is made up of 13 exons and 12 introns spanning 5497 bp. Orthologous genes in *Arabidopsis* and rice have 12 and 13 introns respectively.

Web-based membrane topology prediction programs (TMpred and TopPred, Chapter 2.2.9) predicted that the LpGI protein contains up to 7 transmembrane domains. Four domains with the highest probability score were also conserved in the rice and *Arabidopsis* proteins (Fig 3.7). LpGI is also predicted to be a nuclear protein (PSORT prediction server, Chapter 2.2.9).

In summary, a cDNA and genomic sequences from *LpGI* were aligned and compared. The DNA sequence was confirmed by aligning overlapping sequence on both strands and from using clones from independent PCR reactions to avoid problems of PCR-induced errors in the DNA sequence.

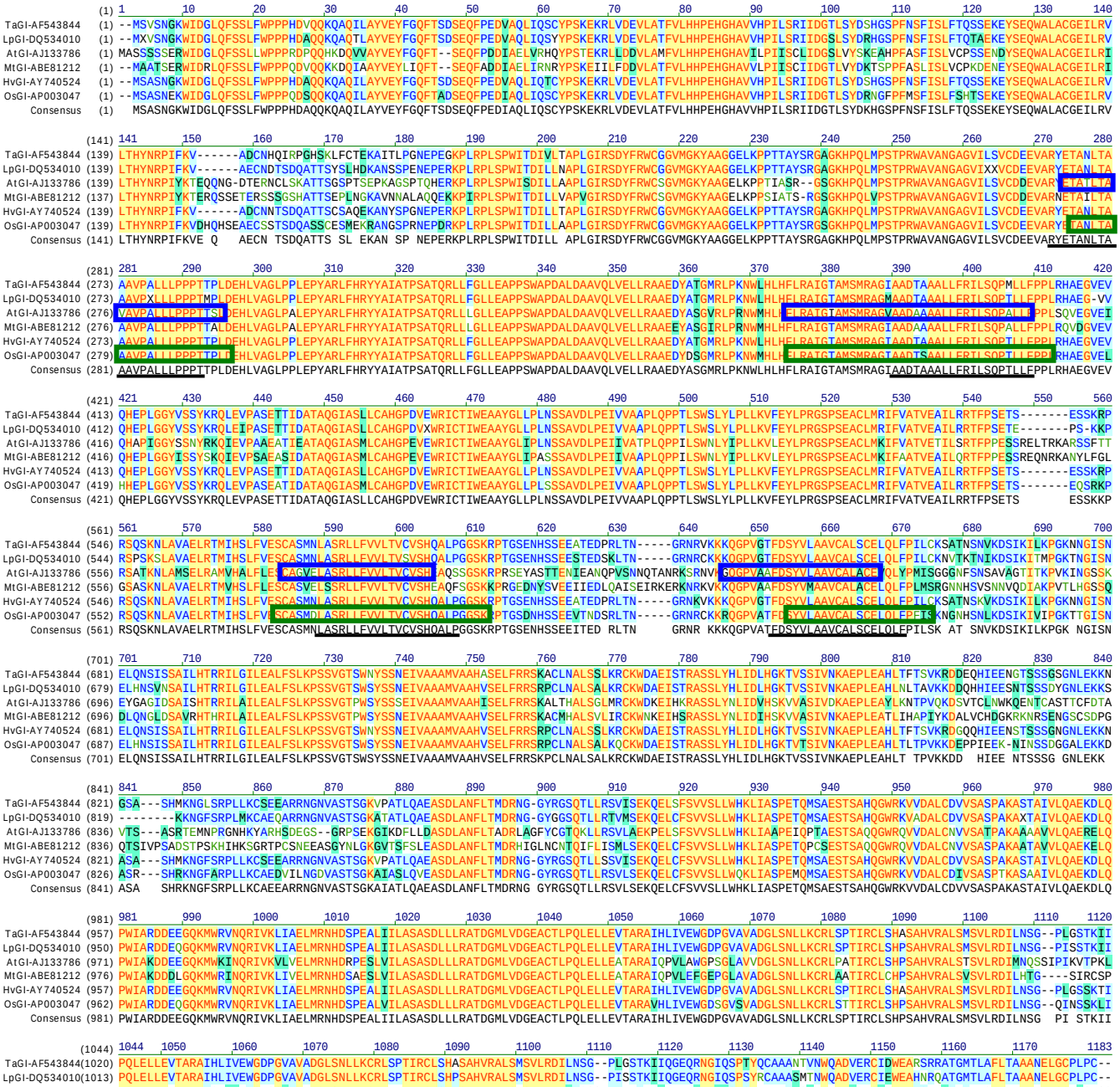


Figure 3.7 GI amino acid sequences similarity and transmembrane domains Ryegrass (Lp), wheat (Ta), barley (Hv), *Arabidopsis* (At), medicago (Mt), rice (Os). Regions in blue squares are transmembrane domains in *Arabidopsis*, green squares are the same regions in rice. Underlined segments are predicted transmembrane domains in ryegrass determined through the TM prediction programs TMpred and TopPred (Chapter 2.2.2.9). Yellow colour represent regions of absolute similarity. Blue represent conserved regions.

3.2.2 Expression analysis of the *LpGI* gene

The expression analysis was performed on a wide range of plants grown under different environmental conditions in order to find gene expression patterns related to commitment to flowering. Commitment period could be defined as a period in which the plants become committed to flowering. Observation on *Lolium perenne* cv. Impact indicates (data not shown) that under inductive conditions, preceded by a 8 week vernalisation treatment commitment is usually 4-7 days after exposure to LD and the elevated temperatures. This observation is in agreement with the previous experiments (Aamlid et al., 2000) which showed that the critical number of inductive days for the secondary induction varied between 1-3 in Mediterranean to 14-16 in Scandinavian accessions.

3.1.2.1 LpGI gene is a circadian clock regulated gene

GI genes including *Arabidopsis AtGI* and rice *OsGI* show a circadian pattern of expression (Fowler et al., 1999; Hayama et al., 2002). To test whether *LpGI* gene is regulated by the circadian clock a free running experiment was performed with the ryegrass plants grown and analysed as described in chapter 2.2.1.2. At least three PCR reactions using the same template were performed to get average values of expression levels. The specific primers MG017/019 (Table 2.1) were used to quantify the expression of the *LpGI* gene (Tm 61°C, extension 1 m, 40 cycles). Housekeeping GAPDH, and ubiquitin genes amplified with the respective primers GIK55/56, and GIK57/58 were used to normalise the results (Chapter 2.2.3, Table 5.2 and 5.3)

The *LpGI* gene was clearly shown to be regulated by the circadian clock (Fig 3.8). In *Arabidopsis* in LD conditions *GI* gene expression peaks 10-12 h after dawn (Fowler et al., 1999). Lp*GI* transcript cycled with the highest level 6-8 h after subjective dawn and the lowest level just after subjective dusk. The amplitude of the cycling subsided in the next two days but the cycling pattern was still clearly visible after 4 days.

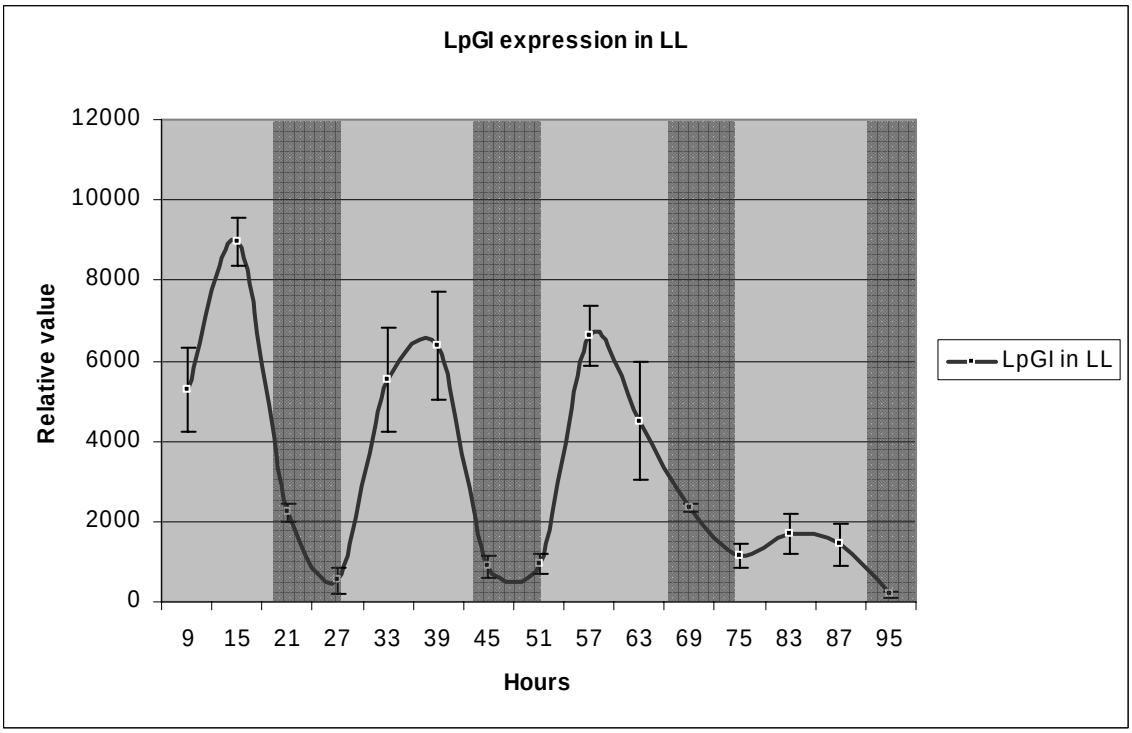


Fig 3.8 Effect of cicrcadian clock on *LpGI* expression.
The plants were fully vernalised, transferred to LD (21°C) for 5 days, and subsequently exposed to the LL conditions. Collection was done every 2 h for 4 days. *LpGI* expression levels were analysed by RT-PCR. The shaded areas represent subjective nights. The results were averaged over three data points and normalised against *Ubi/GAPDH* levels.

3.1.2.2 Diurnal expression of the LpGI gene is altered in LD vs. SD conditions

To further investigate how the *LpGI* expression levels change in respect to the photoperiod conditions the plants were grown under LD and SD conditions and their expression analysed. The mRNA levels of the *LpGI* gene changes significantly during these conditions. In LD conditions *LpGI* peaked ~13 h after dawn with trough levels ~4 h later, while in SD it peaked ~7 h after dawn with trough level ~3 h later. The expression levels are slightly higher in photo inductive conditions, however, there is a phase shift in long days, coincidental with light exposure (Fig 3.9). Such circadian regulation of expression, as well as a phase shift in long days, strongly supports the role of the *LpGI* gene in mediating and controlling the photoperiodic floral response of ryegrass.

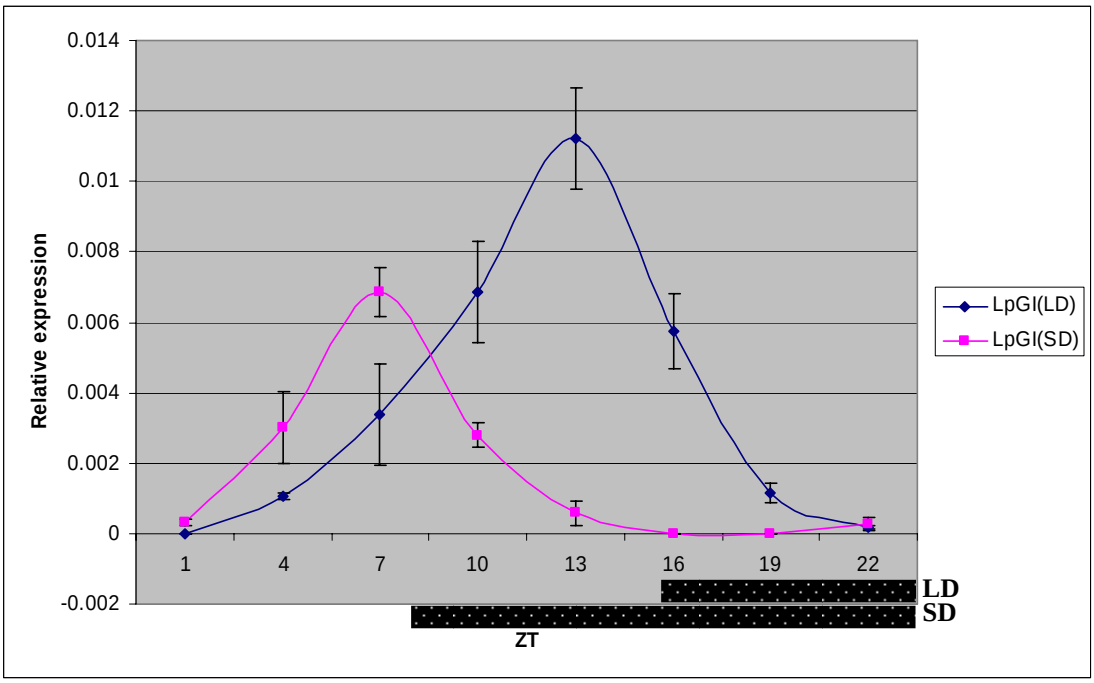


Fig. 3.9 Diurnal expression of the *LpGI* gene as measured by quantitative RT-PCR. *LpGI*(LD) represents samples collected in long day conditions, *LpGI*(SD) samples were harvested in short days. Black lines represent respective night. Results were normalized against GAPDH levels.

3.2.2.3 Vernalisation and LpGI expression

Involvement of the GI gene in the circadian clock regulation and phyB signaling has been extensively documented before (late flowering and long-hypocotyl phenotype in *gi* mutants grown under LD and continuous red light respectively; Fowler et al., 1999; Park et al., 1999; Huq et al., 2000), but the effect of the cold treatment on the expression of the *GI* has never been firmly established. It was shown that the cold stress in *gi-3* mutants delayed flowering significantly longer than in wt plants suggesting that *GI* gene plays a role in freezing tolerance (Cao et al., 2005; Fowler and Thomashow, 2002) and that *GI* expression in trough points is much higher under warmer temperature regimes (Paltiel et al., 2006). In order to check if a correlation between cold temperature and *GI* expression exist in ryegrass the plants were differentially vernalised (1-10 weeks) and transferred to the LD (20ºC) condition. The tissue samples were collected after 7 days, at Zeitgeber time (ZT) 15. Results were analysed by RT-PCR (Table 5.5). Clearly, there was no correlation between vernalisation length and *LpGI* mRNA abundance (Fig 3.10). The expression levels were fairly uniform and there was no fluctuations that could be associated with the vernalisation effect.

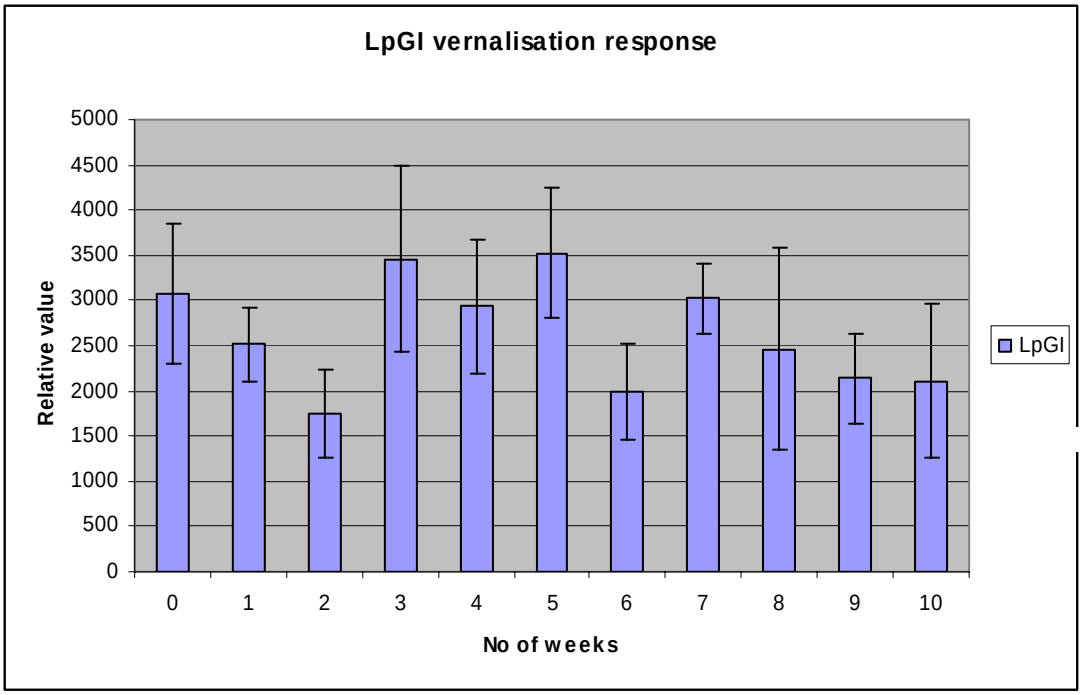


Fig. 3.10 Expression of the *LpGI* gene in plants exposed to different vernalisation treatements. Plants were differentially vernalised (1-10 weeks) and after that shifted to the LD condition. The tissue collection was done after 7 days. Results were normalized against actin levels.

In summary, expression analysis showed that *LpGI* gene is regulated by the circadian clock as shown by the continuous cycling under LL conditions. Photoperiod also had a major impact in the transcript abundance and expression phase regulation, further indicating *LpGI* involvement in the photoperiod regulation of flowering. Finally analysis of the differentially vernalised plants showed that cold treatment did not have any effect on *LpGI* transcript regulation.

3.2.3 Genetic mapping of the *LpGI* gene

Linkage analysis is a method in molecular biology used to establish linkage between genes. Linkage represents the tendency for genes and other genetic markers to be inherited together due to their location near one another on the same chromosome or linkage group (LG). A genetic marker is simply a segment of DNA with an identifiable physical location on a chromosome whose inheritance can be followed. Markers are often used as tools for tracking the inheritance pattern of a gene that has not yet been identified. Statistical estimate of whether two loci are likely to lie near each other on a chromosome and are therefore likely to be inherited together give rise to the possibility of constructing genetic maps based on genetic markers.

Table. 3.2 SNPs of *LpGI* considered for mapping. Three SNPs were discovered during the initial screening. Allele 3 positioned in the intron 10 was chosen for being easly distinguishable and consistent in sequencing signal strength. Gentic conservation of the introns regions where the SNPs were found was low for all three. GI^{Sn} and GI^{Ip} represent SNPs complements found in Samson and Impact varieties respectively

SNPs	GI^{Sn}	GI^{Ip}	gDNA	aa translation	conservation
1	TT	AT	4396	n/a	Low
2	TT	AT	4539	n/a	Low
3	CC	TC	4576	n/a	Low

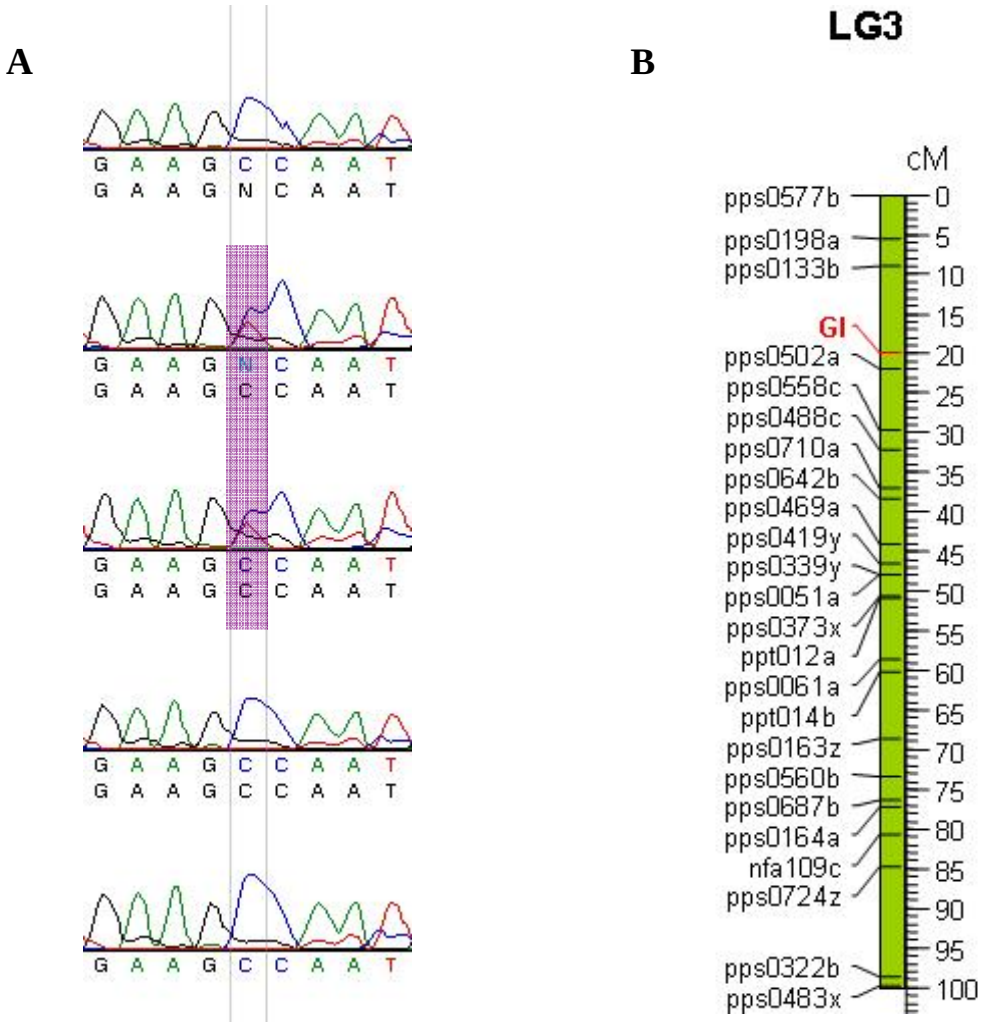


Fig. 3.11 Genetic map of *Lolium perenne* chromosome 3 (LG3)
A. SNP used for the mapping procedure. Shaded genotypes contains C/T heterozygous locus. In all others the locus is homozygous **B.** Analysed ryegrass *LpGI* SNP mapped to a location at position 20cM on the LG3

A genetic linkage map for perennial ryegrass was constructed by a separate group in AgResearch, Grasslands, using a full-sib mapping population consisting of 188 genotypes from the individual heterozygous genotypes cv. "Grasslands Impact" and cv. "Grasslands Samson" pair cross (Chapter 2.2.6).

This map served as the framework for mapping the *LpGI* gene, as a single nucleotide polymorphism (SNP) marker. In general the Samson variety was found to contain less SNPs than Impact. Primers GIK16/GIK17 were selected (Table 2.1) for the mapping experiment. They amplified a single fragment of 689 bp spanning two introns and three exons.

Sequencing of the amplified segment from the parent population revealed several SNPs (Table 3.2), out of which one designated T87, at the position 4576 on the gDNA, was chosen for mapping. This SNP is located within intron 10 as a homozygous allele in Samson variety (C) but heterozygous allele in the Impact variety T/C(Y) (Allele3, Table 3.2) and readily identifiable in the mapping population (Fig 3.10A). Eventually the position was deduced from 140 data points (Table 5.1), and the gene placed on LG3 (Fig 3.10C).

3.2.4 Functional analysis of the *LpGI* gene

To further characterise *LpGI* gene its function was investigated using transgenic approach. *LpGI* was put under the control of 35SCaMV constitutive promoter after which *Arabidopsis gi-3* mutant and wt plants were transformed using *Agrobacterium* mediated transformation. Phenotype change in transformed plants was analysed and results presented.

In order to perform a functional analysis, the complete *LpGI* cDNA was amplified from the Impact variety, using Gateway compatible primers MG046/MG019 (Tm 58°C, extension time 3 m 45 s, 38 cycles, Table 2.1; Fig 3.2). This segment starts at the ATG codon and comprises the complete *LpGI* cDNA including 114 bp of 3'UTR. The PCR product was cloned into the pEntr/D vector (Invitrogen; Fig 3.12A) and sequenced for the potential PCR errors (data not shown). The sequencing primers were the same as used in the Figure 3.1. The restriction digest analysis was also performed as shown on Fig 3.12B confirming the absence of the PCR generated errors. Using LR clonase enzyme mix (Invitrogen, Chapter 2.2.2.2) the insert was subsequently recombined into the pRSh1 binary vector (Fig 3.13). This vector was designed on the backbone of the pArt27 binary vector (Gleave, 1992) containing Spectinomycin bacterial selection, the *BAR* gene as a selectable plant marker (encodes for phosphinothricin acetyl transferase that detoxifies phosphinothricin (PPT), the active ingredient of herbicides such as Basta), and 35SCaMV.

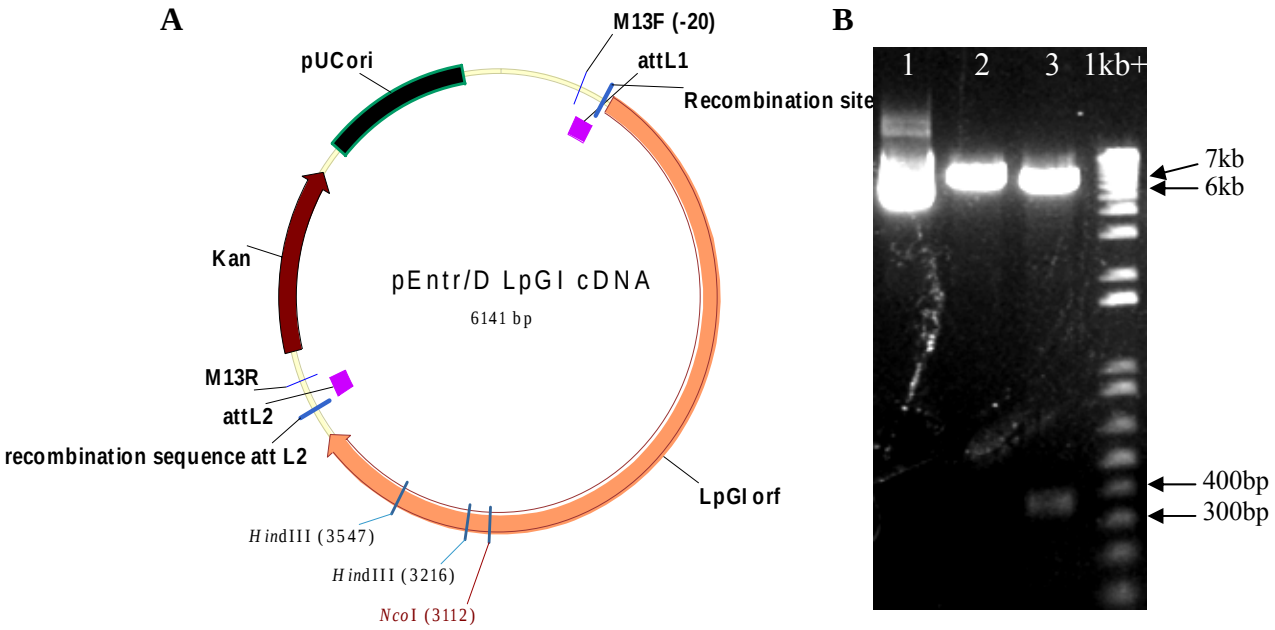


Fig. 3.12 pEntr/D-*LpGI* plasmid and the subsequent restriction enzyme digest.
A. *LpGI* cDNA was amplified using MG046/019 primers and the product cloned into the pEntr/D cloning vector. **B**. Restriction digest with NcoI (lane 2, 1 restriction site), and HindIII (lane 3, 2 restriction sites) enzymes confirmed the presence of the insert. Uncut plasmid is in the lane 1. Samples run alongside 1kb+ DNA marker

3.2.4.1 LpGI overexpression in Arabidopsis gi-3 mutant

To test whether *LpGI* functionally complements the *Arabidopsis gi-3* mutant *35S::LpGI* construct was transferred into *gi-3* mutant *Arabidopsis* plants using pRSh1 vector (Fig 3.13) via *Agrobacterium*-mediated transformation (Chapter 2.2.4.2). The *gi-3* locus has a nucleotide substitution at position 3929 (C-T), introducing premature stop codon and thus producing a truncated protein (Fowler et al., 1999).

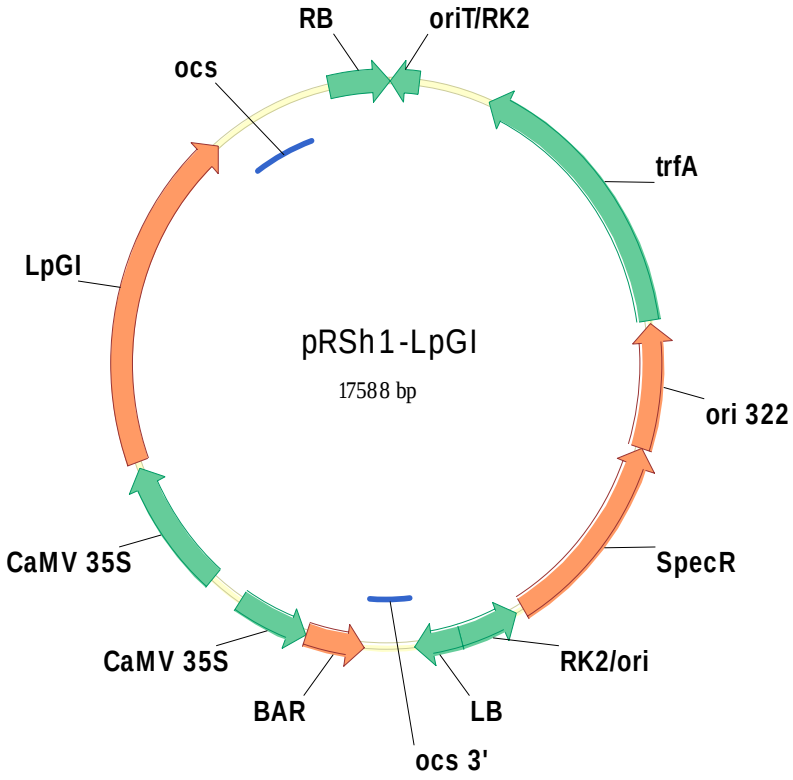


Fig 3.13 Map of the pRSh1-*LpGI* binary vector. The vector has the *LpGI* gene cloned downstream of the constitutive 35S CaMV promoter, bacterial selection marker (SpecR), and plant selection marker (BAR).

From the initial screening >20 lines showed resistance to BASTA with 15 of them carried into the next generation (T2). The progeny of 15 lines was screened on for insert copy number by spraying with BASTA and it was found that 11 lines showed 3:1 segregation ratio which is characteristic for a single copy insert. Plants from these lines were scored for leaf number. From these 11 lines, plants from 3 of them were carried into the next generation (T3) where they were BASTA screened. Progeny of these 3 lines that did not segregate (homozygous) were used for the final leaf count. *Arabidopsis* plants containing *gi-3* allele flower later than the wild type. As shown (Table 3.3; Figure 3.14) the flowering time of *gi-3* plants overexpressing *LpGI* gene is very similar to the wild type plants in LD photoperiod, while *gi-3* mutants, as expected, flower much later when grown in the same conditions. This data demonstrates that expression of *LpGI* promotes flowering and complements *gi-3* mutant phenotype. It is interesting that the T2 transgenic plants (heterozygous for the transgene) flower slightly later than the T3 (homozygous) plants which may indicate the dosage effect or haploinsufficiency. Variation of flowering times was observed in different transformants which is most likely caused by differences in *LpGI* expression levels.

Table. 3.3 35S::*LpGI* complementation of the *gi-3 Arabidopsis* plants.
Plants were transformed via *Agrobacterium*-mediated transformation. Rosette and cauline leaves were counted. Lines from the T3 generation are separately listed. No less than 6 plants from each line used for the leaf count.

Genotype	Generation	No of lines	Homozygous	Leaf count
L*er*	n/a	n/a	n/a	10.2±1
gi-3	n/a	n/a	Y	21.8±0.5
Average *gi-3+35S::LpGI*	T2	11	N	11.5±1
COF92-4(*gi-3+35S::LpGI*)	T3	n/a	Y	6.6±0.6
COF94-1(*gi-3+35S::LpGI*)	T3	n/a	Y	6.4±0.5
COF96-2(*gi-3+35S::LpGI*)	T3	n/a	Y	6.8±0.8
Average *gi-3+35S::LpGI*	T3	3	Y	6.6±0.7

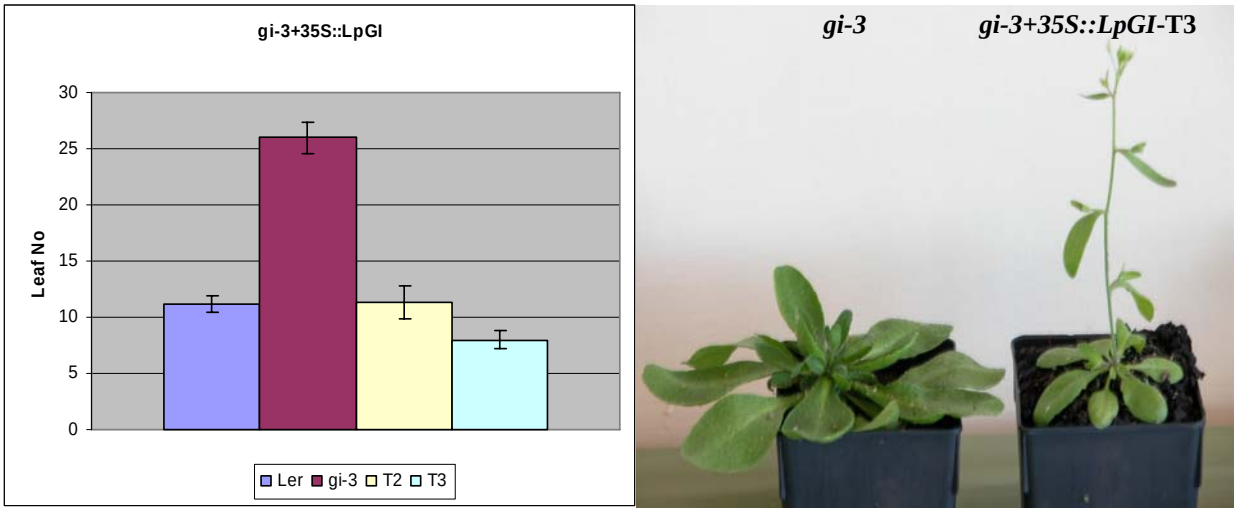


Fig. 3.14 35S::*LpGI* complementation of the *gi-3 Arabidopsis* plants.
Plants were transformed via *Agrobacterium* mediated transformation. Rosette and cauline leaves were counted. Three independent homozygous lines were used in the experiment. No less than 6 plants from each line were used for the leaf count.

