Natural variation in the *Arabidopsis thaliana* circadian clock as a determinant of flowering time: a quantitative genetics and genomics study

Inaugural–Dissertation

zur
Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Muhammad Usman Anwer

aus Faisalabad (Pakistan)

Köln, Juni 2014

Diese Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln, in der Arbeitsgruppe von PD. Dr. Seth J. Davis, Abteilung für Entwicklungsbiologie der Planzen (Direktor: Prof. Dr. George Coupland) angefertigt.

Berichterstatter:	Prof. Dr. George Coupland
	Prof. Dr. Ute Höcker
Prüfungsvorsitzender:	Prof. Dr. Martin Hülskamp
Tag der Disputation:	21 Juni 2012

Dedicated to those raised hands
who always pray for my success

Table of Contents

Chapter One: Introduction……...…………………………………….1

Chapter Two: Materials and methods…………………………………. 25

Table of Contents

Table of Contents

Table of Contents

List of Figures

List of Figures

List of Tables

List of Tables

List of Abbreviations

ACS	AUTOMATIC COFACTOR SELECTION
ANOVA	ANALYSIS OF VARIANCE
BC	CONTINUOUS BLUE LIGHT
bHLH	BASIC-HELIX-LOOP-HELIX
BL	BLUE LIGHT
bZIP	BASIC-REGION LEUCINE-ZIPPER
CAB	CHLOROPHYLL A/B BINDING PROTEIN
CAPS	CLEAVED AMPLIFIED POLYMORPHIC SEQUENCE
CCA1	CIRCADIAN CLOCK ASSOCIATED 1
CCG	CLOCK-CONTROL GENE
CCR2	COLD AND CLOCK REGULATED 2
CO	CONSTANS
COP1	CONSITITUTIVELY PHOTOMORPHOGENIC 1
CRY	CRYPTOCHROME
CT	CIRCADIAN TIME
dCAPS	DERIVED CLEAVED AMPLIFIED POLYMORPHIC SEQUENCE
DD	CONTINUOUS DARK
EE	EVENING ELEMENT
ELF3	EARLY FLOWERING 3
ELF4	EARLY FLOWERING 4
EMS	ETHYL METHANESULFONATE
FHY3	FAR-RED ELONGATED HYPOCOTYL 3
FRC	CONTINUOUS FAR-RED LIGHT
FRL	FAR-RED LIGHT
FT	FLOWERING LOCUS T
GI	GIGANTEA
GLM	GENERAL LINEAR MODEL
HAT	HISTONE-ACETYLTRANSFERASE

List of Abbrevations

IM	INTERVAL MAPPING
LBS	LUX BINDING SITE
LD	LIGHT/DARK CYCLES
LHY	LATE AND LONG HYPOCOTYL
LL	CONTINUOUS LIGHT
LOD	LIKELIHOOD OF ODDS
LUC	LUCIFERASE
LUX	LUX ARRYTHMO
MQM	MULTIPLE QTL MAPPING
NLS	NUCLEAR LOCALIZATION SIGNAL
OX	OVER-EXPRESSOR
PER	PERIOD
PHY	PHYTOCHROME
PIF	PHYTOCHROME INTERACTING FACTOR
PLC1	PHYTOCLOCK
PNB	PHYTOCHROME NUCLEAR BODY
PRC	PHASE RESPONSE CURVE
PRR	PSEUDO RESPONSE REGULATOR
QTL	QUANTITATIVE TRAIT LOCUS
RILS	RECOMBINANT INBRED LINES
SSLP	SIMPLE SEQUENCE LENGTH POLYMORPHISM
TILLING	TARGET INDUCED LOCAL LESIONS IN GENOMES
TIM	TIMELESS
TOC1	TIME OF CAB EXPRESSION
WC	WARM/COOL CYCLES
WC	WHITE COLLAR
ZT	*ZEITGEBER* TIME
ZTL	ZEITLUPE

Abstract

The circadian clock is an endogenous mechanism present in plants that enables anticipation of upcoming environmental changes. In this way, the clock thus facilitates the correct timing of physiological and developmental events. An internal clock synchronized with the external environment ensures that plants flower under favorable environmental conditions, and thus, provides fitness advantages under natural habitats. In Arabidopsis, the molecular framework of the circadian-clock machinery is constituted by a network of transcription/translation-based feedback loops. The main loop that comprises CCA1/LHY and TOC1, is connected with the morning loop comprising PPR7/PRR9 through CCA1/LHY. TOC1 establishes a link of this main loop with an as of yet unknown component (Y) in the evening loop. The flowering-time component GI fulfills some of the requirement of 'Y'. However, functional studies do not fully support the GI to TOC1 link. ELF3 is another clock component whose function is not fully understood. In the past, efforts for functional placement of ELF3 were mainly hampered by the arrhythmic nature of available *elf3* alleles. In chapter 3 of this thesis, I described the characterization of a natural allele of *ELF3* (*ELF3-Sha*) that displays conditional rhythmicity and provides evidences of the ELF3 mode-of-action in the clock. From this, in chapter 4, using quantitative genetics and genomics studies in Arabidopsis, I identified the possible positions of the new allelic variants that connect the speed of circadian oscillation to flowering time.

Adaptation of Arabidopsis accessions to varying environmental conditions of light and temperature at different geographical areas has provided a force for the selection of allelic variants within the circadian clock. Previously, the location of such allelic variants affecting speed of the clock was identified in a Bay x Sha recombinant inbred line (RIL) population. This population was modified with a *CCR2::LUC* promoter-reporter system, which provides robust estimates of the clock activity, and hence, the speed of the clock can be accurately measured. In such experiments, a major-effect periodicity QTL was detected, where the presence of Sha allele at the locus caused a short-period phenotype. I report in chapter 3 the positional isolation of this QTL and revealed that *ELF3* was the underlying gene. ELF3 has been reported to be a nuclear-localized protein required for both the

generation of circadian rhythms and light input to the circadian clock. I found that the casual polymorphism in the *ELF3-Sha* allele was an encoded replacement in highly conserved amino acid that caused the short-period phenotype. Interestingly, this phenotype was light dependant. In constant darkness, a different phenotype was found, and this was a reduction in clock function with a movement towards arrhythmic oscillations. The cellular basis of the ELF3-Sha protein defect correlated with an increase in cytosolic distribution. Taken together, my data suggests that the light and circadian action of ELF3 in the clock is dependent on nuclear recruitment to initiate repressive action in the clock network.

Characterization of *ELF3-Sha* established the notion that natural variation in Arabidopsis accessions provides an extraordinary resource to study complex physiological mechanisms. Therefore, I decided to exploit this natural resource to investigate the relationship between the circadian clock and flowering time, which is reported in chapter 4. For that, I selected three Arabidopsis accessions from geographically diverged areas: two from northern latitudes and one from the equator. Whole genome re-sequencing of these accessions revealed that, on a genome-wide scale, the equatorial accession is highly diverged compare to the northern accessions. This encouraged me to utilize these accessions in quantitative-genetic studies. Pair-wise crosses were made to generate RIL sets that were screened for circadian periodicity and flowering-time traits. Construction of genetic maps followed by QTL mapping revealed several loci controlling the speed of clock and flowering time. Many of these QTLs were localized with the known genes controlling both of these traits. Novel loci were also detected. Validation of a major QTL in a near isogenic line, followed by sequence comparison of the candidate gene in parental accessions, revealed that a single encoded amino-acid change in *GI* potentially causes the periodicity acceleration. Further, identification of flowering-time loci overlapping with the position of known clock genes provided the mechanistic explanation of the relationship of clock with flowering time. Taken together, my research strongly advocates the importance of the natural-variation studies in understanding intricate physiological pathways and their interactions with each other.

Zusammenfasung

Die innere Uhr ist ein endogener Mechanismus in Pflanzen, der es ermöglicht Umweltveränderungen vorauszusagen. Deshalb erleichtert die innere Uhr die korrekte Abstimmung von physiologischen und entwicklungsbiologischen Vorgängen. Läuft die innere Uhr synchron mit den äußeren Umwelbedingungen, wird sichergestellt, dass Pflanzen bei günstigen Umweltbedingunen blühen, wodurch in natürlichen Habitaten ein Fintess-Vorteil erreicht wird. In Arabidopsis besteht das molekulare Netzwerk der inneren Uhr aus transkriptionellen und translationellen *feedback-Schleifen.* Die Hauptschleife bestehend aus CCA1/LHY und TOC1, ist mit der sog. Morgenschleife verbunden, die PRR7/PRR9 und CCA1/LHY miteinander verknüpft. TOC1 ist mit der sog. Abendschleife über eine bisher unbekannte Komponente Y verknüpft. Das Blütezeitgen GI erfüllt einiger der für Y notwendigen Anforderungen. Allerdings konnten funktionelle Studien nicht vollständig die Verbindung zwischen GI und TOC1 belegen. ELF3 ist eine weitere Komponente der inneren Uhr, deren Funktion noch nicht eindeutig geklärt wurde. In der Vergangenheit wurden Versuche, die Funktion und Rolle von ELF3 in der inneren Uhr zu klären durch den arrhythmischen Phänotyp der zur Verfügung stehenden ELF3 Allele behindert. In Kapitel 3 dieser Doktorarbeit beschreibe ich die Charakterisierung eines natürlichen Allels von ELF3 (ELF3-Sha), welches eine gewisse Rhythmik aufweist und daher einige Hinweise auf die Funktionsweise von ELF3 gibt. In Kapitel 4 habe ich beschrieben, wie durch quantitative Genetik und genomische Studien in Arabdiopsis mögliche Funktionen dieser neuen Allelvarianten, die die Geschwindgkeit des circadianen Oszillators mit der Blütezeit verknüpft, identifiziert wurden.

Die Anpassung von Arabidpsis *Accessions* an sich verändernde Licht- und Temperaturbedingungen an verschiedenen geografischen Standorten hat einen Selektionsdruck auf Allelvarianten der inneren Uhr ausgeübt. Zuvor wurde solch eine Allelvariation, die die Geschwindigkeit der inneren Uhr beinflusst, durch eine Bay x Sha rekombinante Inzuchtlinie identifiziert. Diese Population wurde mit einem *CCR2::luc* promoter-reporter System modifiziert, das eine exakte Bestimmung der Aktivität der inneren Uhr ermöglicht, wodurch die Geschwindigkeit genau gemessen werden kann. In

einem dolchen Experiment konnte eine QTL identifiziert werden, die eine großen Einfluss auf die Periodizität hatte und bei der, wenn das Sha Allel an diesem Locus present war, eine kurze Periode der Uhr beobachtet werden konnte. In Kapitel 3 beschreibe ich die positionelle Isolation dieser QTL, die hervorbrachte, dass *ELF3* das zugrundelegende Gen war. Von ELF3 wurde bereits berichtet, dass es ein im Nucleus lokalisiertes Protein ist, das sowohl für die Erzeugung circadianer Rhythmes als auch für die Lichtsignalgebung in die innere Uhr benötigt wird. Ich konnte herausfinden, dass ein gelegentlicher Polymporphismus im ELF3-Sha Allel durch einen Aminosäureaustausch in einer hochkonservierten Sequenz kodiert war, der den Phenotyp der kurzen Periode verursachte. Interessanter Weise war dieser Phänotyp lichtabhängig. In konstanter Dubkelheit konnte ein anderer Phänotyp beobachtet werden, nämlich eine reduzierte Funktion der Uhr, die zu arrhythmischer Oszillation tendierte. Der Defekt des ELF3-Sha Proteins korrelierte mit einer vermehrt zytosolischen Lokalisation des Proteins. Zusammenfassend lassen meine Daten vermuten, dass Lichtinput durch ELF3 und die circadiane Aktivität von ELF3 abhängig sind von einer verstärkte Lokalisierung des Proteins im Kern, wo dann die Reprimierung auf das circadiane Netzwerk ausgeübt werden kann.

Die Charakterisierung von *ELF3-Sha* etablierte die Idee, dass natürliche Variation in Arabiopsis *Accessions* eine hervorragende Möglichkeit bietet um komplexe physiologische Mechanismen zu untersuchen. Deshalb entschied ich mich dazu diese natürliche Resource zu nutzen um die Verknüpfung zwischen der inneren Uhr und der Blütezeit zu untersuchen und beschreibe dies in Kapitel 4. Dazu wählte ich drei Arabidopsis *Accessions* aus, die aus geographisch sehr unterschiedlichen Regionen stamen: zwei von nördlichen Breitengraden und eine vom Äquator. Eine Re-Sequenzierung des gesamten Genoms dieser *Accessions* deckte auf, dass auf genomweiter Skala die äquatoriale *Accession* sich deutlich von den nördlicheren *Accessions* unterscheidet. Dies ermutigte mich dazu, diese *Accessions* für quantitative genetische Studien zu verwenden. Zur erzeugung Rekombinanter Inzuchtlinien wurden paarweise Kreuzungen durchgeführt, die auf circadiane Periodizität und Blütezeit Eigenschaften hin untersucht wurden. Nach erfolgter Konstruktuion genetischer Karten und einem QTL Mapping konnten mehrere Loci indentifiziert werden, die die

Geschwindigkeit der inneren Uhr und die Blütezeit kontrollieren. Viele dieser QTLs waren bei bereits bekannten Genen für diese Eigenschaften lokalisiert worden. Allerdings konnten auch neue Loci identifiziert werden. Eine Validierung der wichtigsten QTL in einer Nahen Isogenen Linie und ein anschließender Sequenzvergleich der Kandidatengene in den elterlichen *Accessions* deckte auf, dass ein einziger Aminosäureaustausch in *GI* vermutlich die Beschleunigung der Periode verursachte. Darüberhinaus wurde durch die Indentifizerung von Blütezeit Loci, die mit der Position bekannter Gene für die Innere Uhr überlappten, eine Erklärung für die Verknüpfung der inneren Uhr mit der Blütezeit gefunden werden. Zusammenfassend lässt sich sagen, dass meine Untersuchungen sehr empfehlen, Studien zur natürlichen Variation zu verwenden um komplexe physiologische Abläufe und deren Zusammenhänge aufzuklären.

Chapter One: **Introduction**

1.1 Synchronization with external environment: a process that provides fitness

Ever since life arose on the earth, organisms have been faced constant, but predictable changes in their external environment. These include daily fluctuations in light and temperature. Consequently, many organisms have acquired the capacity to temporally regulate their biology to coordinate this with their environment in anticipation of these predictably coming changes. This is mainly achieved by an internal time-keeping system known as the circadian clock or biological clock (McClung, 2006). The necessity for this clock system is suggested by its existence in most organisms (Young and Kay, 2001), and its importance is underscored by the discovery that coincidence between endogenous rhythms and environmental cycles increases the biological fitness (Dong and Golden, 2008; McClung, 2011; Yerushalmi and Green, 2009).

The first report that plants exhibit circadian rhythms was provided by de Mairan (1729) when he found that the daily leaf movement of *Mimosa pudica* remains persistent in constant darkness. Later, de Candolle confirmed this observation and determined that the free-running period of rhythms in *Mimosa pudica* was 22-23h. This showed that the existence of an internal time keeping system can deviate from external environmental cycles (de Candolle, 1832). Since then, many authors have reported these kinds of circadian rhythms (McClung, 2006). However, the existing of such endogenous system was not universally accepted until chronological experiments on the fungus *Neurospora crassa* were conducted in space (Sulzman et al., 1984). Those confirmed that rhythms were truly endogenous and not driven by some undetected geophysical factors that are associated with the rotation of the earth.

Circadian rhythms can be mathematically defined (Dunlap et al., 2004, McClung 2006). A complete circadian cycle can be fully described by no less than four parameters: period, phase, amplitude and robustness (Figure 1.1). Period is defined as the time requires to complete one cycle. This is normally measured from peak to peak or trough to trough. Period can be used to estimate the speed of the clock. Phase is the time of particular event

during a circadian cycle relative to a specific time of the day, for instance, peak of an activity relative to dawn. Both period and phase have extensively been studied to understand the controlling features of circadian system (Boikoglou et al., 2011; Darrah et al., 2006; Edwards et al., 2005; Kolmos et al., 2011). A less well utilized parameter is amplitude, which is defined as half of the magnitude of peak to trough in a circadian cycle. Amplitude has a direct relationship with the fourth circadian parameter, robustness, which is define by the precision of the oscillation and is measured as an atypical term, relative amplitude error (R.A.E.). RAE is the output of the Fast Fourier Transform (FFT) analysis that describes how good the observed oscillations cycle fits to a perfect oscillation. Thus, RAE can be used to estimate the overall robustness of the circadian oscillation. Importantly, all of these parameters are genetically tractable, for instance by measuring the leaf movement rhythms (McClung, 2006; McWatters et al., 2007).