3.3 *CONSTANS* (*LpCOL1*)

In order to further dissect the ryegrass photoperiod pathway a search for the ryegrass *CO-like* gene was initiated. At the beginning of this work the *COL* family in ryegrass was poorly characterised with no genes described from this group. The last three years saw progress in research regarding the ryegrass photoperiod pathway with several papers being published on COL genes. A putative *AtCO* orthologue was recently mapped and characterized (Armstead et al., 2004; Martin et al., 2004). In addition to this another LpCOL gene was sequenced in relation to the vernalisation response but was found to be homologous to the group III type of COL proteins and does not have a clear homologue in *Arabidopsis* (Ciannamea et al., 2006).

3.3.1 Characterisation of the *LpCOL1* gene

An initial search of the ryegrass EST library did not reveal any sequences resembling *CO* orthologs from other species. Therefore *CO* multiple nucleotide and protein alignments from wheat, rice, and *Arabidopsis* were used to design degenerate primers from homologous regions, which were then used with ryegrass cDNA to amplify the *CO* segment. Total RNA from ryegrass tissue was isolated using the Trizol protocol (Chapter 2.2.2.1.3). The Dynabeads mRNA purification kit was used for mRNA separation and purification followed by first strand cDNA amplification using Thermoscript RT-PCR system with oligo dT primers (Chapter 2.2.2.5). Several degenerate primers were designed and two of them (MG003/MG005, Table 2.1, Fig 3.15A) produced a distinctive band of expected size of ~ 1100 bp (Tm 63ºC, extension time 3 m, 40 cycles, Fig 3.15B).

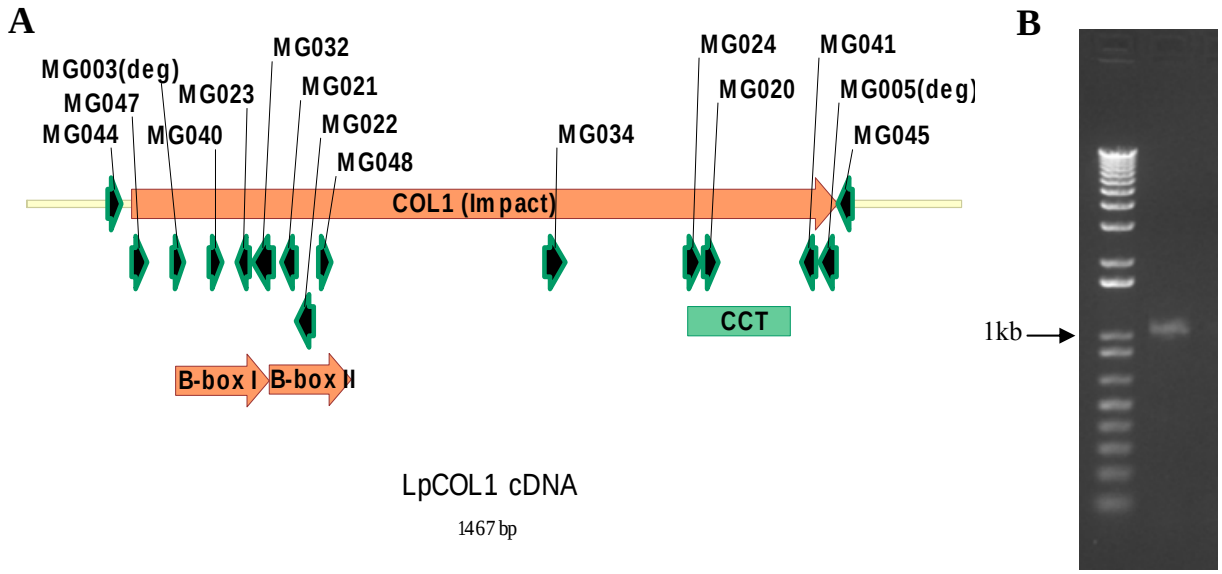


Fig. 3.15 Amplification of *LpCOL1*. **A**. cDNA segment of the *LpCOL1* gene. It includes two zinc-finger regions and one CCT region spanning 1110 bp which encode 369 aa protein. Primers MG003(deg)/MG005(deg) were used to amplify segment of *LpCOL1* gene from ryegrass cDNA. **B**. The amplified segment was ~1100 bp long, and was used for subsequent cloning and sequencing. The sample was run against the 1kb+ DNA ladder

The band was gel purified and cloned into pCR2.1 plasmid (Invitrogen) for further analysis. Sequencing (Chapter 2.2.2.8) revealed extensive similarity of this nucleotide sequence with the *CO* genes from rice, wheat, and *Arabidopsis* especially at the zinc-finger (N-terminus) and CCT (C-terminus) domains (Fig 3.16). It spanned almost the whole length of the predicted ryegrass *CO-like* (*LpCOL1)* gene except for ~50 bp at the 3' and 5-10 bp at the 5' end.

In order to obtain the missing 5' and 3' ends the primers for the 5' and 3' RACE-PCR experiments were designed. Several attempts with different primers failed to amplify either end of the gene, presumably due to its low expression level in all stages of plant development. A possible solution was to try to amplify the *CO* genomic sequence, rather than cDNA using either ryegrass BAC clones or genome walker libraries with the latter being chosen.

A ryegrass genome walker library was made using Clontech GenomeWalker kit (Chapter 2.2.2.7) and genomic DNA extracted from the Impact ryegrass variety (Chapter 2.2.2.1.2). Five libraries of uncloned, adaptor ligated genomic DNA fragments were created according to the recommended protocol. Initially four gene specific primers (MG022, MG023, MG032, and MG021, Table 2.1, Fig 3.17A) were used for the amplification of the 5' end of the gene in combination with the outer adaptor primer (AP1, Table 2.1).

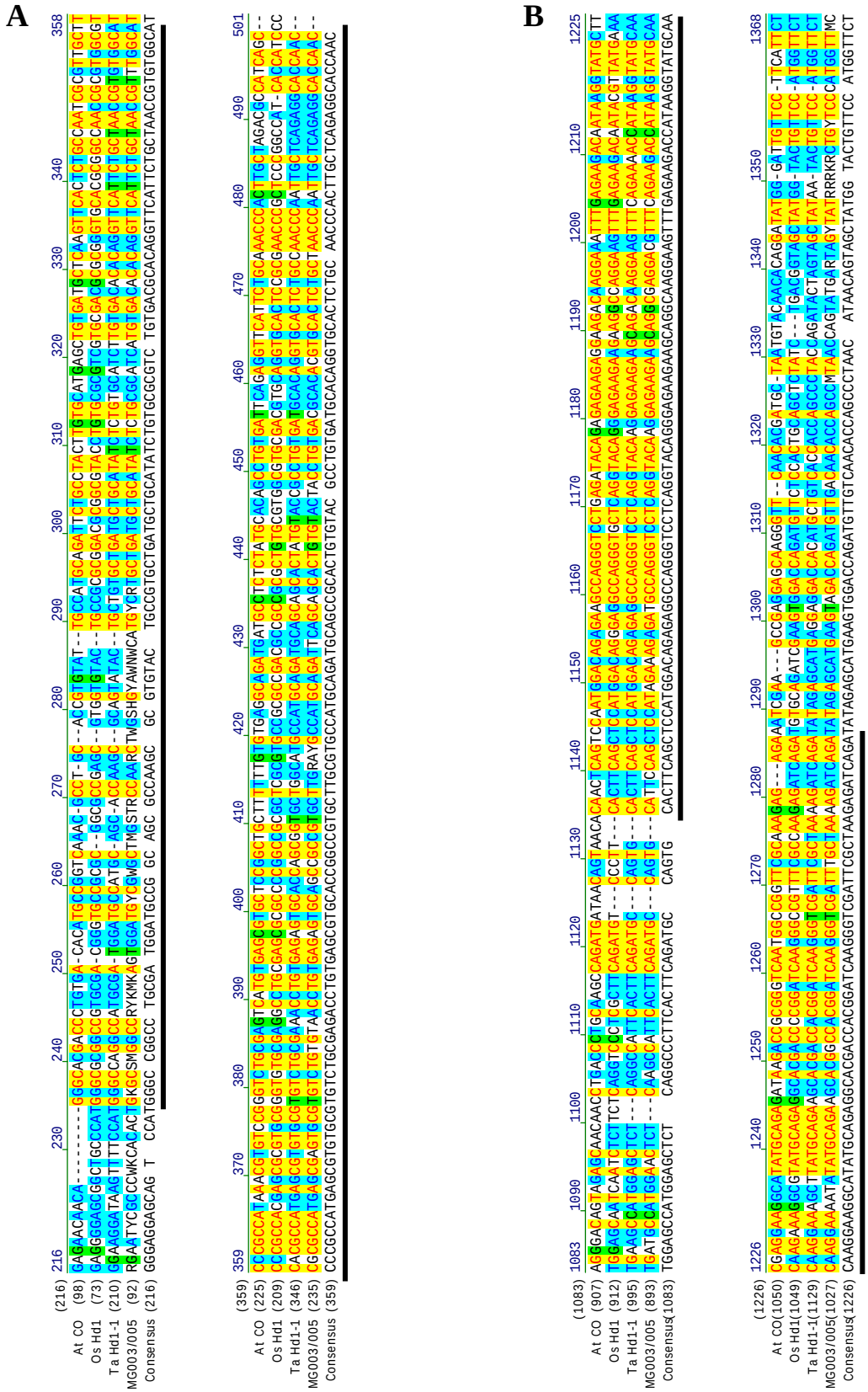


Fig. 3.16 Alignment of the *CO* zinc-finger and CCT regions
Alignment of the zinc-finger regions (**A**) and CCT regions (**B**) from the Arabidopsis *CO* (X94937), rice *Hd1* (AB001882), wheat *Hd1-1* (AB094490) genes, and the putative *LpCOL1* segment (MG003/005). Underlined regions show the zinc-finger and the CCT domains. Colours are as described in Figure 3.7

Touchdown PCR reaction (Chapter 2.2.2.4) was performed for all PCR reactions. The primer pair MG21/AP1 produced distinctive bands in 4 out of 5 reactions and the products from the *EcoRV* and *SmaI/HpaI/HincII* libraries (Fig 3.17A) were purified and used for the nested PCR reactions.

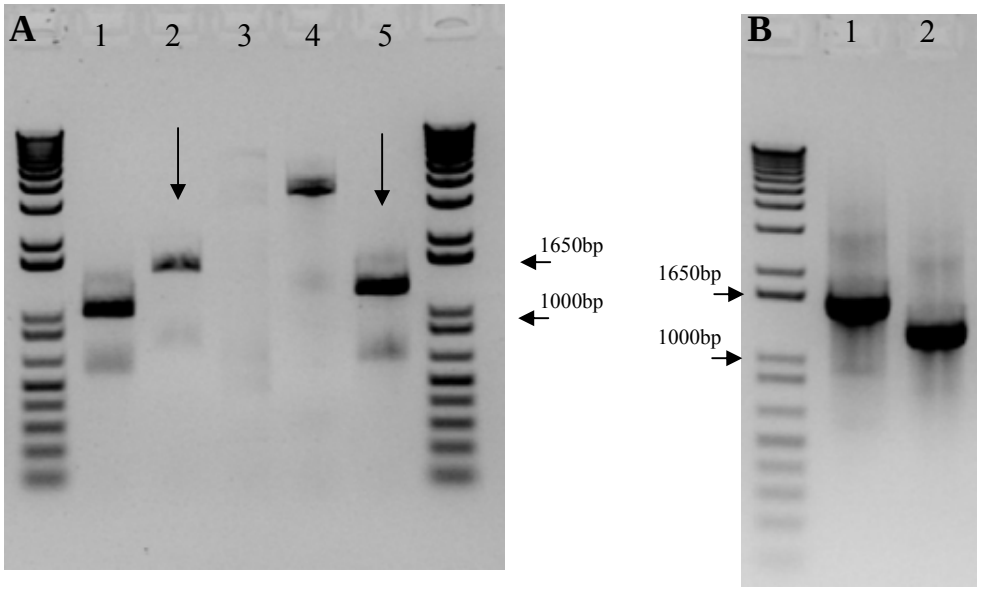


Fig. 3.17 5'*LpCOL1* Genome Walking PCR reaction
PCR was performed using MG021/AP1 primers. DNA templates digested with the following enzymes used in lines 1-5 respectively; *DraI, EcoRV, PvuII, StuI*, and *SmaI/HpaI/HincII*. Arrows point to the PCR products used for the subsequent nested PCR reaction. **B**. Nested PCR reaction using MG023/AP2 primers. In lines 1and 2 DNA template was the product of MG21/AP1 reaction on EcoRV and SmaI/HpaI/HincII library respectively. The samples were run alongside 1kb+ marker.

The nested PCR was performed with the MG023/AP2 primers (Tm 68ºC, extension time 1 m 30 s, 40 cycles, Table 2.1) which produced distinctive bands of ~1.5 kb and ~1.35 kb respectively (Fig 3.17B). The bands were gel purified, and TOPO cloned (data not shown).

A similar amplification procedure was used in order to amplify 3' end of the gene. Three GSP primers (MG020, MG024, and MG034, Table 2.1) were used in combination with AP1, with the pair MG20/AP1 producing distinctive bands with the *PvuII* and *StuI* libraries (Fig 3.18). Product derived from the *PvuII* library was gel purified and TOPO cloned. Upon sequencing both the 5' and 3' gene walker inserts, a contig containing a full *LpCOL1* cDNA was assembled and the open reading frame (orf) determined.

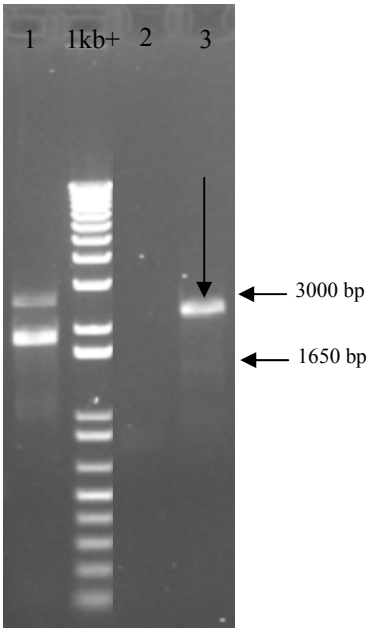


Fig. 3.18 3'*LpCOL1* Genome Walking PCR reaction
PCR was performed using MG020/AP1 primers. The DNA template used was ryegrass cv. Imapct gDNA cut with *StuI* (1), *EcoRV* (2), and *PvuII* (3). Arrows point to the PCR product which was gel purified and used for the subsequent TOPO cloning reaction. The samples were run alongside the 1kb+ marker.

3.3.2 *LpCOL1* sequence analysis

Previous analysis of *CO*-like genes in *Arabidopsis* showed that the family is subdivided into several broad groups regarding the number and structure of zinc-finger domains ranging from genes with two functional domains down to the *COL* genes without zinc-finger domains (Robson et al, 2001; Griffiths et al., 2003; Yan et al., 2004). It has been shown that comparing the *COL* genes using only peptide or nucleotide sequence from the CCT domain can accurately predict the number and functionality of zinc-finger domains in respective genes (Griffiths et al., 2003). Using the CCT domain COL genes from *Arabidopsis*, rice, and barley consistently formed four main clusters. When analysed using fused B-box and CCT domains the same four principal groups were formed with group I being further subdivided into subgroups with higher bootstrap values.

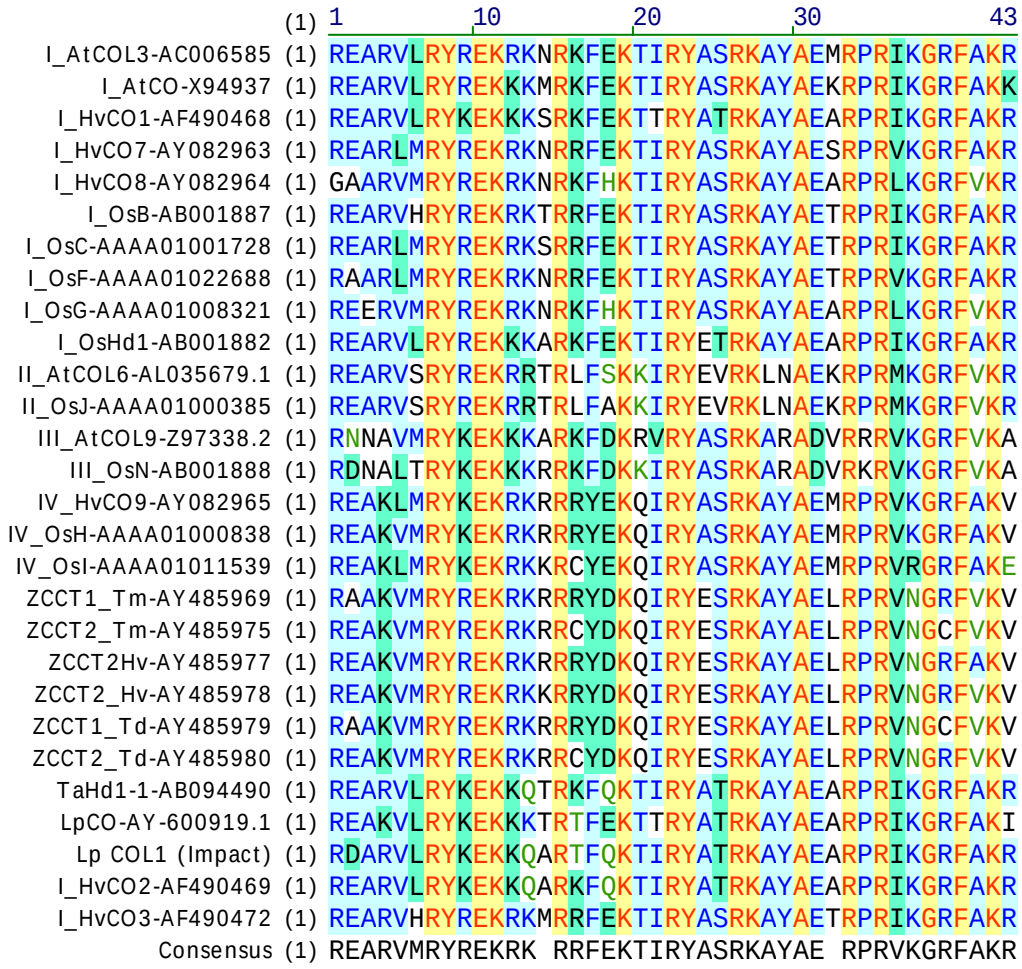


Fig. 3.19 Analysis of the CCT region of different CO, COL, TOC1, and VRN proteins. 28 protein sequences from barley (Hv), rice (Os), *Arabidopsis* (At), wheat (Ta, Tm, Td), and ryegrass (Lp) was aligned. The number next to the gene name represent Genebank accession number. Roman numbers in front of the genes indicate clusters they were grouped in Griffiths et al. (2003). ZCCT genes are a separate group in wheat and barley with similarity to the CCT region. *TaHd1*, *LpCO*, and *LpCOL1* are new genes added to this alignment.

When included in the analysis (Fig 3.19), the CCT region from the putative *LpCOL1* gene appeared to fall within group I (the group with two B-box domains), as expected, with the highest similarity to the barley *HvCO1* and rice *Hd1* genes (Fig 3.20). The presented data suggested that the *LpCOL1* gene is the homologue of the *AtCO* and *OsHd1* gene

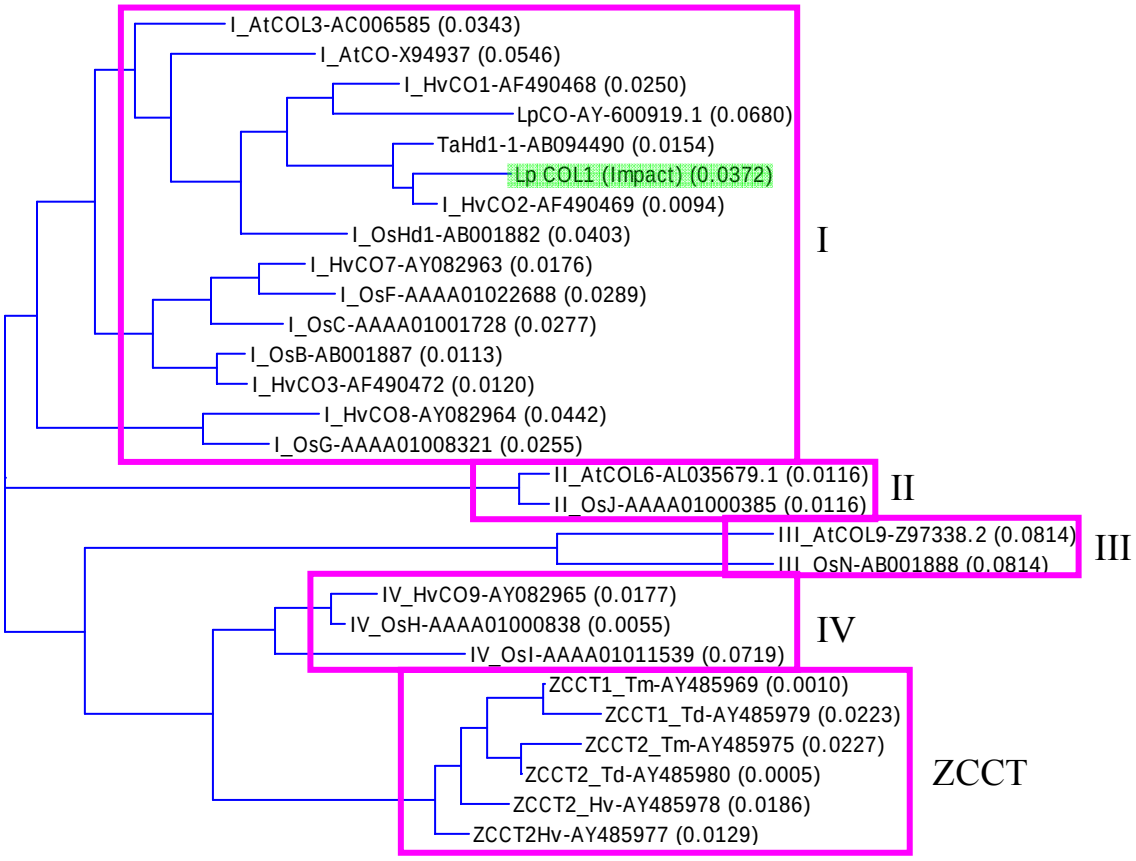


Figure 3.20 Phylogenetic analysis of the *CO* and *COL* genes. The CCT domain analysis indicated that CO and COL proteins are divided into four different groups depending on the number of B-boxes they contain. The putative *LpCOL1* protein was positioned within group I which has proteins with two characteristic B-box domains. The *ZCCT* genes implicated in the vernalisation response formed a separate group. Phylogenetic tree calculation is based on a sequence distance method and utilizes the Neighbor Joining (NJ) algorithm. Abbreviations are as described in Fig 3.19.

The *LpCOL1* cDNA contains 1110 bp encoding a 369 amino acid protein (GeneBank DQ534011; Appendix, Fig 5.5). Amplification of the gDNA revealed one intron positioned at 754 bp which is 575 bp long. The overall length of the gDNA is 1685 bp. Multiple alignment of the predicted LpCOL1 protein shows two zinc-finger domains characterised by the presence of the conserved His and Cys residues that meet the consensus formula CX2CX8CX7CX2CX4HX8H, and the CCT region similar to *Arabidopsis,* wheat, and barley CCT (Fig 3.21). At the protein level the overall similarity of the LpCOL1 protein with *Arabidopsis*, rice, and wheat is 38%, 39%, and 78% respectively but when only B-box regions were compared the similarity increased to 70%, 76%, and 90%. Within the CCT domain the identity rose to 75%, 86%, and 93% for *Arabidopsis*, rice and wheat respectively.

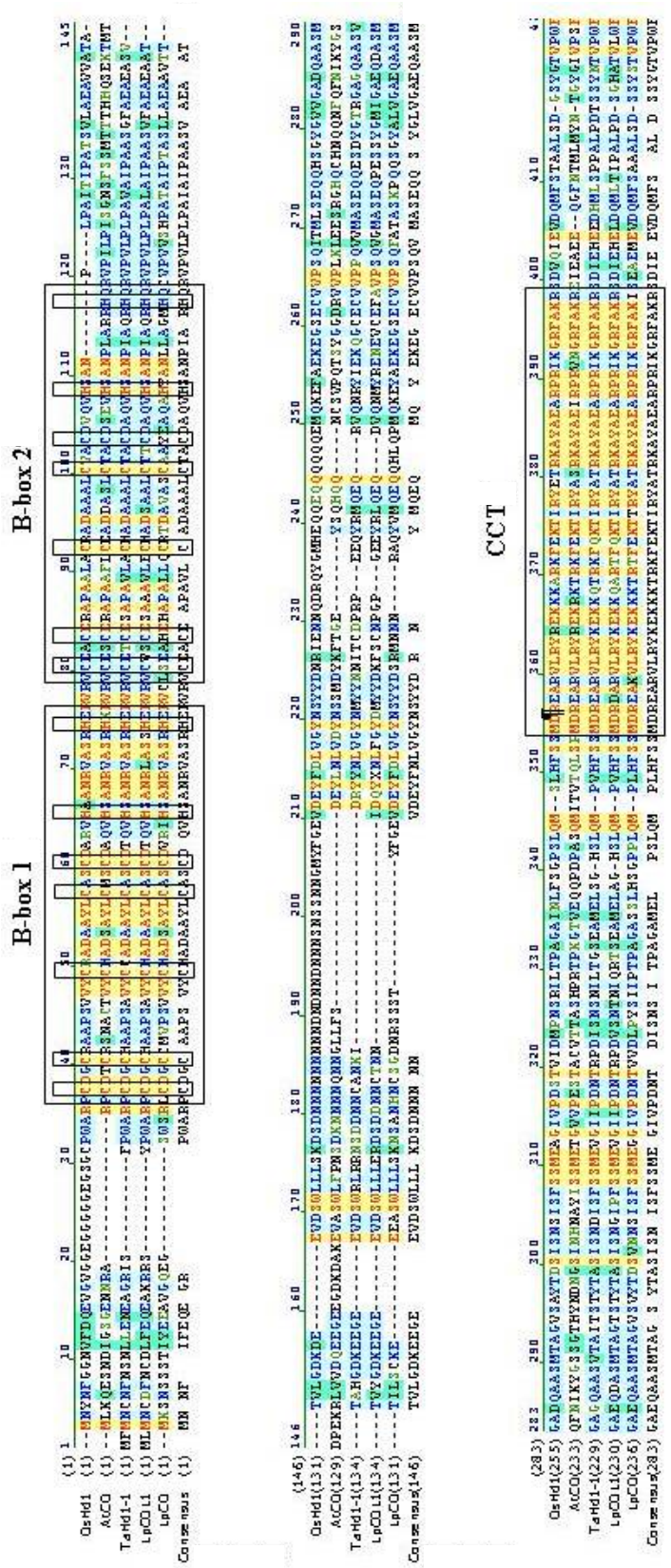
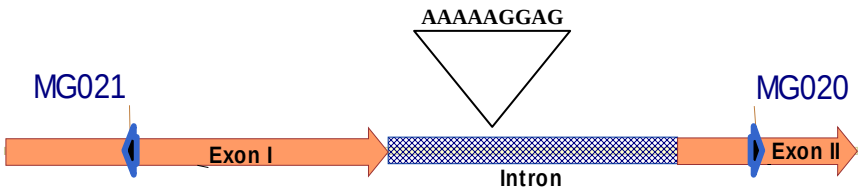


Fig. 3.21 Multiple alignment of the *LpCOL1* gene
It reveals two conserved B-box domains characterized by the presence of 7 conserved His and Cys residues in each. There is also high similarity in the CCT domain between *Arabidopsis*, rice (Os), wheat (Ta), and *LpCOL1* genes.

Closer analysis of the nucleotide sequence of the *LpCOL1* genomic DNA using Transfaq database (Chapter 2.2.2.9) revealed a conserved Dof2 binding domain within the intron region (Fig 3.22). The Dof proteins are a wide family of transcription factors present only in plants and characterized by a strongly conserved 52-amino acid domain encompassing a single CX2CX21CX2C zinc finger (Yanigiasawa, 1995). They are known to be transcriptional activators and repressors, tissue-specific, and light regulated.



LpCOL1 gene

Fig 3.22 Map of the *LpCOL1* gene. The gene contains one intron with a Dof2 binding site. The whole gene comprises 1691 bp.

3.3.2 Expression analysis of the *LpCOL1* gene

3.3.2.1 Circadian regulation of the LpCOL1 gene

Arabidopsis CO and rice *Hd1* genes show circadian pattern of expression (Suarez-Lopez et al., 2001; Hayama et al., 2003). To test whether the *LpCOL1* gene is regulated by the circadian clock ryegrass plants were fully vernalized (SD, 4°C) for 12 weeks, transferred into the glasshouse for 5 days (21ºC, LD) and subsequently shifted into the continuous light (LL) conditions at 20ºC. Samples were collected every 2 h for next 4 days. The specific primers MG060/MG041 (Table 2.1) were used to quantify the *LpCOL1* gene (Tm 60°C, extension time 40 s, 45 cycles). *GAPDH* and *Ubiquitin* genes, were used to normalise the results (Chapter 2.2.3, Table 5.2 and 5.3).

LpCOL1 gene was shown to be regulated by the circadian clock (Fig 3.23). The amplitude of the cycling dramatically subsided after next three days but the cycling pattern was still clearly visible at day 3.

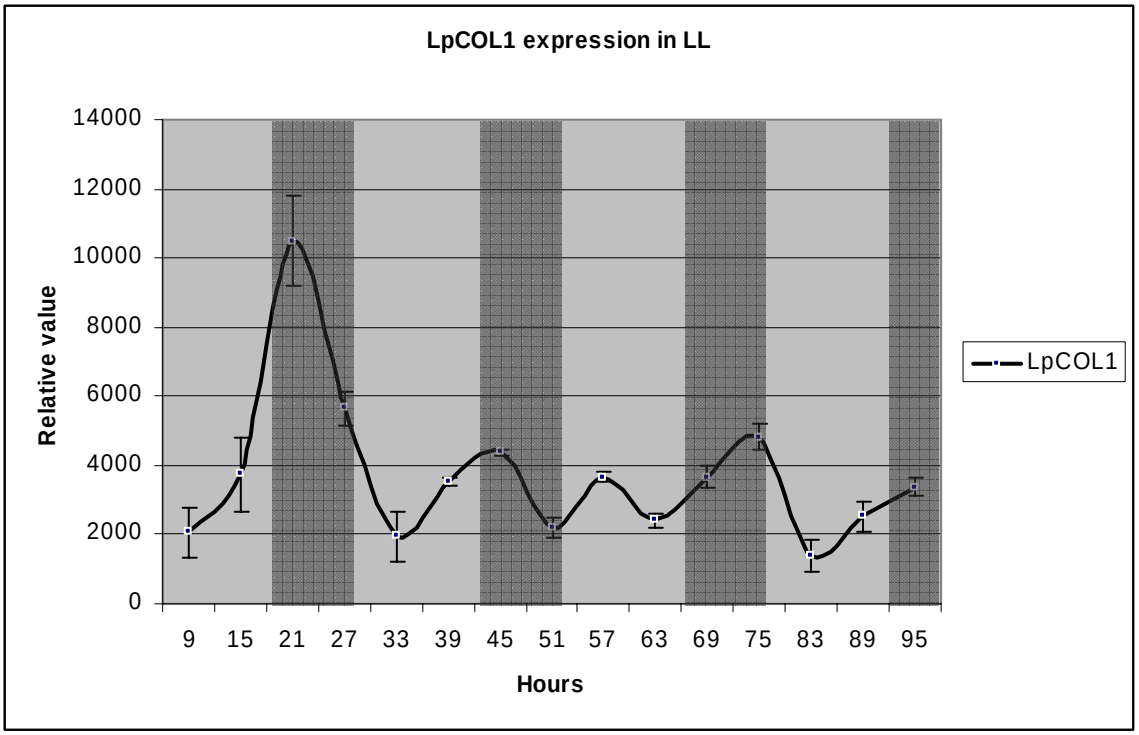


Fig 3.23 Effect of cicrcadian clock on the *LpCOL1* expression.
The plant growing conditions for this experiment were the same as for the *LpGI* free running experiment. *LpCOL1* expression levels were analysed by RT-PCR.The shaded areas represent subjective night. The results were averaged over three data points and normalised against *Ubi/GAPDH* levels.

3.3.2.2 Circadian expression of the LpCOL1 gene in LD vs. SD conditions

To investigate further how the *LpCOL1* expression pattern and level of expression change in respect to the photoperiod conditions the plants were grown under the LD and SD conditions and their expression analysed. Genetically identical ryegrass clones were grown as described in Chapter 2.2.1.1. Diurnal collection was performed after 8 days, every 3 h for 24 h. Expression analysis of the *LpCOL1* gene was assesed by RT-PCR to determine mRNA levels over a 24-hour cycle. The mRNA levels of the *LpCOL1* gene changes significantly during this period in both, LD and SD conditions (Fig 3.24, Table 5.5). *LpCOL1* peaked ~17 h after dawn with trough levels ~3 h later in SD, while in LD conditions it peaked at approximately the same time but with trough levels ~5 h later. The results also show that the expression levels increase ~4 fold in LD conditions. Results from the *LpCOL1* circadian expression were also compared with the circadian expression of the related *LpCO* gene (Martin et al., 2004) under the identical environmental conditions. Results showed that the pattern of expression was very similar in both of them, with peak expression occurring during daylight in LD conditions and during the night in SD conditions (Appendix, Figure 5.6, Table 5.5).

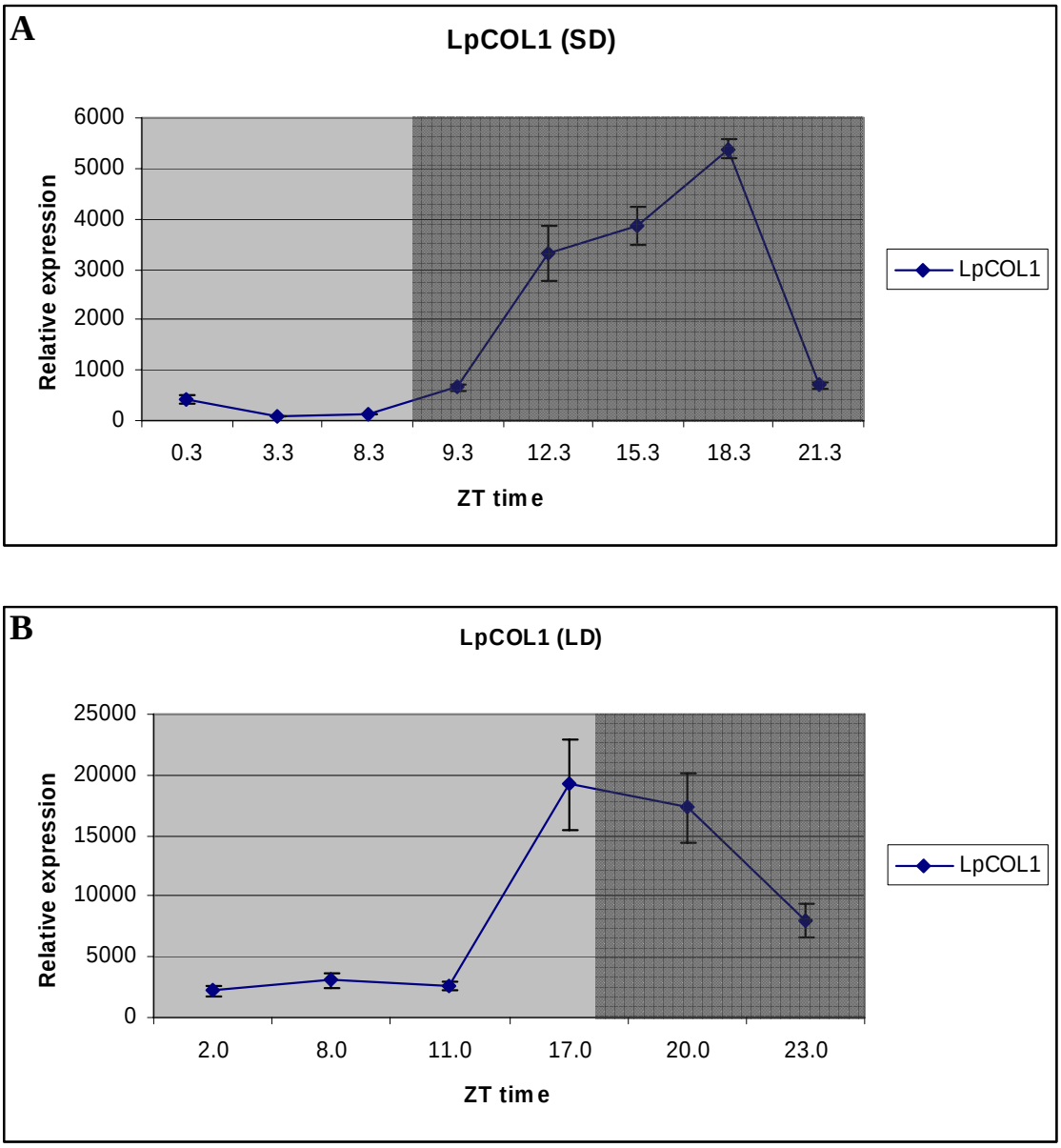


Fig. 3.24 Diurnal expression of the *LpCOL1* gene as measured by quantitative RT-PCR. **A**. *LpCOL1*(SD) represents samples collected in short day conditions, **B**. *LpCOL1*(LD) samples were harvested in long days. Shaded areas represent respective night. Results were normalized against *GAPDH* levels.

3.3.2.3 LpCOL1 expression throughout the vernalisation process

CO is one of the key components of the photoperiod pathway (Putterill et al., 1995), which integrates information from the circadian clock as well as PHYA and CRY2 signalling (Suarez-Lopez et al., 2001; Yanovsky and Kay, 2002; Valverde et al., 2004). The expression pattern of the *LpCOL1* in LD conditions in relation to the vernalisation status was investigated in order to establish if cold treatment prior to the LD exposure had any effect on LpCOL1 expression. Plants were differentially vernalised, transferred into LD conditions, and sampled

after 0, 3, and 7 days just before dusk. Samples were treated and analysed as described in Chapter 2.2.2.1.3. Analysis showed that the expression pattern of *LpCOL1* was unchanged regardless of the vernalisation length (Fig 3.25). During the vernalisation and immediately after the transition to the LD conditions *LpCOL1* expression is low but starts to raise as observed during the subsequent 7 days. This is a consistent pattern which does not change in relation to the vernalisation length. This is in line with expectations since *COL* genes, with the exception of *ZCCT* gene family, have not been linked with the vernalisation pathway so far.

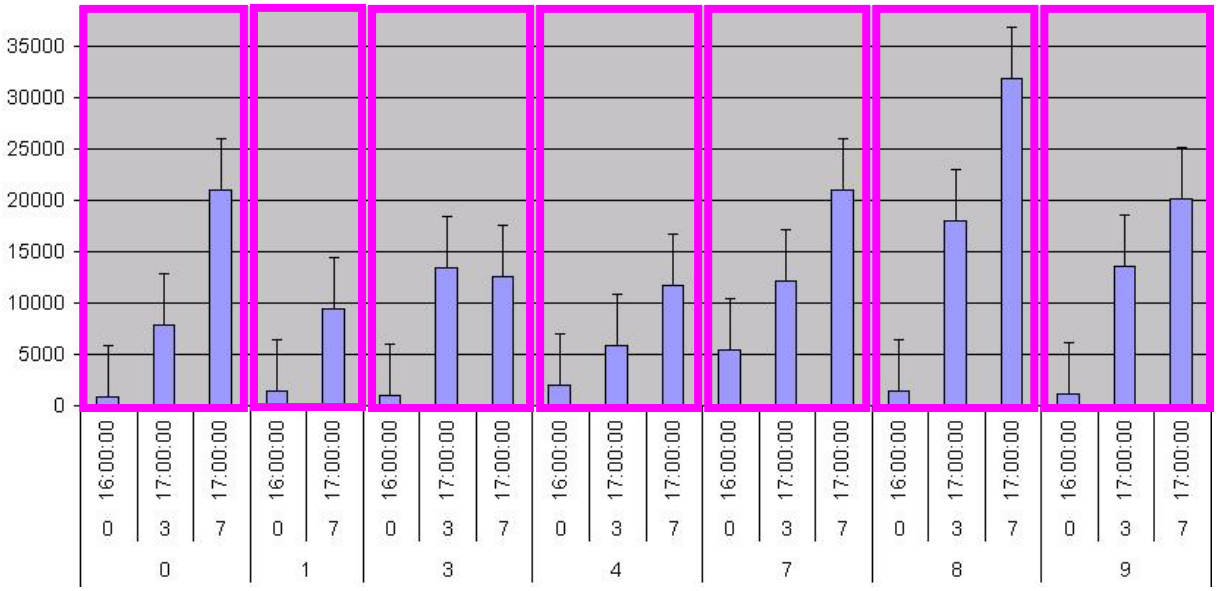


Fig. 3.25 Expression of the *LpCOL1* gene upon shift to the LD after differential vernalisation treatment. Ryegrass plant were differentially vernalised (0-9 weeks) in SD and exposed to LD conditions for 0, 3, and 7 days. The samples were collected at ~ ZT16 and expression of *LpCOL1* analysed. The results were normalized against *GAPDH* levels

3.2.3 Functional analysis of the *LpCOL1* gene

To investigate and understand the functions of *LpCOL1*, transgenic *Arabidopsis* plants carrying sense *LpCOL1* cDNA under the control of CaMV35S promoter were generated and analysed under the LD conditions. Gateway compatible primers MG047/MG045 (Table 2.1) were used to obtain full length of the *LpCOL1* (Tm 58°C, extension time 1 m 30 s, 35 cycles). This segment starts at the ATG codon and comprises complete *LpCOL1* cDNA including 23bp of 3'UTR. The PCR product was cloned into the pEntr/D vector (Invitrogen; Fig 3.26A, 3.26B) and sequenced for the potential PCR errors (data not shown). Restriction digest analysis was also performed as shown on Fig 3.26C. No PCR generated errors were identified and the insert encoded full length LpCOL1.

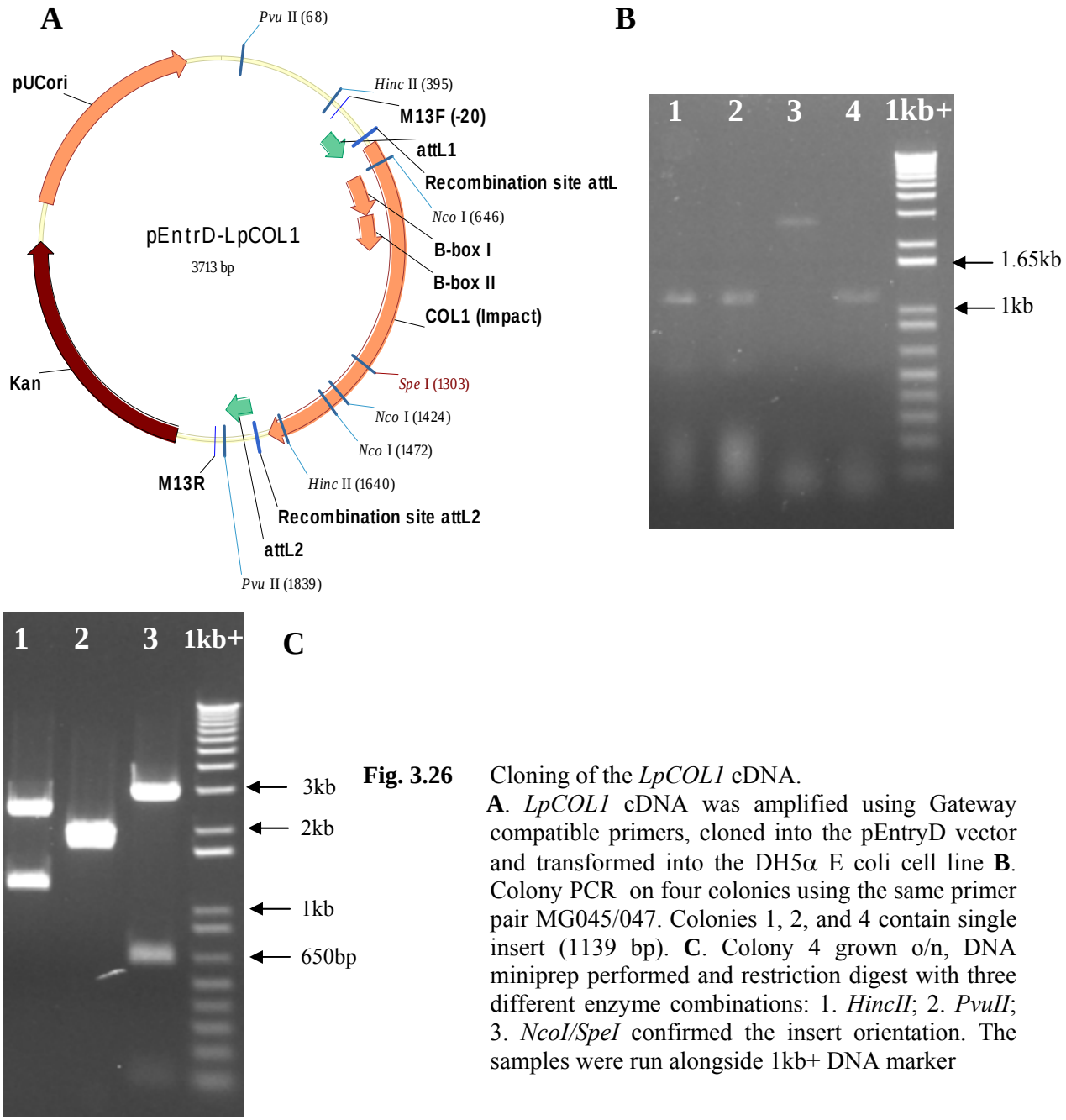


Fig. 3.26 Cloning of the *LpCOL1* cDNA.
A. *LpCOL1* cDNA was amplified using Gateway compatible primers, cloned into the pEntryD vector and transformed into the DH5α E coli cell line **B**. Colony PCR on four colonies using the same primer pair MG045/047. Colonies 1, 2, and 4 contain single insert (1139 bp). **C**. Colony 4 grown o/n, DNA miniprep performed and restriction digest with three different enzyme combinations: 1. *HincII*; 2. *PvuII*; 3. *NcoI/SpeI* confirmed the insert orientation. The samples were run alongside 1kb+ DNA marker

The insert was recombined using LR clonase enzyme mix (Chapter 2.2.2.2) into the pRSh1 binary vector containing the constitutive 35S promoter (Fig 3.27A). This vector has the same features as the one previously used for the *LpGI* functional analysis containing spectinomycin bacterial selection, the *BAR* gene as a selectable plant marker and the CaMV35S promoter. The insert was checked by restriction digest, using two restriction enzymes *SpeI* and *KpnI* (Fig 3.27B) which confirmed its orientation.

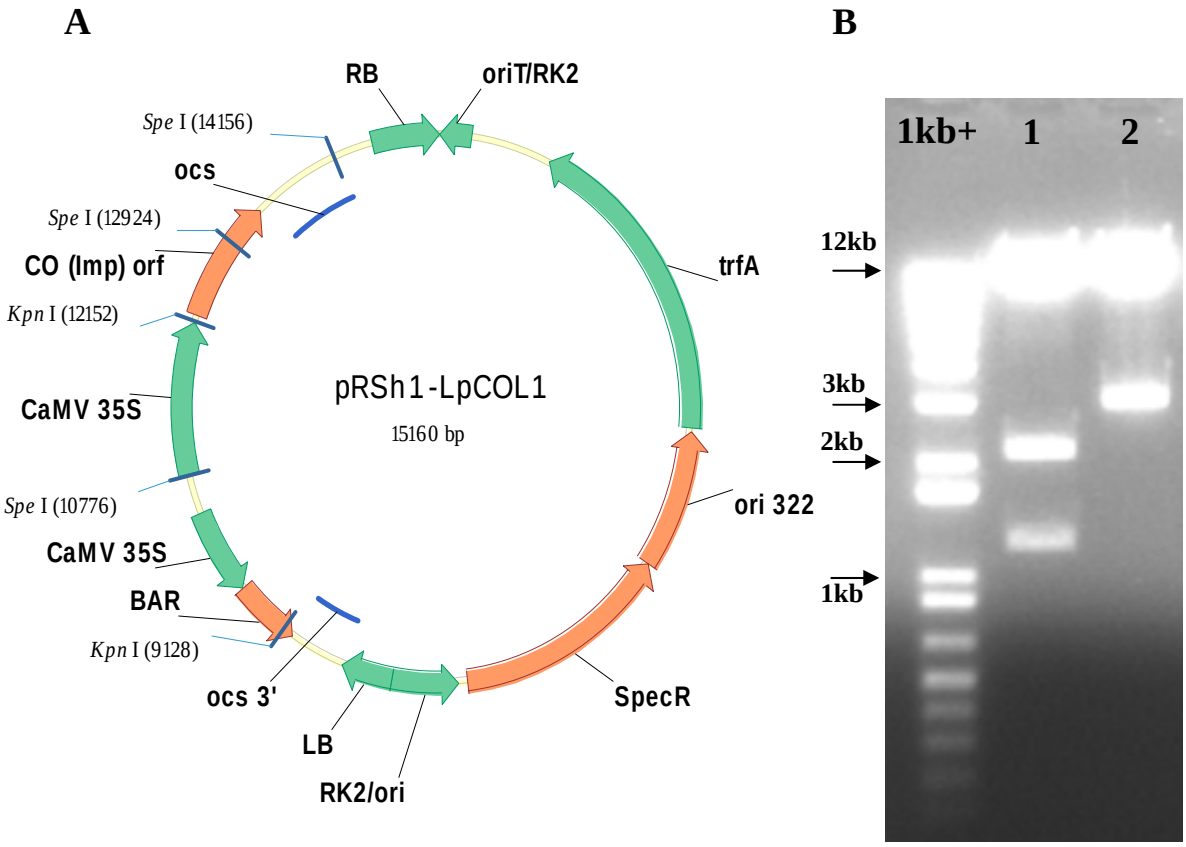


Fig 3.27 Map and the restriction digest of the pRSh1-*LpCOL1* binary vector.
A. The vector has *LpCOL1* gene cloned downstream of the constitutive 35S CaMV promoter **B**. The *LpCOL1* insert was checked with SpeI (1), and KpnI (2) which produced bands of expected sizes. The digests were run alongside 1kb+ DNA marker

3.3.3.1 LpCOL1 gene does not affect flowering in co-2 or wt Arabidopsis plants

To characterise flowering time phenotype of the *Arabidopsis co-2* mutants and Col wild type plants containing *LpCOL1* cDNA fragment under the 35S constitutive promoter these plants were transformed by *Agrobacterioum*-mediated floral dip. Seed was collected and selected for transformants using BASTA selection. Sixteen and four resistant lines were collected from *Col* and *co-2* transformants respectively. Progeny of these lines were sown, BASTA sprayed and survival ratio analysed. Three lines from each background showing a ratio of 3:1 (single insert) were carried over into the next generation (T3), where they were analysed by counting the leaves.

The *co-2* mutation converts an arginine to a histidine towards the carboxy-terminal end of the first B-box rendering the plant late flowering (Robson et al., 2001). If functional, overexpression of the *LpCOL1* gene should rescue or accelerate flowering in *co-2* and wt plants respectively. Surprisingly, overexpression did not produce expected flowering acceleration and complementing phenotype in wt *Col* and *co-2* plants (Table 3.4; Fig 3.28).

Table. 3.4 35S::*LpCOL1* complementation of the *Col (wt), and co-2* mutant *Arabidopsis* plants. Plants were transformed using *Agrobacterium* mediated transformation. Rosette and cauline leaves were counted. No less than 6 plants from each line used for the leaf count.

Genotype	Generation	No of lines	Homozygous	Leaf count
Col	n/a	n/a	n/a	10.8±0.8
L*er*	n/a	n/a	n/a	10.1±1
co-2	n/a	n/a	Y	21.4±0.6
COF113(Col+*35S::LpCOL1*)	T3	n/a	Y	10.0±1.1
COF114(Col+*35S::LpCOL1*)	T3	n/a	Y	10.6±0.5
COF115(Col+*35S::LpCOL1*)	T3	n/a	Y	10.3±0.7
Average Col+*35S::LpCOL1*	T3	3	Y	**10.3±0.8**
COF164(*co-2+35S::LpCOL1*)	T3	n/a	Y	20.9±0.8
COF165(*co-2+35S::LpCOL1*)	T3	n/a	Y	23±0.8
COF166(*co-2+35S::LpCOL1*)	T3	n/a	Y	21.2±0.2
Average *co-2+35S::LpCOL1*	T3	3	Y	**21.7±1.3**

Fig. 3.28 *LpCOL1* overexpression in *Arabidopsis co-2 and* wt (Col) plants.
There was no statistically significant difference in flowering time between wt vs. *35S::LpCOL1 Col* plants (**A**) and co-2 vs. *35S::LpCOL1* co-2 plants. **(B)**. Plants were transformed using *Agrobacterium* mediated transformation.

Wild type Columbia plants containing overexpressed *LpCOL1* gene flowered after approximately 10.3 leaves which is just slightly earlier and statistically not significant compared with the wild type. Similarly, *co-2* mutants containing the same construct flowered after 21.7 leaves, not significantly different from the 21.3 leaves in the *co-2* plants using Students t-test.

3.3.3.2 Overexpression of LpCOL1 rescues Arabidopsis gi-3 mutant

Although the previous experiment indicated that *LpCOL1* does not rescue *co-2* mutant, it has been shown in *Arabidopsis* that *AtCO* acts downstream from *AtGI* and that overexpression of *AtCO* can also correct the late flowering phenotypes in *gi-3* mutant (Suarez-Lopez et al., 2001). To see if this might be the case with *LpCOL* the same construct as from the *co-2* complementation experiment was used (Fig 3.26) in order to generate *LpCOL1* overexpressor lines in the *gi-3* background. Surprisingly, *LpCOL1* managed to completely rescue *gi-3* mutant plants, which flowered significantly earlier than the L*er* wt (Table 3.5; Figure 3.29). Furthermore, using Students t-test ($P<0.05$), the haploinsufficiency phenomenon previously seen in *gi-3+35S::LpGI* overexpressing plants was also observed here when T2 (heterozygous, BASTA surviving) plants were compared with T3 (homozygous) plants of the same genotype. Heterozygotes showed a flowering phenotype somewhere in between the homozygous mutants and wild-type suggesting that heterozygotes do not produce enough CO protein to promote early flowering.

Table. 3.5 35S::*LpCOL1* complementation of the *gi-3* mutant *Arabidopsis* plants.
Plants were transformed using *Agrobacterium* mediated transformation. Rosette and cauline leaves were counted. T3 lines are listed separately. No less than 6 plants from each line used for the leaf count. Asterisk point to statistically significant difference in leaf number found between T2 and T3 plants.

Genotype	**Generation**	**No of lines**	**Homozygous**	**Leaf count**
L*er*	n/a	n/a	n/a	11.4±0.5
gi-3	n/a	n/a	Y	26±1.4
*gi-3+35S::LpCOL1**	T2	5	N	12.2±2.1
COF159-5(*gi-3+35S::LpCOL1*)	T3	n/a	Y	6.4±0.5
COF160-2(*gi-3+35S::LpCOL1*)	T3	n/a	Y	8.1±1.2
COF161-1*(gi-3+35S::LpCOL1*)	T3	n/a	Y	9.2±0.4
Average gi-3+*35S::LpCOL1**	T3	3	Y	8.3±1.5

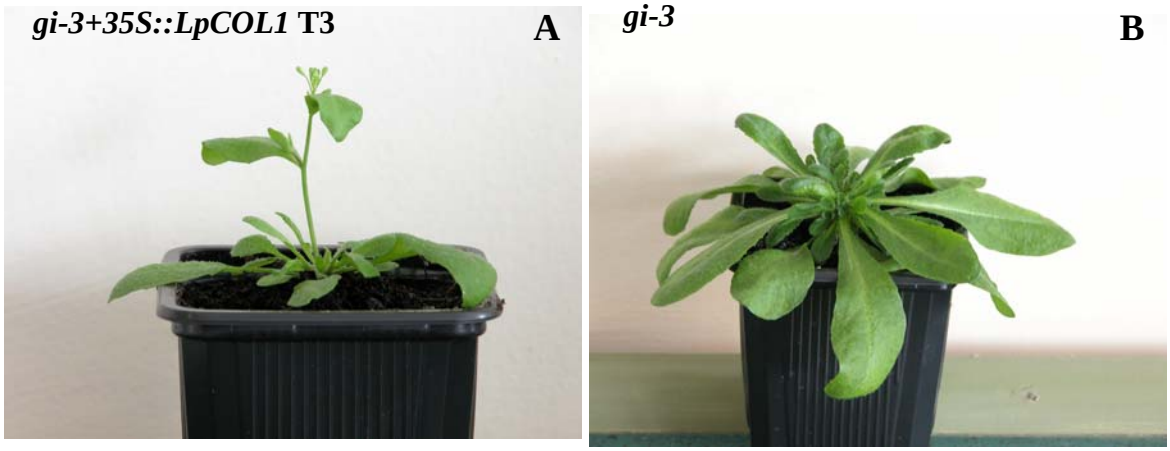


Fig. 3.29 *LpCOL1* complementation of the *gi-3 Arabidopsis* plants.
gi-3 plant with 35S::*LpCOL1* gene (**A**) compared with the *gi-3* mutant plant (**B**).

3.3.4 Mapping of the *LpCOL1* gene

The same linkage map used for mapping *LpGI* gene was used to map the ryegrass *COL1* gene. The map served as the framework for mapping the *LpCOL1* gene, as a single nucleotide polymorphysm (SNP) marker. *LpCOL1* gene was amplified using MG048/MG041 primers (Tm 61°C, extension time 1 m 30 s, 42 cycles; Table 2.1), and SNPs determined as described in Chapter 2.2.6. The allelic status was determined by direct sequencing of the parents (cv. Impact and cv. Samson) DNA. Alignment of the amplified regions revealed 5 candidate SNPs which were suitable (heterozygous in one or both parents, Table 3.6). For 131 out of 188 genotypes readable results were obtained and the SNP at cDNA position 494 produced a clear chromatogram peak distinction and an expected allele frequency (Mendelian segregation ratio 1:1; Appendix, Table 5.1). In order to fully utilize the mapping population and to check the obtained results from the direct sequencing, a dCAPS method was devised for *LpCOL1* mapping (Chapter 2.2.2.10).

Table. 3.6 SNPs of *LpCOL1* considered for mapping. Five SNPs were discovered during the initial screening. SNP 1 was chosen for being easly distinguishable and consistent in sequencing signal strength. $COL1^{Sn}$ and $COL1^{Ip}$ represent SNPs found in Samson and Impact varieties respectively.

SNPs	$COL1^{Sn}$	$COL1^{Ip}$	cDNA pos	aa translation	conservation
1	AA	GA	494	Asp/Gly	Low
2	CC	GC	622	Glu/Gln	Low
3	GG	TG	647	Glu/Val	Low
4	GG	CG	intron	n/a	Low
5	TT	CT	intron	n/a	Low

The SNP used in mapping by direct sequencing was part of a *MboI* restriction site. By amplifying the *LpCOL1* gene and then cutting it with *MboI* (Appendix, Fig 5.7) the presence or absence of the additional band would indicate the allelic status (Fig 3.30A). Analysis of these results and subsequent mapping showed that *LpCOL1* mapped at the end of the LG 6 at a position of 37 cM (Fig 3.30B). When blasted against rice DNA database (www.gramene.org) *LpCOL1* showed 85.2% identity with a 100 bp segment from the clone AP005284 at position 30.3 Mb on rice chromosome 2. This region in rice is predicted to encode a 104 aa hypothetical protein (ID XP_467590.1) as predicted by GlimmerM. Furthermore protein identity between LpCOL1 B-box2 and 77-108 segment of the hypotetical protein was found to be 78.1%. These results indicate high homology between *LpCOL1* and the described region on rice LG2.

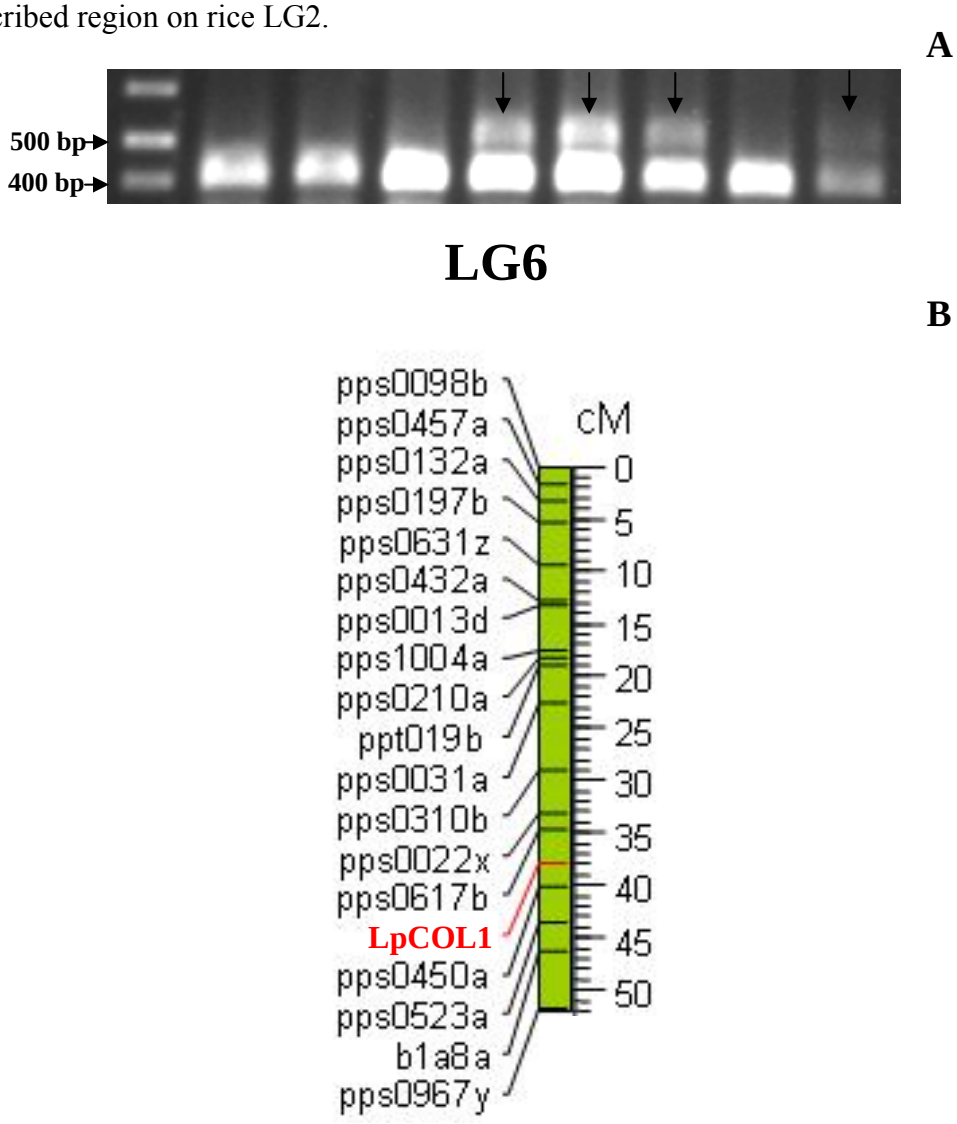


Fig 3.30 Mapping of the *LpCOL1* gene. **A**. Presence (black arrows) or absence of the 500 bp band in the selection of different progeny genotypes indicate allelic status of the *LpCOL1* gene which was used to map it on the linkage map **B**. The gene mapped at the end of LG6 at position of 38 cM

3.3 *FLOWERING TIME 3* (*LpFT3*)

FT and *Hd3a* genes in *Arabidopsis* and rice respectively, have been known as the floral integrators of several different flowering time pathways including photoperiod. In order to obtain the full picture of the photoperiod pathway and to see if ryegrass *FT* has the same function, the ryegrass *FT* gene was characterized and its function, expression, and relationship to the other photoperiod genes investigated.

3.3.1 Identification, sequencing and analysis of the *LpFT3* gene

Previous work by Dr Igor Kardailsky identified three *FT*-like sequences using degenerate oligos based on rice *Hd3a* nucleotide and protein sequences. Sequence comparison identified a segment of the *LpFT3* gene, amplified with degenerate primers GIK18/19 (Table 2.1), as the one with the highest nucleotide identity (85%) to the rice *Hd3a*. In his subsequent work, using multiple RACE-PCRs, Dr Kardailsky revealed missing 5' and 3' segments of the *LpFT3* cDNA. From these sequences the *LpFT3* contig was assembled, and the full length cDNA, gDNA, and protein sequence deduced (Appendix, Figure 5.8). The gene is submitted to the GeneBank database (DQ309592). The reminder of the work was solely done by Milan Gagic
On the protein level *LpFT3* gene is 97% identical with wheat *TaFT* and barley *HvFT1* genes, 89% with rice *Hd3a*, and 70% with *AtFT* (Appendix, Figure 5.9).
The overall structure of the *LpFT3* gene is slightly different when compared with the structure of the rice *Hd3a* and *Arabidopsis AtFT* genes (Fig 3.31). Exons 1 and 2 are of almost identical size in all three genes with the introns of variable size between them. Exons 3 and 4 in *Arabidopsis* and rice (41/224 bp and 41/227 bp respectively) are joined and represented in the *LpFT3* gene by only one exon of 271 bp.

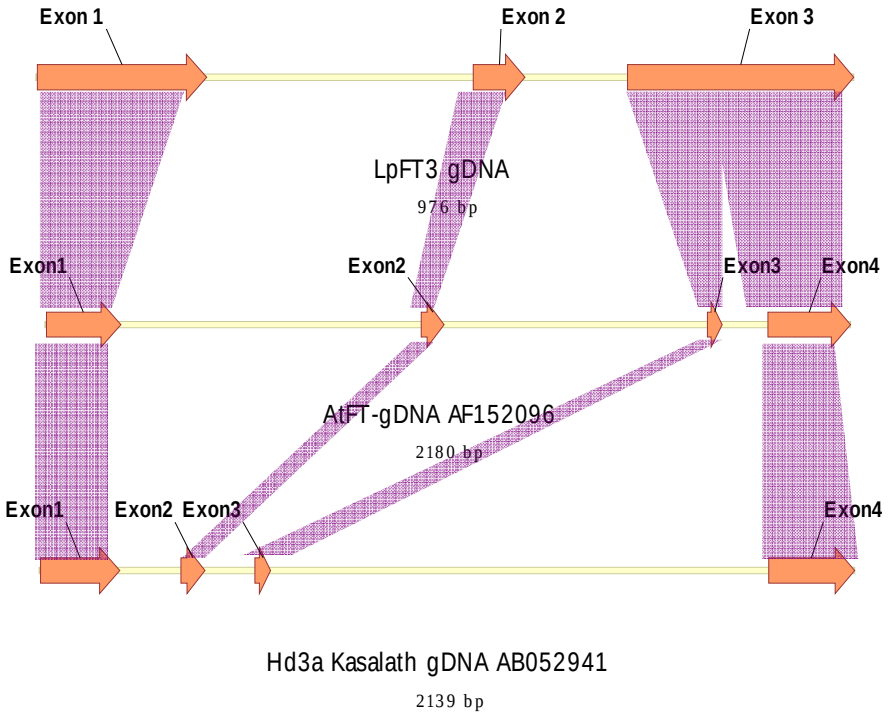


Fig 3.31 Structure of the ryegrass, *Arabidopsis* and rice *FT* genes.

3.3.2 Expression analysis of the *LpFT3* gene

The *FT* gene in *Arabidopsis* is considered to be the most potent activator of flowering (Kardailsky et al., 1999) since overexpression of the *FT* gene accelerates floral development, and normal photoperiodic response is lost. In order to determine its expression pattern plants were cultivated under different environmental conditions, sampled and RNA extracted using Trizol® reagent according to the recommended protocol (Chapter 2.2.2.1.3). Reverse transcription was performed using SuperScript® (Invitrogen) reverse transcription system (Chapter 2.2.2.5). Real time PCR (RT-PCR) was performed with SYBR-green as a fluorescent reporter on i-cycler (Chapter 2.2.3). The primers used were GIK19/20 (Tm 63°C, extension time 25 s, 45 cycles; Table 2.1) and the results were normalized against ubiquitin and *GAPDH* levels (Chapter 2.2.3). At least three PCR reactions using the same template were performed to get average values of expression levels.