External environmental stimuli are required to initiate circadian oscillations and to keep the internal system well synchronized with the external environment. This process is called entrainment and the environmental cues governing this process are termed as *zeitgebers* (from German: time givers). Light and temperature are two predominant *zeitgebers* that resets the clock daily. Light changes during a day-night cycle are pronounced and thus robustly entrain the clock. Temperature changes of as little as Δ0.5°C can also effectively entrain a plant clock. Thus, this system is precise in response to thermal changes (Rensing and Ruoff, 2002). One remarkable property of the clock is that it is also temperature compensatable. This means that the pace of the oscillations remain nearly the same over a range of physiologically relevant temperatures. Thus when plants encounter a period of sudden temperature fluctuations, they can maintain a proper phase with the environment. Furthermore, under natural conditions, changes in both light and temperature during a diurnal cycle are coupled with warmth correlating with light. However, the genetic components of circadian clock that respond to these *zeitgebers* could be different (Boikoglou et al., 2011). This establishes the unique genetic identity of circadian clock in response to these major environmental stimuli.
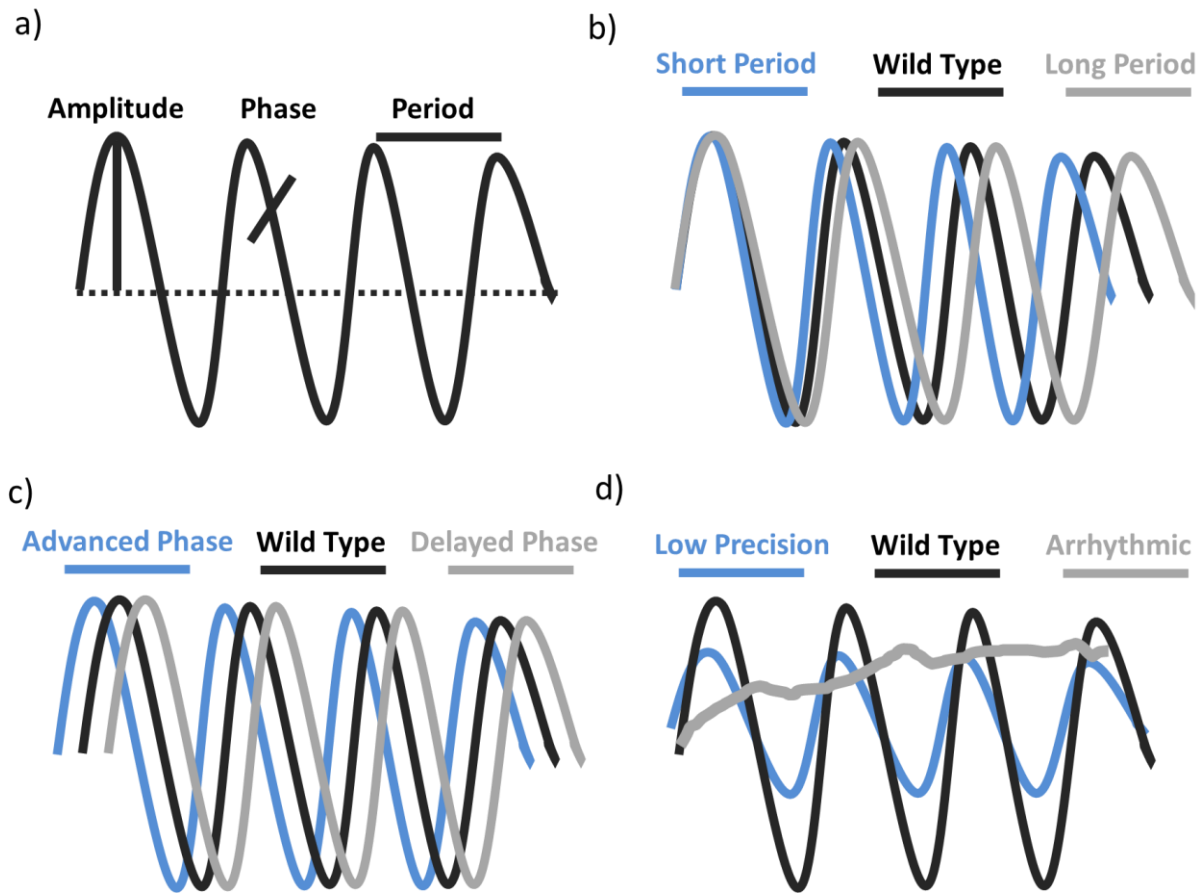
Figure 1.1: Properties of circadian oscillations.

(a) schematic diagram showing three main parameters of a circadian rhythm. Amplitude is half the difference between the highest and the lowest value of the oscillation. Phase refers to the timing of a physiological event within the circadian cycle. Often, it refers to the highest rhythm value within the oscillation (peak). Period represents the time to complete one oscillation, which is normally taken as a time difference between two consecutive peaks. Under diurnal conditions (or entrainment) these parameters normally remains similar for all genotypes, however, under constant conditions (free-running), they vary among different genotypes, depending upon the endogenous clock system. (b-d) shows the hypothetical Period (b) Phase (c) and Precision (d) of three genotypes under free-running conditions.

A well synchronized circadian system that accurately resets with the external environmental enhances fitness (Resco et al., 2009; Yerushalmi and Green, 2009). It has been observed that a plant with the internal-clock length matching with the external environment displayed increase survival and biomass accumulation, which is suggested to be the result of improved chlorophyll content and photosynthesis (Dodd et al., 2005; Yakir et al., 2007a; Yerushalmi et al., 2011). Further, optimized starch utilization during night and better water-use efficiency has also been reported to link with a robust clock (Graf et al., 2010; Takase et al., 2011). From these results, it was concluded that the circadian clock provides adaptive advantage to the plants. This was recently confirmed in a study where a heterogeneous F2 progeny segregating for short-period and long-period mutants were grown under short and long photo cycles (T-cycles 20h Vs. T-cycles 28 h). The seeds from these plants were harvested in these conditions and F3 progeny was tested for circadian periodicity. It was observed that the mean period of F3 plants obtain from F2 plants grown under T-cycle 28 h was longer than the F3 plants obtained from F2 plants grown under T-cycle 20. This showed that there is a strong selection pressure for circadian rhythms that are resonance with the environment (Yerushalmi et al., 2011). Thus, having a circadian clock that matches with the environment is adaptive and increases fitness. Probably this is the reason that circadian clocks are ubiquitous from cyanobacteria to mammals and regulates diverse physiological processes. In its simplest form, such a clock can be created with only one feedback loop and this is seen in *Ostreococcus tauri* (Troein et al., 2011). In complex eukaryotes, for instance, *Arabidopsis thaliana* (Arabidopsis), there is an interconnection of networks that are comprised of transcriptional-translational feedback loops (Zhang and Kay, 2010).

1.2 Genetic architecture of Arabidopsis circadian clock

In Arabidopsis, the circadian clock is proposed to be comprised by three inter-connected transcription-translation feedback loops (Figure 1.2). The central loop, called the core oscillator, was first proposed a decade ago. It is comprised by two MYB transcription

factors, CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) (Wang and Tobin, 1998), LATE ELONGATED HYPOCOTYL (LHY) (Schaffer et al., 1998), and a member of pseudo-response regulator gene family TIMING OF CAB EXPRESSION 1 (TOC1; also known as PRR1) (Alabadí et al., 2001). All these components were identified in genetic screens, and all displayed a short circadian period (Alabadí et al., 2001; Schaffer et al., 1998; Wang and Tobin, 1998). These components proved essential for proper oscillator activity because their constitutive over expressions resulted in arrhythmia (Schaffer et al., 1998; Wang and Tobin, 1998). This loop functions in a way that both CCA1 and LHY express in the morning and act to repress transcription of *TOC1* (Mizoguchi et al., 2002). The mechanism of this repression relies on changes in chromatin structure, favoring Histone 3 deacetylation at the *TOC1* promoter (Perales and Más, 2007). TOC1 in turn represses the expression of *CCA1* and *LHY* in the evening, by direct binding to their promoters via its CCT domain (Gendron et al., 2012).

Mathematical and computational modeling incorporated a morning and an evening loop with the core oscillator (Locke et al., 2006; Zeilinger et al., 2006). Many components of the morning loop have been experimentally verified. These include three TOC1-related PSUEDO-RESPONSE REGULATORS: PRR5, PPR7 and PPR9, whose expressions are positively regulated by CCA1 and LHY, whereas, in turn these components repress *CCA1* and *LHY* (Farre et al., 2005; Harmer and Kay, 2005; Salomé and McClung, 2005). Both repression and activation activities in this loop are a consequence of the binding of the respective protein to the promoters of their target genes (Nakamichi et al., 2010). Further, there is a dual function of PRR5 in the oscillator. First, it acts as a transcriptional repressor of *CCA1* and *LHY*, and secondly, it enhances TOC1 activity by promoting its phosphorylation, nuclear accumulation and recruitment to the nuclear foci (Wang et al., 2010).

The evening loop is more complicated and has yet not been fully defined. Mathematical modeling proposed that TOC1 represses the activity of an unknown factor 'Y' whereas, in turn, 'Y' activates TOC1 to close this loop. Partial 'Y' function was assigned to the nuclear protein GIGENTEA (GI), as its temporal expression pattern matches that of the predicted component (Fowler et al., 1999; Park et al., 1999). However, the relationship
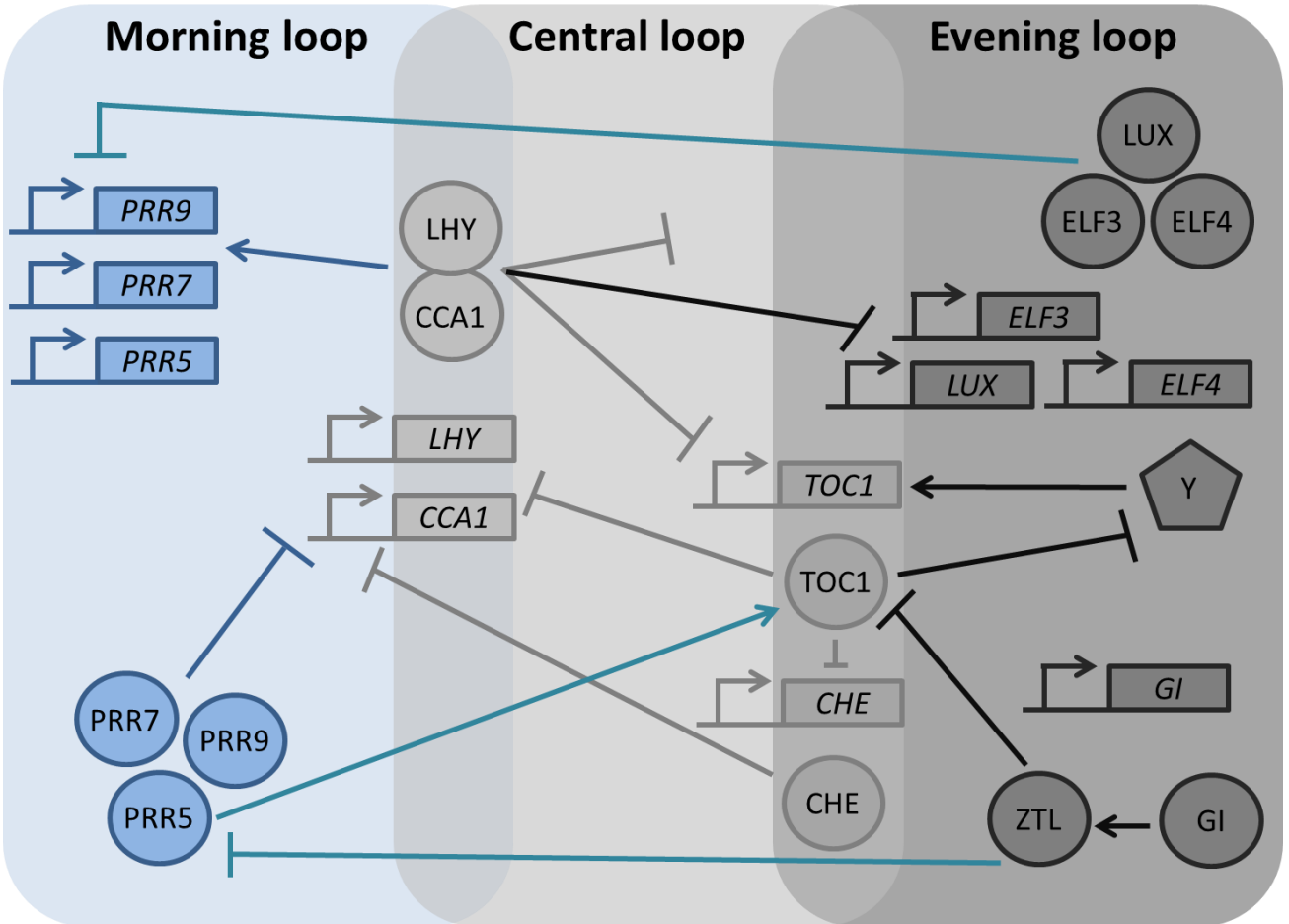
Figure 1.2: A simplified version of Arabidopsis circadian clock.

The Arabidopsis circadian clock consists of three interconnected feedback loops: a morning loop, a central loop and an evening loop. Genes are coded as rectangles and proteins are coded as circles. Gray lines represent regulatory connections among components of central loop. Dark blue lines represent activation and repression of components of morning loop. Black lines represent regulatory connections among components of evening loop. Regulatory connections among morning and evening loop components are shown with light blue lines. Arrows represent activation, whereas, perpendicular lines represent repression. Because the identity of 'Y' remains unknown, it is simply represented as a pentagon.

between TOC1 and GI is complex. Loss of function of *toc1-2* has little effect on *GI* expression (Ito et al., 2008; Martin-Tryon et al., 2007). Over expression of *TOC1* strongly represses *GI* (Makino et al., 2002). *gi* loss of function results in high levels of *TOC1* (Martin-Tryon et al., 2007). Further, double loss of function mutant of *toc1 gi* displays additive phenotypes (Martin-Tryon et al., 2007). Taken together these data suggest that it is probable that repression of *TOC1* to *GI* is indirect.

GI is a multifunctional protein of unknown biochemical activity. Contrary to its proposed function as a positive regulator of *TOC1* (if it works as Y), it has been shown that a major function of GI is to stabilize the F-box photoreceptor ZEITLUPE (ZTL), which mediates proteasomal decay of TOC1 protein in darkness (Kim et al., 2007). However, the formation of GI-ZTL complex requires light for its own stabilization. In darkness, ZTL also mediate the proteasomal degradation of PRR5, a positive regulator of TOC1 (Kiba et al., 2007). This further supports that GI is an indirect repressor of TOC1. Based on these contradictions, a recent model distinguishes GI from the unknown component Y, and rejects its role as a direct activator of TOC1 (Pokhilko et al., 2010).

1.3 Additional components of the clock: *ELF3* a case study

The circadian clock is likely to comprise by additional levels of regulation. A number of components have been identified whose function has not been fully assigned in the clock. These include CASEIN KINASE 2 (CK2) (Sugano et al., 1998), TIME FOR COFFEE (TIC) (Ding et al., 2007), FIONA 1 (FIO1) (Kim et al., 2008), XAP 5 CIRCADIAN TIME KEEPER (XCT), LIGHT REGULATED WD PROTEINS 1 and 2 (LWD1 AND LWD2) (Wu et al., 2008), JUMONJI (JMJ) (Jones et al., 2010), LUX ARRYTHMO (LUX) or PHYTOCLOCK1 (PCL1) (Hazen et al., 2005; Onai and Ishiura, 2005), BROTHER OF LUX ARRYTHMO (BOA) (Dai et al., 2011), EARLY FLOWERING 4 (ELF4) (Doyle et al., 2002) and EARLY FLOWERING 3 (ELF3) (Kim et al., 2005). Placing these genes into the clock has often been complicated by the mutant phenotype. Below ELF3 will be used as a case study to illustrate this point.

The *early flowering 3* (*elf3*) mutant of Arabidopsis was originally isolated in a screen

for altered flowering-time phenotype displaying a photoperiod-independent early flowering, when grown under either long day (LgD) or short day (StD) (Zagotta et al., 1996; Zagotta et al., 1992). Further characterization of *elf3* revealed other severe phenotypes including defects in clock-regulated leaf-movement rhythms, hypocotyl cell elongation and transcript accumulation of the *chlorophyll A/B–binding protein 2 gene* (*CAB2*) promoter. Moreover, the mutant displayed arrhythmic expression of *CAB2* promoter under free-running condition of constant light (LL) and in constant darkness (DD) (Hicks et al., 1996; Reed et al., 2000; Thines and Harmon, 2010). The *elf3* mutant thus appears to lack an oscillating clock.

ELF3 was reported as a nuclear-localized protein with no structural similarity to known functional domains (Herrero et al., 2012; Hicks et al., 2001; Liu et al., 2001). However, based on interactions studies, ELF3 was divided into three domains: an N-terminal domain (residues 1-261), a middle domain (residues 261-484) and a C-terminal domain (residues 485-695) (Herrero et al., 2012; Liu et al., 2001; Yu et al., 2008). Different domains of ELF3 had been found to interact with different proteins, which consolidated the previous hypothesis that ELF3 is a multi-functional protein (Hicks et al., 2001; Liu et al., 2001). Specifically, in the yeast-two-hybrid system (Y2H), the N-terminal domain of ELF3 interacted with both PHYTOCHROME B (PHYB) and CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1). Similarly, the middle-domain provided strong evidences of ELF3 interaction with both ELF4 and GI (Herrero et al., 2012; Liu et al., 2001; Yu et al., 2008). To date, the C-terminal domain has not been shown to interact with any protein. However, it has been demonstrated that ELF3 N-terminal domain binds to itself in the C-terminal region to form a homodimer (Liu et al., 2001). Furthermore, the nuclear localization signal (NLS) is proposed to be located in the C-terminal domain and presence of this domain had shown to be crucial for proper placement of ELF3 in the nucleus (Herrero et al., 2012).