3.3.2.1 LpFT3 gene is under the control of the circadian clock

It was shown in *Arabidopsis* that *AtFT* is regulated by the circadian clock and the aim of this experiment was to see if this is the same case with the ryegrass *LpFT3* gene. Plants for this experiment were grown and sampled under the same conditions as for *LpGI* and *LpCOL1* free running experiments (Chapters 3.1.2.1 and 3.2.2.1).

Results clearly show that *LpFT3* is regulated by the circadian clock (Fig 3.32, Table 5.2 and 5.3) with the peaks close to the subjective night periods and rhytmicity maintained throughout the experiment.

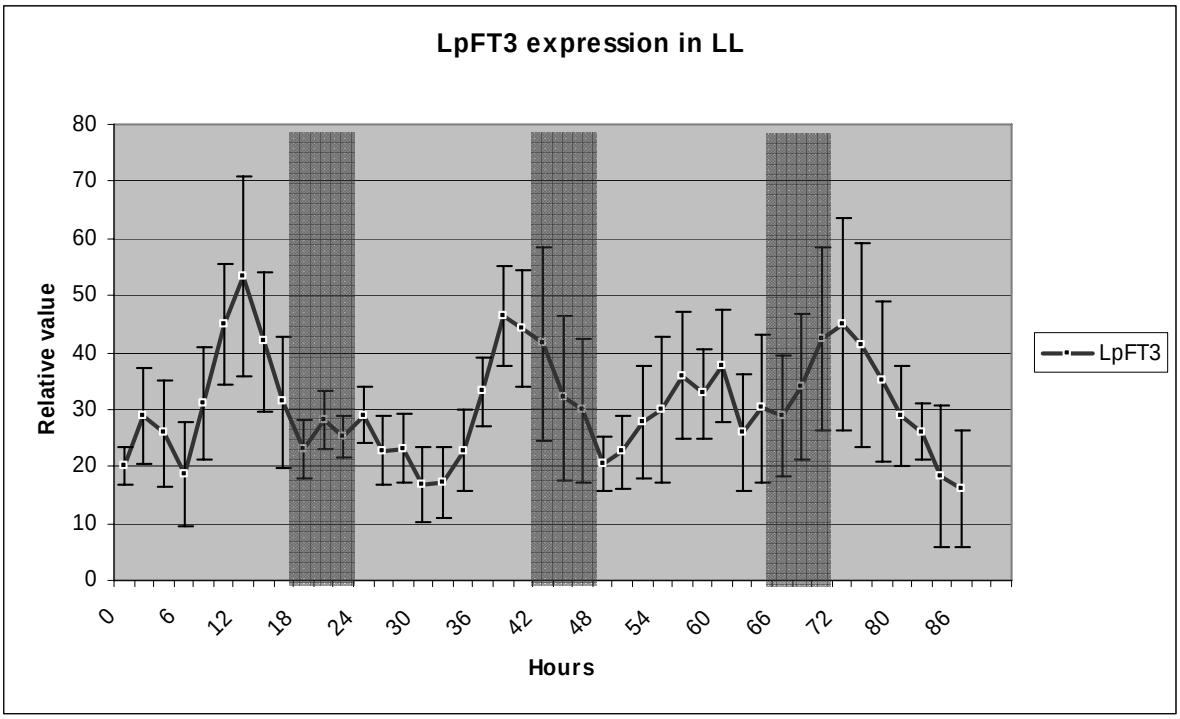


Fig 3.32 Effect of cicrcadian clock on *LpFT3* expression.
The plant growing conditions were the same as described for the *LpGI* and *LpCOL1* free running experiment. *LpFT3* expression levels were analysed by RT-PCR. The shaded areas represent subjective night. Results were normalised against *Ubi/GAPDH* levels.

3.3.2.2 LDs promote flowering in ryegrass through LpFT3 activation

Early flowering in *Arabidopsis* was correlated with *AtFT* mRNA accumulation (Kardailsky et al., 1999: Kobayashi et al., 1999). To extend this knowledge to ryegrass, *LpFT3* expression was correlated with day length, an external cue necessary for the flowering induction. Results presented (Figure 3.33) show an almost 30 fold difference in *LpFT3* expression between plants grown in LD and plants grown in SD conditions. In SD *LpFT3* peaked approximately 7 h after dawn with trough levels ~ 5 h later while in LD plants LpFT3 peaked slightly later but the peak was very wide spanning ~ 5 h with the through level ~3 h later all of which occurred during the daylight. This is consistent with the results seen previously in *Arabidopsis*.

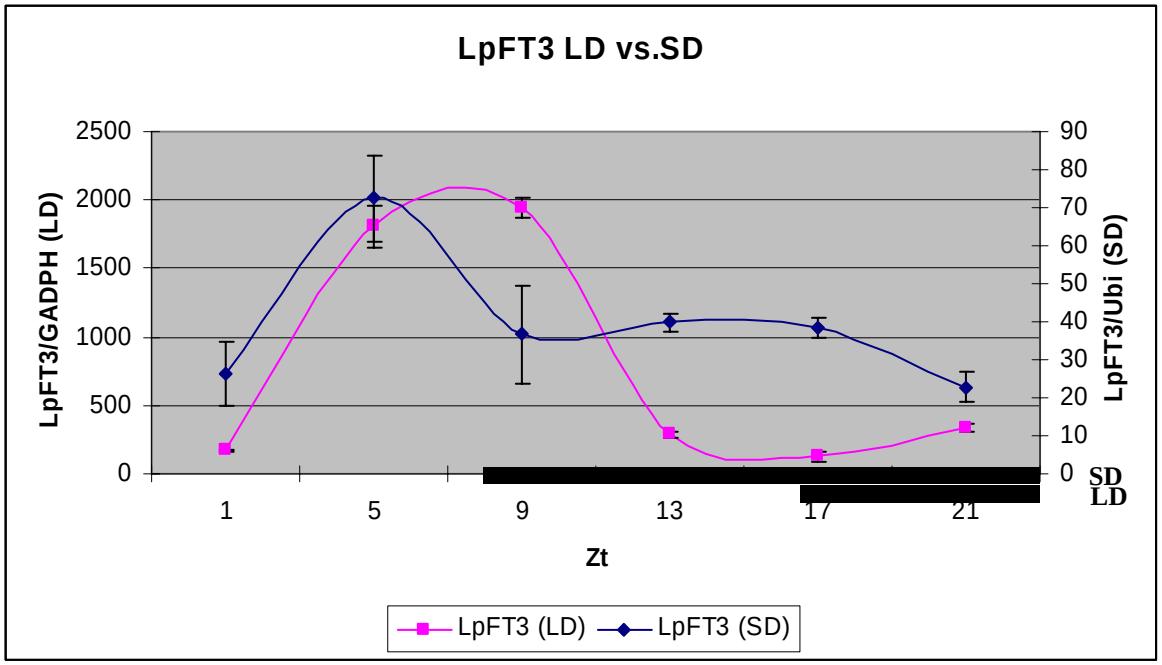


Fig. 3.33 Diurnal expression of the *LpFT3* gene. *LpFT3*-LD represent samples collected in LD conditions, *LpFT3*-SD are the ones from the SD conditions. Note the difference in the scale between two of them. There is almost 30 fold difference in expression levels between LD and SD conditions. The black bars under the graph represent respective night periods. Results were normalized against *GAPDH* levels.

3.3.2.3 LpFT3 acts as a flowering integrator

AtFT was shown to integrate signals from photoperiod, vernalisation, and autonomous pathways. Considering that *Lolium perenne* cv. Impact plants require vernalisation period of at least 8 weeks it was interesting to see if the vernalisation process affects *LpFT3* expression. In order to investigate that ryegrass plants were differentially vernalised, transferred to LDs and grown and sampled as described in Chapter 2.2.1.3.

The expression of the *LpFT3* gene in LDs was closely linked with the duration of the vernalisation period (Fig 3.34). In the first five weeks *LpFT3* gene expression was very low except for week 3 when a significant increase was observed. A steady increase in *FT3* expression was seen after 5 weeks of vernalisation which continued until the end of the experiment.

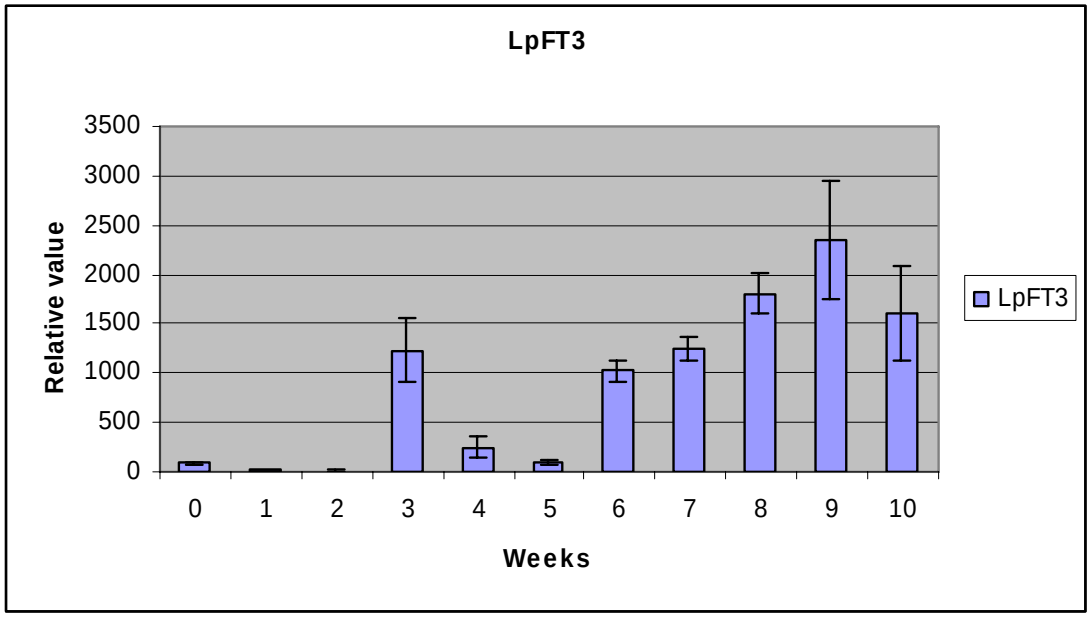


Fig. 3.34 Expression of the *LpFT3* gene in relation to the length of vernalisation treatment. Plants were differentially vernalised (0-10 weeks) and transferred into LD for tissue collection. Collection was done after 7 days at ZT16. The results were normalised against *Ubi*.

In the same context there was a question of how long does it take for ryegrass to initiate flowering sequence of events after being exposed to the floral promotive environmental conditions by analysing expression levels of the floral integrator *LpFT3*.

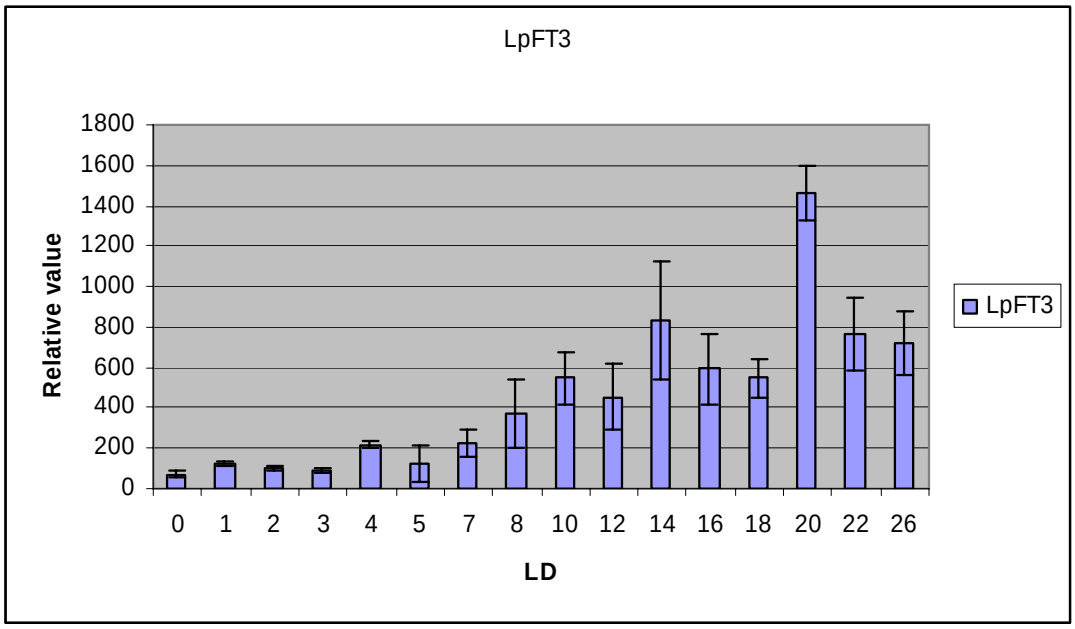


Fig. 3.35 Expression of the *LpFT3* gene in relation to the number of LDs after vernalisation. Plants were fully vernalised (10 weeks) and transferred into the LD for tissue collection. Collection was done every day from day 0-26 at ZT16. The results were normalised against *Ubi/GAPDH*.

Ryegrass plants were grown as described in Chapter 2.2.1.3. The *LpFT3* gene was found to have a very low level of expression in the first 3 days after which the expression started to rise reaching the plateau approximately 14 days later and staying at similar levels thereafter (Fig 3.35).

3.3.3 Functional analysis of the *LpFT3* gene

On the protein level ryegrass LpFT3 shows 89% identity with rice Hd3a, and 70% with AtFT. Overexpression of the *Hd3a* and *AtFT* cause early heading in rice (Kojima et al., 2002), and early flowering in *Arabidopsis* (Kardailsky et al., 1999) which further shows that *Hd3a* promotes floral transition as does *AtFT*. To see if overexpression of ryegrass *LpFT3* in *Arabidopsis* will cause an early flowering phenotype in transformed plants and to better understand the function of *LpFT3*, transgenic *Arabidopsis* plants carrying sense *LpFT3* cDNA under the control of CaMV35S promoter were generated and analysed under the LD conditions. Gateway compatible primers MG058/GIK40 (Table 2.1) were used to obtain full length of the *LpFT3* (Tm 59°C, extension time 50 s, 40 cycles) This segment starts at ATG codon and includes the complete *LpFT3* cDNA including 40 bp of the 3'UTR. The PCR product was cloned into the pEntr/D vector (Invitrogen; Fig 3.36) and transformed into the TOP10 competent cells (Invitrogen). Kanamycin resistant cells were selected, plasmid isolated and insert sequenced for the potential PCR errors (data not shown). The sequence was free of PCR generated point mutations encoding the full length of the LpFT3 protein.

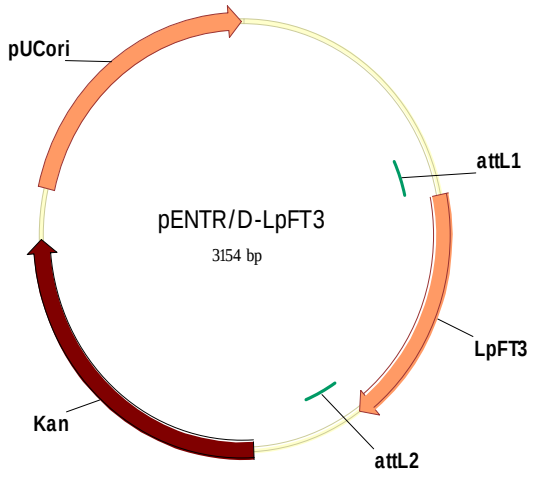


Fig. 3.36 Cloning of the *LpFT3* cDNA. *LpFT3* cDNA was amplified using Gateway compatible primers, cloned into the pEntrD vector and transformed into the TOP10 E coli cell line

The insert was recombined using LR clonase enzyme mix (Invitrogen, Chapter 2.2.2.2) into the pRSh1 binary vector containing constitutive 35S promoter (Fig 3.37A) This vector contains the same features as the one previously used for the *LpGI* and *LpCOL1* functional analysis including spectinomycin bacterial selection, the *BAR* gene as a selectable plant

marker and the CaMV35S. The insert was again checked by restriction digest, using three restriction enzymes *SmaI*, *PvuII,* and *NcoI* (Fig 3.37B).

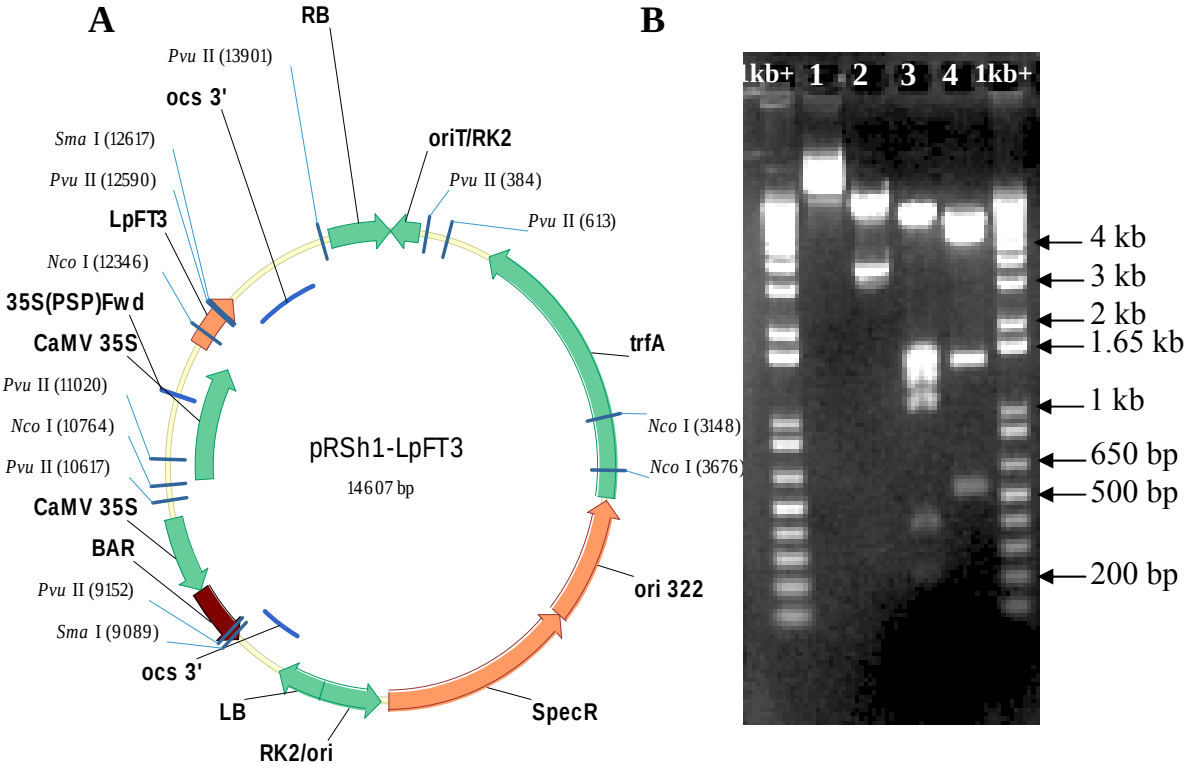


Fig 3.37 Map and the restriction digest of the pRSh1-*LpFT3* binary vector.
A. The vector has the *LpFT3* gene cloned downstream of the constitutive 35S CaMV promoter **B**. The *LpFT3* insert was checked with *SmaI* (2), *PvuII* (3), and *NcoI* (4) which produced bands of expected sizes. Lane 1 is an uncut construct. The digests were run alongside 1kb+ DNA marker

3.3.3.1 Overexpression of the LpFT3 gene in Arabidopsis

To characterise the flowering time phenotype of the *Arabidopsis ft-1* mutants and Col wt plants containing *LpFT3* cDNA fragment under the 35S constitutive promoter, *Arabidopsis* plants of respective backgrounds were transformed (Chapter 2.2.4.2), and >20 BASTA resistant plants (T1) selected. In the T2 generation the lines were again selected by BASTA resistance and 15 T2 lines that segregated 3:1 (single transgene) were selected and the leaf number counted. In the T3 generation another round of BASTA selection revealed 5 homozygous lines (no segregation for the transgene) with 3 lines were selected and the leaves counted. The *ft-1* mutation converts a glycine to a glutamine in the position 171 producing late flowering plants (Kardailsky et al., 1999). Expression of the *LpFT3* gene should rescue and accelerate flowering in *ft-1* and wt plants respectively. In our experiments transformed *ft-1 Arabidopsis* plants were completely rescued and flowered even earlier than the wt plants

(Table 3.7). Transgenic 35S::*LpFT3* plants in *ft-1* background flowered after 6.5 leaves, compared with 24.5 leaves in non-transgenic *ft-1* plants and 10.8 leaves in wt plants (Fig 3.37A, B, and C).

Similarly to *LpGI* and *LpCOL1*, effect of *LpFT3* was stronger in homozygous lines than in the heterozygous ones indicating the existence of the dosage effect.

Table. 3.7 35S::*LpFT3* complementation of the *ft-1* mutant *Arabidopsis* plants.
Plants were transformed using *Agrobacterium* mediated transformation. Rosette and cauline leaves were counted. Separate T3 lines are listed. No less than 6 plants from each line used for the leaf count.

Genotype	**B Generation**	**No of lines**	**Homozygous**	**C Leaf count**
L*er*	n/a	n/a	n/a	10.8±0.9
ft-1	n/a	n/a	Y	24.5±1
Average *ft-1+35S::LpFT3*	T2	15	N	8.5±1
COF131-2 (*ft-1+35S::LpFT3*)	T3	n/a	Y	6.6±0.7
COF138-1 (*ft-1+35S::LpFT3*)	T3	n/a	Y	6.7±0.5
COF143-2 (*ft-1+35S::LpFT3*)	T3	n/a	Y	6.0±0.5
Average *ft-1+35S::LpFT3*	T3	3	Y	6.4±0.6

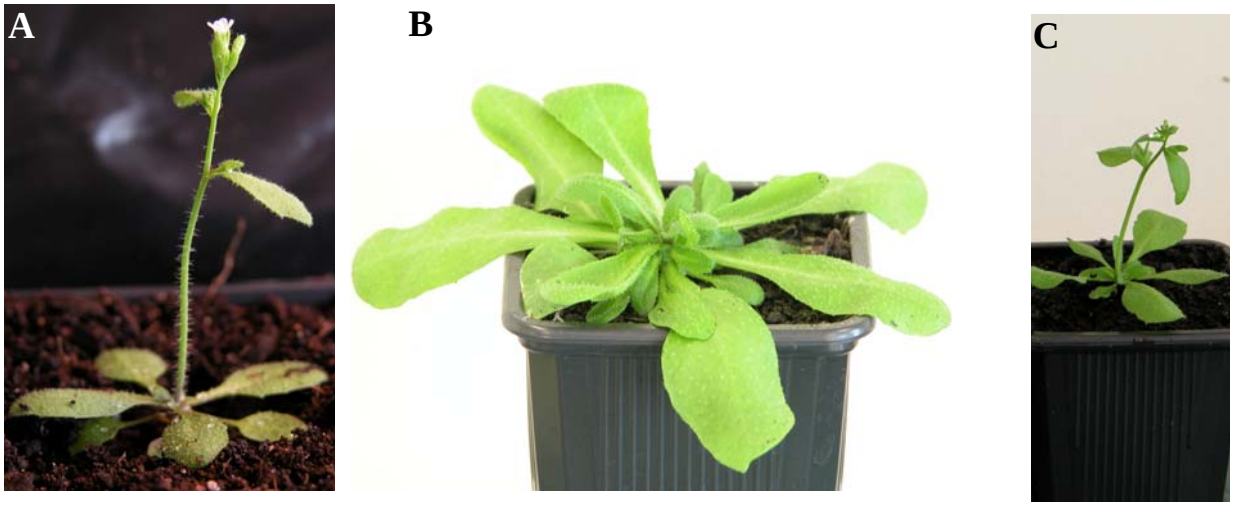


Fig. 3.38 35S::*LpFT3* complementation of the *ft-1 Arabidopsis* plants.
A.The *35S::LpFT3* construct fully rescued the late flowering phenotype of the *ft-1* plants flowering approximately after 6.4 leaves compared with 24.5 in *ft-1* mutants, and 10.8 in L*er* wt plants. **B**.. *ft-1* **C**. L*er* (wt). .

These results were corroborated with the experiments in which *Arabidopsis* wild type plants (Col) were used as the vehicles for *LpFT3* overexpression. The plants were transformed with the same construct and the number of leaves at flowering were counted. It was found that *LpFT3* significantly accelerated wild type plants so that in T3 generation they flowered approximately after 7.8 leaves in contrast to the untransformed ones which flowered after 10.8 leaves (Table 3.8; Fig 3.39).

Table. 3.8 35S::*LpFT3* complementation of the *Col wt Arabidopsis* plants.
Plants were transformed using *Agrobacterium* mediated transformation. Rosette andcauline leaves were counted. No less than 6 plants from each line used for the leaf count.

Genotype	**Generation**	**No of lines**	**Homozygous**	**Leaf count**
Col	n/a	n/a	n/a	10.8±0.8
COF168-1 (Col+*35S::LpFT3*)	T3	n/a	Y	6.5±0.5
COF169-3 (Col+*35S::LpFT3*)	T3	n/a	Y	8.5±0.6
COF170-3 (Col+*35S::LpFT3*)	T3	n/a	Y	8.7±0.5
Average Col+*35S::LpFT3*	T3	3	Y	7.8±1.2

Fig. 3.39 *LpFT3* overexpression in *Arabidopsis* (Col) wt plants.
A. Col wt plant **B**. *35S::LpFT3* construct accelerated flowering of the wild type Col plants which flowered approximately after 7.8 leaves compared with 10.8 in Col wt plants.

3.3.3.2 AtFT expression in LpGI, LpCOL1, and LpFT3 transgenenic plants

It has been well established that the *FT* gene acts as a flowering integrator, downstream of *GI* and *CO* genes (Kardailsky et al., 1999; Kobayashi et al., 1999). Any change upstream of the *FT* gene would be reflected on the *FT* expression levels. In order to see if the sequence of events in the ryegrass photoperiod pathway is similar to the *Arabidopsis* one and to investigate the unexpected results regarding *LpCOL1* gene function, the expression of the *Arabidopsis AtFT* gene was measured in different genetic backgrounds and the results were correlated with the flowering time measurements in respective lines. To perform this experiment *Arabidopsis* plants of different genotypes were grown in LD conditions for 14 days and sampled them at ZT0 when the expression of *FT* is thought to be the lowest, and ZT16 when it is at its highest. The samples were analysed by RT-PCR (Chapter 2.2.3) using GIK05/06 primer pair (Table 2.1; Tm 56°C,extension time 20 s, 45 cycles) expected to amplify specific 140 bp segment from *AtFT* cDNA. The number of leaves at flowering was counted within the analysed lines and correlated with the expression levels. Results were listed in the table 3.6.

Initially, the *AtFT* expression level was measured in *Arabidopsis* Col and L*er* wild type plants, *gi-3*, *co-2*, and *ft-1* mutants. As expected, in Col and L*er* plants expression was low in the morning (ZT0), and high in the afternoon (ZT16; Table 3.9). In contrast to this in *gi-3* and *co-2* mutants the *FT* expression was consistently low (Table 3.9). In *ft-1* mutants transcript of the mutated *FT* gene was cycling as in the wt plants, but the non-functional protein was causing the late flowering phenotype. Leaf numbers from these lines were consistent with the expression levels. Col and Ler wt plants flowered after 15.6 and 11.2 average leaves respectively while the mutant lines flowered much later.

To see if *LpGI* functions within the same photoperiod hierarchy *AtFT* expression in transgenic *Arabidopsis* plants containing 35S::*LpGI* construct in the *gi-3* background was investigated. A change in the *AtFT* expression would indicate a change in the upstream regulatory mechanism. The expression level of three transgenic lines is presented in the Table 3.6 (Lines COF92-4, COF 94-1, and COF96-2). The level of the *AtFT* increased when compared to the *AtFT* level in the *gi-3* mutant. All three of them had expression levels similar to each other, producing the phenotypes that were similar in relation to the numbers of leaves at flowering.

The *35S::LpCOL1* construct did not accelerate flowering in Col wt plants (Fig 3.27). To find out if *AtFT* levels were consistent with this observation its expression in the Col plants overexpressing *LpCOL1* was checked. It was found that the *AtFT* levels were just slightly

elevated in all three independent lines (COF113, COF114, and COF115) which eventually did not reflect on the overall flowering times measured by the number of leaves before flowering. The plants flowered after approximately the same number of leaves compared with the Col wt plants.

Table. 3.9 Expression of *AtFT* in various transgenic lines.
Plants of respective backgrounds were transformed with constructs overexpressing *LpGI*, *LpCOL1*, and *LpFT3* genes using *Agrobacterium* mediated transformation. Expression of the endogenous AtFT gene was measures by the RT-PCR. Results were normalised against *Ubi*. Analysed lines were allowed to flower and rosette and cauline leaves were counted.

Label	B/ground	Construct	ZT	FT_norm	FT_err	Mean Leaf No
Col	Col	n/a	0	87	7%	15.6
Col	Col	n/a	16	1412	5%	
L*er*	L*er*	n/a	0	119	7%	11.2
L*er*	L*er*	n/a	16	1749	5%	
gi-3	*gi-3*	n/a	0	202	6%	26
gi-3	*gi-3*	n/a	16	99	14%	
co-2	*co-2*	n/a	0	99	16%	24.3
co-2	*co-2*	n/a	16	85	10%	
ft-1	*ft-1*	n/a	0	98	13%	24.5
ft-1	*ft-1*	n/a	16	1870	15%	
COF113	Col	35S::LpCOL1	0	88	3%	15.3
COF113	Col	35S::LpCOL1	16	1664	14%	
COF114	Col	35S::LpCOL1	0	139	16%	14.8
COF114	Col	35S::LpCOL1	16	1701	8%	
COF115	Col	35S::LpCOL1	0	175	13%	14
COF115	Col	35S::LpCOL1	16	1482	19%	
COF138-1	*ft-1*	35S::LpFT3	0	189	10%	6.7
COF138-1	*ft-1*	35S::LpFT3	16	4480	19%	
COF159-5	*gi-3*	35S::LpCOL1	0	300	10%	6.6
COF159-5	*gi-3*	35S::LpCOL1	16	1186	7%	
COF160-1	*gi-3*	35S::LpCOL1	0	245	10%	7.7
COF160-1	*gi-3*	35S::LpCOL1	16	804	10%	
COF168-1	Col	35S::LpFT3	0	332	13%	6.5
COF168-1	Col	35S::LpFT3	16	3263	29%	
COF169-3	Col	35S::LpFT3	0	300	8%	8.5
COF169-3	Col	35S::LpFT3	16	1334	25%	
COF170-3	Col	35S::LpFT3	0	427	17%	8.6
COF170-3	Col	35S::LpFT3	16	1157	19%	
COF92-4	*gi-3*	35S::LpGI	0	145	13%	6.7
COF92-4	*gi-3*	35S::LpGI	16	623	28%	
COF94-1	*gi-3*	35S::LpGI	0	147	11%	6.4
COF94-1	*gi-3*	35S::LpGI	16	1567	8%	
COF96-2	*gi-3*	35S::LpGI	0	94	19%	6.8
COF96-2	*gi-3*	35S::LpGI	16	1176	12%	

Contrary to this *35S::LpCOL1* construct fully compensated for the flowering defect in *gi-3* mutant plants as shown by the functional analysis on transgenic *Arabidopsis* plants (Chapter 3.3.3.2). Expression analysis of the endogenous *Arabidopsis AtFT* gene in *gi-3* plants overexpressing *LpCOL1* showed its expression rising to the levels seen in wt plants which indicate that *LpCOL1* regulates expression of the *FT* gene either directly or through some of the possible intermediaries.

The *35S::LpFT3* construct managed not just to completely rescue *ft-1* mutant but cause extremely precocious flowering. Similarly, the same construct caused *Arabidopsis* (Col) wild type plants to flower earlier (Lines COF168-1, COF169-3, and COF170-3, Table 3.9).

It was also interesting to see whether the constitutive expression of *LpFT3* caused any change in expression of the endogenous *AtFT* gene. When overexpressed in *ft-1* background *LpFT3* caused 2.5 fold increase in endogenous *AtFT* expression (Line COF138-1, Table 3.9) although *AtFT* gene itself, due to its mutation, could not cause premature flowering. Similarly in the *Col* wt background overexpression of the *LpFT3* gene in 3 different lines (COF 168-1, COF169-3, and COF170-3) caused plants to flower earlier with only one line (COF 168-1) showing expression levels of the endogenous *AtFT* elevated in comparison to the L*er* wt plant.

3.3.4 Mapping of the *LpFT3* gene

The same linkage map used for mapping *LpGI and LpCOL1* gene was used to map the ryegrass *FT3* gene. The map served as the framework for mapping the *LpFT3* gene, as a single nucleotide polymorphism (SNP) marker. Initially, the *LpFT3* gene from both parents (cv. Samson and cv. Impact) was amplified using GIK31/GIK32 primers (Tm 50°C, extension time 1 m, 38 cycles; Table 2.1). The PCR products were sequenced and aligned in order to find SNPs suitable as markers for gene mapping. Five SNPs were discovered (Table 3.10) out of which Imp9 on position 180 on the cDNA was chosen as most easily distinguishable (Allele 4, Fig 3.40A). Following the same procedure 188 progeny genotypes were analysed and the informative results were obtained for 159 of them (Table 5.1) giving sufficient data to map the *LpFT3* gene to the chromosome 7 (LG 7) at the position of 12cM (Fig 3.40B).

Table. 3.10 SNPs of *LpFT3* considered for mapping.
Six SNPs were discovered during the initial screening. Allele 4 was chosen for being easly distinguishable and consistent in sequencing signal strength. $FT3^{Sn}$ and $FT3^{Ip}$ represent SNPs found in Samson and Impact varieties respectively

SNPs	$FT3^{Sn}$	$FT3^{Ip}$	cDNA pos	aa translation	conservation
1	AA	GA	32	Val/Gly	High
2	TT	GT	89	Val/Gly	High
3	CC	GC	103	Arg/Gly	Low
4	GG	CG	180	Glu/Asp	Low
5	GG	AG	242	Asn/Ser	Low

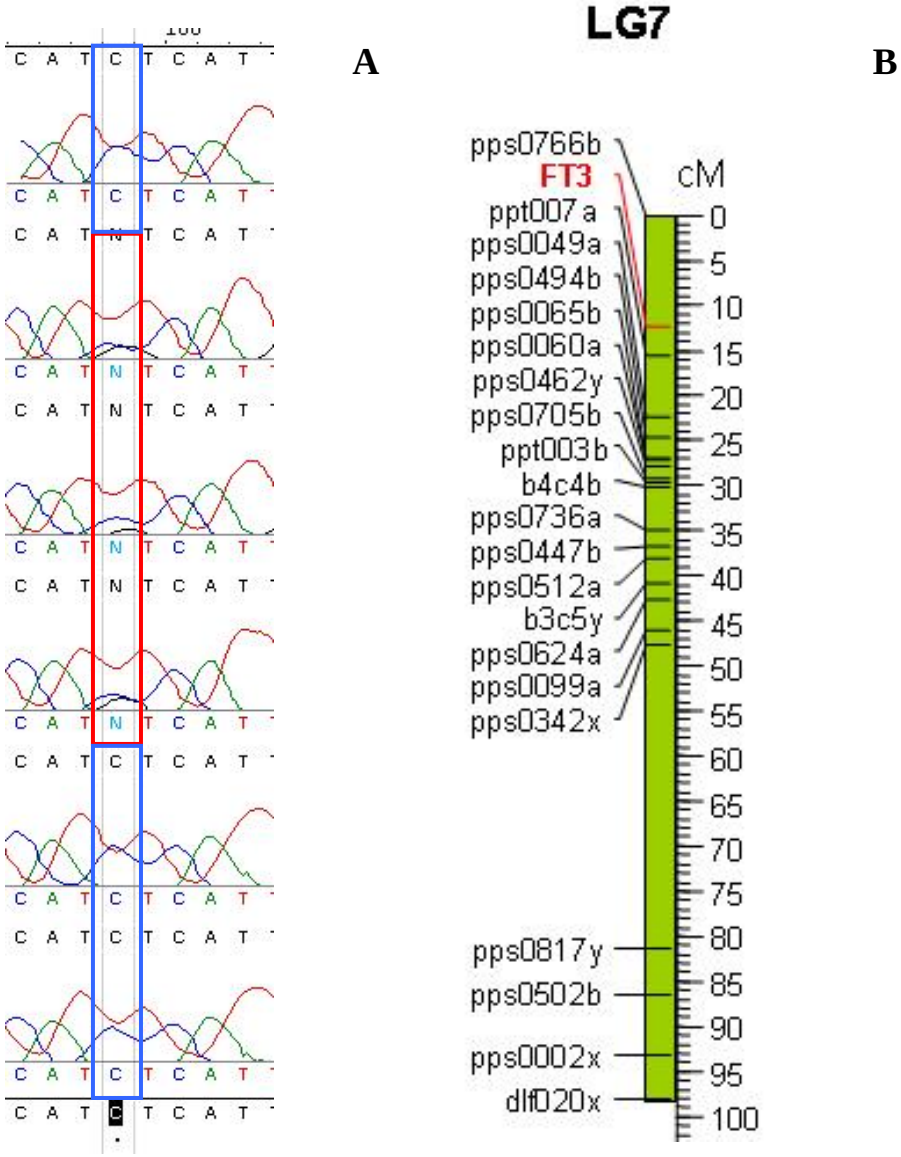


Fig 3.40 Mapping of the *LpFT3* gene.
A. Presence (red box) or absence (blue box) of the G peak on the position 297 indicated allelic status of the *LpFT3* used to map it on the linkage map. **B.** The *LpFT3* gene was mapped at the LG6 on the position 12cM.

4 CONCLUDING DISCUSSION

The model plants *Arabidopsis* and rice have provided a wealth of information about genes and genetic pathways controlling the flowering process, but little is known about the corresponding pathways in forage plants. Ryegrass (*Lolium perenne*) is one of the most important temperate forage grasses but the lack of molecular information about ryegrass flowering genes has prevented direct comparison with other species. To address this problem expressed sequence tags and genome sequence databases were searched and three flowering photoperiod-related genes from ryegrass identified. Here, their structure, role within the photoperiod pathway and implications for future study is discussed.

4.1 Conservation of the genetic structure

4.1.1 The *LpGI* gene is an ortholog of the *Arabidopsis GI* gene

An extensive search through the ryegrass EST libraries using the *AtGI* and *OsGI* sequences revealed a partial ryegrass *GI* cDNA highly homologous to the *OsGI* and *AtGI* genes. Since the sequence comprised only part of the gene, subsequent degenerate PCR and RACE-PCR revealed the *LpGI* gene which is 3447 bp long, encoded a 1149 aa protein with a molecular mass of 127 kDa. Compared with the *Arabidopsis* and rice (1173 aa), barley and wheat (1155 aa) GI proteins, the ryegrass GI has the shortest amino acid sequence (1149 aa). On the protein level it has high identity to the other GI proteins (*Arabidopsis* 65%, rice 85%, and wheat 91%) and the level of similarity is almost identical on both nucleotide and protein levels. The *LpGI* gene shows a structure almost identical to the *Arabidopsis* and rice *GI* genes. The *LpGI* is 5497 bp long, which is somewhere in between *AtGI* (5050 bp) and *OsGI* (6710 bp). It is interesting that although phylogenetically ryegrass is closer to rice, the genetic structure of *LpGI* is more similar to the *Arabidopsis AtGI*; it contains 13 exons as opposed to

rice *OsGI* with 14. This reduction in the number of exons came from fusion of exons 4 and 5 in rice into a single exon 4 in *Arabidopsis* and ryegrass.

The web-based membrane topology prediction programs TMpred and TopPred predicted that the *Arabidopsis* GI protein contains up to 11 transmembrane domains and that the rice protein contains up to six (Fowler et al., 1999). Four of these were highly conserved in both species. Up to 13 transmembrane helices were found in LpGI when analysed using topology prediction programs. Among these 13 there were four strongly preferred models which correlated well with the four conserved regions in *Arabidopsis* and rice (Fig 3.5). These results indicate a high level of sequence conservation in the *GI* genes especially in the N-terminal part of the protein where the conserved transmembrane domains are positioned. To corroborate this data the LpGI protein sequence was analysed with the PSORT (Prediction of Protein Localization Sites) program which indicated nucleus or plasma membrane localization. Previous work on AtGI subcellular localization showed that *Arabidopsis* GI protein is localized exclusively in the cell nucleus (Huq et al., 2000). Furthermore using several different GUS-GI constructs they showed that region of GI between residues 543 and 783 is required for the nuclear localization.

4.1.2 *LpCOL1* is a member of the *CONSTANS-LIKE* gene family

The *Arabidopsis* photoperiod pathway is currently the most completely understood aspect of flowering time control with the same pathway also highly conserved in rice. This study describes the ryegrass *LpCOL1* gene, homologue of the *Arabidopsis AtCO*, and rice *OsHd1*. Homology in this context represent "relationship of two characters that have descended, usually with divergence, from a common ancestral character" (Fitch, 2000). The deduced LpCOL1 protein has 369 amino acid residues with a molecular mass of 41kDa. At the protein level LpCOL1 is most similar to the wheat TaHd1 protein (78%), followed by rice OsHd1 (56%), and *Arabidopsis* AtCO (37%). The overall nucleotide structure of *LpCOL1* is very similar to the three previously mentioned genes having two exons of 759 and 357 bp respectively, separated by a 575 bp intron located between the B-box and CCT domain. This structure is by no means ubiquitous among *CO-like* genes. *CO* and *CO-like* genes may contain 1 intron (Group I and II; *CO*, *COL1* to *COL8*), three introns (Group III; *COL9* to *COL15*) or no introns at all (*HvCO8*, *BnCOL1*; Griffiths et al., 2003; Shavorskaya and Lagercrantz, 2006).

CO-like genes are characterized by the presence of two highly conserved regions; a B-box domain and a CCT box. To examine the relationship between genes, the conserved regions

were used to determine genetic distances and construct phylogenetic trees. Using the CCT domain alone the *CO* and *CO-like* genes grouped into four principal clusters (Griffiths et al., 2003; Yan et al., 2004). Group I comprised the most CO-like genes with two functional B-box domains. Group II comprised the *Arabidopsis* and rice genes with a single B-box, while group III comprised genes with two zinc finger domains, with the second one diverged from the CO-type B-box. Barley *HvCO9* had the most diverged CCT domain of the barley genes and was placed in the group IV with two related rice genes *OsI*, and *OsH* whose predicted peptides had no B-box. This group had no counterpart in *Arabidopsis*. Yan et al. (2004) included in this phylogenetic three additional group consisting of ZCCT genes implicated in the vernalisation response which contain only the conserved CCT region. The ryegrass *LpCOL1* gene contains two B-box domains which fit the consensus structure meaning that both of them have the conserved residues that would make them fully functional.

Rapid gene discovery brought about knowledge regarding the different species and a pattern started to emerge that it is not necessary to have both B-boxes in order to produce a functional protein. The HvCO1 peptide lacks three highly conserved residues that would be predicted to abolish the B-box 2 function (Griffiths et al., 2003) while in ryegrass the putative *AtCO* orthologue *LpCO* lacks the consensus B-box motif in B-box2 (Martin et al., 2004; Armstead et al., 2005). Even *HvCOL2*, which mapped to the homologous *LpCOL1* chromosome and with protein sequence highly similar to LpCOL1, contains two non conserved amino acids within the B-box2 region. Although the relevance of the B-box structure is still not fully understood it is possible that B-box2, although less important in grasses and cereals, is nevertheless present in some of its functional genes as part of a different evolutionary path.

Additional analysis using the TRANSFAC database revealed the presence of Dof2 (DNA binding with One Finger); a highly conserved nucleotide sequence involved in protein binding within the intron region of the *LpCOL1* gene. This site contained the sequence AAAAAGGAG, which is a core sequence of the many plant promoters (Yanigisawa and Sheen, 1998). Dof 1 and Dof 2 have been shown to act as a transcriptional activator and repressor respectively, binding selectively to their targets and in a sequence specific manner also thought to be regulated by light. These findings fit within the current concept of the *CO* regulatory pathways. Recent findings implicated *FKF1* in the process of day-length discrimination and control of *CO* expression (Imaizumi et al., 2003) through the cyclic degradation of *DOF FACTOR 1* (*CDF1*; Imaizumi et al., 2005) which is thought to be a transcriptional repressor of *CO*. The mechanism of this reaction is still not fully understood

but the results indicate that CDF1 directly binds to the Dof binding sites in the *CO* promoter region to regulate CO transcription. Although promoter sequences were not found within the intron region of *LpCOL1* where the Dof binding site is Dof1 could still regulate transcription by being part of the enhancer element which activate or repress nearest promoter which is a feature commonly seen in enhancers. This has been previously described for example in the rice *RFL* gene, an *AtLFY* orthologue, where intron 2 functions as a transcription enhancer in concert with far-upstream sequences (Prasad et al., 2003)

4.1.3 *LpFT3* shows high similarity to the wheat and barley orthologues

Sequencing analysis showed that *LpFT3* has three exons encoding a 177 aa protein. This gene structure and protein length is virtually identical to its orthologues in wheat and barley, TaFT and HvFT, respectively. In contrast, all other FT and FT-like genes described so far have four exons. This difference arose from the fusion of exons 3 and 4 into one single exon. Exons 1 and 2 are of almost identical size in *Hd3a*, *AtFT*, and *LpFT3* with introns of the variable size between them. Exons 3 and 4 in *Arabidopsis* and rice (41/224 bp and 41/227 bp respectively) are joined and represented in the *LpFT3* gene by only one exon of 271 bp.

Two main models for intron loss have been proposed (reviewed by Roy and Gilbert, 2006). Reverse transcription of a spliced mRNA transcript deleting one or more introns seems to be more plausible than the genomic deletion mechanism. Genomic deletion often delete adjacent coding sequence or leave segments of the introns sequence which is not the case here since the bordering exons are of almost identical size. Loss of *FT* intron was found in several barley cultivars, wild barley accessions and wheat (Faure et al., 2007) suggesting it may be a general feature of temperate grasses.

The overall genetic structure of *LpFT3* is reflected at the protein level when compared with other FT proteins. The ryegrass LpFT3 shows 97% identity with both TaFT, and HvFT but when compared with rice Hd3a, and *Arabidopsis* AtFT, LpFT3 showed identity levels of 89%, and 70%, respectively. Analysis of the protein sequence showed that LpFT3 is similar to that of Raf Kinase Inhibitor Proteins (RKIP) and its bacterial/archeal homologue PhosphatidylEthanolamine-Binding Proteins (PEBP), which belong to a highly conserved family of phospholipid-binding proteins represented in all three major phylogenetic divisions (eukaryotes, bacteria, archaea) with no significant sequence homology to other proteins. In angiosperms, members of the PEBP gene family have been found to play important roles in the control of flowering (reviewed by Hecht et al., 2005; Faure et al., 2007).

4.2 Circadian clock influences expression of the photoperiod genes

The circadian clock has been previously implicated in photoperiodic time measurement (reviewed by Hotta et al., 2007) and genetic experiments in *Arabidopsis* identified mutants that disrupt both circadian clock-controlled responses and photoperiodic response. Fully vernalised ryegrass plants exposed to LD conditions for 5 days and transferred into continuous light (LL) conditions exhibited the *LpGI* cycling pattern of expression for the next 4 days (Fig 3.7) regardless of the absence of day/night shift. This evidence strongly supports the previous results from *Arabidopsis*, wheat, and rice (Fowler et al., 1999; Hayama et al., 2002; Zhao et al., 2005), indicating circadian clock control over the *GI* gene. The cycling amplitude was highest during the first day but slowly subsided over the subsequent three days. Stable circadian rhythms were also observed in *OsGI* mRNA levels under both constant light and constant dark with similar expression patterns; high in the first day and slowly subsiding afterwards (Hayama et al., 2002). Inability to cycle indefinitely with the same amplitude indicate that the circadian clock is just one of the factors regulating *GI* expression. This is in agreement with previous experiments regarding the separable roles of the AtGI protein which found that AtGI was implicated in PHYB signaling during seedling deetiolation in red light (Huq et al., 2000). In addition to this it was also found that PHYB does not reduce *AtCO* transcription (Cerdan and Chory, 2003).

A slightly different expression pattern was seen in the *LpCOL1* gene, although our results demonstrate that one of the *LpCOL1* regulators is the circadian clock. After LD entrainment the *LpCOL1* transcript continued to cycle even after the plants were transferred into LL conditions (Fig 3.22), a clear sign of the circadian clock involvement in the gene regulation. Cycling was not so robust after the first day but the cycling trend was clearly visible. This pattern of expression was previously described in the cycling of the *Arabidopsis AtCOL9* transcript (Cheng and Wang, 2005) where the amplitude of *AtCOL9* was much lower after the first day under continuous light.

It has been thought that circadian clock controls the timing of products of *CO* expression through the *CCA1* and *LHY* genes by controlling peak of *GI* expression (Mizoguchi et al., 2005). Recent experiments implicated *PSEUDO-RESPONSE REGULATOR* (*PRR*) gene family in circadian-clock associated functions (Nakamichi et al., 2007). It has been found that *PRR9*, *PRR7*, and *PRR5* redundantly activate *CO* expression in the daytime under LD acting

antagonistically to *CCA1/LHY*. This mechanism of control does not involve *GI* as an intermediary but rather *CDF1* gene, a DNA-binding transcriptional repressor (Imazumi et al., 2005). The relationship between these genes is yet to be investigated in ryegrass.

Both the *Arabidopsis AtFT* and rice *Hd3a* genes are regulated by the circadian clock (Kojima et al., 2002). Similarly, the *LpFT3* gene transcript showed robust cycling over 4 days when placed in LL conditions (Fig 3.32). Towards the end of the experiment the cycling started to loose periodicity but the amplitude of expression stayed the same.

4.3 Exogenous factors control expression of the photoperiod genes

To respond to environmental changes, plants must be able to detect light duration and change in temperature and link this to the cycling of the circadian clock. Balance between clock components may play an important role in the process of temperature compensation. Recently Gould et al. (2006) showed that *GI*, *LHY*, *TOC1*, and *CCA1* modulate their expression in response to different temperatures in order to maintain the rhythmicity of the circadian clock. At higher temperatures levels of *LHY* mRNA transcript decrease while the levels of *GI* and *TOC1* increase whereas at low temperature *CCA1* expression becomes crucial for the process of temperature compensation. Expression analysis is crucial in establishing the mechanism of this response. It has been known that change in the gene expression pattern may in turn bring about change in meristem identity and subsequent commitment to flowering. The mRNA abundance of *LpGI* over 24 h under SD and LD conditions was exeamined. The results of the experiment showed different expression patterns in SD conditions where *LpGI* transcript peaked ~ 7 h after dawn with the trough levels ~ 5 h later, compared with LD where it peaked after ~ 13 h with trough levels ~ 6 h later (Fig 3.9). This is very similar to the expression pattern of *GI* in *Arabidopsis* where transcript levels also oscillate in a circadian fashion that differs in LD and SD conditions (Fowler et al., 1999). In SD the peak came earlier and was of shorter duration reaching trough levels after transition to darkness as opposed to the LD conditions where the peak occurred 2 h later with trough levels before the onset of night. Rice, wheat, and barley expression patterns were very similar to the *Arabidopsis* ones (Hayama et al., 2002; Zhao et al., 2005). These results indicate existence of similar mechanisms for the photoperiod regulation in all previously mentioned plants.

It is clear that the timing of *GI* expression is of crucial importance. An expression peak during the day acts as a positive flowering signal whereas if it happens just before night or during the

night such signal is abolished. This hypothesis has been substantiated in *Arabidopsis* where it was found that the abundance of the GI protein is directly responsive to light/dark transitions (David et al., 2006), with the mechanism involving GI proteolysis by the 26S proteasome. *Arabidopsis* transcriptome profiling during cold acclimation revealed that transcripts for *GI* increased 5 to 10-fold in response to low temperature (Fowler and Thomashow, 2002). In addition an increased sensitivity of *gi-3* plants to freezing stress was observed (Cao et al., 2005) suggesting that the *GI* gene positively regulates freezing tolerance via a CBF-independent pathway. On the basis of these results and considering that ryegrass is a vernalisation sensitive plant which flowers much faster after being exposed to cold, it was necessary to see if this vernalisation requirement would be reflected on the *LpGI* transcription levels. Experiments showed that the expression levels fluctuated but not in the manner that could explain vernalisation requirements (Fig 3.10). There was no steady change of *LpGI* levels which would be expected if vernalisation affected expression. It is highly likely that the converging point of the vernalisation and the photoperiod pathways is somewhere downstream from the *GI* gene rendering it non responsive to the periods of low temperature.

Under LD conditions *LpCOL1* peaks just before the onset of darkness while in SD conditions the peak happens during the night period (Fig 3.24). This pattern of diurnal oscillation is similar to the *AtCO* and *OsHd1* genes, under both LD and SD conditions which supports the coincidence model suggesting that the perception of photoperiod and activation of the floral genes occur when the peak of expression coincides with the inductive period of the day (Samach et al., 2000). Rice *Hd1* is thought to regulate flowering by both repression and activation of the downstream genes in LD and SD respectively (Yano et al., 2000) contrary to *AtCO* which acts as an activator in LD but has no effect in SD. In LD plants the activation happens during LD conditions when *CO* expression is highest before night. It has been shown in *Arabidopsis* that light stabilizes the nuclear CO protein before sunset whereas during the night the protein is degraded by the proteasome (Valverde et al., 2004). The circadian expression pattern of *LpCOL1* was also similar to the expression of the previously published *LpCO* (Armstead et al., 2005). Both genes peaked before the sunset with *LpCOL1* peaking slightly earlier under LD conditions, while under SD conditions both reached maximum expression during the night period (Fig. 5.6). The expression of *LpCOL1* in LD overlaps with the light period for more than 4 h as opposed to SD conditions where it is entirely expressed during the night. This makes LpCOL1 a likely subject for post-transcriptional degradation as observed for the *Arabidopsis* AtCO.

In *Arabidopsis*, under SD conditions *FT* is barely detectable whereas under LD conditions *FT* mRNA accumulates during the light period of the day with a peak at dusk (Suarez-Lopez et al., 2001). A similar pattern of expression can be seen in rice with the inductive conditions during the SD (Kojima et al., 2002). In ryegrass, expression of *LpFT3* between inductive and non-inductive conditions differ in both expression patterns and expression levels (Fig 3.33). Under SD conditions the transcript levels reach their peak ~ 7 h after dawn, just before night with the trough levels ~ 4 h later, while in LD conditions the peak is reached ~ 8 h after dawn with the through levels ~ 8 h later; still during the daytime. Transcript abundance at the peak of expression in the afternoon is ~ 30 fold higher during the inductive conditions. Although there is still no evidence all this might suggest that overall light/dark transitions regulate *LpFT3* in a manner similar to that already described for *AtGI* (David et al., 2005), and *AtCO* (Valverde et al., 2004), involving ubiquitination and degradation of the protein in darkness by the 26S proteasome.

It has been shown that the rice *Hd3a* abundance in SD starts to rise immediately after sowing and reaches its peak 38 days after, contrary to the non-inductive LD conditions where it stays flat most of the time (Kojima et al., 2002). In ryegrass *LpFT3* expression stays low for the first 5 days after transfer to LD conditions after which it starts to rise reaching the plateau ~ 15 days later. These results do not agree with the previously published data on *Lolium temulantum LtFT* expression which shows an increase of >80 fold upon exposure to a single long day (King et al., 2006). Explanation for this discrepancy could be assigned to different cultivars/strains used. It was shown that critical number of LDs needed for secondary induction in ryegrass can vary between 0 and 16 in some Mediterranean and Scandinavian varieties respectively (Aamlid et al., 2000). Observation that vernalised ryegrass plants (cv. Impact) grown under LD conditions for more than 5 days flower normally even if transferred into the SD conditions afterwards (data not shown) just substantiate the *LpFT3* expression data.

Flowering in winter annual species is controlled by both vernalisation and photoperiod. The results of this study have shown that *LpFT3* expression levels rise with the length of the vernalisation process in plants grown under LDs (Fig. 3.34), indicating possibility that *LpFT3* may act as a floral integrator. The function of *FT* in the regulation of flowering time is most closely associated with the promotion of flowering in response to inductive photoperiods. Nevertheless, integration of flowering signals from the photoperiod and vernalisation pathway occurs at least in part through the regulation of FT (Boss et al., 2004; Michaels et al., 2004).

This results indicates the possibility that *FT* may be used as a flowering marker since a rise in its expression generally indicates the beginning of the flowering process.

4.4 Mapping confirmed orthologues nature of the photoperiod genes

A separate group within AgResearch has developed a genetic linkage map for perennial ryegrass using 165 EST-SSR markers developed from a pair cross between individual heterozygous genotypes from cv. 'Grasslands Impact' and cv. 'Grasslands Samson'. This enabled the use of SNPs and dCAPS among the parents to distinguish between the different progeny genotypes and allowed the genetic mapping of the three photoperiod genes. Using SNPs *LpGI* was mapped on the short arm of ryegrass chromosome 3H (Fig 4.1) at the 20 cM position. Analysis of the rice genomic sequence using the www.gramene.com database revealed a single *GI* homologue found at the short arm of chromosome 1 (1S; AP003047). Previous studies showed that the ryegrass short arm on chromosome 3 is colinear with the short arm of rice chromosome 1 (Jones et al., 2002; Faville, 2005). Furthermore, the rice 1S is colinear with the short arms of wheat and barley group 3 chromosomes (Moore et al., 1995) where only a single copy of the *GI* gene was detected (Dunford et al., 2005), and chromosomes 3 and 8 of maize (Van Deynze et al., 1995) where southern hybridisation indicated the presence of two *GI* sequences (Dunford et al., 2005). These results consistently show that the *LpGI* gene positioned on ryegrass chromosome 3 is a true orthologue of the *Arabidopsis GI* gene. Orthology has been described as homology subtype with "a relationship where sequence divergence follows speciation and where the common ancestor of the two genes lies in the most recent common ancestor of the taxa from which the two sequences were obtained" (Fitch, 2000). Bennetzen at al., (2004) have found that comparative sequence analysis of homologous genomic regions, conserved at the same locations is very reliable indicator of orthologous nature of two genes.

Both *CO* and *CO*-like genes have been identified as being of high importance in the floral induction pathways of both dicot and monocot species. Genetic mapping using a SNP based approach coupled with the dCAPS method placed the ryegrass *LpCOL1* gene on to LG6. Comparative analysis with other *Poaceae* genomes showed that LG6 in ryegrass, although containing a larger non-syntenic region, still shows partial synteny with the rice chromosome 2 region (Fig. 4.1; Jones et al., 2002; Faville, 2005), which has been shown to contain heading date QTL region *Hd7* (Yano et al., 2001). In rice the major heading date QTL named *Hd1* has

been placed on chromosome 6 (Yano et al., 2000). The ryegrass *LpCO* gene, putative *AtCO* orthologue, was found to be positioned on ryegrass LG7 (Armstead et al., 2005). Comparative mapping has shown that the region of ryegrass chromosome to which *LpCO* maps has a degree of conserved synteny with the rice chromosome 6 region associated with *OsHd1* (Armstead et al., 2005).

A study on the wheat *TaHd1* gene suggested its orthologues function similar to the rice *Hd1* on the basis of sequence similarity and functional complementation of an *Hd1*-deficient rice line, despite the fact it was assigned to the wheat LG6 (Nemoto et al., 2003). The authors tried to explain the inconsistency between the expected position on homologous group and the observed position on group 6 by possible chromosomal rearrangements. Wheat *TaHd1* and ryegrass *LpCOL1* share high sequence similarity between them (78%) and are highly similar to the barley *HvCO2* (92% and 77% respectively) with all three positioned on homologous group 6.

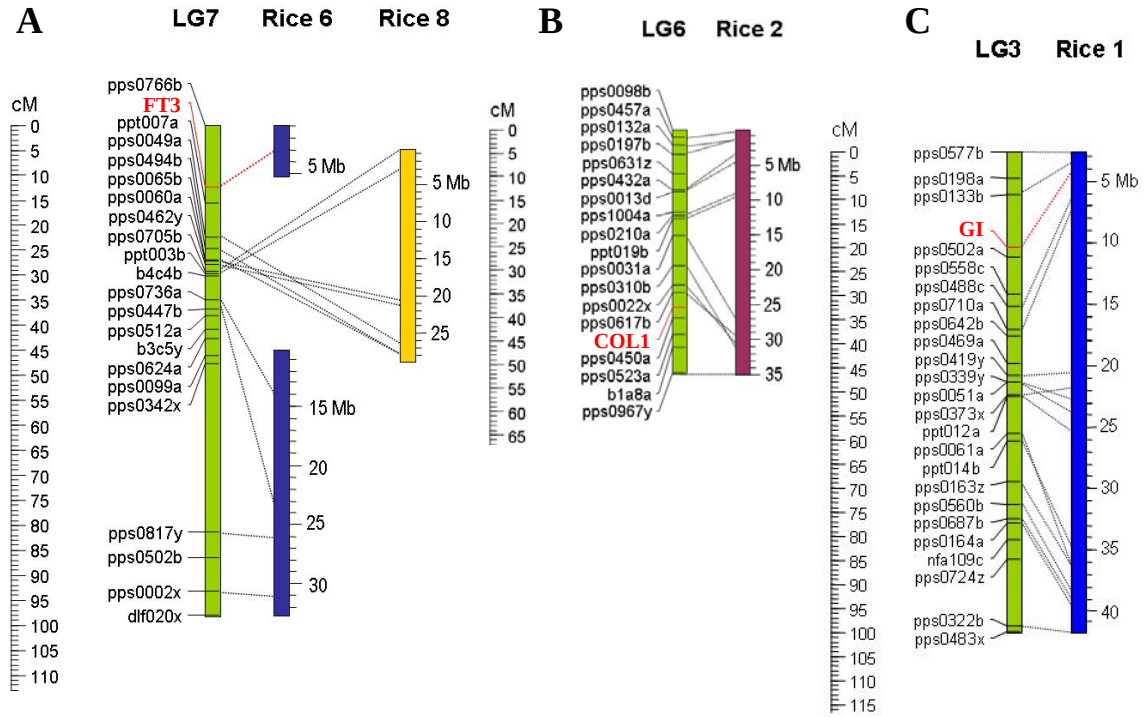


Fig. 4.1 Genetic map of *Lolium perenne* chromosomes 7, 6, and 3. Three ryegrass photoperiod genes *LpGI* (**A**), *LpCOL1* (**B**), and *LpFT3* (**C**), were mapped to the ryegrass LG 7, 6, and 3 with orthologous regions of rice plotted alongside (comparative maps from Faville, 2005).

These findings suggest gene duplication occurred followed by a rearrangement event in the common ancestor that led to the present existence of a homologous set of genes on chromosomes 6 and 7.

Using a SNP based approach, the *LpFT3* gene was mapped to the short arm of chromosome 7, colinear to homologous region 6 in rice (Fig. 4.1). A previous study found the ryegrass LG7 region containing *LpFT3* to be colinear with the short arm of chromosome 6 in rice and the short arm of chromosome 7 in wheat (Jones et al., 2002; Faville, 2005). The LpFT3 protein is also highly similar to the wheat, barley, and rice FT orthologues (97%, 97%, and 89% respectively). In addition to this the rice *Hd3a* was mapped to the region within the short arm of chromosome 6, syntenic to homologous groups 7 in barley and wheat on which their respective *FT* orthologues were mapped (Faure et al., 2007; Yan et al., 2006). The short arm of chromosome 6 in rice is where the heading date QTL *Hd3* was mapped to (Yamamoto et al., 1998). Additional studies showed that this region contains two tightly linked genes with related functions namely *Hd3a*, and *Hd3b* (Monna et al., 2002) with *Hd3a* promoting flowering in inductive conditions (Kojima et al., 2002). This tandem duplication event probably happened after divergence from *Triticae* since Southern blot analysis using *HvFT* as a probe in barley resulted in a single hybridization band with *genomic* DNA (Yan et al., 2006; Faure et al., 2007).

4.5 Eudicot and monocot plants show functional conservation of the photoperiod pathway

4.5.1 *LpGI* functionally complements *Arabidopsis gi-3* mutants

The *GI* gene has been implicated in photoperiodic flowering because the flowering time of *gi* mutants is severely delayed in long day conditions (Fowler et al., 1999; Park et al., 1999; Hayama et al., 2003). Moreover expression of *AtGI* from the cauliflower mosaic virus 35S promoter causes early flowering in *Arabidopsis* (Mizoguchi et al., 2005). This system has worked even in heterologous species such as wheat where *TaGI* was shown to be functionally active when overexpressed in *Arabidopsis* (Zhao et al., 2005). Similarly expression of an antisense *AtGI* in radish caused significant delay in both bolting and flowering (Curtis et al., 2002). Ryegrass shows photoperiod adaptations similar to *Arabidopsis*, flowering under a long day photoperiod. The work outlined in this thesis provides evidence of *LpGI* involvement in photoperiodic flowering based on complementation analysis of *LpGI* in *Arabidopsis* plants (Fig 3.14). The *LpGI* gene does not just rescue the *gi-3* mutant but causes plants to flower significantly earlier than the wild type. This is further proof that genes within the photoperiod pathway are conserved not just phylogenetically but functionally as well. T2

transgenic plants, a mixture of hemizygous and homozygous plants, flowered slightly later than the T3 (homozygous) plants. This may indicate the existence of a dosage effect previously described for the *Arabidopsis CONSTANS* gene (Putterill et al., 1995). This work suggested that the heterozygote does not produce enough protein to produce extremely early flowering, as opposed to the homozygous plants which have two copies of the same gene. In *Arabidopsis*, *35S::GI* plants show a severe early flowering phenotype and enhanced expression of *CO* and *FT* (Mizoguchi et al., 2005). *Arabidopsis 35S::LpGI gi-3* plants show an increase in abundance of endogenous *AtFT* when compared with the *gi-3* mutant plants (Table 3.9), indicating a functional conservation of the photoperiod pathway. It is interesting that the *35S::LpGI gi-3* plants flowered earlier than the wild type plants although the expression levels of *AtFT* were comparable with the wild type L*er* plants. This may indicate existence of downstream targets other than that of *CO* and *FT*.

4.5.2 *LpCOL1* accelerates flowering through an as yet unknown mechanism

To assess the regulatory role in floral induction, *LpCOL1* was over-expressed under the control of the constitutive CaMV 35S promoter in the wild type Col, and *co-2* mutant plants. It was found that *LpCOL1* did not promote flowering in wild type plants nor did it rescue the late flowering phenotype of *co-2* mutants (Table 3.4). This is not unusual since *Arabidopsis COL* genes are not just promoters of flowering; some of them have been shown to have little effect on flowering time (Ledger et al., 2001) and even acted as repressors of flowering (Cheng and Wang, 2005). Interestingly when the functionality of *LpCOL1* in *gi-3* mutants was investigated, it was found that *LpCOL1* not just completely rescued *gi-3*, but also accelerated flowering even further (Table 3.5). This mode of action has not been described before within *CO-like* genes. *Arabidopsis* served as a transformation vehicle for several *CO-like* genes from dicotyledonous species such as *Brassica napus* (*BnCOa1*), and *Pharbitis nil* (*PnCO*; Robert et al., 1998; Liu et al., 2001). Only recently a *CO* monocot homolog from ryegrass *LpCO*, has been transformed into the *Arabidopsis co-2* mutant (Martin et al., 2004). *LpCO* contains two B-box elements: conserved B-box1 and the less conserved B-box2 in contrast to *LpCOL1* which contains two fully conserved B-box domains. However only *LpCO* rescued *co-2* mutant.

In order to establish *LpCOL1* mode of action the expression of the *AtFT* in *35S::LpCOL1* transgenic plants was followed and no sign of change was found compared with the wild type

plants, consistent with the flowering times between these two. It has been suggested that B-boxes may function to orientate coiled-coil domains that are the site of interaction with other proteins (Torok and Etkin, 2001). In the case of *LpCOL1* this interaction may be missing because of inadequate binding sites within the B-boxes. The mode of action and position within the photoperiodic hierarchy of *LpCOL1* is hard to determine given the present level of knowledge.