ELF3 is required to be nuclear localized to initiate its repressive action in the clock (Herrero et al., 2012; Liu et al., 2001). In the nucleus ELF3 forms distinct nuclear bodies that co-localize with ELF4, GI and COP1 (Herrero et al., 2012; Yu et al., 2008). These nuclear

bodies have been proposed to be the point of interaction of binding partners (Mas et al., 2000). Consistently, genetic misplacement of ELF3 at a sub-cellular level resulted in aberrant clock function in the *elf3-12* mutant (Kolmos et al., 2011). One known component that modulates ELF3 sub-cellular localization is ELF4. ELF4 had been proposed to provide an anchor for ELF3 in the nucleus, and thus results in increased accumulation of ELF3 in the nucleus (Herrero et al., 2012; Kolmos et al., 2009). Genetically, ELF4 is downstream of ELF3 in mediating clock action as the *elf4-1* arrhythmic phenotype was rescued in *elf4-1 ELF3-OX* mutant. Conversely, *elf3-4 ELF4-OX* was arrhythmic (Herrero et al., 2012). Collectively, ELF3 is a multifunctional protein with multiple binding partners and redistribution of ELF3 to the nucleus is required for the clock to run.

Light is another factor that stabilizes ELF3 (Liu et al., 2001; Nusinow et al., 2011; Yu et al., 2008). An increased accumulation of ELF3 was detected when plants were transferred to LL after initial entrainment under LD. Consistently, ELF3 protein degraded rapidly when the plants were shifted from LD to DD (Liu et al., 2001; Yu et al., 2008). This rapid decay in DD had been explained by COP1-mediated proteasomal degradation of ELF3 (Yu et al., 2008). In this study it was also demonstrated that both COP1 and ELF3 play a pivotal role in the cyclic accumulation of GI protein by mediating its targeted degradation in the darkness (Yu et al., 2008). Notably, ELF3 mediates the proteasomal decay of GI, as in the absence of ELF3, GI level remained high in the darkness (Yu et al., 2008). Both ELF3 and GI are important in a wide range of plant physiological pathways, including circadian clock, hypocotyl length and flowering time. Genetically, GI is downstream to ELF3 in flowering time pathway, whereas, ELF3 is downstream to GI in control of hypocotyl length and circadian clock (Fowler et al., 1999; Huq et al., 2000; Lu et al., 2012; Martin-Tryon et al., 2007).

ELF3 was found to initiate repressor action by assembling into a complex with functionally required components. For instance, recent work identified that LUX act as a repressor of the morning gene *PRR9*, possibly by direct binding to the newly identified DNA binding site in the *PRR9* promoter (Helfer et al., 2011). A similar mechanism of action has

been discovered for ELF3, which associate in a time-dependent manner to the promoter of *PRR9* to regulate its expression (Dixon et al., 2011). Another study extended these results and showed that LUX, ELF3 and another component ELF4 makes a complex that associate to *PRR9* promoter to initiate repression (Herrero et al., 2012). Further, it was found that LUX-ELF3-ELF4 repressive complex regulates plant growth by directly binding to the promoters of the growth-related transcription factors PHYTOCHROMES INTERACTING FACTOR 4 and 5 (PIF4 and PIF5). An important finding in this was that ELF3 is essential for the formation of LUX-ELF3-ELF4 complex (Nusinow et al., 2011).

1.4 Mechanisms of light and temperature perception of circadian clock

Light and temperature are the main cues that synchronize the clock with the external environment. Comparatively, light is the stronger and well characterized input signal. In diurnal organisms such as Arabidopsis, light input to the circadian clock follows Aschoff's Rule: Increasing light intensities result in a shortening of the free-running period (Aschoff, 1979; Devlin and Kay, 2000). Light input to the clock is under the action of photoreceptors. Three types of photoreceptors have been identified. These include red light sensing PHYTOCHROMES (PHYA-PHYE), blue light responsive CRYPTOCHROMES (CRY1 and CRY2) and PHOTOTROPINS (Millar, 2004). Although, photoreceptors provide input to the clock, they are also a part of the output clock pathway as the transcript abundance of *phyb, cry1* and *cry2* was found to be rhythmic (Tóth et al., 2001). Further, the clock modulates its own sensitivity to light in a time-dependent manner in a process called gating (McWatters et al., 2000). *elf3* loss-of-function mutants displayed severe gating defect as strong phase shifts were observed when light pulses were given during night when the gate is normally closed (Covington et al., 2001; Kolmos et al., 2011) . Similar types of gating defects were also observed in the *tic* mutant (Ding et al., 2007; Hall et al., 2003). This establishes that both *ELF3* and *TIC* are part of light input to clock.

Other components that strongly respond to light signals include CCA1, LHY, PRR7

and PRR9. Light affects the post-transcriptional stability of *CCA1* transcript (Yakir et al., 2007b) as well as the translation of LHY (Kim et al., 2003). Mutations in *PRR7* and *PRR9* lengthen the period of *CCR2::LUC* under light, whereas, period remained unaltered in darkness (Farre et al., 2005). Moreover, post-translational stability of many of clock proteins is light dependent. Specifically, the protein accumulation of GI, ELF4 and ELF3 was found to be considerably increased under light, whereas, a rapid decrease was observed in darkness (Liu et al., 2001; Nusinow et al., 2011; Yu et al., 2008). Further, in the light, GI forms a complex with blue-light photoreceptor ZTL that promotes its accumulation. Conversely, dissociation of this complex occurs in the darkness (Kim et al., 2007). This light dependent mechanism explains why the ZTL protein is rhythmic while its mRNA is constitutively expressed (Somers et al., 2000).

Less well understood is the mechanism by which temperature entrains the clock. It is known that temperature cycles of as little as 4°C can entrain the Arabidopsis clock (McClung, 2006). However, the identity of primary temperature sensing mechanism is not yet established. Two components, PRR7 and PRR9, were found to be required for the clock to entrain to thermal cycles, as the double mutant *prr7prr9* failed to entrain to 22°C:12°C thermal cycles and displayed arrhythmia under constant conditions (Salomé and McClung, 2005). Interestingly, in another experiment, this double mutant could entrain to a different thermal cycle of 28°C:22°C (Salome et al., 2010), revealing that other factors are also involved in temperature input to the circadian clock.

A recent study dissected the differential behavior of photic and thermal perception of the clock using a quantitative-trait-loci approach. The results showed that periodicity response of a recombinant inbred population to photic versus thermal cues was different. Consequently, same and different QTLs were detected establishing the notion that circadian clock perceive both these inputs differently (Boikoglou et al., 2011).

1.5 Clock regulated flowering time mechanism in Arabidopsis

Circadian clock plays an important role in control of flowering time. The clock

ensures that plants produce seeds at the right time of the year and thus provides fitness advantages (Yerushalmi and Green, 2009). In Arabidopsis, the circadian clock regulates the induction of flowering time by measuring day length (de Montaigu et al., 2010). As Arabidopsis is facultative long-day plant, the clock ensures that important components for floral induction are maximally expressed under conductive long day (LgD) conditions. This is done by clock regulation of *CONSTANS (CO)* gene expression, and further, by monitoring its protein stability in day-length-dependent manner (Kobayashi and Weigel, 2007). Transcriptional and post-translational regulation of CO ensures that its target *FLOWERING LOCUS T (FT)* is specifically induced under LgD (Suarez-Lopez et al., 2001). FT protein is synthesized in leaves and then translocated to the shoot apical meristem where it interacts with the FLOWERING LOCUS D (FD) to initiate flowering (Wigge et al., 2005).

Clock regulated components FLAVIN KELCH F BOX (FKF1), GI, and ZTL ensure that CO peak of expression occurs before dusk, which is a prerequisite for floral induction under LgD. To achieve this, GI interacts with FKF1 under light, forms a complex that mediate the target degradation of CYCLING DOF FACTOR 1 (CDF1) and CDF2, two transcriptional repressors of CO (Fornara et al., 2009) and this leads to the release of repression of *CO* before dusk. This activates FT to trigger flowering. Besides CDFs, which are repressors of CO, it was recently identified that FLOWERING BHLH 1 (FBH1), FBH2, FBH3 and FBH4 induce FT expression by activating CO (Ito et al., 2012). The clock thus works at many levels to regulate the timing of flowering.

Transcriptional regulation of *CO* does not fully explain the CO mediated flowering-time mechanism, as *CO* mRNA is present under both LgD and short days (StD). Here post-translational stability of CO protein defines the dynamics of flowering time. In darkness, CO protein is unstable because a complex between SUPPRESSOR OF PHYA 1 (SPA1) and COP1 mediate its degradation by 26S proteasome. SPA1-COP1 complex is light sensitive that requires darkness to function (Laubinger et al., 2006; Valverde et al., 2004). In this way, CO protein produced in StD gets degraded before it can induce *FT* expression. In contrast, a peak of *CO* expression before darkness under LgD results in functional CO that induces *FT*

expression and hence flowering is initiated. Here photoreceptors plays important role by rescuing CO from the SPA1-COP1 mediated proteasomal decay (Zuo et al., 2011).

Several other clock-regulated transcription factors that affect FT expression independent of CO have also been identified. These include TEMPRANILLO 1 (TEM1), TEM2, TARGET OF EAT1 (TOE1), TOE2, SCHLAFMUTZE (SMZ), SCHNARCHZAPFEN (SNZ), SHORT VEGETATIVE PHASE (SVP), and FLOWERING LOCUS C (FLC). These all seem to function as FT repressors (Imaizumi, 2010). FLC and SVP are MADS-box transcription factors that directly repress FT expression in leaves (Li et al., 2008; Searle et al., 2006). CCA1 and LHY decrease SVP protein stability, possibly through protein–protein interaction via ELF3 under continuous light conditions (Yoshida et al., 2009). *FLC* works in vernalization pathway where FRIGIDA (FRI) promotes its expression to repress FT. The mechanism by which FRI regulates expression of *FLC* is still not fully understood, although it was recently shown that FRI protein interacts with the cap binding complex (CBC) through its two coiled coil domains, and that this interaction is essential for FRI function (Geraldo et al., 2009). After prolonged vernalization, *FLC* becomes epigenetically silenced and FT repression is released resulting vegetative to reproductive transition (Shindo et al., 2006).

1.6 Natural variation

Arabidopsis has a worldwide distribution. It grows in Europe, Asia, Africa, America, Japan and Australia (Alonso-Blanco and Koornneef, 2000). This covers an impressive latitudinal range from 68°N (North Scandinavia) to 0° (mountains of Tanzania and Kenya) and an extraordinary altitudinal range from sea level (Netherlands) to 4000 meters (Himalayan) (Koornneef et al., 2004). This implies that Arabidopsis has a huge capacity to adapt to ecological habitats, and thus, reflects the plastic capacity of its genome. It further manifest that these accessions are a colossal genetic resource that can be used to study complex ecological questions. To facilitate such studies, a collection of more than 2000 distinct Arabidopsis accessions from diverge geographical locations has been assembled over the past few years (Méndez-Vigo et al., 2011; Platt et al., 2010). Enormous variation

for several plant mechanisms has been detected in these accessions. These include developmental traits such as, flowering time (de Montaigu et al., 2010; Srikanth and Schmid, 2011), hypocotyl elongation (Coluccio et al., 2011), plant size (Jimenez-Gomez et al., 2010; Weinig et al., 2002), trichome density (Hilscher et al., 2009) and physiological traits such as seed dormancy (Alonso-Blanco et al., 2003), plant responses to light and hormone treatment and circadian-clock function (Boikoglou et al., 2011; Darrah et al., 2006). Based on natural variation studies in Arabidopsis, new data explaining genetic and molecular basis of these mechanisms, is being generated at a tremendous pace, probably because of the availability of new high-throughput phenotyping facilities and next generation sequencing platforms.

1.6.1 Tools to study natural variation: QTL mapping, a classical approach

Arabidopsis accessions vary in number of physiological and developmental traits. Most of these traits are quantitative. This means that the phenotypic trait values vary continuously in a segregating population. This depicts that the overall effect of the trait is determined by several loci (multigenic) instead of a single locus (monogenic). Further complex genetic and environmental interactions influence the trait performance. Thus the genotype at the loci underlying these cannot be directly known from the single phenotypic value of a plant determined by the various loci and the environment. This can only be inferred indirectly from linked marker loci. One of the most utilized methods to study quantitative traits is quantitative trait loci (QTL) mapping. QTL mapping comprises four basic steps that include (a) generation of suitable mapping populations (b) define their genotyping with markers covering complete genome (c) phenotype for the trait of interest and finally (d) statistical analysis to detect potential loci that associate with the trait under study.

1.6.1.1 Generation of mapping populations

In principal, any segregating population can be used for QTL mapping. However, there are certain pros and cons associated with each population. For instance, F2

populations can be generated in a very short time span, but because of their heterogeneity, such a population cannot be used multiple times, under several environmental conditions/replications. This dilemma can be removed by self-fertilizing the F2 population for several generations by single seed descent (SSD) until near homozygous mosaics are obtained. These lines are called recombinant inbred lines (RILs) (Figure 1.3).

RILs offer several advantages over F2. Primarily, as these lines are practically homozygous for every locus, they provide a homogenous genetic resource that can be maintained ''indefinitely'' by selfing, without changing the genetic identity of each line. In this way, genotypic map performed on a RIL becomes an ''eternal identity''. The RIL then can be used for phenotyping of several traits under many environmental conditions. Thus environmental effects can be tested many times on such a unique genetic identity. Further, the phenotypic values can be obtained under multiple replicates that reduce the environmental effect and thus, increase the power to detect QTLs.

Although RILs already contain more recombination events than F2, advanced-cross RILs can be made to further increase the recombination rates that will ultimately increase the mapping resolution. Back cross inbred lines are one such kind of lines that can be obtained by crossing the F1 population to one of the parent followed by selfing for several generations. Another type of advanced-intercrossed RILs are generated by crossing several F2 lines to each other before starting the inbreeding. In this way, several new recombination events are ''re-mixed'' and these are then fixed by selfing. Thus, RILs can provide a unique advantage as they are likely to reveal novel gene interactions that are not found in wild.

Multiple RIL populations generated in a systematic way, such that each RIL set contains one common parent, can greatly increase the power to detect QTLs. This process can be greatly facilitated by using a common genetic map, which can be obtained by genotyping all RIL populations with a common set of markers. Using phenotypic data of all RIL populations in a single QTL analysis will reveal novel QTL and QTL interactions, which cannot be detected by single bi-parental RIL analysis. This has been shown in maize, where

a joint QTL analysis of 6 connected populations revealed several novel QTLs and QTL interactions that otherwise was not appeared in the single RIL analysis (Blanc et al., 2006).

One principle drawback associated to RILs, is the time require to achieve certain level of homozygosity in the genome. Recently, new methods were introduced that allows the facile production of double haploid plants from recombinant population (Ravi and Chan, 2010; Seymour et al., 2012). This way lengthy process of inbreeding for several generations can be bypassed. However, the RILs generated will have fewer recombination events because of the elimination of normal meiotic events by self-fertilization.
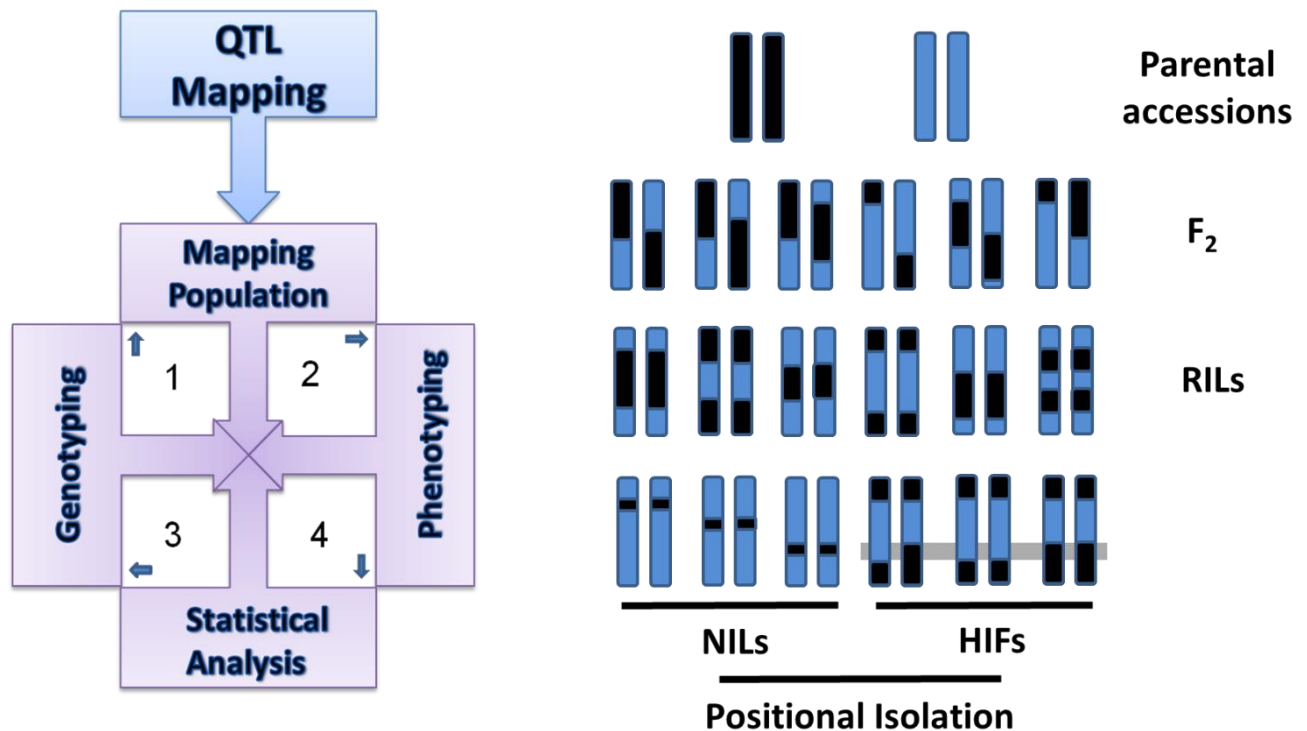
Figure 1.3: Schematic diagram showing steps involved in QTL mapping.