Based on the fact that it can rescue *Arabidopsis gi-3* mutants, it is likely that *LpCOL1* acts downstream of the *GI* gene. In *gi-3* mutants *AtFT* abundance was very low and without visible cycling patterns. By measuring *AtFT* expression in mutated *gi-3 Arabidopsis* plants transformed with *35S::LpCOL1* an increase of the *AtFT* transcript was noticed to the levels comparable with the wild type lines. This was a clear sign that the induction of flowering by *LpCOL1* involves the *FT* gene which in turn activates downstream meristem identity genes (Fig 4.2). Although only 4 lines were obtained in *35S::LpCOL1 co-2* experiment which raise the possibility that some of the rescued lines may have been missed, this was offset by 16 rescued lines in *35S::LpCOL1* Col experiment that did not show any change in flowering time. Further gene expression experiments are also needed to confirm that the *co-2* mutants express the transgene

It is still not known whether *LpCOL1* interacts with other genes in order to activate expression of *FT* and what these genes are. Inability of *LpCOL1* to promote flowering in *co-2* and wild type plants is still very puzzling. At present transformation of ryegrass is possible but still laborious and with mixed results. Overexpression and silencing of *LpCOL1* in ryegrass would give us an insight into its real function and position within the pathway. Expression analysis of other, previously described photoperiod genes, in *LpCOL1* transgenic plants would enable us to establish relationship between them. Given the present level of knowledge it would be interesting to identify *CDF1* transcription factor in ryegrass alongside genes from the *PRR* gene family and their potential effect on *LpCOL1*. Further experiments on *LpCOL1* investigating expression of related genes under variety of environmental conditions are necessary to elucidate its mechanism of action and position within the photoperiod pathway.

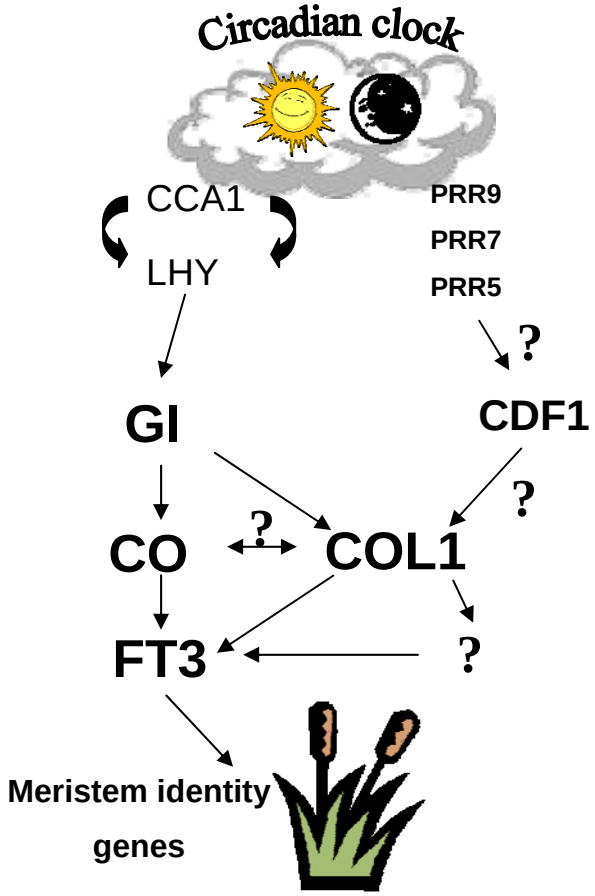


Fig. 4.2 Model of the photoperiod pathway in *Lolium perenne*
Genetic structure, functional analysis, and gene mapping confirmed *LpGI* and *LpFT3* as orthologues to *AtGI*, and *AtFT* respectively. The function and position of *LpCOL1* within the photoperiod pathway is still unclear.

4.5.2 The *LpFT3* gene rescues *ft-1* mutants and accelerates flowering in wild type Col plants

Orthologues of the *FT* gene accelerate flowering in the LD plant *Arabidopsis* and SD plant rice (Kardailsky et al., 1999; Kojima et al., 2002). To test whether overexpression of *LpFT3* promotes flowering the overexpression construct was introduced into *Arabidopsis* and it was found that *LpFT3* corrected the late flowering phenotype of the *ft-1* mutant and accelerated flowering in wt Col plants (Table 3.7, 3.8). Acceleration in the wt plants was not so profound as seen in *Arabidopsis* when *AtFT* was overexpressed (Kardailsky et al., 1999; Kobayashi et al., 1999). This could be attributed the non-homologous nature of the overexpressed sequences. Overexpression of the same construct in the *ft-1 Arabidopsis* mutant plants did not only fully rescue the mutants but accelerated flowering even further. This supports the existence of the common genetic elements within the *FT* locus in both monocots and dicots required for promotion of flowering.

The *FT* gene is a direct regulatory target of *CO* (Samach et al., 2000), and through *CO* it is also regulated by *GI*, another upstream gene within the photoperiod chain (Suarez-Lopez et al., 2001). It was found that the expression of the endogenous *AtFT* gene, in both *ft-1* and wild type Columbia plants, was upregulated upon transformation with *35S::LpFT3* overexpression construct. In *ft-1* mutants *AtFT* showed a diurnal cycling pattern producing a non-functional FT protein. In a single analysed line COF138-1 overexpression of *LpFT3* in an *ft-1* background caused *AtFT* transcript to increases 4 fold in the afternoon while in the morning it stayed at the same levels (Table 3.6). This may indicate a positive feedback loop which acts through the *AtFT* gene itself or through some of the upstream intermediaries. A similar effect was seen in one of the three wt Col lines transformed with *35S::LpFT3*. The existence of the positive feedback loop in *AtFT* has been implied before (Huang et al., 2005; Bohlenius et al., 2007) but additional experiments are needed to determine major genes within the loop.

4.6 Conclusion

Flowering time is important when adapting crop or forage plants to different environments such as location at different latitudes and cropping seasons. Ryegrass generally require a prolonged period of low temperatures (vernalisation), followed by an increase in photoperiod (long days) to induce flowering. This dual requirement ensures that flowering and seed production is aligned to the favorable conditions of spring and summer. Although high feeding quality of forage grasses is improved by repression of flowering, flowering is necessary to induce grass seed production which makes these two highly conflicting goals for grass breeders. Therefore, the identification and characterization of the genes controlling flowering time in forage grasses, including perennial ryegrass is of great interest.

There is a high degree of conservation between genes crucial for the transition to flowering between *Arabidopsis* and monocot species, although their modes of action may be different. In monocot species, other than rice, the data concerning the genes responsible for controlling heading date are not so abundant and sufficient to produce a complete picture although there are some interesting results coming from research done on wheat and barley.

Here three genes thought to be extremely important for flowering initiation in ryegrass have been described. It has been shown that they bear extensive genetic and functional similarity to the orthologous *Arabidopsis* genes, although the similarity was much higher with the orthologous genes in monocot species such as rice, wheat, and barley. The novel *CO*-like gene *LpCOL1* was shown to be slightly different to the *Arabidopsis* and rice homologues with some features not previously described in any other *CO* and *CO*-like genes.

The ability to associate sequences from different species on the basis of comparative genomics and sequence homology, as shown here, would make interpreting the observations made in other plant species easier, and hopefully bring a new level to our understanding of universality and flexibility of the genetics involved in flowering control in plants in general. This research could also provide insight into the number of alleles of each gene implicated in the photoperiod pathway. Combined with expression date and functional analysis, this would assist plant breeders in selecting varieties with adaptations to existing environmental conditions or for new ones arising from continual climate change.

The work described in this thesis could be seen as the first step towards QTL mapping and identifying genetic markers which in return offer the opportunity to increase the precision of selection and therefore the rate of improvement in breeding programmes.

REFERENCES

Aamlid TS, Heide OM, Boelt B: Primary and Secondary Induction Requirements for Flowering of Contrasting European Varieties of *Lolium perenne*. Ann Bota 86: 1087-1095 (2000).

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T: FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science 309: 1052-6 (2005).

Ahmad M, Cashmore AR: *HY4* gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. Nature 366: 162-6 (1993).

Alabadi D, Oyama T, Yanovsky MJ, Harmon FG, Mas P, Kay SA: Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. Science 293: 880-3 (2001).

Alm V, Fang C, Busso CS, Devos KM, Vollan K, Grieg Z, Rognli OA: A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species. Theor Appl Genet 108: 25-40 (2003).

Amasino RM: Vernalization and flowering time. Curr Opin Biotechnol 16: 154-8 (2005).

An H, Roussot C, Suarez-Lopez P, Corbesier L, Vincent C, Pineiro M, Hepworth S, Mouradov A, Justin S, Turnbull C, Coupland G: *CONSTANS* acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of *Arabidopsis*. Development 131: 3615-26 (2004).

Andersen JR, Jensen LB, Asp T, Lubberstedt T: Vernalization Response in Perennial Ryegrass (Lolium perenne L.) Involves Orthologues of Diploid Wheat (Triticum monococcum) *VRN1* and Rice (Oryza sativa) *Hd1*. Plant Mol Biol 60: 481-94 (2006).

Armstead IP, Skot L, Turner LB, Skot K, Donnison IS, Humphreys MO, King IP: Identification of perennial ryegrass (*Lolium perenne* (L.)) and meadow fescue (*Festuca pratensis* (Huds.)) candidate orthologous sequences to the rice *Hd1*(*Se1*) and barley *HvCO1 CONSTANS*-like genes through comparative mapping and microsynteny. New Phytol 167: 239-47 (2005).

Armstead IP, Turner LB, Farrell M, Skot L, Gomez P, Montoya T, Donnison IS, King IP, Humphreys MO: Synteny between a major heading-date QTL in perennial ryegrass (*Lolium perenne* L.) and the *Hd3* heading-date locus in rice. Theor Appl Genet 108: 822-8 (2004).

Aukerman MJ, Hirschfeld M, Wester L, Weaver M, Clack T, Amasino RM, Sharrock RA: A deletion in the *PHYD* gene of the *Arabidopsis* Wassilewskija ecotype defines a role for phytochrome D in red/far-red light sensing. Plant Cell 9: 1317-26 (1997).

Bagnall DJ, King RW, Whitelam GC, Boylan MT, Wagner D, Quail PH: Flowering responses to altered expression of phytochrome in mutants and transgenic lines of *Arabidopsis thaliana* (L.) Heynh. Plant Physiol 108: 1495-503 (1995).

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J, Weigel D: The *PHYTOCHROME C* photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*. Nat Genet 38: 711-5 (2006).

Barnes RF: Importance and problems of tall fescue. Boca Raton, FL: CRC Press (1990).

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C: Vernalization requires epigenetic silencing of *FLC* by histone methylation. Nature 427: 164-7 (2004).

Bennetzen JL, Coleman C, Liu R, Ma J, Ramakrishna W: Consistent over-estimation of gene number in complex plant genomes. Curr Opin Plant Biol 7: 732-6 (2004).

Bernier G: The Control of Floral Evocation and Morphogenesis. Ann Rev Plant Phy Plant Mol Biol 39: 175-219 (1988).

Blazquez MA: Plant science. The right time and place for making flowers. Science 309: 1024-5 (2005).

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

Blazquez MA, Weigel D: Independent regulation of flowering by phytochrome B and gibberellins in *Arabidopsis*. Plant Physiol 120: 1025-32 (1999).

Blazquez MA, Weigel D: Integration of floral inductive signals in *Arabidopsis*. Nature 404: 889-92 (2000).

Bohlenius H, Eriksson S, Parcy F, Nilsson O: Retraction. Science 316: 367 (2007).

Borden KL: RING fingers and B-boxes: zinc-binding protein-protein interaction domains. Biochem Cell Biol 76: 351-8 (1998).

Boss PK, Bastow RM, Mylne JS, Dean C: Multiple pathways in the decision to flower: enabling, promoting, and resetting. Plant Cell 16 Suppl: S18-31 (2004).

Buckler ESt, Thornsberry JM, Kresovich S: Molecular diversity, structure and domestication of grasses. Genet Res 77: 213-8 (2001).

Bünning E: Die endogene Tagesrhytmik als Grundlage der photo-periodischen Reaktion. Ber Dtch Bot Ges 54: 590-607 (1936).

Cao S, Ye M, Jiang S: Involvement of *GIGANTEA* gene in the regulation of the cold stress response in Arabidopsis. Plant Cell Rep 24: 683-90 (2005).

Cerdan PD, Chory J: Regulation of flowering time by light quality. Nature 423: 881-5 (2003).

Chandler, J. and Dean, C: Factors influencing the vernalization response and flowering time of late flowering mutants of *Arabidopsis thaliana*. J Exp Bot 45: 1279-1288 (1994).

Cheng XF, Wang ZY: Overexpression of *COL9*, a CONSTANS-LIKE gene, delays flowering by reducing expression of *CO* and *FT* in *Arabidopsis thaliana*. Plant J 43: 758-68 (2005).

Christie JM: Phototropin blue-light receptors. Annu Rev Plant Biol 58: 21-45 (2007).

Ciannamea S, Busscher-Lange J, de Folter S, Angenent GC, Immink RG: Characterization of the vernalization response in Lolium perenne by a cDNA microarray approach. Plant Cell Physiol 47: 481-92 (2006).

Clough SJ, Bent AF: Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735-43 (1998).

Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ: Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. J Exp Bot (2007).

Corbesier L, Coupland G: The quest for florigen: a review of recent progress. J Exp Bot 57: 3395-403 (2006).

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

Cornish MA, Hayward MD, Lawrence MJ: Self incompatibility in ryegrass. I. Genetic control in diploid *Lolium perenne* L. Heredity: 95-106 (1979).

Curtis IS, Nam HG, Yun JY, Seo KH: Expression of an antisense *GIGANTEA* (*GI*) gene fragment in transgenic radish causes delayed bolting and flowering. Transgenic Res 11: 249-56 (2002).

David KM, Armbruster U, Tama N, Putterill J: *Arabidopsis* GIGANTEA protein is post-transcriptionally regulated by light and dark. FEBS Lett 580: 1193-7 (2006).

Devlin PF, Robson PR, Patel SR, Goosey L, Sharrock RA, Whitelam GC: Phytochrome D acts in the shade-avoidance syndrome in *Arabidopsis* by controlling elongation growth and flowering time. Plant Physiol 119: 909-15 (1999).

Dewhurst RJ, Scollan ND, Lee MR, Ougham HJ, Humphreys MO: Forage breeding and management to increase the beneficial fatty acid content of ruminant products. Proc Nutr Soc 62: 329-36 (2003).

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A: Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT*-like gene expression independently of *Hd1*. Genes Dev 18: 926-36 (2004).

Dunford RP, Griffiths S, Christodoulou V, Laurie DA: Characterisation of a barley (*Hordeum vulgare* L.) homologue of the *Arabidopsis* flowering time regulator *GIGANTEA*. Theor Appl Genet 110: 925-31 (2005).

Dunlap JC: Molecular bases for circadian clocks. Cell 96: 271-90 (1999).

Edwards KD, Lynn JR, Gyula P, Nagy F, Millar AJ: Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. Genetics 170: 387-400 (2005).

Eimert K, Wang SM, Lue WI, Chen J: Monogenic Recessive Mutations Causing Both Late Floral Initiation and Excess Starch Accumulation in *Arabidopsis*. Plant Cell 7: 1703-1712 (1995).

Endo M, Mochizuki N, Suzuki T, Nagatani A: CRYPTOCHROME2 in Vascular Bundles Regulates Flowering in *Arabidopsis*. Plant Cell 19: 84-93 (2007).

Eriksson S, Bohlenius H, Moritz T, Nilsson O: GA4 is the active gibberellin in the regulation of *LEAFY* transcription and *Arabidopsis* floral initiation. Plant Cell 18: 2172-81 (2006).

Faure S, Higgins J, Turner A, Laurie DA: The *FLOWERING LOCUS T*-Like Gene Family in Barley (*Hordeum vulgare*). Genetics 176: 599-609 (2007).

Faville MJ, Vecchies AC, Schreiber M, Drayton MC, Hughes LJ, Jones ES, Guthridge KM, Smith KF, Sawbridge T, Spangenberg GC, Bryan GT, Forster JW: Functionally associated molecular genetic marker map construction in perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 110: 12-32 (2004).

Faville MJ: An in silico DNA sequence comparison of the perennial ryegrass and rice genomes. Proceedings of the 4th International Symposium on the Molecular Breeding of Forage and Turf 181 (2005).

Fitch WM: Homology a personal view on some of the problems. Trends Genet 16: 227-31 (2000).

Fowler S, Lee K, Onouchi H, Samach A, Richardson K, Morris B, Coupland G, Putterill J: *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. EMBO 18: 4679-88 (1999).

Fowler S, Thomashow MF: *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-90 (2002).

Fu D, Dunbar M, Dubcovsky J: Wheat *VIN3*-like PHD finger genes are up-regulated by vernalization. Mol Genet Genomics 277: 301-13 (2007).

Gaudin V, Libault M, Pouteau S, Juul T, Zhao G, Lefebvre D, Grandjean O: Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in *Arabidopsis*. Development 128: 4847-58 (2001).

Gazzani S, Gendall AR, Lister C, Dean C: Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions. Plant Physiol 132: 1107-14 (2003).

Ge Y, Cheng X, Hopkins A, Wang ZY: Generation of transgenic Lolium temulentum plants by Agrobacterium tumefaciens-mediated transformation. Plant Cell Rep 26: 783-9 (2007).

Gendall AR, Levy YY, Wilson A, Dean C: The *VERNALIZATION 2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. Cell 107: 525-35 (2001).

Gleave AP: A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. Plant Mol Biol 20: 1203-7 (1992).

Gocal GF, Poole AT, Gubler F, Watts RJ, Blundell C, King RW: Long-day up-regulation of a *GAMYB* gene during *Lolium temulentum* inflorescence formation. Plant Physiol 119: 1271-8 (1999).

Gould PD, Locke JC, Larue C, Southern MM, Davis SJ, Hanano S, Moyle R, Milich R, Putterill J, Millar AJ, Hall A: The molecular basis of temperature compensation in the *Arabidopsis* circadian clock. Plant Cell 18: 1177-87 (2006).

Griffiths S, Dunford RP, Coupland G, Laurie DA: The evolution of *CONSTANS*-like gene families in barley, rice, and *Arabidopsis*. Plant Physiol 131: 1855-67 (2003).

Gubler F, Kalla R, Roberts JK, Jacobsen JV: Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter. Plant Cell 7: 1879-91 (1995).

Guo H, Yang H, Mockler TC, Lin C: Regulation of flowering time by *Arabidopsis* photoreceptors. Science 279: 1360-3 (1998).

Halliday KJ, Whitelam GC: Changes in photoperiod or temperature alter the functional relationships between phytochromes and reveal roles for phyD and phyE. Plant Physiol 131: 1913-20 (2003).

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K: Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422: 719-22 (2003).

Hepworth SR, Valverde F, Ravenscroft D, Mouradov A, Coupland G: Antagonistic regulation of flowering-time gene *SOC1* by *CONSTANS* and *FLC* via separate promoter motifs. EMBO 21: 4327-37 (2002).

Hotta CT, Gardner MJ, Hubbard KE, Baek SJ, Dalchau N, Suhita D, Dodd AN, Webb AA: Modulation of environmental responses of plants by circadian clocks. Plant Cell Environ 30: 333-49 (2007).

Humphreys MW, Yadav RS, Cairns AJ, Turner LB, Humphreys J, Skot L: A changing climate for grassland research. New Phytol 169: 9-26 (2006).

Huq E, Tepperman JM, Quail PH: GIGANTEA is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. Proc Natl Acad Sci U S A 97: 9789-94 (2000).

Imaizumi T, Kay SA: Photoperiodic control of flowering: not only by coincidence. Trends Plant Sci 11: 550-8 (2006).

Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA: FKF1 F-box protein mediates cyclic degradation of a repressor of *CONSTANS* in *Arabidopsis*. Science 309: 293-7 (2005).

Imaizumi T, Tran HG, Swartz TE, Briggs WR, Kay SA: *FKF1* is essential for photoperiodic-specific light signalling in *Arabidopsis*. Nature 426: 302-6 (2003).

Jacobsen SE, Olszewski NE: Mutations at the *SPINDLY* locus of *Arabidopsis* alter gibberellin signal transduction. Plant Cell 5: 887-96 (1993).

Jensen LB, Andersen JR, Frei U, Xing Y, Taylor C, Holm PB, Lubberstedt T: QTL mapping of vernalization response in perennial ryegrass (*Lolium perenne* L.) reveals co-location with an orthologue of wheat *VRN1*. Theor Appl Genet 110: 527-36 (2005).

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C: Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290: 344-7 (2000).

Johnson IR, Parsons AJ: Use of a model to analyse the effects of continuous grazing managements on seasonal patterns of grass production. Grass Forage Sci 40: 449-458 (1985).

Johnson E, Bradley M, Harberd NP, Whitelam GC: Photoresponses of Light-Grown *phyA* Mutants of *Arabidopsis* (Phytochrome A Is Required for the Perception of Daylength Extensions). Plant Physiol 105: 141-149 (1994).

Jones MB, Lazenby A: The Grass Crop. Chapman and Hall Ltd (1988).

Jones S, Dupal P, Dumsday L, Hughes J, Forster W: An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 105: 577-584 (2002).

Kania T, Russenberger D, Peng S, Apel K, Melzer S: *FPF1* promotes flowering in *Arabidopsis*. Plant Cell 9: 1327-38 (1997).

Kellogg EA: Relationships of cereal crops and other grasses. Proc Natl Acad Sci U S A 95: 2005-10 (1998).

King RW, Moritz T, Evans LT, Junttila O, Herlt AJ: Long-day induction of flowering in *Lolium temulentum* involves sequential increases in specific gibberellins at the shoot apex. Plant Physiol 127: 624-32 (2001).

King RW, Moritz T, Evans LT, Martin J, Andersen CH, Blundell C, Kardailsky I, Chandler PM: Regulation of flowering in the long-day grass *Lolium temulentum* by gibberellins and the *FLOWERING LOCUS T* gene. Plant Physiol 141: 498-507 (2006).

Klar T, Pokorny R, Moldt J, Batschauer A, Essen LO: Cryptochrome 3 from Arabidopsis thaliana: structural and functional analysis of its complex with a folate light antenna. J Mol Biol 366: 954-64 (2007).

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M: Hd3a, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. Plant Cell Physiol 43: 1096-105 (2002).

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

Kotake T, Takada S, Nakahigashi K, Ohto M, Goto K: *Arabidopsis TERMINAL FLOWER 2* gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes. Plant Cell Physiol 44: 555-64 (2003).

Kouzarides T: Histone methylation in transcriptional control. Curr Opin Genet Dev 12: 198-209 (2002).

Kurepa J, Smalle J, Van Montagu M, Inze D: Polyamines and paraquat toxicity in *Arabidopsis thaliana*. Plant Cell Physiol 39: 987-92 (1998).

Laurie DA: Comparative genetics of flowering time. Plant Mol Biol 35: 167-77 (1997).

Laurie DA, Pratchett N, Bezant JH, Snape JW: RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter X spring barley (*Hordeum vulgare* L.) cross. Genome 38: 575-585 (1995).

Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C: Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. Science 297: 243-6 (2002).

Lifschitz E, Eshed Y: Universal florigenic signals triggered by *FT* homologues regulate growth and flowering cycles in perennial day-neutral tomato. J Exp Bot 57: 3405-14 (2006).

Lin C: Photoreceptors and regulation of flowering time. Plant Physiol 123: 39-50 (2000).

Lin C, Yang H, Guo H, Mockler T, Chen J, Cashmore AR: Enhancement of blue-light sensitivity of *Arabidopsis* seedlings by a blue light receptor cryptochrome 2. Proc Natl Acad Sci U S A 95: 2686-90 (1998).

Martin J, Storgaard M, Andersen CH, Nielsen KK: Photoperiodic regulation of flowering in perennial ryegrass involving a *CONSTANS*-like homolog. Plant Mol Biol 56: 159-169 (2004).

Martin-Tryon EL, Kreps JA, Harmer SL: GIGANTEA acts in blue light signaling and has biochemically separable roles in circadian clock and flowering time regulation. Plant Physiol 143: 473-86 (2007).

Mas P, Devlin PF, Panda S, Kay SA: Functional interaction of phytochrome B and cryptochrome 2. Nature 408: 207-11 (2000).

Michaels SD, Amasino RM: *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11: 949-56 (1999).

Michaels SD, Amasino RM: Loss of *FLOWERING LOCUS C* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell 13: 935-41 (2001).

Michaels SD, Ditta G, Gustafson-Brown C, Pelaz S, Yanofsky M, Amasino RM: *AGL24* acts as a promoter of flowering in *Arabidopsis* and is positively regulated by vernalization. Plant J 33: 867-74 (2003).

Michaels SD, He Y, Scortecci KC, Amasino RM: Attenuation of *FLOWERING LOCUS C* activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. Proc Natl Acad Sci U S A 100: 10102-7 (2003a).

Mittag M, Kiaulehn S, Johnson CH: The circadian clock in *Chlamydomonas reinhardtii*. What is it for? What is it similar to? Plant Physiol 137: 399-409 (2005).

Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song HR, Carre IA, Coupland G: *LHY* and *CCA1* are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. Dev Cell 2: 629-41 (2002).

Mizoguchi T, Wright L, Fujiwara S, Cremer F, Lee K, Onouchi H, Mouradov A, Fowler S, Kamada H, Putterill J, Coupland G: Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in *Arabidopsis*. Plant Cell 17: 2255-70 (2005).

Mockler T, Yang H, Yu X, Parikh D, Cheng YC, Dolan S, Lin C: Regulation of photoperiodic flowering by *Arabidopsis* photoreceptors. Proc Natl Acad Sci U S A 100: 2140-5 (2003).

Mockler TC, Guo H, Yang H, Duong H, Lin C: Antagonistic actions of *Arabidopsis* cryptochromes and phytochrome B in the regulation of floral induction. Development 126: 2073-82 (1999).

Monte E, Alonso JM, Ecker JR, Zhang Y, Li X, Young J, Austin-Phillips S, Quail PH: Isolation and characterization of phyC mutants in *Arabidopsis* reveals complex crosstalk between phytochrome signaling pathways. Plant Cell 15: 1962-80 (2003).

Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I: The *SOC1* MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*. Plant J 35: 613-23 (2003).

Mylne JS, Barrett L, Tessadori F, Mesnage S, Johnson L, Bernatavichute YV, Jacobsen SE, Fransz P, Dean C: LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc Natl Acad Sci U S A 103: 5012-7 (2006).

Nakamichi N, Kita M, Niimura K, Ito S, Yamashino T, Mizoguchi T, Mizuno T: *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 (2007).

Neff MM, Chory J: Genetic interactions between phytochrome A, phytochrome B, and cryptochrome 1 during *Arabidopsis* development. Plant Physiol 118: 27-35 (1998).

Nemoto Y, Kisaka M, Fuse T, Yano M, Ogihara Y: Characterization and functional analysis of three wheat genes with homology to the *CONSTANS* flowering time gene in transgenic rice. Plant J 36: 82-93 (2003).

Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, Kim HJ, Kay SA, Nam HG: Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis GIGANTEA* gene. Science 285: 1579-82 (1999).

Petersen K, Didion T, Andersen CH, Nielsen KK: MADS-box genes from perennial ryegrass differentially expressed during transition from vegetative to reproductive growth. J Plant Physiol 161: 439-47 (2004).

Prasad K, Kushalappa K, Vijayraghavan U: Mechanism underlying regulated expression of *RFL*, a conserved transcription factor, in the developing rice inflorescence. Mech Dev 120: 491-502 (2003).

Putterill J, Laurie R, Macknight R: It's time to flower: the genetic control of flowering time. Bioessays 26: 363-73 (2004).

Putterill J, Robson F, Lee K, Simon R, Coupland G: The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80: 847-57 (1995).

Quail PH: Phytochrome photosensory signalling networks. Nat Rev Mol Cell Biol 3: 85-93 (2002).

Radojevic I. SRJ, St. John J.A., Humphreys M.O.: Chemical Composition and in vitro Digestibility of Lines of *Lolium perenne* Selected for High Concetrations of Water-soluble Carbohydrate. Australian Journal of Agricultural Research 45: 901-912 (1994).

Raikhel N: Nuclear Targeting in Plants. Plant Physiol 100: 1627-1632 (1992).

Ratcliffe OJ, Nadzan GC, Reuber TL, Riechmann JL: Regulation of flowering in *Arabidopsis* by an *FLC* homologue. Plant Physiol 126: 122-32 (2001).

Robert LS, Robson F, Sharpe A, Lydiate D, Coupland G: Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in Brassica napus. Plant Mol Biol 37: 763-72 (1998).

Robson F, Costa MM, Hepworth SR, Vizir I, Pineiro M, Reeves PH, Putterill J, Coupland G: Functional importance of conserved domains in the flowering-time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28: 619-31 (2001).

Roden LC, Song HR, Jackson S, Morris K, Carre IA: Floral responses to photoperiod are correlated with the timing of rhythmic expression relative to dawn and dusk in *Arabidopsis*. Proc Natl Acad Sci U S A 99: 13313-8 (2002).

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G: Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. Science 288: 1613-6 (2000).

Sambrook J, Russel DV: Molecular cloning. Third edition, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York (2001).

Schaffer R, Ramsay N, Samach A, Corden S, Putterill J, Carre IA, Coupland G: The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. Cell 93: 1219-29 (1998).

Searle I, Coupland G: Induction of flowering by seasonal changes in photoperiod. EMBO 23: 1217-22 (2004).

Searle I, He Y, Turck F, Vincent C, Fornara F, Krober S, Amasino RA, Coupland G: The transcription factor *FLC* confers a flowering response to vernalization by repressing meristem competence and systemic signaling in *Arabidopsis*. Genes Dev 20: 898-912 (2006).

Sheldon CC, Finnegan EJ, Dennis ES, Peacock WJ: Quantitative effects of vernalization on *FLC* and *SOC1* expression. Plant J 45: 871-83 (2006).

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES: The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). Proc Natl Acad Sci U S A 97: 3753-3758 (2000).

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES: The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11: 445-58 (1999).

Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C: Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. Plant Physiol 138: 1163-73 (2005).

Simpson GG, Dean C: Arabidopsis, the Rosetta stone of flowering time? Science 296: 285-9 (2002).

Smith KF, Reed KFM, Foot JZ: An assessment of the relative importance of specific traits for the genetic improvement of nutritive value in dairy pasture. Grass Forage Sci 52: 167-175 (1997).

Somers DE: The physiology and molecular bases of the plant circadian clock. Plant Physiol 121: 9-20 (1999).

Somers DE, Webb AA, Pearson M, Kay SA: The short-period mutant, *toc1-1* alters circadian clock regulation of multiple outputs throughout development in *Arabidopsis thaliana*. Development 125: 485-94 (1998).

Strayer C, Oyama T, Schultz TF, Raman R, Somers DE, Mas P, Panda S, Kreps JA, Kay SA: Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. Science 289: 768-71 (2000).

Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G: *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature 410: 1116-20 (2001).

Sung S, Amasino RM: Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein *VIN3*. Nature 427: 159-64 (2004).

Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM: Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. Nat Genet 38: 706-10 (2006).

Taiz L, Zeiger E: Plant physiology. Sinauer Associates, Sunderland, Massachusetts, USA (2002).

Takada S, Goto K: Terminal flower2, an *Arabidopsis* homolog of heterochromatin protein1, counteracts the activation of flowering locus T by *CONSTANS* in the vascular tissues of leaves to regulate flowering time. Plant Cell 15: 2856-65 (2003).

Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K: Hd3a protein is a mobile flowering signal in rice. Science 316: 1033-6 (2007).

Thomas B: Light signals and flowering. J Exp Bot 57: 3387-93 (2006).

Thomas B, Vince-Prue B: Photoperiodism in Plants, second edition. San Diego, CA; Academic Press (1997).

Torok M, Etkin LD: Two B or not two B? Overview of the rapidly expanding B-box family of proteins. Differentiation 67: 63-71 (2001).

Tranquilli G, Dubcovsky J: Epistatic interaction between vernalization genes *Vrn-Am1* and *Vrn-Am2* in diploid wheat. J Hered 91: 304-6 (2000).

Tseng TS, Salome PA, McClung CR, Olszewski NE: *SPINDLY* and *GIGANTEA* interact and act in *Arabidopsis thaliana* pathways involved in light responses, flowering, and rhythms in cotyledon movements. Plant Cell 16: 1550-63 (2004).

Turner A, Beales J, Faure S, Dunford RP, Laurie DA: The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. Science 310: 1031-4 (2005).

Tzeng TY, Hsiao CC, Chi PJ, Yang CH: Two lily *SEPALLATA*-like genes cause different effects on floral formation and floral transition in *Arabidopsis*. Plant Physiol 133: 1091-101 (2003).

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G: Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science 303: 1003-6 (2004).

van der Maas HM, de Jong ER, Rueb S, Hensgens LA, Krens FA: Stable transformation and long-term expression of the gusA reporter gene in callus lines of perennial ryegrass (*Lolium perenne* L.). Plant Mol Biol 24: 401-5 (1994).

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME: Comparative mapping in grasses. Wheat relationships. Mol Gen Genet 248: 744-54 (1995).

Weller JL, Murfet IC, Reid JB: Pea Mutants with Reduced Sensitivity to Far-Red Light Define an Important Role for Phytochrome A in Day-Length Detection. Plant Physiol 114: 1225-1236 (1997).

Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G: CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. Plant Cell 18: 2971-84 (2006).

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D: Integration of spatial and temporal information during floral induction in *Arabidopsis*. Science 309: 1056-9 (2005).

Wilkins PW, Allen DK, Mytton LR: Differences in the nitrogen use efficiency of perennial ryegrass varieties under simulated rotational grazing and their effects on nitrogen recovery and herbage nitrogen content. Grass Forage Sci 55. 14 (2000).

Wilson RN, Heckman JW, Somerville CR: Gibberellin Is Required for Flowering in *Arabidopsis* thaliana under Short Days. Plant Physiol 100: 403-408 (1992).

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA: The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc Natl Acad Sci U S A 103: 14631-6 (2006).

Wratt GS, Smith HC: Plant Breeding in New Zealand. Butterworths, Wellington, New Zealand (1983).

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J: The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-4 (2004).

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J: Positional cloning of the wheat vernalization gene *VRN1*. Proc Natl Acad Sci U S A 100: 6263-8 (2003).

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T: *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12: 2473-2484 (2000).