Four basic steps of QTL mapping are (i) generation of suitable mapping populations (ii) their genotyping with molecular markers (iii) and phenotyping for the trait of interest and (iv) the statistical analysis to find the association between the genotypic and phenotypic data that will disclose the region of the genome linked to the phenotypic variation (right panel). Right panel is showing the hypothetical structure of the two standard mapping populations F2 and RILs, along with the near isogenic lines, HIFs and NILs. HIFs and NILs can be used to generate large recombinant population that is used for the fine mapping/positional isolation.

1.6.1.2 Construction of linkage maps

Construction of linkage maps is one important step that defines the overall accuracy of QTL mapping (Lister and Dean, 1993). Since Arabidopsis has been accepted as a model organism, serious efforts have been employed to develop robust, efficient, and reproducible molecular markers that can be used to generate reliable genetics maps. The first map of F2/F3 population was generated by Koornneef M. in 1983 (Koornneef, 1983). This effort was followed by the generation of linkage map of the first RIL population using restriction fragment length polymorphism (RFLP) markers (Lister and Dean, 1993). Though, RFLP are co-dominant and are reliable markers, genotyping large sets of population with RFLP markers was laborious. Polymerase chain reaction (PCR) based markers that include random amplified polymorphic DNAs (RAPD), amplified fragment length polymorphism (AFLP), cleaved amplified polymorphic sequences (CAPS) and simple sequence length polymorphism (SSLP) substantially reduced the efforts required for genotyping and the amount of DNA required. Because RAPD were dominant markers and CAPS and AFLP require restriction digestion, SSLP have emerged as the simplest markers for use. Loudet et al. successfully demonstrated the usefulness of SSLP markers by generating the linkage map of the Bay X Sha RIL population using 38 SSLP markers (Loudet et al., 2002). With the availability of complete genome sequences of many of the Arabidopsis accessions, SSLP markers are joined with the frequent and robust InDels (insertions or deletions) and SNP (single nucleotide polymorphism) markers. Recently, these markers were used to generate high density maps of many RIL populations, and thus, confirmed the accuracy and efficiency of these markers (Giakountis et al., 2010; Kover et al., 2009; Păcurar et al., 2012; Salomé et al., 2011; Weigel, 2012).

1.6.1.3 QTL mapping: an overview of circadian clock and flowering time trait studied

Once linkage maps are generated and the phenotypic evaluation of the RILs is performed, the next step in the QTL mapping is to perform statistical analysis that detects the association of particular genomic region to the phenotypic variation observed in the

trait. This process detects regions of QTLs. In the past, QTL mapping has been successfully employed to identify loci affecting the circadian clock and flowering-time traits (Boikoglou et al., 2011; Giakountis et al., 2010). Two clock parameters studied as quantitative traits were phase and period of free-running oscillations (Darrah et al., 2006; Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999). In all of these studies, a common QTL at chromosome 1, co-localizing with GI was identified. Two of these studies validated the QTL identity in NILs (Edwards et al., 2005; Swarup et al., 1999), and a further study disclosed the polymorphisms in *GI* coding sequence in the parental accessions (Edwards et al., 2005). This advocated strongly that *GI* is the candidate gene for this region that affects speed of the clock under free-running conditions. However, a QTL detected on the same region for circadian phase, when characterized with NILs, eliminated *GI* as a candidate and supported the existence of a novel gene (Darrah et al., 2006).

Another locus that was commonly detected in QTL studies of clock behavior was the region at the bottom of chromosome 5 (Darrah et al., 2006; Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999). This region contained many circadian clock genes that include *PRR3*, *SIGNALING IN RED LIGHT REDUCED 1 (SRR1)*, *ZTL* and *TOC1*. One of the studies validated the QTL in NILs and with further sequence comparison of the candidate genes of the region concluded that *ZTL* is a plausible candidate for this locus (Edwards et al., 2005). Contrarily, a different study excluded *ZTL* as the candidate by phase analysis of the NILs, and with sequence analysis of candidates, proposed that *SRR1* and *PRR3* are possible candidates associated to this chromosome 5 QTL (Darrah et al., 2006). Nevertheless, even though the candidate genes might vary, the involvement of both chromosome 1 and chromosome 5 loci in natural variation of clock parameters is confirmed.

In all of the above mentioned studies, either the phase of *CAB2::LUC* (Darrah et al., 2006) or the free-running period of leaf movement was estimated after entraining the plants to Light-Dark (LD) cycles (Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999). As mentioned before, temperature can also entrain the circadian clock. A recent

study compared the free-running period of *CCR2::LUC* in a modified Cvi X Ler (CxL) population after LD and Warm Cool (WC) cycles. Depending upon the entrainment conditions, the individual lines behaved differently under constant conditions of light and temperature. Consistently, similar and different QTLs were detected after LD and WC. The previously detected QTL at chromosome 1, co-localizing with *GI*, was also detected in this population. However, this QTL was only detected after WC entrainment, but not after LD entrainment. Further, the bottom of the chromosome 5 was also found to be associated with the speed of oscillation, after either LD and WC entrainment (Boikoglou et al., 2011).

In a similar study, the comparative effect of photic (LD) and thermal (WC) entrainment on speed of circadian clock was observed in Bay x Sha (BxS) population that was modified with *CCR2::LUC* (Boikoglou, 2008). For this, a core set of 69 BXS RILs [as described in (Loudet et al., 2002)] were individually transformed with *CCR2::LUC* constructs and multiple transformants were selected. The free-running period was then estimated in T2 progenies after LD and WC entrainment. For that, the individual RILs were first entrained under the respective (LD or WC) entrainment conditions for 7 days and then the period was estimated under constant light and temperature. This ensured that period differences obtained for individual RILs were associated to the preceding entraining conditions. The linkage map for this BxS population was already available (Loudet et al., 2002). Therefore, QTL mapping was performed separately for LD and for WC period data. A QTL at chromosome 1 near the *GI* position was detected after LD. Two additional QTLs, one each at chromosome 2 and chromosome 4, were detected after both LD and WC entrainments. The chromosome 2 QTL was a highly significant main-effect QTL that explained ~25% of the variance for a P<0.001. The allelic interaction results demonstrated that the presence of Sha allele at this locus causes ~1 hour periodicity acceleration (Boikoglou, 2008). This QTL was identified at 63.7 cM, where *ELF3*, a known circadian clock and flowering-time regulator, is localized. In chapter 3 of this thesis, I will describe the positional isolation and functional characterization of this locus.

Unlike the circadian clock, where not many dedicated natural variations studies

have been implemented, flowering time mechanism has been intensively studied. The first natural allele identified in Arabidopsis using molecular markers was *FRI (Clark et al., 2007)*. The first QTL mapped in Arabidopsis were also ones controlling flowering time (Clarke et al., 1995). Since then, more than 100 genes in Arabidopsis have been found to affect flowering time (Srikanth and Schmid, 2011). The identification of two large effect loci that resulted in the isolation of two novel genes *FRI* and *FLC* strongly advocated the importance of natural variation studies (Johanson et al., 2000; Michaels and Amasino, 1999), as these alleles could not be found using laboratory accession because many of the laboratory accession carry loss-of-function alleles of either *FRI* or *FLC* (Koornneef et al., 2004). It has been estimated that without vernalization both *FRI* and *FLC* account for more than 75% of the flowering time variation among the accessions (Salomé et al., 2011). However, their effect is reduced after vernalization and other natural allele become more prominent. Notably, functionally distinct alleles of *FLOWERING LOCUS M* (*FLM*) or *MADS AFFECTING FLOWERING 1* and *PHYTOCHROME C* appear to be quite frequent in natural accessions (Balasubramanian et al., 2006; Caicedo et al., 2009).

A number of genes have been found to be QTL for flowering time. *FLM* is one such gene and it belongs to a family of transcription factors that include *FLC* and four related MAF proteins, MAF2 to MAF5 (Scortecci et al., 2003). This family of transcription factors is polymorphic between accessions and recently found to be involved in floral induction (Caicedo et al., 2009). Some rare alleles at *PHYA*, *PHYB*, *PHYD*, *CRY2*, *FT*, and *HUA2* also contribute in flowering time variation in natural accessions (Koornneef et al., 2004; Weigel, 2012). A recent study based on the flowering time analysis of 17 connected F2 populations revealed that the majority of flowering-time QTL cluster in as few as five genomic regions that include bottom of chromosome 1, middle part of chromosome 2, top of chromosome 4, and top and bottom of chromosome 5. However rare alleles were also identified that contributed to some extent to the flowering time variation observed (Salomé et al., 2011). Conclusively, the major part of natural variation in flowering time could be assigned to the five main genomic regions mentioned above.

1.6.1.4 From gene location to gene function

The availability of nearly limitless molecular markers and the technical development of high-throughput phenotyping facilities have made the QTL mapping quite facile. However, the identification of the molecular variation underlying QTLs and their functional characterization still remains as a major challenge (Koornneef et al., 2004; Mackay et al., 2009; Weigel and Nordborg, 2005). Many hundreds of genes normally reside in the original region defined by a QTL confidence interval. Moving from the QTL to the causal polymorphism is tedious; therefore, the identity of QTL must has to be validated. Isogenic lines in the form of heterogeneous inbred family (HIF) that carry a small portion of the genome segregates for the two parental alleles can be used in gene identification (Coluccio et al., 2011; Tuinstra et al., 1997). A disadvantage of this approach is that each HIF has a unique genome composition. Therefore, one cannot place several QTL in one common genetic background. Near isogenic lines (NILs) that carry only a small introgression of one parental allele at the QTL interval, while otherwise in a homogenous background of the other parent, can eliminate this drawback (Eshed and Zamir, 1995). These NILs can confirm the identity of the QTL. As described in the last section, several clock related QTLs have been validated using NILs (Darrah et al., 2006; Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999)

Both HIFs and NILs provide additional benefit as they can be subsequently used to generate a large population of plants that carry several recombination events in the region of the QTL. This population can then be used for positional cloning of the QTL with principles similar to map-based cloning of genes identified by induced forward screens. It basically consists in the identification of the closely linked recombination events, requiring the analysis of the large segregating progeny with the molecular markers covering the region of interest at high density. Phenotypic analysis is then carried out in parallel, and the region of confidence is gradually narrowed down. Usually the number of plant analyzed is defined by the resolution of fine mapping. For instance on, average, 1000 Arabidopsis gametes will provide a resolution of 50Kb, which will contain around ten genes (Koornneef

et al., 2004). If a clear candidate exists, sequence comparison of this gene should reveal polymorphism(s) between the parental accessions. Final proof for that gene underlying the QTL can be obtained from complementation experiments by plant transformation. The transfer of the either allele in a null background should show the predicted effect of the QTL, and if so, the gene can be designated as quantitative trait gene (QTG) (Koornneef et al., 2004).

Insights into the mechanisms of maintenance of quantitative-genetic variation and the evolution of quantitative traits require that we understand the causal molecular variants (quantitative trait nucleotides; QTN). SNPs that generate single encoded amino-acid substitutions and small insertions or deletions (InDels) are the most frequent QTNs that have been assigned to different flowering-time QTLs. For instance, a single encoded amino-acid substitution from Methionine to Valine in the *CRY2* coding region was associated to the *EDI* early flowering QTL (El-Din El-Assal et al., 2001). Similarly an encoded methionine to threonine substitution was associated to the differential response of Lm-2 accession to far-red signaling (Maloof et al., 2001). Natural deletions in *FRI*, *PHYD*, and *FLC* have been reported to result in early flowering (Aukerman et al., 1997; Gazzani et al., 2003). To date, no natural allele of a *bona fide* clock gene has been mapped to this level of resolution. However, recently, the *LIGHT5* QTL detected in hypocotyl length screen of B x S RIL population led to the identification of a zinc knuckle protein TZP that carries a pre-mature stop codon introduced by an 8bp insertion in the coding region. TZP acts downstream of the circadian clock and photoreceptor signaling pathways to directly control genes responsible for growth (Loudet et al., 2008).

1.6.2 Quantitative genetics in the era of next generation sequencing

In the past, most linkage mapping studies have been hampered by the lack of dense polymorphic molecular marker maps. The advent of efficient, high-throughput, and cost effective next-generation sequencing (NGS) methods have empowered individual laboratories to sequence parental strains of interest to rapidly identify marker loci and

design custom genotyping arrays (Mackay et al., 2009). Further assistance in genotyping was provided by the advancement in technologies that allowed the low coverage sequencing of a pool of DNA from several lines followed by genotyping calling and recombination breakpoint determination using a sliding window approach, with an average resolution of 40Kb (Huang et al., 2009). These methods along with the high quality reference genomes already available, the phenotype association to the polymorphisms is becoming remarkably easier.

With SNP-tiling arrays (AtSNPtile1) that provide probe sets for 248,584 SNPs covering the complete genome sequences of 20 natural accessions and Affymetrix genotyping array that provides, on average, one SNP every 500bp, genome wide association mapping (GWA) appeared next as an alternative for fine-mapping genomic regions to a resolution of ~10Kb (Bergelson and Roux, 2010). GWA mapping uses natural linkage disequilibrium (LD) to identify polymorphisms that are associated with phenotypic variation. Compared to RILs, the resolution to fine map is greatly enhanced in GWA, as it takes the advantage of recombination events that have accumulated over thousands of years. However, few limitations accompany GWA. One important limitation arises from population structure, that is, not all the accessions used in the study are equally distantly related to each other. An over-representation of a SNP in the population thus appears as a potential candidate even though it is not linked to the phenotype. Eventhough, powerful methods have been developed to correct of the population structure, the higher false positive rate remains to be associated with GWA. Thus traditional QTL mapping using RILs alongside GWA has been proposed as one solution to estimate the reliability of detected allele.

Chapter Two: **Materials and methods**

2.1 Materials

2.1.1 Plant material

2.1.1.1 Accessions

Following accessions were used as wild type or modified with promoter::reporter constructs (Table 2.1).

Table 2.1: Details of accession used

Accessions	Seed Source	Modifications (Promoter::LUC[+])	Integration method
Strand	M.Koornneef[a]	CCR2	Floral dipping
Ws-2	Doyle et al., 2002	CCR2, LHY	Floral dipping
Tnz-1	M.Koornneef[a]	CCR2	Floral dipping
Sha	Boikoglou 2008	CCR2	Floral dipping
Bay-0	Boikoglou 2008	CCR2, CCA1, LHY, TOC1, GI, PRR7, PRR9	Floral dipping

(a) Donated by Maarten Koornneef.

2.1.1.2 Recombinant Inbred Lines

Three independent RILs sets were generated by making pair-wise crosses among three parental accessions: Strand, Ws-2 and Tnz-1. These RIL sets included (1) Strand X Ws-2 (hereafter Strand X Ws) (2) Ws-2 X Tnz-1 (hereafter Ws X Tnz) and (3) Strand X Tnz-1 (hereafter Strand X Tnz) (Table 2.2-2.4). The detail procedure for generation of these RILs is described in methods section of this chapter (2.2.2). For *ELF3-Sha* positional isolation, modified BxS RILs used to detect chromosome 2 QTL were generated by Boikoglou 2008.