Yano M, Kojima S, Takahashi Y, Lin H, Sasaki T: Genetic control of flowering time in rice, a short-day plant. Plant Physiol 127: 1425-9 (2001).

Yanovsky MJ, Kay SA: Molecular basis of seasonal time measurement in *Arabidopsis*. Nature 419: 308-12 (2002).

Yoo SK, Chung KS, Kim J, Lee JH, Hong SM, Yoo SJ, Yoo SY, Lee JS, Ahn JH: *CONSTANS* Activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to Promote Flowering in *Arabidopsis*. Plant Physiol 139: 770-8 (2005).

Yu H, Xu Y, Tan EL, Kumar PP: *AGAMOUS-LIKE 24*, a dosage-dependent mediator of the flowering signals. Proc Natl Acad Sci U S A 99: 16336-41 (2002).

Zeevaart JAD: Physiology of flower formation. Ann Rev Plant Phys 27. 5 (1976).

Zhao XY, Liu MS, Li JR, Guan CM, Zhang XS: The wheat *TaGI1*, involved in photoperiodic flowering, encodes an *Arabidopsis GI* ortholog. Plant Mol Biol 58: 53-64 (2005).

APPENDIX

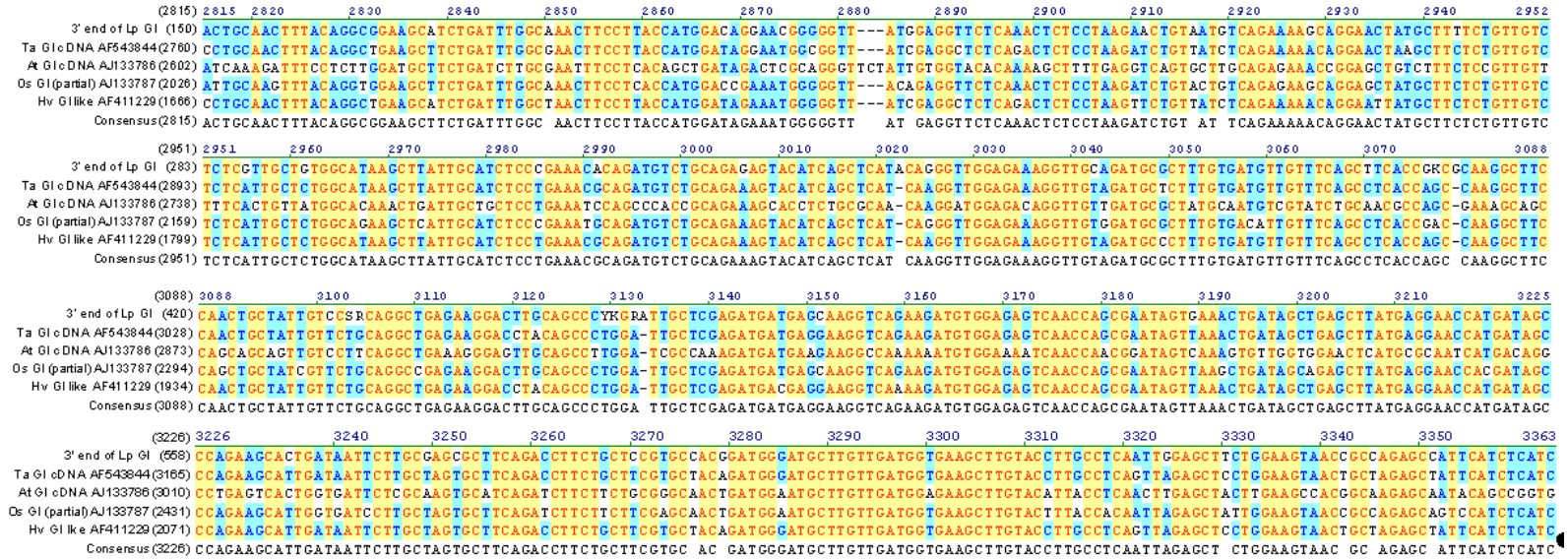


Figure 5.1 Comparison of the partial nucleotide sequences from wheat (AF543844), *Arabidopsis* (AJ133786), rice (AJ133787) and barley (AF411229) *GI* genes and *L. perenne* contig derived from the EST clones containing putative *GI* gene. The numbers in the brackets represents Genbank accession numbers.

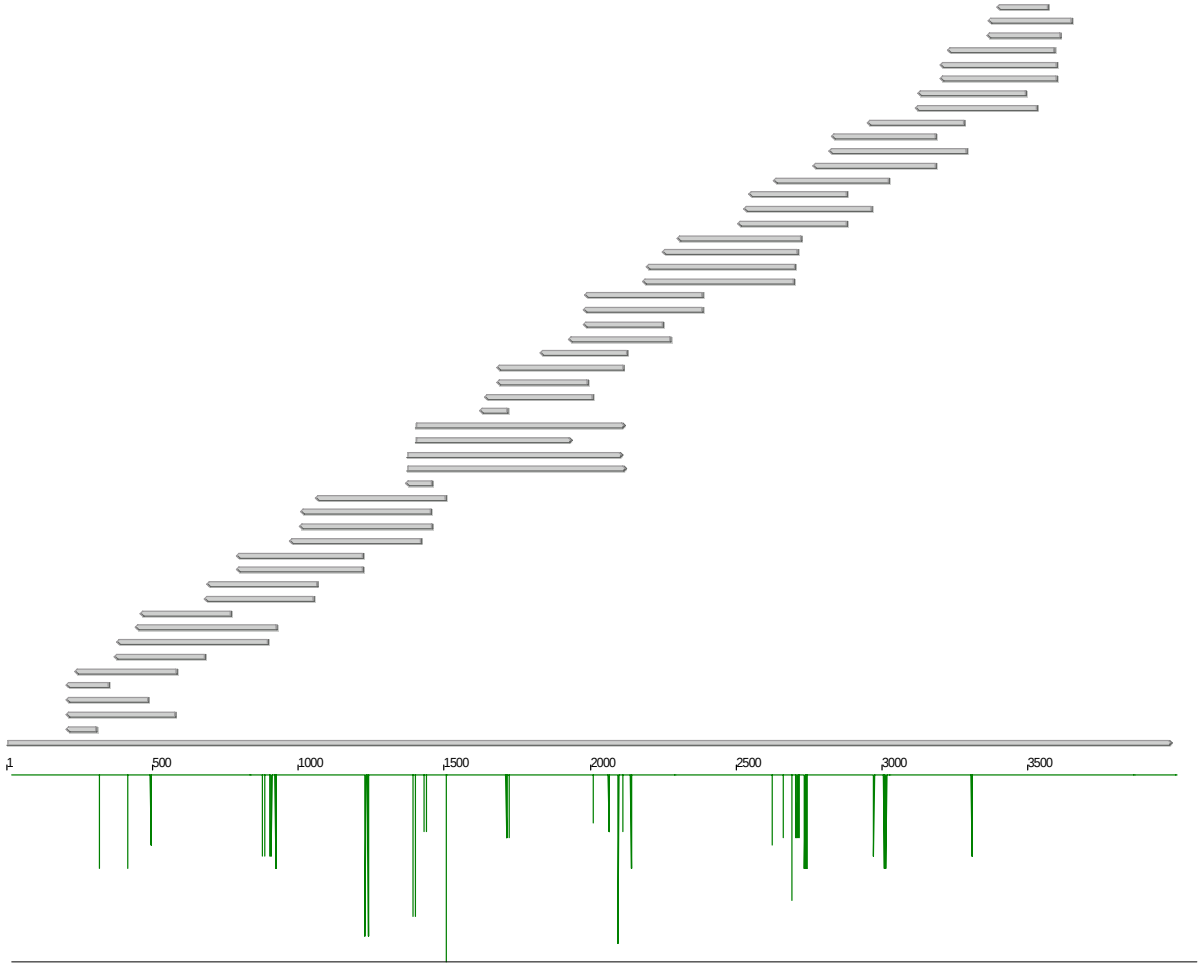
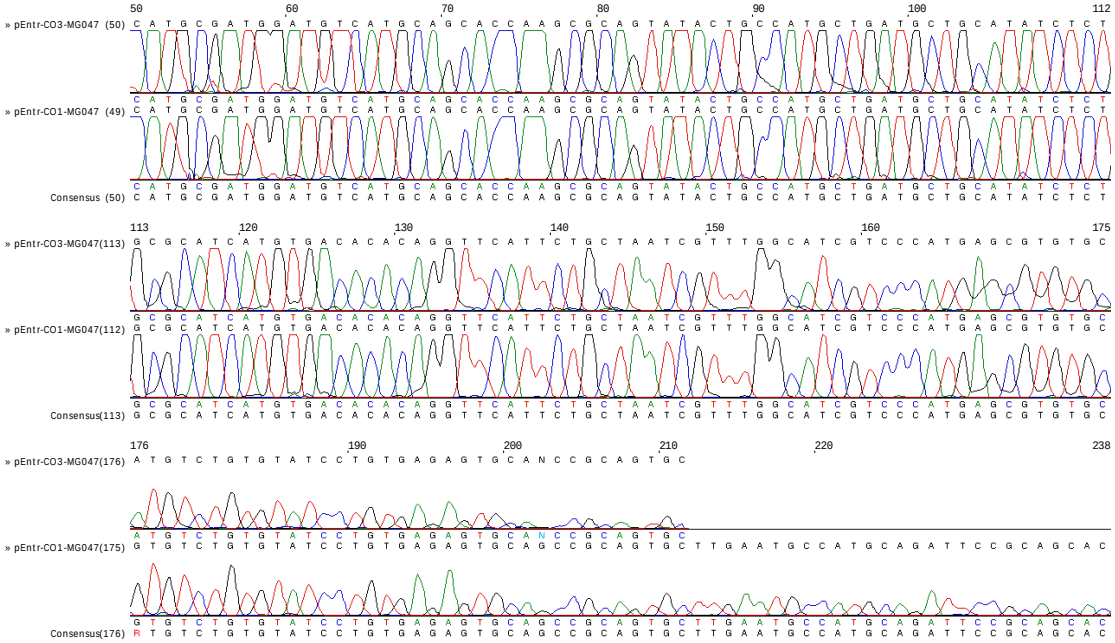


Figure 5.2 Alignment of the *LpGI* contig using Vector NTI ContigExpress module. 14 different primers (MG019, MG037, MG063, MG064, MG065, MG066, MG067, MG068, MG069, MG070, MG071. MG072, MG073, and MG074) were used to sequence the *LpGI* cDNA from the TOPO pCR2.1® cloning vector. Vertical green bars represent SNPs and/or regions with poor sequencing signal in some clones. Any given region within sequence is covered with at least 4 different clones.

```
MXVSNGKWIDGLQFSSLFWP PPHDAQQKQA QTLAYVEYFG
QFTSDSEQFPEDVAQLIQSY YPSKEKRLVD EVLATFVLHH PEHGHAVVHP
ILSRIIDGSLSYDRHGSPFN SFISLFTQTA EKEYSEQWAL ACGEILRVLT
HYNRPIFKVAECNDTSDQAT TSYSLHDKAN SSPENEPERK PLRPLSPWIT
DILLNAPLGIRSDYFRWCGG VMGKYAAGGE LKPPTTAYSR GAGKHPQLMP
STPRWAVANGAGVIXXVCDE EVARYETANL TAAAVPXLLL PPPTMPLDEH
LVAGLPPLEPYARLFHRYYA IATPSATQRL LFGLLEAPPS WAPDALDAAV
QLVELLRAAEDYATGMRLPK NWLHLHFLRA IGTAMSMRAG MAADTAAALL
FRILSQPTLLFPPLRHAEGV VQHEPLGGYV SSYKRQLEIP ASETTIDATA
QGIASLLCAHGPDVXWRICT IWEAAYGLLP LNSSAVDLPE IVVAAPLQPP
TLSWSLYLPLLKVFEYLPRG SPSEACLMRI FVATVEAILR RTFPSETEPS
KKPRSPSKSLAVAELRTMIH SLFVESCASM NLASRLLFVV LTVSVSHQAL
PGGSKRPTGSENHSSEESTE DSKLTNGRNR CKKKQGPVGT FDSYVLAAVC
ALSCELQLFPILCKNVTKTN IKDSIKITMP GKTNGISNEL HNSVNSAILH
TRRILGILEALFSLKPSSVG TSWSYSSNEI VAAAMVAAHV SELFRRSRPC
LNALSALKRCKWDAEISTRA SSLYHLIDLH GKTVSSIVNK AEPLEAHLNL
TAVKKDDQHHIEESNTSSSD YGNLEKKSKK NGFSRPLMKC AEQARRNGNV
ASTSGKATATLQAEASDLAN FLTMDRNGGY GGSQTLLRTV MSEKQELCFS
VVSLLWHKLIASPETQMSAE STSAHQGWRK VADALCDVVS ASPAKAXTAI
VLQAEKDLQPWIARDDEQGQ KMWRVNQRIV KLIAELMRNH DSPEALIILA
SASDLLLRATDGMLVDGEAC TLPQLELLEV TARAIHLIVE WGDPGVAVAD
GLSNLLKCRLSPTIRCLSHP SAHVRALSMS VLRDILNSGP ISSTKIIQGE
QRNGIQSPSYRCAAASMTNW QADVERCIEW EAHNRQATGM TLAFLTAAAN
ELGCPLPC*
```

Figure 5.3 Complete predicted protein sequence encoded by the *LpGI* gene (GB-DQ534010)

A



B

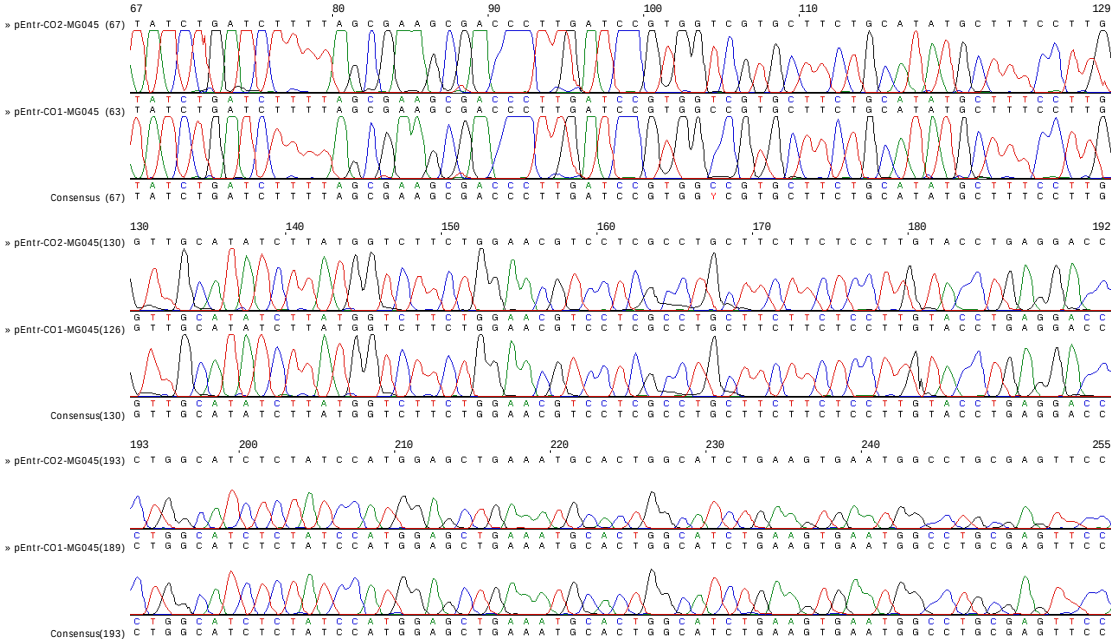


Figure 5.4 *LpCOL1* sequencing electropherogram. A. Electropherogram of the B-box regions B. Electropherogram showing the CCT region of the gene

MLMNCDFNCDLFEQEAKRRSYPWARPCDGCHAAPSAVYCHADAAYLCASCDTQVH
SANRLASSHERVRVCVSCESAAAVLECHADSAALCTTCDAQVHSANPIAQRHQRVP
VLPLPALAIPAASVFAEAEAATTVYGDKEEGEEVDSWLLLERDSDDNNCTNNIDQYX
NLFGYDMYYDKFSCNPGPGEEYRLQEQDVQNMYRENEVCEFAVPSQVGMASEQPES
SYGMIGAEQDASMTAGTSTYTASISNGIPFSSMEVGIIPDNTRPDVSNTNIQRTSEAME
LAGHSLQMPVHFSSMDRDARVLRYKEKKQARTFQKTIRYATRKAYAEARPRIKGRF
AKRSDIEHELDQMLTIPALPDSGHATVLWF

Fig. 5.5 Complete predicted protein sequence encoded by the *LpCOL1* gene (*GB*-DQ534011)

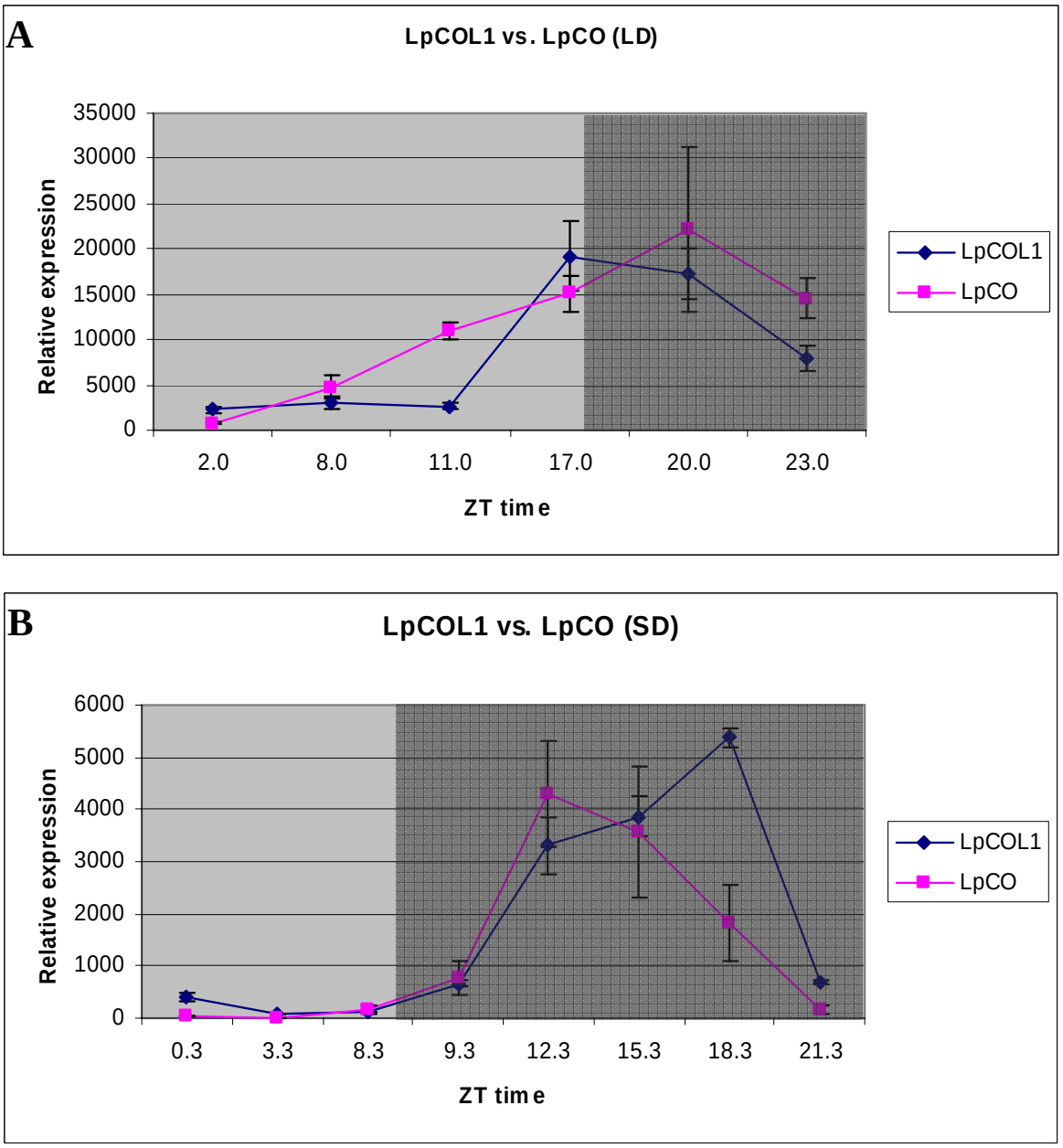


Fig. 5.6 Circadian expression of the *LpCOL1* vs. *LpCO* gene
A. In LD conditions expression of both genes peak late in the afternoon with *LpCO* slightly later than *LpCOL1* **B**. In SD conditions both genes peak during the night. Shaded areas represent night periods. Results were normalized against GAPDH levels.

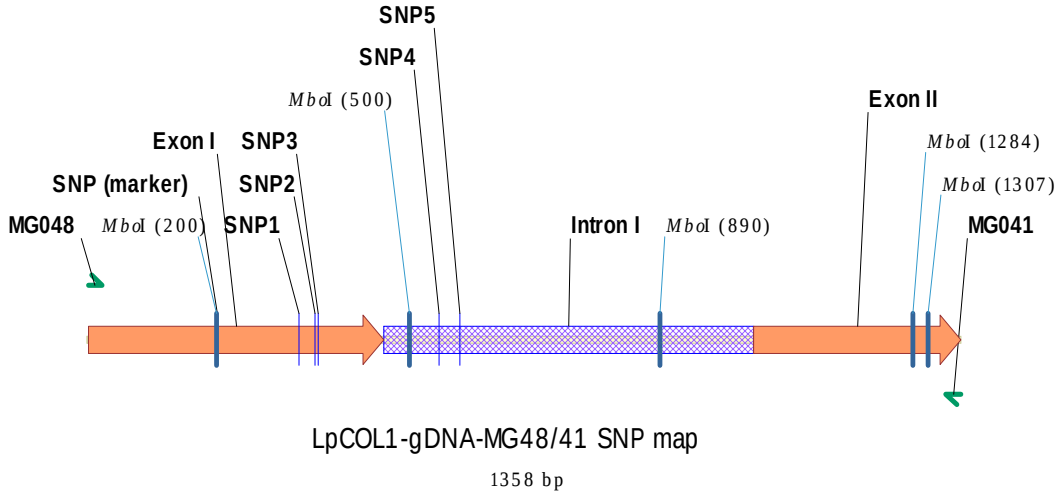


Fig. 5.7 dCAPS method for *LpCOL1* mapping .
SNP from *LpCOL1* (MG048/041) amplicon used for the mapping experiment was at position 201 and part of the *MboI* restriction site. Allelic status was reflected in presence or absence of additional 499 bp band on the gel.

MSINIRDPLIVSRVVGDVLDPFNRSITLKVTYGQREVTNGLDLRPSQVQNKPRVEIGGE
DLRNFYTLVMVDPDVPSPSNPHLREYLHWLVTDIPATTGTTFGNEIVCYENPSPTAGI
HRVVFILFRQLGRQTVYAPGWRQNFNTREFAEIYNLGLPVAAVFYNCQRESGCGGRR
L

Fig. 5.8 LpFT3 protein sequence (GB-DQ309592).

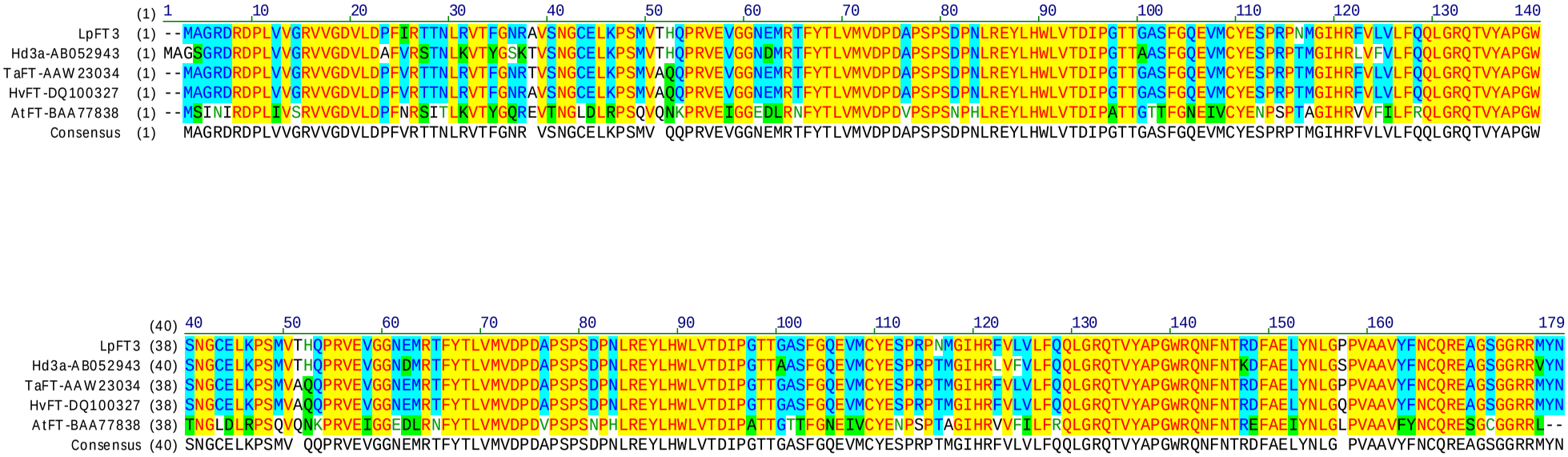


Fig 5.9 Comparison of the FT protein sequences. Proteins from rice (Hd3a), barley (HvFT), wheat (TaFT), and Arabidopsis (AtFT) were used for multiple alignment. Numbers behind the gene names represent respective GeneBank numbers. Yellow blocks represent conserved regions.

Table 5.1 Scoring table for the SNPs used for mapping LpGI, LpCOL1, and LpFT3 genes. 188 genotypes, progeny of cv. Impact and cv. Samson, were available. Some reactions failed to produce reliable results and were therefore excluded from the analysis (empty cells). Y (T/C), S (G/C), R (G/A).

	LpGI		LpCOL1				LpFT3	
platel	**position**	**allele**	**position**	**Seq**	**dCAPS**	**Cons**	**position**	**allele**
1-a	4576	Y	494		A	A	180	
1-b	4576	Y	494		A	A	180	S
1-c	4576		494	A	A	A	180	C
1-d	4576	C	494		R	R	180	
1-e	4576	Y	494				180	
1-f	4576	C	494		R	R	180	
1-g	4576	C	494				180	
1-h	4576	C	494				180	
2-a	4576		494		A	A	180	
2-b	4576	Y	494				180	C
2-c	4576	Y	494		A	A	180	S
2-d	4576	C	494		R	R	180	C
2-e	4576	C	494		A	A	180	C
2-f	4576	C	494		A	A	180	S
2-g	4576		494	A	A	A	180	S
2-h	4576	Y	494		R	R	180	
3-a	4576	Y	494	R	R	R	180	C
3-b	4576	Y	494	R	R	R	180	C
3-c	4576	C	494		A	A	180	C
3-d	4576	Y	494		R	R	180	
3-e	4576	Y	494		R	R	180	
3-f	4576		494		A	A	180	
3-g	4576		494		R	R	180	C
3-h	4576		494		R	R	180	S
4-a	4576	Y	494	R	R	R	180	C
4-b	4576	Y	494	R	R	R	180	C
4-c	4576	C	494		A	A	180	C
4-d	4576	Y	494		R	R	180	C
4-e	4576		494	R	R	R	180	C
4-f	4576	C	494		A	A	180	C
4-g	4576	Y	494		A	A	180	S
4-h	4576	C	494	R	R	R	180	C
5-a	4576		494				180	S
5-b	4576		494	A		A	180	S
5-c	4576		494	R		R	180	S
5-d	4576		494	A		A	180	S
5-e	4576		494	R	R	R	180	C
5-f	4576		494	A		A	180	C
5-g	4576		494	R		R	180	S
5-h	4576		494	A	A	A	180	S
6-a	4576		494				180	

6-b	4576		494				180	C
6-c	4576		494	R	R	R	180	C
6-d	4576		494	A	A	A	180	S
6-e	4576		494				180	S
6-f	4576		494	R		R	180	S
6-g	4576		494				180	C
6-h	4576		494				180	S
7-a	4576	Y	494	R		R	180	S
7-b	4576	Y	494	R		R	180	C
7-c	4576	C	494	A		A	180	C
7-d	4576		494	R	R	R	180	S
7-e	4576	C	494	R		R	180	S
7-f	4576	C	494	R	R	R	180	C
7-g	4576	Y	494	A		A	180	C
7-h	4576	C	494	R	R	R	180	S
8-a	4576	C	494	A		A	180	S
8-b	4576	C	494	A		A	180	S
8-c	4576	Y	494	A	A	A	180	S
8-d	4576	C	494	R	R	R	180	S
8-e	4576	Y	494	A	A	A	180	C
8-f	4576	C	494	R	R	R	180	
8-g	4576	Y	494	A	A	A	180	S
8-h	4576	C	494	A		A	180	S
9-a	4576	C	494				180	S
9-b	4576	Y	494	A		A	180	
9-c	4576	C	494	R		R	180	C
9-d	4576	C	494	R		R	180	S
9-e	4576	Y	494	A		A	180	S
9-f	4576	C	494			R	180	S
9-g	4576	Y	494	A		A	180	S
9-h	4576	C	494	R		R	180	S
10-a	4576	Y	494	A		A	180	S
10-b	4576	Y	494	R		R	180	S
10-c	4576	Y	494	A	A	A	180	
10-d	4576	Y	494	R		R	180	
10-e	4576	Y	494				180	S
10-f	4576	Y	494	R		R	180	C
10-g	4576	Y	494	A		A	180	C
10-h	4576	Y	494	A	A	A	180	
11-a	4576	Y	494	R		R	180	
11-b	4576	Y	494	A		A	180	
11-c	4576	C	494	A		A	180	S
11-d	4576	C	494	A		A	180	
11-e	4576	Y	494	R		R	180	C
11-f	4576	Y	494	R		R	180	C
11-g	4576	Y	494	A		A	180	
11-h	4576	C	494	R		R	180	C
12-a	4576	Y	494			R	180	

12-b	4576	Y	494				180	S
12-c	4576	C	494				180	
12-d	4576	C	494			A	180	C
12-e	4576	C	494			R	180	S
12-f	4576	C	494				180	
12-g	4576	C	494			A	180	C
12-h	4576	C	494			R	180	
plate II								
1-a			494		A	A	180	
1-b			494	R	R	R	180	C
1-c			494	A	A	A	180	
1-d			494	A	A	A	180	C
1-e			494	R	R	R	180	
1-f			494	A	R	R	180	
1-g			494	R	R	R	180	S
1-h			494	A	A	A	180	C
2-a			494	A	A	A	180	C
2-b			494	A	A	A	180	C
2-c			494	A	A	A	180	S
2-d			494	A	A	A	180	S
2-e			494	R	R	R	180	
2-f			494	A	A	A	180	C
2-g			494	A	A	A	180	S
2-h			494	A	A	A	180	
3-a			494		A	A	180	S
3-b			494	R	R	R	180	C
3-c			494	R	R	R	180	C
3-d			494	R	R	R	180	C
3-e			494	R	R	R	180	S
3-f			494	R	R	R	180	C
3-g			494	R	R	R	180	C
3-h			494	R	R	R	180	S
4-a			494	R	R	R	180	C
4-b			494	A	A	A	180	C
4-c			494	A	A	A	180	C
4-d			494	A	A	A	180	S
4-e			494	R	R	R	180	C
4-f			494	R	R	R	180	C
4-g			494	R	R	R	180	C/G
4-h			494	A	A	A	180	C
5-a			494	R	R	R	180	C
5-b			494	R	R	R	180	C
5-c			494	A	A	A	180	C
5-d			494		A	A	180	S
5-e			494	R	R	R	180	C
5-f			494	R	R	R	180	S
5-g			494	A	A	A	180	C
5-h			494	R	R	R	180	

6-a			494		A	A	180	
6-b			494	A	A	A	180	
6-c			494	A	A	A	180	S
6-d			494	A	A	A	180	S
6-e			494	R	R	R	180	S
6-f			494	R	R	R	180	S
6-g			494	R	R	R	180	C
6-h			494	A	A	A	180	
7-a			494	A	A	A	180	S
7-b			494	A	A	A	180	
7-c			494	R	R	R	180	S
7-d			494	R	R	R	180	C
7-e			494	R	R	R	180	C
7-f			494	A	A	A	180	C
7-g			494	R	R	R	180	S
7-h			494	A	A	A	180	S
8-a			494	R	R	R	180	S
8-b			494	A	A	A	180	
8-c			494	R	R	R	180	S
8-d			494	R	A	A	180	S
8-e			494		A	A	180	S
8-f			494	R	R	R	180	S
8-g			494	A	A	A	180	S
8-h			494	R	R	R	180	C
9-a			494				180	
9-b			494	R	R	R	180	
9-c			494	R	R	R	180	S
9-d			494	A	A	A	180	S
9-e			494		R	R	180	C
9-f			494	R	R	R	180	C
9-g			494	R	R	R	180	S
9-h			494	R	R	R	180	C
10-a			494	R	R	R	180	C
10-b			494		R	R	180	S
10-c			494	A	A	A	180	C
10-d			494		R	R	180	
10-e			494	A	A	A	180	S
10-f			494	A	R	R	180	C
10-g			494		R	R	180	S
10-h			494		R	R	180	S
11-a			494	R	R	R	180	
11-b			494	R	R	R	180	S
11-c			494	A	A	A	180	C
11-d			494	A	A	A	180	C
11-e			494	R	R	R	180	S
11-f			494		R	R	180	S
11-g			494	A	A	A	180	S
11-h			494	A	A	A	180	S

12-a			494	R	R	R	180	
12-b			494	A	A	A	180	C
12-c			494	R	R	R	180	C
12-d			494	R	R	R	180	C
12-e			494	R	R	R	180	S
12-f			494	A	A	A	180	C
12-g			494	A	A	A	180	C
12-h			494	A	A	A	180	C

Table. 5.2 RT-PCR raw data for the *LpGI*, *LpCOL1*, and *LpFT3* free running experiments (FR). TP- time point after start of the experiment.

Gene	TP	Ct	Gene	TP	Ct	Gene	TP	Ct
LpGI	0	29.20142	LpCOL1	0	34.65695	LpFT3	0	36.86822
LpGI	0	28.70645	LpCOL1	0	35.33246	LpFT3	0	37.09233
LpGI	2	29.88723	LpCOL1	0	33.55763	LpFT3	2	35.82327
LpGI	2	30.33571	LpCOL1	2	30.81466	LpFT3	2	36.09296
LpGI	2	30.4737	LpCOL1	2	30.85065	LpFT3	4	36.30673
LpGI	4	27.30044	LpCOL1	2	31.22773	LpFT3	4	37.99375
LpGI	4	26.90876	LpCOL1	4	31.86862	LpFT3	6	38.12424
LpGI	6	27.64253	LpCOL1	4	31.55414	LpFT3	8	37.01216
LpGI	6	28.04835	LpCOL1	4	31.71188	LpFT3	8	37.41956
LpGI	6	27.75476	LpCOL1	6	32.44195	LpFT3	10	36.40298
LpGI	8	30.14175	LpCOL1	6	31.68966	LpFT3	10	35.66256
LpGI	8	29.83772	LpCOL1	6	31.96955	LpFT3	12	35.92771
LpGI	8	29.6544	LpCOL1	8	30.94926	LpFT3	12	36.14834
LpGI	10	31.05495	LpCOL1	8	31.39132	LpFT3	14	37.12515
LpGI	10	30.91571	LpCOL1	8	30.17978	LpFT3	14	36.02779
LpGI	10	30.99717	LpCOL1	10	32.05358	LpFT3	16	37.33452
LpGI	12	31.07957	LpCOL1	10	32.39334	LpFT3	16	37.2868
LpGI	12	31.30153	LpCOL1	10	32.73078	LpFT3	18	37.27916
LpGI	12	31.61136	LpCOL1	12	29.98079	LpFT3	18	37.78962
LpGI	14	32.00474	LpCOL1	12	30.64024	LpFT3	20	37.05093
LpGI	14	32.00856	LpCOL1	12	30.50592	LpFT3	20	37.58479
LpGI	14	32.07184	LpCOL1	14	30.98535	LpFT3	22	35.88663
LpGI	16	32.0124	LpCOL1	14	30.80053	LpFT3	22	36.01953
LpGI	16	32.03094	LpCOL1	14	30.69845	LpFT3	24	38.07585
LpGI	16	32.53551	LpCOL1	16	31.66434	LpFT3	24	38.22891
LpGI	18	31.37477	LpCOL1	16	30.77233	LpFT3	26	36.98162
LpGI	18	31.62307	LpCOL1	16	31.11798	LpFT3	26	36.38594
LpGI	18	31.84082	LpCOL1	18	32.59257	LpFT3	28	34.6046
LpGI	20	31.54458	LpCOL1	18	31.5553	LpFT3	28	34.59309
LpGI	20	31.60516	LpCOL1	18	31.08045	LpFT3	30	38.79268
LpGI	20	31.64446	LpCOL1	20	31.04704	LpFT3	30	38.6775
LpGI	22	30.79175	LpCOL1	20	31.62478	LpFT3	32	37.93322
LpGI	22	30.86325	LpCOL1	20	31.02261	LpFT3	34	38.03184
LpGI	22	30.66289	LpCOL1	22	33.80769	LpFT3	34	37.39683
LpGI	24	27.17736	LpCOL1	22	33.35276	LpFT3	36	37.13835
LpGI	24	27.19758	LpCOL1	22	33.68268	LpFT3	36	37.18605
LpGI	24	26.84034	LpCOL1	24	32.98037	LpFT3	38	36.21658

LpGI	26	29.11515	LpCOL1	24	32.97328	LpFT3	38	36.01824
LpGI	26	28.96941	LpCOL1	24	32.84651	LpFT3	40	35.7223
LpGI	26	30.43638	LpCOL1	26	30.31956	LpFT3	40	35.97551
LpGI	28	30.64144	LpCOL1	26	30.66135	LpFT3	42	37.67923
LpGI	28	31.54153	LpCOL1	26	30.67052	LpFT3	42	37.2053
LpGI	30	28.68588	LpCOL1	28	31.23222	LpFT3	44	36.05838
LpGI	30	28.75829	LpCOL1	28	30.6405	LpFT3	46	36.96432
LpGI	30	29.20568	LpCOL1	28	31.14866	LpFT3	46	36.9703
LpGI	32	30.13643	LpCOL1	30	33.42557	LpFT3	48	38.19784
LpGI	32	29.62327	LpCOL1	30	32.61435	LpFT3	48	38.11903
LpGI	32	30.57244	LpCOL1	30	32.20994	LpFT3	50	37.67284
LpGI	34	29.54732	LpCOL1	32	31.3177	LpFT3	50	38.02236
LpGI	34	29.72115	LpCOL1	32	31.48281	LpFT3	52	35.94925
LpGI	34	29.76778	LpCOL1	32	30.71257	LpFT3	52	36.21226
LpGI	36	30.82832	LpCOL1	34	30.45986	LpFT3	54	37.14247
LpGI	36	31.14842	LpCOL1	34	31.76313	LpFT3	54	36.55729
LpGI	36	30.50688	LpCOL1	34	32.4342	LpFT3	56	37.73665
LpGI	38	31.32155	LpCOL1	36	31.92568	LpFT3	56	36.41212
LpGI	38	31.31645	LpCOL1	36	31.86107	LpFT3	58	36.28336
LpGI	38	31.4264	LpCOL1	36	32.03822	LpFT3	58	36.17374
LpGI	40	31.14541	LpCOL1	38	31.95123	LpFT3	60	37.6001
LpGI	40	31.10528	LpCOL1	38	31.00516	LpFT3	60	37.58936
LpGI	40	31.0646	LpCOL1	38	30.85587	LpFT3	62	36.63209
LpGI	42	31.41545	LpCOL1	40	32.41858	LpFT3	64	37.04675
LpGI	42	31.29646	LpCOL1	40	30.55712	LpFT3	64	37.84023
LpGI	42	31.30071	LpCOL1	40	30.15011	LpFT3	66	36.36503
LpGI	44	30.7615	LpCOL1	42	33.67232	LpFT3	66	36.0324
LpGI	44	31.10693	LpCOL1	42	32.50088	LpFT3	68	36.99625
LpGI	44	30.69619	LpCOL1	42	32.58789	LpFT3	68	36.48653
LpGI	46	30.77807	LpCOL1	44	32.00183	LpFT3	70	34.583
LpGI	46	30.28704	LpCOL1	44	32.75799	LpFT3	70	34.16631
LpGI	46	30.29855	LpCOL1	44	31.66598	LpFT3	72	36.46865
LpGI	48	29.93385	LpCOL1	46	31.73498	LpFT3	72	35.96917
LpGI	48	30.74558	LpCOL1	46	31.95319	LpFT3	76	37.02389
LpGI	48	30.29003	LpCOL1	46	31.8541	LpFT3	76	36.02684
LpGI	50	29.6115	LpCOL1	48	30.73477	LpFT3	80	35.52249
LpGI	50	29.85512	LpCOL1	48	30.93637	LpFT3	80	35.86723
LpGI	50	30.17192	LpCOL1	48	30.64975	LpFT3	82	36.97897
LpGI	52	30.13183	LpCOL1	50	34.23698	LpFT3	82	37.23535
LpGI	52	30.73239	LpCOL1	50	34.25014	LpFT3	84	37.74477
LpGI	52	30.4378	LpCOL1	50	34.74499	LpFT3	84	37.7915
LpGI	54	28.91053	LpCOL1	52	31.64681	LpFT3	86	34.21162
LpGI	54	28.80837	LpCOL1	52	30.54959	LpFT3	86	33.9421
LpGI	54	28.31056	LpCOL1	52	31.55094	LpFT3	88	35.98853
LpGI	56	28.82633	LpCOL1	54	33.65055	LpFT3	88	37.8628
LpGI	56	28.39995	LpCOL1	54	33.06518	LpFT3		
LpGI	56	28.42485	LpCOL1	54	32.85623	LpFT3		
LpGI	58	30.47492	LpCOL1	56	29.9071	LpFT3		
LpGI	58	30.17249	LpCOL1	56	31.03347	LpFT3		
LpGI	58	30.92055	LpCOL1	56	29.60071	LpFT3		
LpGI	60	30.86923	LpCOL1	58	30.96141	LpFT3		
LpGI	60	30.74063	LpCOL1	58	31.79563	LpFT3		
LpGI	60	30.38805	LpCOL1	58	31.74843	LpFT3		
LpGI	62	31.10328	LpCOL1	60	30.76029	LpFT3		

LpGI	62	30.95194	LpCOL1	60	30.77823	LpFT3		
LpGI	62	30.59123	LpCOL1	60	30.75519	LpFT3		
LpGI	64	29.61965	LpCOL1	62	32.08159	LpFT3		
LpGI	64	30.23888	LpCOL1	62	31.93048	LpFT3		
LpGI	64	30.51685	LpCOL1	62	29.86879	LpFT3		
LpGI	66	29.6942	LpCOL1	64	31.65844	LpFT3		
LpGI	66	30.23489	LpCOL1	64	31.01805	LpFT3		
LpGI	66	30.23032	LpCOL1	64	30.87254	LpFT3		
LpGI	68	30.97868	LpCOL1	66	32.36806	LpFT3		
LpGI	68	30.31265	LpCOL1	66	31.97646	LpFT3		
LpGI	68	30.91219	LpCOL1	66	31.83131	LpFT3		
LpGI	68	30.44987	LpCOL1	68	33.83968	LpFT3		
LpGI	70	32.06916	LpCOL1	68	34.87745	LpFT3		
LpGI	70	31.61523	LpCOL1	68	32.81101	LpFT3		
LpGI	70	32.0272	LpCOL1	70	30.4691	LpFT3		
LpGI	72	30.53106	LpCOL1	70	30.50045	LpFT3		
LpGI	72	30.48935	LpCOL1	70	30.00116	LpFT3		
LpGI	72	30.82782	LpCOL1	72	33.02548	LpFT3		
LpGI	76	29.25103	LpCOL1	72	33.03856	LpFT3		
LpGI	76	29.54864	LpCOL1	72	32.71232	LpFT3		
LpGI	76	29.31152	LpCOL1	76	32.2377	LpFT3		
LpGI	80	28.28523	LpCOL1	76	31.22604	LpFT3		
LpGI	80	28.6003	LpCOL1	76	31.93	LpFT3		
LpGI	80	28.18997	LpCOL1	80	32.65336	LpFT3		
LpGI	82	29.41255	LpCOL1	80	31.55697	LpFT3		
LpGI	82	29.64198	LpCOL1	80	31.68381	LpFT3		
LpGI	82	29.89928	LpCOL1	82	33.39966	LpFT3		
LpGI	84	30.02607	LpCOL1	82	33.22667	LpFT3		
LpGI	84	29.23879	LpCOL1	82	32.39016	LpFT3		
LpGI	84	30.49842	LpCOL1	84	29.80491	LpFT3		
LpGI	86	34.22867	LpCOL1	84	30.74022	LpFT3		
LpGI	86	33.54253	LpCOL1	84	30.67544	LpFT3		
LpGI	86	34.46868	LpCOL1	86	31.62358	LpFT3		
LpGI	88	28.59502	LpCOL1	86	31.96788	LpFT3		
LpGI	88	28.27408	LpCOL1	86	31.67129	LpFT3		
LpGI	88	27.80336	LpCOL1	88	32.87704	LpFT3		
LpGI			LpCOL1	88	32.19174	LpFT3		
LpGI			LpCOL1	88	31.70752	LpFT3		

Table. 5.3 RT-PCR raw data for *GAPDH* and *Ubi* genes in FR experiment

Gene	TP	Ct	Gene	TP	Copy
GAPDH	0	21.35549	Ubi	0	23.30669
GAPDH	0	21.90062	Ubi	0	23.78284
GAPDH	2	22.08551	Ubi	2	23.57261
GAPDH	2	22.01882	Ubi	2	23.83317
GAPDH	4	21.68832	Ubi	4	23.6269
GAPDH	4	21.989	Ubi	4	23.79087
GAPDH	6	21.72165	Ubi	6	23.47724
GAPDH	6	22.07133	Ubi	6	23.58073
GAPDH	8	22.3648	Ubi	8	23.76127
GAPDH	8	22.15628	Ubi	8	23.79532
GAPDH	10	22.74835	Ubi	10	24.1584
GAPDH	10	22.72113	Ubi	10	24.3483
GAPDH	12	22.34461	Ubi	12	24.02381
GAPDH	12	22.4992	Ubi	12	24.2256
GAPDH	14	22.7495	Ubi	14	24.46957
GAPDH	14	22.73713	Ubi	14	24.42409
GAPDH	16	22.56561	Ubi	16	24.37229
GAPDH	16	22.51336	Ubi	16	24.16924
GAPDH	18	22.27238	Ubi	18	23.94015
GAPDH	18	22.33326	Ubi	18	24.19004
GAPDH	20	22.22975	Ubi	20	24.14674
GAPDH	20	22.27147	Ubi	20	23.99286
GAPDH	22	21.85644	Ubi	22	23.41345
GAPDH	22	22.00174	Ubi	22	23.84822
GAPDH	24	21.89765	Ubi	24	23.55236
GAPDH	24	21.8618	Ubi	24	23.81903
GAPDH	26	22.19331	Ubi	26	23.93685
GAPDH	26	22.50827	Ubi	26	24.2235
GAPDH	28	23.10362	Ubi	28	24.14339
GAPDH	28	23.14203	Ubi	28	24.30012
GAPDH	30	22.67325	Ubi	30	24.27471
GAPDH	30	22.75732	Ubi	30	24.37907
GAPDH	32	22.98478	Ubi	32	24.63845
GAPDH	32	22.98199	Ubi	32	24.48217
GAPDH	34	22.52821	Ubi	34	23.91528
GAPDH	34	22.09766	Ubi	34	24.118
GAPDH	36	22.43146	Ubi	36	24.35354
GAPDH	36	22.63462	Ubi	36	24.23492
GAPDH	38	22.49922	Ubi	38	24.15416
GAPDH	38	22.51233	Ubi	38	24.3409
GAPDH	40	22.23689	Ubi	40	24.39165
GAPDH	40	22.34118	Ubi	40	24.019
GAPDH	42	22.37856	Ubi	42	24.28108
GAPDH	42	22.0178	Ubi	42	24.05524
GAPDH	44	21.9495	Ubi	44	24.16598
GAPDH	44	22.09369	Ubi	44	23.89153
GAPDH	46	22.27218	Ubi	46	24.4642
GAPDH	46	22.08354	Ubi	46	24.1399
GAPDH	48	22.22954	Ubi	48	24.28543

GAPDH	48	22.58901	Ubi	48	24.1547
GAPDH	50	22.28217	Ubi	50	24.32625
GAPDH	50	22.03379	Ubi	50	24.15317
GAPDH	52	22.32416	Ubi	52	23.48796
GAPDH	52	21.87159	Ubi	52	23.07727
GAPDH	54	22.20142	Ubi	54	24.10075
GAPDH	54	22.11136	Ubi	54	24.13655
GAPDH	56	21.86384	Ubi	56	24.01384
GAPDH	56	22.01532	Ubi	56	23.61568
GAPDH	58	22.55058	Ubi	58	24.3622
GAPDH	58	22.8432	Ubi	58	24.36004
GAPDH	60	22.28525	Ubi	60	24.16527
GAPDH	60	22.55419	Ubi	60	24.23677
GAPDH	62	22.63251	Ubi	62	24.2831
GAPDH	64	22.47611	Ubi	64	24.17493
GAPDH	64	21.95915	Ubi	64	23.59541
GAPDH	66	22.07447	Ubi	66	23.96535
GAPDH	66	21.89652	Ubi	66	23.7064
GAPDH	66	21.26616	Ubi	66	23.66555
GAPDH	68	22.26563	Ubi	68	24.27122
GAPDH	68	22.1176	Ubi	68	24.35157
GAPDH	70	22.82439	Ubi	70	24.48253
GAPDH	70	22.83913	Ubi	70	24.3937
GAPDH	72	22.54264	Ubi	72	24.3897
GAPDH	72	22.38053	Ubi	72	24.21202
GAPDH	76	22.15697	Ubi	76	24.1372
GAPDH	76	22.49383	Ubi	76	24.12232
GAPDH	80	21.40818	Ubi	80	23.38705
GAPDH	80	22.02161	Ubi	80	22.18239
GAPDH	82	22.60652	Ubi	82	24.28644
GAPDH	82	22.81515	Ubi	82	24.40142
GAPDH	84	23.06208	Ubi	84	24.17305
GAPDH	84	23.06036	Ubi	84	23.92059
GAPDH	86	24.38391	Ubi	86	24.46289
GAPDH	86	24.07667	Ubi	86	24.58136
GAPDH	88	21.83788	Ubi	88	22.25171
GAPDH	88	21.46171	Ubi	88	22.19819

Table. 5.4 RT-PCR raw data for *LpGI* and *GAPDH* genes in the diurnal expression experiment (CL and CS)

Gene	Zt	Exp	Ct	Gene	Zt	Exp	Ct
LpGI	0	LD	32.16801	GAPDH	0	LD	24.23534
LpGI	0	LD	32.0114	GAPDH	0	LD	24.22419
LpGI	0	LD	31.23809	GAPDH	3	LD	23.45892
LpGI	3	LD	30.32231	GAPDH	3	LD	23.49013
LpGI	3	LD	29.42914	GAPDH	6	LD	23.63351
LpGI	3	LD	28.94007	GAPDH	6	LD	23.5219
LpGI	6	LD	28.82306	GAPDH	9	LD	23.79578
LpGI	6	LD	28.61051	GAPDH	9	LD	23.80578
LpGI	6	LD	28.21851	GAPDH	12	LD	22.82533
LpGI	9	LD	28.19335	GAPDH	12	LD	23.1647
LpGI	9	LD	27.92425	GAPDH	15	LD	24.25768
LpGI	12	LD	28.13231	GAPDH	15	LD	24.2656
LpGI	12	LD	28.06988	GAPDH	18	LD	23.40715
LpGI	12	LD	27.06501	GAPDH	18	LD	23.47845
LpGI	15	LD	32.42337	GAPDH	21	LD	23.75066
LpGI	15	LD	31.89363	GAPDH	21	LD	24.10338
LpGI	15	LD	31.72632	GAPDH	0	SD	25.37787
LpGI	18	LD	33.91653	GAPDH	0	SD	25.50205
LpGI	18	LD	33.49259	GAPDH	3	SD	24.12085
LpGI	21	LD	34.82578	GAPDH	3	SD	24.23532
LpGI	0	SD	34.75185	GAPDH	6	SD	23.93244
LpGI	0	SD	35.58306	GAPDH	6	SD	23.89335
LpGI	3	SD	34.07813	GAPDH	9	SD	24.14715
LpGI	3	SD	33.47499	GAPDH	9	SD	24.28006
LpGI	3	SD	30.7573	GAPDH	12	SD	25.60389
LpGI	6	SD	30.55893	GAPDH	12	SD	25.43552
LpGI	6	SD	29.82961	GAPDH	15	SD	25.03848
LpGI	9	SD	29.38018	GAPDH	15	SD	24.80002
LpGI	9	SD	29.1621	GAPDH	18	SD	25.73789
LpGI	12	SD	32.28198	GAPDH	18	SD	25.62487
LpGI	12	SD	32.19982	GAPDH	21	SD	24.53243
LpGI	12	SD	31.94145	GAPDH	21	SD	24.56338
LpGI	15	SD	34.56743				
LpGI	15	SD	33.22564				
LpGI	18	SD	33.45214				
LpGI	18	SD	33.58879				
LpGI	21	SD	32.29856				
LpGI	21	SD	32.68923				

Table. 5.5 RT-PCR raw data for *LpGI* and *GAPDH* genes in the vernalisation response (GV) experiment

Gene	Week	Ct	Gene	Week	Ct
LpGI	0	30.12219	GAPDH	0	22.28328
LpGI	0	30.02236	GAPDH	0	22.32569
LpGI	0	29.97279	GAPDH	0	22.49123
LpGI	1	29.65896	GAPDH	1	21.77589
LpGI	1	29.5325	GAPDH	1	21.66766
LpGI	1	29.59882	GAPDH	1	21.80007
LpGI	2	29.63658	GAPDH	2	21.36985
LpGI	2	29.91239	GAPDH	2	21.33615
LpGI	2	29.30571	GAPDH	2	21.34425
LpGI	3	28.63258	GAPDH	3	20.88895
LpGI	3	28.82967	GAPDH	3	20.73898
LpGI	3	28.19751	GAPDH	3	21.14274
LpGI	4	29.56698	GAPDH	4	21.74123
LpGI	4	29.24603	GAPDH	4	21.67443
LpGI	4	29.95098	GAPDH	4	21.97645
LpGI	5	28.23659	GAPDH	5	20.56236
LpGI	5	28.06722	GAPDH	5	20.84809
LpGI	5	28.67151	GAPDH	5	20.2024
LpGI	6	29.88965	GAPDH	6	21.70012
LpGI	6	29.96044	GAPDH	6	21.71538
LpGI	6	29.93269	GAPDH	6	21.71421
LpGI	7	29.12569	GAPDH	7	21.54289
LpGI	7	28.8558	GAPDH	7	21.48119
LpGI	7	29.24692	GAPDH	7	21.53705
LpGI	8	29.32568	GAPDH	8	21.55692
LpGI	8	30.40576	GAPDH	8	21.52525
LpGI	8	28.8614	GAPDH	8	21.57771
LpGI	9	30.03152	GAPDH	9	21.71242
LpGI	9	30.27454	GAPDH	9	21.47905
LpGI	9	30.02256	GAPDH	9	21.52397
LpGI	10	29.95874	GAPDH	10	21.00236
LpGI	10	30.01535	GAPDH	10	21.33485
LpGI	10	29.62069	GAPDH	10	20.69382

Table. 5.6 RT-PCR raw data for *LpCOL1*, *LpCO* and *GAPDH* genes in diurnal expression experiments (CL and CS). Empty cells represent failed reactions

Exp	Zt	Gene	Ct	Gene	Ct	Gene	Ct
LD	2	LpCOL	33.50114	LpCO	35.19153	GAPDH	23.59346
LD	2	LpCOL	33.85698	LpCO		GAPDH	23.77856
LD	2	LpCOL	34.11303	LpCO	35.45143	GAPDH	23.79307
LD	8	LpCOL	33.40542	LpCO	32.10592	GAPDH	23.71261
LD	8	LpCOL		LpCO	32.55623	GAPDH	23.56897
LD	8	LpCOL	33.63812	LpCO	32.9492	GAPDH	23.46188
LD	11	LpCOL	32.01011	LpCO	30.76983	GAPDH	22.70663
LD	11	LpCOL	32.23269	LpCO	30.52367	GAPDH	22.88896
LD	11	LpCOL	33.55932	LpCO	30.47702	GAPDH	23.00074
LD	17	LpCOL	29.11156	LpCO		GAPDH	22.45891
LD	17	LpCOL	29.66915	LpCO	29.62958	GAPDH	22.32934
LD	17	LpCOL	28.90073	LpCO	29.29635	GAPDH	22.50868
LD	20	LpCOL	29.81233	LpCO	29.207	GAPDH	24.31619
LD	20	LpCOL		LpCO	29.88569	GAPDH	23.85632
LD	20	LpCOL	31.09258	LpCO	31.53705	GAPDH	23.5536
LD	23	LpCOL	30.12256	LpCO	29.55369	GAPDH	22.56987
LD	23	LpCOL	30.62717	LpCO	29.62958	GAPDH	22.65455
LD	23	LpCOL	31.58246	LpCO	29.29635	GAPDH	22.86413
SD	0.3	LpCOL	37.33648	LpCO	35.26531	GAPDH	23.51236
SD	0.3	LpCOL	38.46596	LpCO	35.59033	GAPDH	23.47959
SD	0.3	LpCOL	36.34442	LpCO	34.9527	GAPDH	23.6669
SD	3.3	LpCOL	38.53685	LpCO	36.8159	GAPDH	21.53672
SD	3.3	LpCOL	38.56978	LpCO	36.99128	GAPDH	21.85632
SD	3.3	LpCOL	38.69658	LpCO	37.83829	GAPDH	22.49789
SD	8.3	LpCOL	38.52189	LpCO	32.89631	GAPDH	22.86523
SD	8.3	LpCOL	38.45117	LpCO	32.51491	GAPDH	22.51072
SD	8.3	LpCOL	38.73232	LpCO	33.66081	GAPDH	23.23376
SD	9.3	LpCOL		LpCO	31.11913	GAPDH	22.89963
SD	9.3	LpCOL	36.58435	LpCO	31.56497	GAPDH	22.98997
SD	9.3	LpCOL	36.77856	LpCO	30.62422	GAPDH	22.93852
SD	12.3	LpCOL	32.19748	LpCO	28.39761	GAPDH	22.69817
SD	12.3	LpCOL	33.11278	LpCO	28.25193	GAPDH	22.87562
SD	12.3	LpCOL	33.81397	LpCO	27.9214	GAPDH	23.33344
SD	15.3	LpCOL	32.62802	LpCO	28.83825	GAPDH	22.61268
SD	15.3	LpCOL	32.26985	LpCO	28.56929	GAPDH	22.78936
SD	15.3	LpCOL	32.16729	LpCO	28.00323	GAPDH	23.30121
SD	18.3	LpCOL	33.03328	LpCO	30.52953	GAPDH	23.72292
SD	18.3	LpCOL	32.23641	LpCO	30.11523	GAPDH	23.96524
SD	18.3	LpCOL	31.91257	LpCO	29.59308	GAPDH	24.3353
SD	21.3	LpCOL	35.98524	LpCO	34.07065	GAPDH	23.55603
SD	21.3	LpCOL	35.98589	LpCO	33.62398	GAPDH	23.32656
SD	21.3	LpCOL	36.06236	LpCO	33.04623	GAPDH	23.20189

Table. 5.7 RT-PCR raw data for *LpCOL1* and *GAPDH* genes in the vernalisation vs. photoperiod response experiment (GV)

Vern.	LDs	Gene	Ct	Gene	Ct
0	0	LpCOL1	29.4837	GAPDH	21.92208
0	0	LpCOL1	30.89218	GAPDH	22.36186
0	3	LpCOL1	28.68389	GAPDH	22.36958
0	3	LpCOL1	29.90747	GAPDH	22.36331
0	7	LpCOL1	28.59315	GAPDH	21.39905
0	7	LpCOL1	28.33658	GAPDH	21.25961
1	0	LpCOL1	30.46426	GAPDH	22.28328
1	0	LpCOL1	30.37393	GAPDH	22.49123
1	7	LpCOL1	28.79748	GAPDH	21.7399
1	7	LpCOL1	29.05941	GAPDH	21.84616
3	0	LpCOL1	29.28483	GAPDH	21.72279
3	0	LpCOL1	30.54017	GAPDH	21.66766
3	3	LpCOL1	28.35899	GAPDH	21.80007
3	3	LpCOL1	28.64327	GAPDH	21.89822
3	7	LpCOL1	27.49349	GAPDH	24.05762
3	7	LpCOL1	27.45236	GAPDH	23.37737
4	0	LpCOL1	32.90261	GAPDH	21.83416
4	0	LpCOL1	32.90524	GAPDH	21.2619
4	3	LpCOL1	30.05466	GAPDH	21.57978
4	3	LpCOL1	30.12896	GAPDH	21.33615
4	7	LpCOL1	29.25074	GAPDH	21.34425
4	7	LpCOL1	28.57324	GAPDH	21.66499
7	0	LpCOL1	28.79971	GAPDH	23.86751
7	0	LpCOL1	29.66772	GAPDH	23.70707
7	3	LpCOL1	27.41648	GAPDH	21.90493
7	3	LpCOL1	28.09173	GAPDH	21.49909
7	7	LpCOL1	26.91269	GAPDH	21.3113
7	7	LpCOL1	27.46489	GAPDH	20.73898
8	0	LpCOL1	32.02553	GAPDH	21.14274
8	0	LpCOL1	31.89652	GAPDH	22.20183
8	3	LpCOL1	27.84834	GAPDH	21.61659
8	3	LpCOL1	27.37364	GAPDH	22.76131
8	7	LpCOL1	26.33112	GAPDH	23.01118
8	7	LpCOL1	26.98315	GAPDH	22.51316
9	0	LpCOL1	31.92663	GAPDH	22.78557
9	0	LpCOL1	32.49901	GAPDH	21.67443
9	3	LpCOL1	28.54783	GAPDH	21.97645
9	3	LpCOL1	28.07627	GAPDH	21.92831
9	7	LpCOL1	27.57929	GAPDH	21.80454
9	7	LpCOL1	27.35652	GAPDH	21.88726

Table. 5.8 RT-PCR raw data for *LpFT3* and *Ubi* genes in the vernalisation response experiment (GV).

Week	Gene	Ct	Gene	Ct
0	LpFT3	38.5279	Ubi	23.50972
0	LpFT3	38.47612	Ubi	23.40458
1	LpFT3	38.90939	Ubi	22.42768
1	LpFT3	38.78536	Ubi	22.42342
2	LpFT3	40.04804	Ubi	21.7797
2	LpFT3	41.21225	Ubi	21.58422
3	LpFT3	32.05114	Ubi	22.10235
3	LpFT3	32.27134	Ubi	22.03615
4	LpFT3	36.34324	Ubi	22.99751
4	LpFT3	35.79975	Ubi	22.85194
5	LpFT3	36.67396	Ubi	22.18055
5	LpFT3	36.74895	Ubi	22.13077
6	LpFT3	32.79422	Ubi	22.32226
6	LpFT3	32.68569	Ubi	22.5868
7	LpFT3	32.23739	Ubi	22.38537
7	LpFT3	32.31852	Ubi	22.02936
8	LpFT3	31.87476	Ubi	22.39122
8	LpFT3	32.04088	Ubi	22.57347
9	LpFT3	31.80737	Ubi	23.27004
9	LpFT3	32.02356	Ubi	23.0743
10	LpFT3	32.37542	Ubi	23.12529
10	LpFT3	32.07172	Ubi	22.78415

Table. 5.9 RT-PCR raw data for *LpFT3* and *Ubi* genes in the diurnal expression experiment (CL and CS).

Zt	Exp	Gene	Ct	Gene	Ct
1	CL	LpFT3	36.52566	GAPDH	22.69817
1	CL	LpFT3	36.08322	GAPDH	22.61268
1	CL	LpFT3	36.34802	GAPDH	22.51072
5	CL	LpFT3	35.69106	GAPDH	21.53672
5	CL	LpFT3	34.59422	GAPDH	23.72292
5	CL	LpFT3	34.97822	GAPDH	23.6669
9	CL	LpFT3	34.11469	GAPDH	23.20189
9	CL	LpFT3	34.79343	GAPDH	23.35399
9	CL	LpFT3	34.51243	GAPDH	23.33344
13	CL	LpFT3	35.59787	GAPDH	23.30121
13	CL	LpFT3	35.3003	GAPDH	23.23376
13	CL	LpFT3	35.8703	GAPDH	22.98997
17	CL	LpFT3	37.43183	GAPDH	22.49789
17	CL	LpFT3	37.53403	GAPDH	24.3353
17	CL	LpFT3	37.86343	GAPDH	22.93852
21	CL	LpFT3	36.79333	GAPDH	21.92208
21	CL	LpFT3	36.45283	GAPDH	22.36186
21	CL	LpFT3	36.81273	GAPDH	22.36958
1	CS	LpFT3	38.79065	GAPDH	22.36331
1	CS	LpFT3	37.55776	GAPDH	21.39905
1	CS	LpFT3	38.56329	GAPDH	21.25961
5	CS	LpFT3	37.67548	GAPDH	22.28328
5	CS	LpFT3	36.11812	GAPDH	22.49123
5	CS	LpFT3	36.21985	GAPDH	21.7399
9	CS	LpFT3	37.18358	GAPDH	21.84616
9	CS	LpFT3	36.96622	GAPDH	21.72279
9	CS	LpFT3	37.82659	GAPDH	21.66766
13	CS	LpFT3	37.59605	GAPDH	21.80007
13	CS	LpFT3	37.52067	GAPDH	22.89822
13	CS	LpFT3	37.82395	GAPDH	24.05762
17	CS	LpFT3	39.14064	GAPDH	21.37737
17	CS	LpFT3	37.58628	GAPDH	21.83416
17	CS	LpFT3	38.58228	GAPDH	21.2619
21	CS	LpFT3	38.91045	GAPDH	21.57978
21	CS	LpFT3	38.52398	GAPDH	21.33615
21	CS	LpFT3	39.01293	GAPDH	21.34425

Table. 5.10 RT-PCR raw data for *LpFT3* and *Ubi* genes in the floral induction experiment (LD).

LD	Gene	Ct	Gene	Ct
0	LpFT3	38.08381	Ubi	24.0549
0	LpFT3	39.23468	Ubi	23.62368
1	LpFT3	37.79783	Ubi	23.9851
1	LpFT3	38.55463	Ubi	24.26474
2	LpFT3	37.95625	Ubi	23.47867
2	LpFT3	37.65603	Ubi	23.62686
3	LpFT3	36.95643	Ubi	23.21151
3	LpFT3	37.74946	Ubi	23.17205
4	LpFT3	36.70036	Ubi	24.23609
4	LpFT3	37.31949	Ubi	24.05089
5	LpFT3	39.17657	Ubi	24.32874
5	LpFT3	37.69032	Ubi	24.37572
7	LpFT3	37.14322	Ubi	24.24203
7	LpFT3	36.85693	Ubi	24.34279
8	LpFT3	34.94706	Ubi	23.21856
8	LpFT3	35.75523	Ubi	23.3857
10	LpFT3	36.40576	Ubi	24.79328
10	LpFT3	36.57925	Ubi	25.50021
12	LpFT3	36.20679	Ubi	24.43479
12	LpFT3	35.67611	Ubi	24.24522
14	LpFT3	35.20388	Ubi	24.5574
14	LpFT3	35.71149	Ubi	24.7224
16	LpFT3	35.40834	Ubi	24.40644
16	LpFT3	35.78166	Ubi	24.58139
18	LpFT3	34.278	Ubi	23.83629
18	LpFT3	34.15285	Ubi	23.00197
20	LpFT3	34.34205	Ubi	25.45044
20	LpFT3	34.86811	Ubi	24.71303
22	LpFT3	35.12246	Ubi	24.53444
22	LpFT3	35.34356	Ubi	24.72801
24	LpFT3	35.45121	Ubi	26.00921
24	LpFT3	35.86864	Ubi	25.97369
26	LpFT3	36.23675	Ubi	24.81626
26	LpFT3	35.34007	Ubi	24.81887