2.1.1.3 Heterogeneous Inbred Families, Near Isogenic Lines and mutants

HIFs and NIL-S were generated by successive backcrossing using Bay-0 as a recurrent parent. For HIFs, three independent RILs 57, 92 and 343 were selected from a modified Bay X Sha RIL population [(described in (Boikoglou, 2008)] and were

Table 2.2: Strand X Ws RIL set

RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name
1	F5-WNW-36F	51	F5-WNW-15A	101	F6-WN-3-III	151	F5-WNW-22B
2	F5-WNW-28E	52	F6-WN-18A-I	102	F5-WNW-39D	152	F6-WN-7A-II
3	F6-WN-22B-II	53	F5-WNW-13E	103	F5-WNW-5A	153	F6-WN-14-I
4	F5-WNW-16E	54	F5-WNW-4E	104	F5-WNW-35A	154	F5-WNW-1D
5	F6-WN-20B-I	55	F5-WNW-4D	105	F6-WN-14-II	155	F6-WN-16-III
6	F5-WNW-36C	56	F6-WN-2-II	106	F6-WN-10A-I	156	F5-WNW-55E
7	F5-WNW-28A	57	F5-WNW-25C	107	F5-WNW-13A	157	F5-WNW-25F
8	F6-WN-13B-I	58	F6-WN-5A-I	108	F5-WNW-30E	158	F5-WNW-23A
9	F5-WNW-29A	59	F6-WN-11-I	109	F5-WNW-17D	159	F6-WN-7B-I
10	F5-WNW-39F	60	F5-WNW-24D	110	F6-WN-13B-II	160	F5-WNW-38F
11	F5-WNW-36B	61	F5-WNW-17C	111	F5-WNW-18B	161	F6-WN-23-II
12	F6-WN-1-II	62	F5-WNW-28D	112	F6-WN-15-II	162	F5-WNW-41F
13	F6-WN-7B-II	63	F5-WNW-17E	113	F5-WNW-31D	163	F5-WNW-29D
14	F5-WNW-36E	64	F6-WN-15-I	114	F6-WN-15B-I	164	F5-WNW-13D
15	F5-WNW-00	65	F5-WNW-27B	115	F5-WNW-28E	165	F5-WNW-8C
16	F6-WN-22A	66	F5-WNW-11E	116	F5-WNW-20D	166	F6-WN-18A-II
17	F6-WN-3-II	67	F5-WNW-24F	117	F6-WN-4-I	167	F5-WNW-5E
18	F6-WN-24A-II	68	F5-WNW-16A	118	F5-WNW-14E	168	F5-WNW-38E
19	F5-WNW-36A	69	F5-WNW-6E	119	F6-WN-12-III	169	F6-WN-13A-I
20	F5-WNW-1B	70	F5-WNW-38B	120	F5-WNW-8A	170	F5-WNW-31F
21	F5-WNW-32A	71	F6-WN-16-II	121	F5-WNW-38A	171	F6-WN-5A-II
22	F5-WNW-39B	72	F5-WNW-23B	122	F5-WNW-24B	172	F613BX10B-BI
23	F6-WN-11-III	73	F5-WNW-19D	123	F5-WNW-33A	173	F5-WNW-22E
24	F5-WNW-30B	74	F5-WNW-18A	124	F5-WNW-30C	174	F6-WN-1-I
25	F5-WNW-37F	75	F5-WNW-30A	125	F5-WNW-39E	175	F5-WNW-5D
26	F6-WN-22A-I	76	F5-WNW-24E	126	F6-WN-5B-I	176	F5-WNW-39AI
27	F5-WNW-36D	77	F5-WNW-55C	127	F5-WNW-38D	177	F5-WNW-32E
28	F5-WNW-27E	78	F6-WN-22B-I	128	F5-WNW-11D	178	F5-WNW-22F
29	F6-WN-10B-I	79	F5-WNW-18E	129	F5-WNW-6D	179	F6-WN-8-I
30	F5-WNW-50B	80	F5-WNW-22D	130	F5-WNW-3A	180	F5-WNW-4C
31	F5-WNW-2B	81	F5-WNW-32B	131	F6-WN-24A-I	181	F5-WNW-33D
32	F6-WN-8-II	82	F6-WN-10A-II	132	F5-WNW-55B	182	F5-WNW-8D
33	F5-WNW-31C	83	F5-WNW-44B	133	F5-WNW-20C	183	F6-WN-12-I
34	F6-WN-18B-I	84	F5-WNW-18D	134	F5-WNW-25B	184	F5-WNW-21F
35	F5-WNW-3C	85	F5-WNW-40D	135	F5-WNW-11A	185	F5-WNW-27A
36	F5-WNW-35B	86	F6-WN-16-I	136	F5-WNW-27C	186	F6-WN-8-III
37	F6-WN-20A-II	87	F5-WNW-6B	137	F5-WNW-31B	187	F5-WNW-20B
38	F6-WN-24B-I	88	F5-WNW-19A	138	F5-WNW-55D	188	F5-WNW-1E
39	F5-WNW-35E	89	F5-WNW-27F	139	F5-WNW-19B	189	F5-WNW-16C
40	F5-WNW-39A	90	F5-WNW-5B	140	F5-WNW-15D	190	F5-WNW-1DI
41	F6-WN-3-I	91	F5-WNW-14A	141	F5-WNW-5C	191	F5-WNW-24DI
42	F5-WNW-37E	92	F5-WNW-26D	142	F6-WN-20A-I	192	F6-WN-23-I
43	F6-WN-10B-II	93	F5-WNW-23F	143	F5-WNW-26B	193	3X22B A-1
44	F5-WNW-20A	94	F5-WNW-11C	144	F6-WN-11-II	194	3X22B A-2
45	F5-WNW-35C	95	F6-WN-24B-II	145	F5-WNW-40A	195	3X22B B-2
46	F6-WN-13A-II	96	F5-WNW-14B	146	F5-WNW-14D	196	3X22B C-1
47	F5-WNW-1F	97	F5-WNW-22C	147	F6-WN-12-II	197	3X22B C-2
48	F5-WNW-17F	98	F5-WNW-32F	148	F5-WNW-41E	198	3X22B D-1
49	F5-WNW-21B	99	F6-WN-17-II	149	F6-WN-4-III	199	3X22B D-2
50	F5-WNW-6A	100	F6-WN-5B-II	150	F5-WNW-16B	200	3X22B E-1

Table 2.2 Strand X Ws RIL set (continue)

RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name
201	3X22B I-1	244	11x12 Db	287	17X10B Db	330	20x11 E-1
202	3X22B I-2	245	11x12Dc	288	17X10B I	331	20x11 E-2
203	3X22B II-1	246	11x15 Aa	289	17X10B II-1	332	20x11 E-3
204	3X22B III-1	247	11x15 Ab	290	17X10B II-2	333	20X13 A-1
205	3X22B III-2	248	12X5B Ab	291	17X22B Aa	334	20X13 B-1
206	3X22B IV-1	249	12X5B Ba	292	17X22B Ab	335	20X13 B-2
207	3X22B IV-2	250	12X5B Bb	293	17X22B Ba	336	22X4 A
208	5Ax21 B	251	12X5B Bc	294	17X22B Bb	337	22X4 B
209	5Bx21 A	252	12X5B Ca	295	17X22B Ca	338	22X4 C
210	5Bx21 C	253	12X5B Cb	296	17X22B Cb	339	22aX4 IV
211	7X1 Aa	254	12X5B D-1	297	17X22B Da	340	22x18 A
212	7X1 Ab	255	12X5B E-1	298	17X22B Db	341	22x18 B
213	7X1 Bb	256	12X5B E-2	299	19x10 Bb	342	22x18 C
214	7X1 Ca	257	13BX10B A-1	300	19x10 F-1	343	22x18 D
215	7X1 Cb	258	13BX10B A-2	301	19x10 F-2	344	22Ax20Ab
216	7X1 Da	259	13BX10B B-2	302	19x20 Aa	345	22Ax20Ac
217	9x10 B	260	13BX10B C-1	303	19x20 Ab	346	23X24 A-1
218	9x10 C	261	13BX10B C-2	304	19x20 Ba	347	23X24 A-2
219	9x10 D	262	13BX10B D-1	305	19x20 Ca	348	23X24 B-1
220	9x10 I	263	13BX10B D-2	306	19x20 Cb	349	23X24 B-2
221	9x10 II	264	13BX10B E-1	307	19x20 D-1	350	23X24 C-1
222	9x10 III	265	13BX10B E-2	308	19x20 D-2	351	23X24 C-2
223	10X15 Aa	266	15X16 Ba	309	19x20 E-1	352	24BX2 A-1
224	10X15 Ab	267	15X16 Bb	310	19x20 E-2	353	24BX2 A-2
225	10X15 Ac	268	15X16 Da	311	20x1 Ab	354	24BX2 B-1
226	10X15 B	269	15Bx18Ac	312	20x1 Ba	355	24BX2 B-2
227	10X15 C-2	270	15Bx21 D	313	20x1 Bb	356	24Bx20B A
228	10X15 D-2	271	15X24 A-2	314	20x1 Ca	357	24Bx20B C
229	11x5 Aa	272	15X24 B-1	315	20x1 Cb	358	24Bx20B D
230	11x5 Ba	273	15X24 B-2	316	20x1 Da	359	24Bx20B E
231	11x5 Bc	274	15X24 C-1	317	20x11 Ab		
232	11x5 C-a	275	15X24 C-2	318	20x11 Ad		
233	11x5 C-b	276	15X24 D-1	319	20x11 Ae		
234	11x5 Da	277	15X24 D-2	320	20x11 Ba		
235	11x5 Db	278	15X24 E-1	321	20x11 Bb		
236	11x5 II-1	279	15X24 E-2	322	20x11 Bc		
237	11x5 II-2	280	17X10B A-1	323	20x11 Bd		
238	11x12 Aa	281	17X10B A-2	324	20x11 C-1		
239	11x12 Ab	282	17X10B Ba	325	20x11 C-2		
240	11x12 Ba	283	17X10B Bb	326	20x11 Da		
241	11x12 Bb	284	17X10B C-1	327	20x11 Db		
242	11x12 Cb	285	17X10B C-2	328	20x11 Dc		
243	11x12 Da	286	17X10B Da	329	20x11 Dd		

Table 2.3: Strand X Tnz RIL set

RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name
1	F5-NT-1	32	F5-NT-37	63	F5-NT-78	94	F5-NT1T1-3
2	F5-NT-2	33	F5-NT-39	64	F5-NT-79	95	F5-NT1T1-5
3	F5-NT-3	34	F5-NT-40	65	F5-NT-80	96	F5-NT1T1-6
4	F5-NT-4	35	F5-NT-41	66	F5-NT-81	97	F5-NT1T1-7
5	F5-NT-6	36	F5-NT-42	67	F5-NT-82	98	F5-NT1T1-8
6	F5-NT-8	37	F5-NT-45	68	F5-NT-83	99	F5-NT1T1-9
7	F5-NT-10	38	F5-NT-46	69	F5-NT-84	100	F5-NT1T1-10
8	F5-NT-11	39	F5-NT-48	70	F5-NT-85	101	F5-NT1T1-11
9	F5-NT-12	40	F5-NT-49	71	F5-NT-86	102	F5-NT1T1-12
10	F5-NT-13	41	F5-NT-50	72	F5-NT-87	103	F5-NT1T1-13
11	F5-NT-14	42	F5-NT-51	73	F5-NT-88	104	F5-NT1T1-14
12	F5-NT-15	43	F5-NT-52	74	F5-NT-90	105	F5-NT1T1-15
13	F5-NT-16	44	F5-NT-54	75	F5-NT-92	106	F5-NT1T1-17
14	F5-NT-17	45	F5-NT-55	76	F5-NT-93	107	F5-NT1T1-18
15	F5-NT-18	46	F5-NT-56	77	F5-NT-94	108	F5-NT1T1-19
16	F5-NT-20	47	F5-NT-57	78	F5-NT-95	109	F5-NT1T1-21
17	F5-NT-21	48	F5-NT-58	79	F5-NT-96	110	F5-NT1T1-22
18	F5-NT-23	49	F5-NT-60	80	F5-NT1N-3	111	F5-NT1T1-23
19	F5-NT-24	50	F5-NT-61	81	F5-NT1N-5	112	F5-NT1T1-24
20	F5-NT-25	51	F5-NT-63	82	F5-NT1N-6	113	F5-NT1T1-27
21	F5-NT-26	52	F5-NT-67	83	F5-NT1N-7	114	F5-NT1T1-28
22	F5-NT-27	53	F5-NT-68	84	F5-NT1N-8	115	F5-NT1T1-31
23	F5-NT-28	54	F5-NT-69	85	F5-NT1N-11	116	F5-NT1T1-33
24	F5-NT-29	55	F5-NT-70	86	F5-NT1N-12	117	F5-NT1T1-35
25	F5-NT-30	56	F5-NT-71	87	F5-NT1N-14	118	F5-NT1T1-36
26	F5-NT-31	57	F5-NT-72	88	F5-NT1N-15	119	F5-NT1T1-37
27	F5-NT-32	58	F5-NT-73	89	F5-NT1N-17	120	F5-NT1T1-38
28	F5-NT-33	59	F5-NT-74	90	F5-NT1N-22	121	F5-NT1T1-39
29	F5-NT-34	60	F5-NT-75	91	F5-NT1N-23		
30	F5-NT-35	61	F5-NT-76	92	F5-NT1N-25		
31	F5-NT-36	62	F5-NT-77	93	F5-NT1T1-2		

backcrossed twice with Bay-0. From the BC_2F_2 generation, lines carrying the Bay allele (HIF-Bay) or Sha allele (HIF-Sha) at QTL interval were selected and designated as HIF 57, HIF 92 and HIF 343, respectively. To obtain NIL-S, HIF-343 was further backcrossed thrice to Bay while selecting for Sha interval at the QTL confidence interval. BC_5F_2 NIL was used for the physiological experiments. Different *CCG::LUC* (clock control genes promoter fused with luciferase) constructs in NIL-S were obtained by crossing NIL-S thrice to the Bay expressing these constructs (Table 2.5). Similarly *elf3-1* and *elf3-4* mutants were generated by successive backcrossing of these mutants with Bay harboring *CCG::LUC* (Table 2.5). After each backcross, *elf3-1* and *elf3-4* mutants were selected based on their long-hypocotyl phenotype and were confirmed for the *CCG::LUC* transgene by imaging.

Table 2.4: Ws X Tnz RIL set

RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name	RIL No.	RIL Name
1	F6-WT1W-1	21	F6-WT1W-22	41	F6-T1W-8	61	F6-T1W-34
2	F6-WT1W-3	22	F6-WT1W-23	42	F6-T1W-11	62	F6-T1W-35
3	F6-WT1W-4	23	F6-WT1W-24	43	F6-T1W-12	63	F6-T1W-36
4	F6-WT1W-5	24	F6-WT1W-25	44	F6-T1W-13	64	F6-T1W-37
5	F6-WT1W-6	25	F6-WT1W-26	45	F6-T1W-14	65	F6-T1W-41
6	F6-WT1W-7	26	F6-WT1W-27	46	F6-T1W-15	66	F6-T1W-42
7	F6-WT1W-8	27	F6-WT1W-28	47	F6-T1W-16	67	F6-T1W-43
8	F6-WT1W-9	28	F6-WT1W-29	48	F6-T1W-17	68	F6-T1W-44
9	F6-WT1W-10	29	F6-WT1W-30	49	F6-T1W-18	69	F6-T1W-45
10	F6-WT1W-11	30	F6-WT1W-31	50	F6-T1W-19	70	F6-T1W-48
11	F6-WT1W-12	31	F6-WT1W-32	51	F6-T1W-20	71	F6-T1W-15
12	F6-WT1W-13	32	F6-WT1W-33	52	F6-T1W-21	72	F6-T1W-1
13	F6-WT1W-14	33	F6-WT1W-34	53	F6-T1W-22	73	F6-T1W-9
14	F6-WT1W-15	34	F6-T1W-1	54	F6-T1W-23	74	F6-T1W-17
15	F6-WT1W-16	35	F6-T1W-2	55	F6-T1W-26	75	F6-T1W-19
16	F6-WT1W-17	36	F6-T1W-3	56	F6-T1W-27	76	F6-T1W-25
17	F6-WT1W-18	37	F6-T1W-4	57	F6-T1W-28	77	F6-T1W-30
18	F6-WT1W-19	38	F6-T1W-5	58	F6-T1W-31	78	F6-T1W-37
19	F6-WT1W-20	39	F6-T1W-6	59	F6-T1W-32		
20	F6-WT1W-21	40	F6-T1W-7	60	F6-T1W-33		

Table 2.5: Mutants and NIL-S with different luciferase markers.

LUC marker	Genetic Background (Bay)	Selection marker	Integration method
CCR2::LUC+	HIFs[a], NIL-S, *elf3-1*, *elf3-4*	Hygromycin (*HPT* gene)	Crossing
CCA1::LUC+	NIL-S, *elf3-1*, *elf3-4*	Hygromycin	Crossing
LHY::LUC+	NIL-S, *elf3-1*, *elf3-4*	Hygromycin	Crossing
GI::LUC+	NIL-S, *elf3-1*, *elf3-4*	PPT (Bar gene)	Crossing
TOC1::LUC+	NIL-S, *elf3-1*, *elf3-4*	Gentamicin (*aacC1* gene)	Crossing
PRR7::LUC+	NIL-S, *elf3-1*, *elf3-4*	Gentamicin	Crossing
PRR9::LUC+	NIL-S, *elf3-1*, *elf3-4*	Gentamicin	Crossing

(a) HIFs also contained few genomic regions of Sha besides QTL interval.

2.1.1.4 *ELF3-Sha* and *ELF3-Bay* Transgenic lines

Transgenic lines used for *ELF3* complementation were obtained by floral dipping (Table 2.6) (Davis et al., 2009). Complete procedure for this is described in the methods section of this chapter (2.2.3).

Table 2.6: *ELF3* transgenic lines

T-DNA insertion	Genetic Background (Ws-2)	Selection Marker	Integration method
ELF3-Bay	*elf 3-4 LHY::LUC*	Kanamycin (*nptII*)	Floral dipping
ELF3-A362V	*elf 3-4 LHY::LUC*	Kanamycin (*nptII*)	Floral dipping
ELF3-Bay::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
ELF3-Sha::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
SpSc::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
SpBc::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
SpSv2a::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
SpBa2v::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
BpSc::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
BpBc::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
BpSv2a::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping
BpBa2v::YFP	*elf 3-4 LHY::LUC, CCR2::LUC*	Kanamycin (*nptII*)	Floral dipping

2.1.2 Oligonucleotides

2.1.2.1 Primers used for genotyping of RILs.

Complete list of primers used for RILs genotyping is given in appendix 1.

Table 2.7: Primers used for genotyping of luciferase marker lines.

Gene	Sequence 5´-> 3´	Gene	Sequence 5´-> 3´
CCA1-f	CCATTTCCGTAGCTTCTGGT	*PRR7-f*	CGTCGTCGTGGCTATCTACA
LHY-f	TGTGGACCACCACACTCACT	*PRR9-f*	TGTGTGTTTTAGGCTTTGTTGG
TOC1-f	CAAGATGACGTGGCCTTTTC	*CCR2-f*	ATCGTCGTCTTTCCATTTGC
GI-f	TTCGCATTGTGATTCCAAAA	*LUC-r**	GCCTTATGCAGTTGCTCTCC

* LUC-r was used as a reverse primer for all listed markers

Table 2.8: Primers used for screening of recombinants and genotyping of HIFs and NIL.

Type	Oligonucleotide Name and Sequence		Restriction Enzyme
SSLP	elf3-51		
5´->3´ f	ACTTGATGAAAAACAGATCCAAGAAAAC		N/A*
5´->3´ r	AACTTGTCTCAAAATCCTTGAATCTCTC		
SSLP	elf3-53		
5´->3´ f	AGAGACAGCTCCCTGAAGAGATGAG		N/A*
5´->3´ r	TTAAACAAAATTACAAAACTGCCATTTC		
SSLP	elf3-54		
5´->3´ f	TTTCAGTTATTGAGATTGACTTAATTTAGTTTT		N/A*
5´->3´ r	CGGAACCTAGGTGTTAGCACATTATTTA		
SSLP	at Elf3		
5´->3´ f	ATGATGCCCACCATAATGAACCCATATTG		N/A*
5´->3´ r	AAAGGACTTGCTACCAGAGATTCCCTGTG		
CAPS	elf1001L		
5´->3´ f	AAAGACTGACCCAAGAGGATAATG		BamHI
5´->3´ r	AGAAGAGTGTGAAGAGAGCCAAAT		
CAPS	elf1001R		
5´->3´ f	TCACGATTTTGACATTTTTAATGG		ApoI
5´->3´ r	AAGTGGTGATGGAATTAGATGAGG		
dCAPS	U15elf60		
5´->3´ f	AAGGTTCGGATCCGTATATCT		PstI
5´->3´ r	CTTGAATGGGCCACAAAGAT		
dCAPS	U21elf40		
5´->3´ f	ATGAACCCACAAACTCAGAAGC		AluI
5´->3´ r	CGTCTCTTCCAGATCCATTCTC		
dCAPS	U25elf20		
5´->3´ f	TGGACAACATTCACTGCCTGCA		PstI
5´->3´ r	GCACCCACCACTTGAAAAAT		
dCAPS	U35elf60		
5´->3´ f	GATGGGATCTGGAGGACTCA		XmnI
5´->3´ r	AAGTCAATGATGAATTCGTGAGA		
dCAPS	U39elf40		
5´->3´ f	GTACTCTCTCTGGCTATCATGG		HaeIII
5´->3´ r	GCACAGTCTCACAAAAGAGCA		
dCAPS	U43elf20		
5´->3´ f	TGCTTCGAAATCTTCTTCTTCTC		HinfI
5´->3´ r	GACTCTTTTCCCGACAGCATAC		
dCAPS	U51elf5		
5´->3´ f	ATCATTCACAAGAGGGTCACTG		PstI
5´->3´ r	TTACTTGGCAATGGCTTTCC		
dCAPS	U65elf8		
5´->3´ f	GTCCAATACGAGTTTGTTTTTA		MseI
5´->3´ r	GGAAGTAAACTTGTAGAGCCATCA		

* SSLP markers

Table 2.9: Primers used for *ELF3* transgenic lines and site directed mutagenesis.

Line	Oligonucleotide Sequence	Restriction
ELF3-Bay		**Enzyme**
5′ -> 3′ f 5 -> 3′ r	GCGCGCCCCGGGAAAAACCCAATAAAAACCACGATCCATTTT GCGCGCCCTGCAGGTCTTCGTCTCCGCCCTATATTAATCGATGT	Xma1/Sbf1
ELF3-A362V point mutation		
5′-> 3′ f 5′-> 3′ r	TAATCAGGTTCAAAAACTTATTG**T**TGCATCACCGGATCTCTTGCTCG CGAGCAAGAGATCCGGTGATGCA**A**CAATAAGTTTTTGAACCTGATTA	N/A
ELF3-V362A point mutation		
5′-> 3′ f 5′-> 3′ r	TAATCAGGTTCAAAAACTTATTG**C**TGCATCACCGGATCTCTTGCTCG CGAGCAAGAGATCCGGTGATGCG**G**CAATAAGTTTTTGAACCTGATTA	N/A
ELF3 promoter gateway (pDONR4-1R)		
5′-> 3′ f 5′-> 3′ r	GGGGACAACTTTGTATAGAAAAGTTGCTAAAAACCCAATAAAAACCAC GGGGACTGCTTTTTTGTACAAACTTGCCACTCACAATTCACAACC	N/A
ELF3 coding gateway (pDONR201)		
5′-> 3′ f 5′-> 3′ r	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAAGAGAGGGAAAGAT GGGGACCACTTTGTACAAGAAAGCTGGGTAAGGCTTAGAGGAGTCATA	N/A

N/A, not applicable, the red letters indicate the induced nucleotide change

Table 2.10: Primers used for qRT-PCR.

Gene	Oligonucleotide Sequence
CCA1	
5′->3′ f 5′->3′ r	TCTGTGTCTGACGAGGGTCGAATT ACTTTGCGGCAATACCTCTCTGG
LHY	
5′->3′ f 5′->3′ r	CAACAGCAACAACAATGCAACTAC AGAGAGCCTGAAACGCTATACGA
TOC1	
5′->3′ f 5′->3′ r	ATCTTCGCAGAATCCCTGTGATA GCACCTAGCTTCAAGCACTTTACA
GI	
5′->3′ f 5′->3′ r	CTGTCTTTCTCCGTTGTTTCACTGT TCATTCCGTTCTTCTCTGTTGTTGG
PRR7	
5′->3′ f 5′->3′ r	TGAAAGTTGGAAAAGGACCA GTTCCACGTGCATTAGCTCT
PRR9	
5′->3′ f 5′->3′ r	GCACAGAGAAACCAAAGGAA CTTTCACTCGAGGACGTTGT
ELF3	
5′->3′ f 5′->3′ r	GATGCCCACCATAATGAACC TTGCTCGCGGATAAGACTTT
ELF4	
5′->3′ f 5′->3′ r	CGACAATCACCAATCGAGAATG AATGTTTCCGTTGAGTTCTTGAATC

2.1.3 List of Reagents used

Table 2.11: Reagents

Tris (hydroxymethyl) aminomethane Hydrochloride, Tris HCl (Roth, #5429.3)	Bromophenol blue (Sigma, #47522)
Ethylenediaminetetraacetic acid, EDTA (Merck, #944)	Glycerol (Roth, #7530.1)
KLORIX®, commercial sodium hypochlorite solution	Ethidium bromide (Sigma; #46067)
Triton-X100 (Roth, #3051)	Glycine (Roth, #3908.2)
Murashige and Skoog media, MS (Sigma, #M5524 and Duchefa, #M0221)	Formaldehyde (Merck, #1.040003)
2-(N-morpholino)ethanesulfonic acid, MES (Duchefa, #M1503)	Hygromycin (Duchefa, #H0192)
Sucrose (Roth, #4621)	D-Luciferin (Synchem, #S039)
Phytoagar (Duchefa, #P0001)	Na_2HPO_4 (Sigma, #S0876)
Bacto-tryptone (BD, #211705)	NaH_2PO_4 (Merck, #1.06346)
Yeast extract (BD, #212750)	Magnesium sulphate, $MgSO_4$ (Duchefa, #110513)
Sodium chloride, NaCl (Mecrk, #1.37017)	Phenol /Chloroform (Roth, #A156.1)
Bactoagar (BD, #214040)	Chloroform (Merck, #1.02445)
Beef extract (BD, #212303)	Isopropanol (Appli. Chem., #A0900)
Gentamicin sulfate (Sigma, #G-3632)	Ethanol (J.T.Baker, #8006)
Carbenicillin (Sigma, #C-1389)	Sodium Acetate (Merck, #1.06268)
Kanamycin sulfate (Duchefa, #K4378)	Lithium chloride (Li Cl) (Roth, #3739.1)
Chloramphenicol (Sigma, #C-0378)	Sodium deoxycholate (Fluka, #30970)
Rifampicin (Sigma, #83907)	Silwet L-77 (Lehle seeds, #Vis-02)
Spectinomycin (Sigma, #S-9007)	Lithium Acetate (Sigma, #L-5750)
Streptomycin (Sigma, #S-9137)	Polyethylene glycol, PEG 3350 (Sigma, # P-3640)
DL-Phosphinothricin, PPT (Duchefa, #P0159)	Dimethyl sulfoxide, DMSO (Sigma, #P8418)

2.1.4 Media

2.1.4.1 Plant growth media

Muraskige and Skoog Basal Salt (MS), containing 3% sucrose (MS3)

1 liter MS3:

∗ 4.4g Muraskige and Skoog Basal Salt (MS)

∗ 30g sucrose

∗ 0.5g 2-(N-morpholino) ethanesulfonic acid (MES)

∗ 15g/L Phytoagar

∗ Adjust pH to 5.7 with potassium hydroxide (KOH)

∗ Autoclave at 120°C for 20 minutes

∗ Add appropriate antibiotics for selection.

This media was used for all plant growth in all luciferase imaging experiments, unless otherwise stated

¼ Muraskige and Skoog Basal Salt (MS) (MS0), liquid medium

1 liter ¼ MS0:

∗ 1.1g of Muraskige and Skoog Basal Salt (MS)

∗ 0.5g MES

∗ Adjust pH to 5.7 with KOH

∗ Autoclave at 120°C for 20 minutes

∗ Add appropriate antibiotics for selection.

This media was used for growing RILs for selfing by single seed descent (SSD) method.

2.1.4.2 Bacterial growth media

Luria-Bertani (LB)

1 liter LB:

* 5g yeast extract

* 10g bactotryptone

* 10g sodium chloride (NaCl)

* Adjust pH to 7.5 with KOH

* Autoclave in 120°C for 20 minutes

* Add appropriate antibiotics.

This media was used for growth of E. coli strain DH5α.

YEBS

1 liter YEBS:

* 1g yeast extract

* 5g beef extract

* 5g peptone

* 5g sucrose

* 0.5g MgSO4

* Adjust pH to 7.0 with sodium hydroxide (NaOH)

* Autoclave in 120°C for 20 minutes

* Add appropriate antibiotics.

This media was used for growth of Agrobacterium tumefaciens strain ABI.

2.1.4.3 Seed Sterilization reagents

Bleach solution

100ml bleach:

* 33ml of Klorix Bleach (commercial sodium hypochlorite solution)

* 20μL Triton X-100

Agar Water

100ml:

* 0.01g of phytoagar

2.1.4.4 Reagents and materials used for luciferase imaging

0.1M Triphosphate Buffer for Luciferin

200ml:

* 3.56g of Na2HPO4

* 2.76g of NaH2PO4

* Adjust pH to 8.0 with NaH2PO4

50mM Luciferin stock solution

50mM Luciferin:

* Dissolve 1g of firefly D-luciferin

* In 71.3mL of 0.1M of the triphosphate Buffer

* Aliquot and Store at -80°C

0.01% Triton X-100

1 liter:

* Add 100µL of Triton X-100 liquid solution

* in one liter of distilled water

5mM Luciferin working solution

* Add 1.5mL of luciferin stock (50mM) into 13.5mL of Triton-X solution

* Filter sterilize

Materials used for Imaging

- ∗ Packard Top Count Scintillation Counter

- ∗ Sterile OptiPlate-96F (Black, 96-well, pinch bar design) (6005270)

- ∗ MS agar containing 3% sucrose media

- ∗ Light-Emitting Diodes (LEDs) panel with Red, Blue, and Far-red LEDs (MD ELECTRONICS, UK) (This panel was used to provide continuous light on the Top Count for periodicity assays)

- ∗ Custom made reflector plates used to mirror light from the LEDs to the plants (used for periodicity assays)

- ∗ Top Seal Microplate Press-On Adhesive Sealing Film (Packard 6005185)

2.1.5 Software and internet resources

Arabidopsis sequence database, search and alignment websites

TAIR: http://www.arabidopsis.org

NCBI: http://www.ncbi.nlm.nih.gov/

Polymorph: http://polymorph-clark20.weigelworld.org/

Genome visualization software

IGV 2.0.10: http://www.broadinstitute.org/igv/home

Primer selection and designing tools

AMP: http://amp.genomics.org.cn/

Primer 3.0: http://frodo.wi.mit.edu/primer3/

dCAPS finder 2.0: http://helix.wustl.edu/dcaps/dcaps.html

Software for the calculation of genetics linkage maps and QTL detection

JoinMap® 4:   http://www.kyazma.nl/index.php/mc.JoinMap

MapQTL® 6:   http://www.kyazma.nl/index.php/mc.MapQTL/

Statistical Analysis

SPSS 17.0: http://www.spss.co.in/

2.2 Methods

2.2.1 Selection of parental accessions for RILs

Based on geographical location and climatic differences, three Arabidopsis accessions were selected: Strand, Ws-2 and Tnz-1. These accessions grow at broadly distant geographical coordinates, and under contrasting environmental conditions (Table 2.12). Specifically, the latitudinal coordinates for Strand, WS-2 and Tnz-1 were 68.7°N, 51°N and 2.9°S, respectively. Light and temperature are two primary environmental factors that follow a considerably different yearly pattern at these locations. Notably, day length remains almost the same throughout the year near the equator where Tnz-1 originated. Contrarily, Strand and Ws-2 being northern accessions are derived from locations with extreme yearly day-length differences of ~20 hours (Table 2.13). Similarly to light, temperature variations present at the origin locations of these accessions is also considerably different. The northern accessions Strand and Ws experience a temperature difference of around 40°C throughout a year, Tnz-1 being an equator accession is from a region where only a mild temperature variation of ~18°C is present in a given year a year (Table 2.13).

Table 2.12: Geographical position of the parental accessions.

Accession	Country	Place of origin	Latitude (°N)	Longitude (°E)	Altitude (m)
Strand	Norway	Strand	68.7	15.45	40
Ws-2	Russia	Vasil'yevskiy	51.51	30	100-200
Tnz-1	Tanzania	Kitumbeine Forest Reserve	-2.9	36.23	2441

As light and temperature are two primary input factors that defines the oscillation properties of circadian clock, and also determine the flowering-time architecture (Boikoglou et al., 2011; de Montaigu et al., 2010; Eckardt, 2005; Giakountis et al., 2010; Millar, 2004), one could expect variation in genetic architecture of these accession, for these physiological mechanisms, because of adaptation to contrasting environmental conditions. Hence, I reasoned that these accessions would serve as a valuable genetic resource for the investigation of circadian-clock mediated flowering-time pathways.

Table 2.13: Climatic information of parental accessions growth habitat.

	Day length (h)		Summer		Winter	
Accession	**Max**	**Min**	**Warmest month**	**Avg. Temp. (°C)**	**Coldest month**	**Avg. Temp. (°C)**
Strand	22	2	July	14.5	February	-5.4
Ws-2	17	5	July	24.8	January	-32
Tnz-1	13.5	11	February	23.1	July	4.7

2.2.2 Generation of RILs

Transformation of parental accessions Strand, Ws-2 and Tnz-1 with *CCR2::LUC* promoter::reporter construct was the first step in the generation of RIL sets. After initial transformation, a single 'T$_1$' transformant for each accession expressing *CCR2::LUC* was selected and backcrossed to the respective parent (wild type). This ''cleared'' the genome for multiple transgene insertions and residual mutations. A single insert line for each parent was then isolated and employed in a full diallel crossing strategy (Blanc et al., 2006; Rebai and Goffinet, 1993), to generate three independent RIL sets. For that, pair-wise crosses were made among all parental accessions in all possible combinations (Figure 2.1). The resultant RIL sets obtained were (i) Strand X Ws (ii) Strand X Tnz, and (iii) Ws X Tnz.
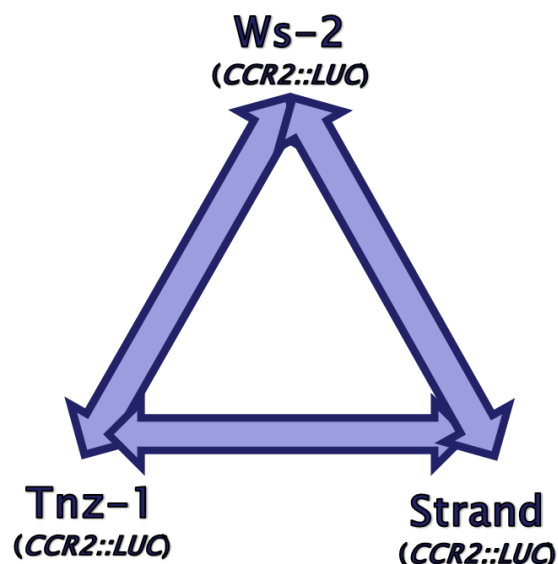
Figure 2.1: A triangle crossing strategy used to generate RILs.

Each of these RIL sets was further comprised of various sub-populations (Table 2.14). The sub-populations were categorized based on the crossing approach followed. The crossing methods used were (i) backcross (BC), (ii) F_2, and (iii) F_2 inter-matting (F_2X) (Kover et al., 2009; Lee et al., 2002). The detail procedure adopted to generate RIL sets based on these crossing methods is as follows (Figure 2.2):

BC: in BC sub-population, the parent one (P1-CCR2: pollen donar parent harboring *CCR2::LUC*) was crossed to parent two (P2-WT: wild type without *CCR2::LUC*) to obtain F_1 (harboring *CCR2::LUC*). Many of these F_1 plants were then backcrossed to parent two (P2-WT: wild type without *CCR2::LUC*) that resulted in a BC_1 population. The progeny of BC_1 produced BC_1F_2 plants, from those several lines were selected and self-fertilized for 3-4 generation by SSD to obtain BC_1F_5 or BC_1F_6 lines. After each round of selfing, the plants were confirmed for the presence of *CCR2::LUC* by selection on antibiotics and by imaging for bioluminescence. The resultant BC_1F_5 or BC_1F_6 lines were designated as RILs because of their expected genome homozygosity of ~95%. Sub-populations WSW, ST1T1, ST1S, and WT1W were generated following this crossing scheme (Table 2.14, Figure 2.2)

Table 2.14: Details of sub-population of RIL sets.

RIL Set	Sub-populations		
	Backcross(BC)	**F_2**	**F_2X**
Strand X Ws	WSW	WS	35WS
Strand X Tnz	ST1T1, ST1S	ST1	--
Ws X Tnz	WT1W	T1W	--

F_2: in F_2 sub-population, the parent one (P1-CCR2: pollen donar parent harboring *CCR2::LUC*) was crossed to parent two (P2-WT: wild type without *CCR2::LUC*) to obtain F_1 (harboring *CCR2::LUC*). Many of these F_1 plants were then selfed to obtain F2 segregating progeny. Several F2 lines were selected and self-fertilized for 3-4 generation by SSD with selection for *CCR2::LUC*. The resultant F_5 or F_6 lines were designated as RILs. The sub-populations WS, ST1 and T1W were generated using this crossing scheme (Table 2.14, Figure 2.2).

F_2X: in F_2X sub-population, the parent one (P1-CCR2: pollen donar parent harboring *CCR2::LUC*) was crossed to parent two (P2-WT: wild type without *CCR2::LUC*) to obtain F_1 (harboring *CCR2::LUC*). Many of these F_1 plants were selfed to obtain F2 segregating progeny. Several F2 lines were then randomly selected and crosses were made between them to create their F3F1 progeny. These F3F1 progeny lines were further self-fertilized in SSD strategy. The resultant population was named as F_3F_5 and designated as RILs. The sub-population 35WS was generated using this crossing scheme (Table 2.14, Figure 2.2).

2.2.3 Agrobacterium mediated transformation of Arabidopsis (Floral Dipping)

All Arabidopsis transformations were performed using a simplified floral-dipping protocol described by Davis et al., 2009. Briefly, *Agrobacterium* strain ABI (Schomburg et al., 2001) bearing the required transgene was streaked on YEBS growth media containing appropriate antibiotics. After two days, a starter culture of 25ml YEBS with appropriate antibiotics was inoculated with 2-3 colonies. The starter culture was then incubated at 28°C
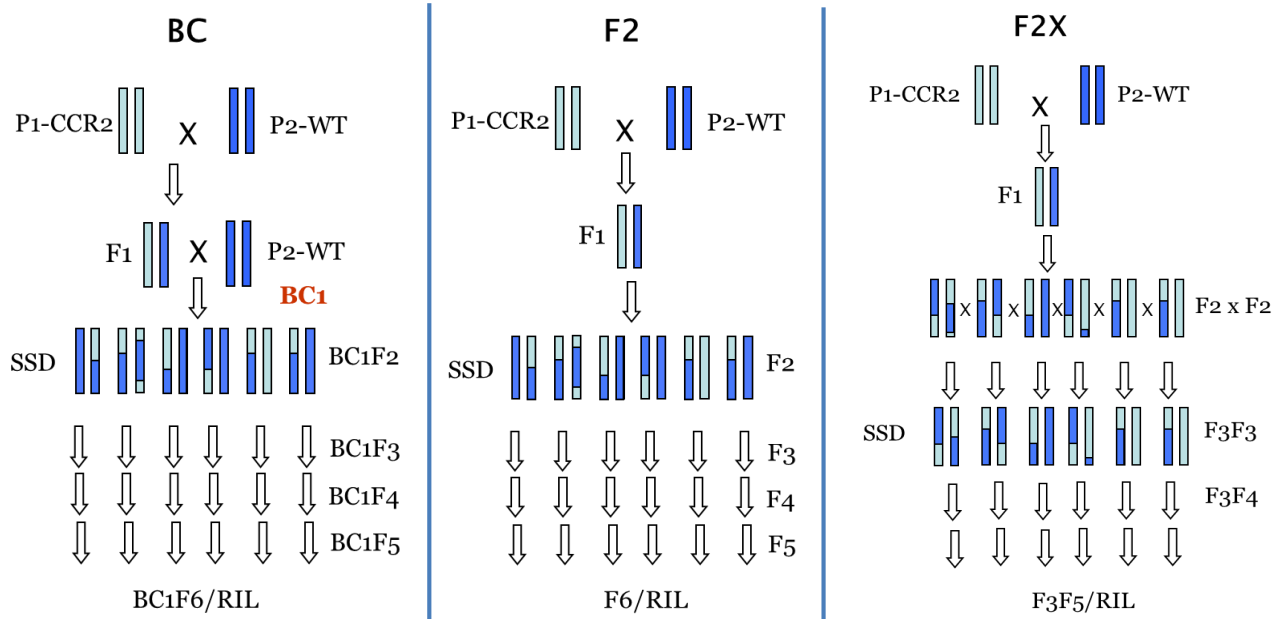
Figure 2.2: Schematic diagram representing the crossing strategy used to generate RILs sub-populations.

for two more days, until optimum bacterium growth was achieved. On the day of the transformation, the starting culture was diluted in 500 mL of YEBS and grown for another 6-8 hours. Then, 80 μl of Silwet L-77 was added to the culture. Arabidopsis plants at flowering-initiation stage were submerged for ~20 seconds in the bacterial culture. After dipping, plants were wrapped with plastics bags for 12-18 hours and then transferred to the greenhouse until seeds were matured. T1 transgenic plants were selected MS plates supplemented with appropriate antibiotics.

2.2.4 DNA extraction, PCR amplification and polymorphism detection for genotyping

For all DNA extractions, a BioSprint® robotic workstation (Qiagen, 9000852) was used. The DNA was isolated using BioSprint® 96 DNA plant kit (Qiagen 941558) following the manufacturer's protocol and the purified DNA solution was diluted to 1:10 for PCR. For all SSLP, CAPS, dCAPS and InDel markers, PCR were performed using commercial Taq-polymerase (Genaxxon) and PCR-product differences were resolved using either DNA

screening kit or DNA high resolution kit on QIAxcel® system (Qiagen, 9001421). For all SNP markers, melt-curve genotyping was performed on Roche LightCycler®480II (Roche Applied Science) using EvaGreen fluorescence dye (Biotium) and a custom made buffer. Gene scanning plug-in of LightCycler®480 SW1.5 software was used to analyze the melting peak differences. PCR conditions were 1X (95°C, 3 min), 45X (95°C, 15 sec: 52°C, 30 sec), melt curve, 1X (95°C, 01 sec: 40°C, 1 min: 60°C, 01 sec: 95°C, continuous).

Reagents	Molarity	Eva Basic Buffer (EBB) Volume (ml)	Recipe (g/ml H2O)
KCl	1	8	3.02/40
Tris	1.5	1.33	0.90/5
MgCl2.6H2O	1	0.5	1.01/5
H2O		0.17	
Total		**10**	
Filter sterile: 0.22uM filter			

2.2.5 Construction of Linkage Maps

Linkage maps were generated using JoinMap 4.0 (Kyazma) mapping software. Unlike its predecessor JoinMap3.0 that cannot handle advance backcrossed populations, JoinMap 4.0 can handle both RIx (selfed from F2 to nth generation) and BCpxFy (advanced backcross populations selfed to nth generation) populations. The largest RIL sub-populations (*e.g.* In Strand X Ws-2, it was WSW) was initially used to calculate the map, and was further confirmed using the complete RIL set under mapping population type RIx. Similar results were obtained for both analyses. The linkage groups were defined at a minimum LOD value of 3.0. The Haldane mapping function was used to calculate map distances.

After assigning initial map positions to the markers at chromosome 3, their positions were further adjusted using their already known positions on other genetic maps. In rare cases where the marker was not mapped in a population, the adjusted position of

that marker was calculated using the average recombination rate of the population and then it was converted from the physical distance to the genetic distance.

2.2.6 Periodicity assay

2.2.6.1 Entrainment conditions

Around 100 seeds were sterilized and plated on the MS3 media containing appropriate antibiotics. The seeds were then stratified for 3 days at 4° and transferred to the growth cabinet under respective environmental conditions: either light dark (LD) or warm cool (WC). For LD entrainment, the growth cabinet was set for 12 hour light and 12 hour darkness (12h:12h LD) cycles under constant 22°C. For WC entrainment the cabinet was set to temperature cycles of 12 hour at 16°C and 12 hour at 22°C (12h:12h WC) under constant white light (Figure 2.3). The plants were entrained for seven days under these conditions. On day seven, plants were prepared for imaging.
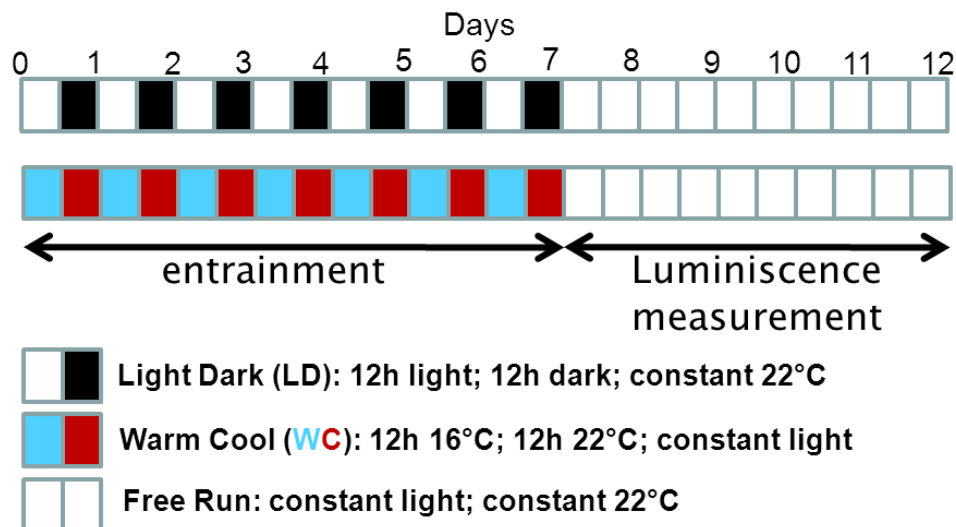
Figure 2.3: Periodicity assay conditions

Plants were entrained for 7 days under LD or WC conditions before transferring for bioluminescence measurements. On day 8, plants were transferred to TOPCOUNT and bioluminescence was measured for at least 5 days under free-running conditions. Consecutive white and black boxes represent LD entrainment. Consecutive blue and red boxes represent WC entrainment. Continuous white boxes represent free-run conditions.

2.2.6.1 Bioluminescence imaging

On day 7, seedlings were transferred to black 96-well Microplates (OPTIPLATE TM-96F, PerkinElmer) containing 200 µl of MS3 agar. Then, 15 µl of 5mM Luciferin was added to each well and the plates were sealed with transparent film (Packard Topseal). Finally, each well was perforated using a needle. Plates were transferred back to their respective cabinet for an additional day of entrainment before plates were transferred to the TOPCOUNT® scintillation counter (PerkinElmer), at subjective dusk. Luminescence values were recorded as the average count of 5 second and monitored every 30-60 min for 5-6 days. When using constant-light conditions, reflector plates were placed in between the seedling plates, and an additional count delay of one minute was applied before the start of the luminescence measurements. The light source was tri-chromatic LED panels (Mark Darby, MD Electronics, UK) attached to the TOPCOUNT® stackers. A minimum of 24 plants per genotype was used for each experiment.

2.2.6.2 Period estimation

Luminescence values obtained by TOPCOUNT were visualized using EXCEL macro TOPTEMP II (http://millar.bio.ed.ac.uk/Downloads.html). Rhythmic traces were analyzed by the Biological Rhythms Analysis Software System (BRASS) macro (Southern and Millar, 2005) in EXCEL that uses the FAST FOURIER TRANSFORMATION NONLINEAR LEAST SQUARES (FFT-NLLS) method to estimate period (Plautz et al., 1997). A 90 h window starting from the beginning of the free-running condition was selected to calculate period. All period values with a weighted real amplitude error (RAE) below 0.45 were considered. R.A.E. defines the extent to which the mathematic model of FFT-NLLs analysis for a perfect curve fits to the actual data. Hence, R.A.E estimates the precision of rhythmicity ranging from 0 (a perfect oscillation) to 1 (arrhythmic oscillation).

2.2.7 Flowering time assay

All flowering-time experiments for RILs were performed in temperature-controlled

greenhouse supplemented with artificial lighting. Experiments were started in January 2010 at Max Planck Institute for Plant Breeding Research (MPIPZ). For this, seeds were first sown on sand supplied with ¼ MS0 media containing hygromycin antibiotics. After stratification for 3 days at 4°C, the plants were grown in the growth cabinet under long day (LgD: 16 hour light and 8 hour darkness). Seven days old seedlings were transferred to the greenhouse under LgD. Nine plants per pot were transferred in 9X9 cm square pots. Number of days to 1 cm bolt was counted as flowering time. Data was taken every consecutive day.

For the *ELF3-Sha* project, flowering time was measured in same green house under both LgD and short day (StD: 8 hour light and 16 hour darkness). Plants were transferred to the greenhouse in May 2011. Number of days to bolting was counted for 24 plants grown in single 7X7 cm pots.

2.2.8 QTL mapping

QTL mapping was performed using MapQTL 6.0 software. Initially, interval mapping (IM) was used to detect putative QTLs, which was then followed by multiple QTL mapping (MQM) using markers as cofactors. Largest effect markers nearest to QTLs detected in IM was initially selected as cofactors. Then, automatic cofactor selection function of MapQTL 6.0 that uses a backward-elimination method was employed to detect screen potential set of markers. Automatic cofactor selection procedure was performed repeatedly, until a final fixed set of markers was obtained. These markers were then used in subsequent MQM analysis as cofactors to detect QTLs. An average genome-wide LOD threshold was estimated by performing permutation test thrice at 1000 iterations.

2.2.9 Statistical analysis

Correlation, Q-Q plots, single factor ANOVA and univariate GLM were performed using the statistical package SPSS 17. Correlation between circadian periodicity and flowering time was calculated using Pearson correlation coefficient. Univariate GLM analysis was used to check the two-way interactions between the markers closely

associated to the QTLs. These were the same markers that were selected by automatic cofactor-selection method and were used as cofactors during MQM analysis. For GLM analysis, the dependent variable assigned was either Period (LD or WC) or flowering time, whereas, the markers were used as fixed factors. A custom model using type III sum of squares was used to calculate main effect and interaction between markers. The significance level used was 0.05. All trait values that correspond to heterozygous genotype were removed from the analysis.

2.2.10 PCR amplification for ELF3 cloning

Proof-reading enzyme Pfu II Ultra HF DNA polymerase (Stratagene) was used to perform all PCRs for *ELF3* amplifications. Primers used are listed in Table 2.9. PCR conditions were as follows. 1X (95°C, 2 min), 30X (95°C, 20 sec: 55°C, 45 sec: 72°C, 4 min), 1X (72°C, 10 min). Correct band size was confirmed on 1% agarose gel. After size confirmation, a given PCR product was purified with 30% PEG8000/ 30mM MgCl2 (Invitrogen) according to manufacturer's protocol. The purified product was then used in BP reaction in Gateway® cloning.

2.2.11 MultiGateway® cloning

BP reaction was performed to recombine purified PCR products into pDONR201 (*ELF3 coding*), and pDONR4R-1PR (*ELF3 promoter*) (Invitrogen). The *YFP* clone was already available in pDONR2-3PR (Herrero et al., 2012). The BP reaction was set up in 1.5 mL eppendorf and was left overnight at 20 °C. Then, 0.5 µl of Proteinase K (Invitrogen) was added to the BP reaction followed by incubation for 10 min at 37°C. Tubes were then kept on ice for five minutes. Finally, 1- 4 µl of BP reaction was used for transforming 50 µl DH5α cells.

Cloning cassettes from pDONR201, pDONR4R-1PR, and pDONR2-3PR were recombined to pDEST (pPZP211R4R3) vector by LR reaction. LR reaction was set up in a fresh 1.5 mL eppendorf and was left overnight at 20 °C. Then, 0.5 µl of Proteinase K

(Invitrogen) was added to the LR reaction followed by incubation for 10 min at 37°C. Tubes were then kept on ice and transformation was performed using 1-5 µl of LR reaction.

Both BP and LR reactions were performed according to manufacturer's recommendations.

2.2.12 Site-directed mutagenesis

pDONR201 containing *ELF3* coding region cloned from Bay or Sha was used as a template to induce point mutation at target sites by the Quikchange method (Stratagene). Primers used to introduce mutations are listed in Table 2.9. The cycling conditions were; 1X (95°C, 30 sec), 13X (95°C, 30 sec: 55°C, 1 min: 68°C, 20 min). The resultant product was then digested using 1µl *Dpn*I enzyme (2 hour at 37°C). Finally, 7 µl reaction was used to transform 50µl DH5α cells. *ELF3-A362V* and *ELF3-V362A* were obtained using this method.

2.2.13 Quantitative real time PCR

After seed stratification for 3 days, plants were grown in 12h:12h LD cabinets for 7 days. The following day at ZT0, plants were transferred under continuous white light for 2 days. The next day starting at ZT0, which is 48 hours under constant light, replicate tissue samples were harvested every 4 hours. After sample collection, RNA extraction was performed with the QIAGEN RNeasy® Plant Mini Kit (Qiagen, 74904), following the manufacturer's protocol. cDNA was obtained using SuperscriptII. qRT PCR was performed on Roche LightCycler480® using Sybergreen fluorescence dye. PCR conditions were. 1X (95°C, 3 min), 45X (95°C, 30 sec: 58°C, 30 sec: 72°C, 30 sec), melt curve, 1X (95°C, 05 sec: 55°C, 1 min: 97°C, continuous).

2.2.14 Confocal microscopy

For all microscopic work, Zeiss LSM700 confocal microscope from Carl Zeiss was used. The plants were grown on MS medium containing 1.5% sucrose. Following ~3 days stratification at 4°C, seedlings were entrained for 6 days under 12L:12D cycles (~100µE light) with constant temperature of 22°C. The following day, the plants were put under

constant light (LL: ~100μE light) for one day. On the next day, at ZT16, the plants were scanned and the photographs were taken. For comparison of *ELF3 Bay-YFP* and *ELF3 Sha-YFP* lines, one slide from each line was prepared and both the slides were put together in the microscope. The plants from each slide were then sequentially scanned within one hour. The microscope settings were as follows (Image size: x = 512, y = 512, z = 20, Channels: 3, 8-bit, Zoom = 0.5, Objective: Plan-Aprochromat 63x/1.40 Oil, Pixel dwell: 2.55μs, Master gain: ch1= 1096, ch2= 928, ch3= 393, Digital gain: ch1 = 1.20, ch2 = 0.61, ch3 = 1.40, Digital offset = 0.0, Pinhole = 156μm and laser: 488nm with 10.0% strength). The same microscope settings were used for all pictures.

Chapter Three: **Positional isolation and Characterization of**
ELF3-Sha

3.1 Introduction

ELF3 is a fundamental component of Arabidopsis circadian clock (Dixon et al., 2011; Thines and Harmon, 2010). The proper functioning of ELF3 is essential for the generation of oscillating rhythms, as well as light-input to the circadian oscillator (Hicks et al., 2001; Kolmos et al., 2011; Liu et al., 2001; McWatters et al., 2000; Yu et al., 2008). Any defect in ELF3 results in aberrant clock function and abnormal growth behavior (Hicks et al., 1996; Nozue et al., 2007; Nusinow et al., 2011; Yoshida et al., 2009; Zagotta et al., 1996; Zagotta et al., 1992). Several *elf3* mutants have been described in the literature with severe to mild phenotypes. Mutants with severe phenotypes include loss-of-function mutants *elf3-1* and *elf3*-4. These mutants displayed clock arrhythmia under constant conditions and defects in normal growth. The growth defects were observed as a photoperiod-independent early flowering and uncontrolled hypocotyl elongation (Dixon et al., 2011; Herrero et al., 2012; Liu et al., 2001; Thines and Harmon, 2010). The arrhythmic phenotype of these mutants had previously precluded the overall understanding of ELF3 function in the past. Recently, *elf3-12* was identified as a reduced-functional allele of ELF3. This allele exhibited light-dependent short period phenotype with overall robust circadian oscillations and wild-type growth behavior (Kolmos et al., 2011). The genetic resources generated in this mutant's background provided strong evidences that ELF3 repressive action in circadian clock is modulated by phytochromes (Kolmos et al., 2011). The characterization of *elf3-12*, thus established the notion that availability of new alleles could be helpful to better understand the overall mechanism of ELF3 action.

In this chapter, I shall describe the identification and characterization of *ELF3-Sha:* a natural allele of *ELF3* that exhibits a light-dependent short-period phenotype. This study was an extension of the previous work where a modified Bay X Sha RIL population (Loudet et al., 2002) was analyzed for free-running circadian periodicity and a major effect QTL was detected at Chromosome 2 (Boikoglou, 2008). This was the most significant QTL that explained 27.3% of the total variation observed. Allelic analysis revealed that presence of Sha allele at this locus results in acceleration of circadian periodicity (Boikoglou, 2008). A

known circadian clock gene *ELF3* co-localizes at the QTL confidence interval. *ELF3* has been established as a core-clock component and a major determinant of flowering time (Liu et al., 2001; Zagotta et al., 1996). As finding connections between circadian clock and flowering time was the main objective of my study, therefore, I decided to clone this QTL. To achieve this goal, I initially validated the QTL in heterogeneous inbred families (HIFs) and near isogenic lines (NILs). Further, using classical, positional isolation and a transgenic-complementation approach, I cloned the underlying gene as *ELF3*. This was done to a resolution of a single nucleotide. With sub-cellular localization studies, I further demonstrated that the single nucleotide mutation in *ELF3-Sha* leads to sub-cellular localization defect of ELF3 protein. Based on my results, I concluded that this localization defect associates to *ELF3-Sha* short-period phenotype.

3.2 Results

3.2.1 Validation of Chromosome 2 QTL

Isogenic lines carrying different parental allele at QTL confidence interval while otherwise having same genetic makeup throughout the genome, could be used to confirm the allelic effect of the locus (Koornneef et al., 2004; Tuinstra et al., 1997; Weigel, 2012). HIFs and NILs are such kind of isogenic lines. To validate the effect of Sha allele at chr2 locus, I generated three HIFs (57, 92 and 343), each carrying either Bay allele (hereafter HIF-Bay) or Sha allele (hereafter HIF-Sha) at QTL confidence interval (see materials and methods). Analysis of free-running period of these liens revealed that all HIF-Sha lines displayed ~1.5h shorter period as compared to the HIF-Bay (Figure 3.1 a). Though, the only difference between the HIF-Sha and HIF-Bay was at the region of the QTL, the effect of interacting loci cannot be ignored in these lines, given their heterogeneous background. To eliminate the possibility of such interactions, a NIL with small introgression of Sha at QTL interval (hereafter NIL-S) in an otherwise homogeneous Bay genetic background was generated. Consistent with HIFs, the comparison of periodicity analysis of NIL-S and Bay

revealed ~2h difference, where NIL-S exhibited short-period (Figure 3.1 b). Interestingly, this period shortening effect of Sha allele was only observed under LL. In darkness, both NIL-S and Bay displayed a similar period (Figure 3.1 b). Moreover, under LL, the Sha wild type, which donated the introgression region to the NIL-S, displayed similar extended period as that of Bay, which was longer than the NIL-S (Figure 3.1 b). Taken together, these result confirmed the existence and the allelic effect of chr2 QTL.
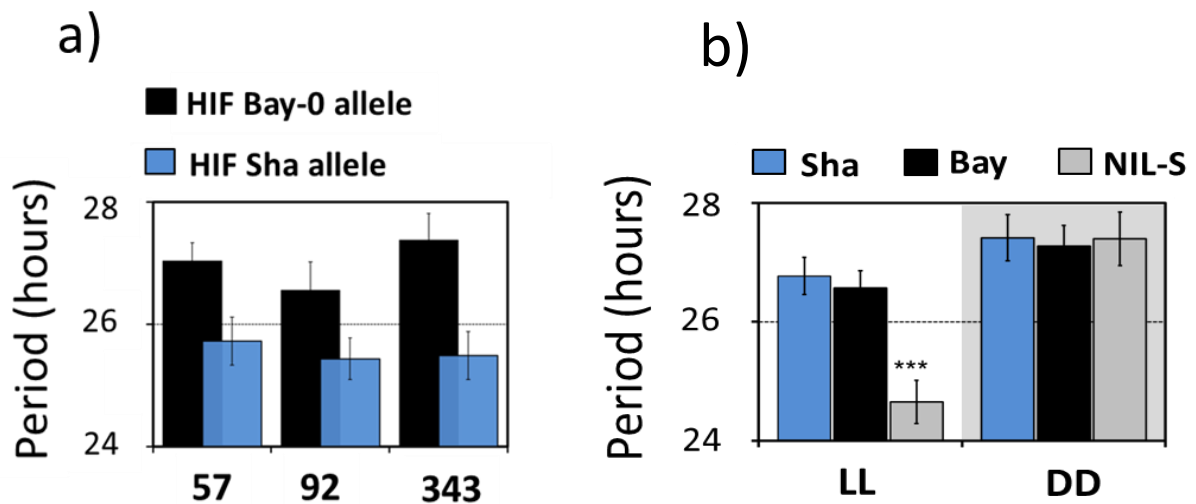
Figure 3.1: Validation of chromosome 2 QTL.

Period estimates of *CCR2::LUC* expression in (a) three independent HIFs (57, 92, and 343) harboring either the Bay-0 or the Sha allele at QTL confidence interval (b) parental accessions and NIL-S with introgression of Sha at QTL confidence interval in otherwise homogenous Bay-0 background, under LL and in DD. All error bars indicates the S.E.M., where n ≥ 24. Mean values that are significantly different from Bay-0 wild type are indicated by *, **, or *** for P-values (ANOVA) < 0.05, 0.01, or 0.001, respectively.

3.2.2 Positional isolation revealed ELF3 as a strong candidate

NIL-S confined the QTL region from several thousand genes to few hundred genes. To further narrow down the region defined by the NIL-S, I followed the classical recombination cloning approach. In a gradual screen of 1100 plants by genotyping and phenotyping, I selected fourteen plants that carried a recombination event in the confidence interval. Further screening of these plants revealed two potential lines: 89-S

and 539-B. These lines had recombination events between the markers U43-elf20 and U25-elf20, a 40Kb region, with nine annotated genes. *ELF3* (*AT2G25930.1*) was the only gene in this interval with known function in the circadian clock. Notably, the only difference between these recombinants was the presence of either Sha or Bay allele between the recombination events, for 89-S and 539-B, respectively (Figure 3.2 a). Analysis of the free-running period of *CCR2::LUC* under LL revealed that the recombinant line 89-S displays ~1.5h shorter period as compared to the line 539-B. However, in DD, both recombinant lines exhibited similar period (Figure 3.2 b, c). These observations were consistent with the period estimates of HIFs and NILs (Figure 3.1 a, b). Collectively, these results linked a 40kb region to the Chr2 QTL, with *ELF3* as a potential candidate.

3.2.3 Identification of ELF3 as a quantitative trait gene (QTG)

Although, positional cloning provided strong evidences in favor of *ELF3* as a gene responsible for the short-period-phenotype observed in NIL-S, one cannot ignore the involvement of any of the other nine genes located in the 40kb region defined by the recombinants 89-S and 539-B. To eliminate these other genes as candidate and to confirm *ELF3* as quantitative trait gene, I cloned *ELF3* along with its own promoter from Bay and Sha, fused it with *YFP*, and transformed these constructs in null mutant *elf3-4* that was already harboring stable *CCR2::LUC* or *LHY::LUC* reporter constructs (see materials and methods). The free-running period analysis of these reporter genes was carried out in T2 transgenic lines expressing *ELF3-Sha-YFP* or *ELF3-Bay-YFP*. Consistent with the belief that *ELF3-Sha* is not a loss-of-function mutant, both *ELF3-Sha-YFP* and *ELF3-Sha-YFP* complimented the arrhythmic phenotype of *elf3-4*, as robust rhythms were detected for both *CCR2::LUC* and *LHY::LUC* under LL (Figure 3.3 a, b). Furthermore, *ELF3-Sha-YFP* exhibited a consistent short-period compared to *ELF3-Bay-YFP,* for both luciferase reporters tested (Figure 3.3 c). Thus, these complementation experiments confirmed that Sha allele of *ELF3* underlies the Chr2 QTL and causes light-dependent periodicity acceleration. Moreover, unlike *elf3-4,* this allele maintains partial activity.

a)

b)   c)
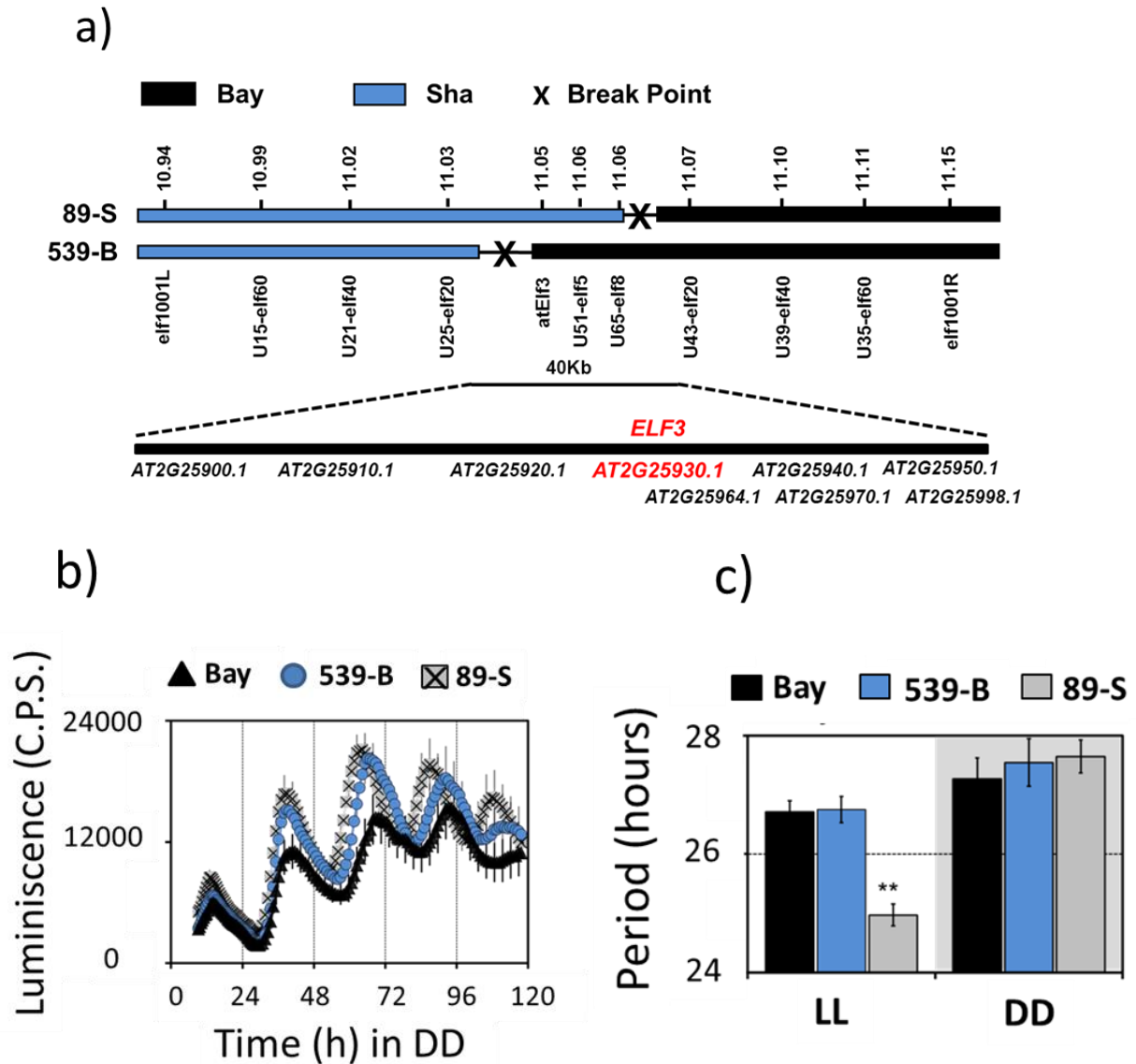
Figure 3.2: Positional isolation revealed *ELF3* as a strong candidate.

(a) Schematic diagram showing the fine mapping strategy of chr2 locus. Black and grey bars represent Bay-0 and Sha genotypes, respectively. The names below the bars represent the molecular markers used for genotyping and the numbers above correspond to their physical position on the genome. The crosses represent the position of recombination event. Two recombinants 89-S and 539-B were identified to carry a recombination event surrounding a region of 40KB, where 9 annotated genes are located, as indicated below the solid bar. (b) Free-running profile of *CCR2::LUC* expression in recombinants 89-S, 539-B, and Bay-0 under continuous red and blue light (LL). (c) Period estimates of rhythm shown in (b). All error bars indicates the S.E.M., where n ≥ 24. Mean values that are significantly different from Bay-0 wild type are indicated by *, **, or *** for P-values (ANOVA) < 0.05, 0.01, or 0.001, respectively.

a)

b)

c)
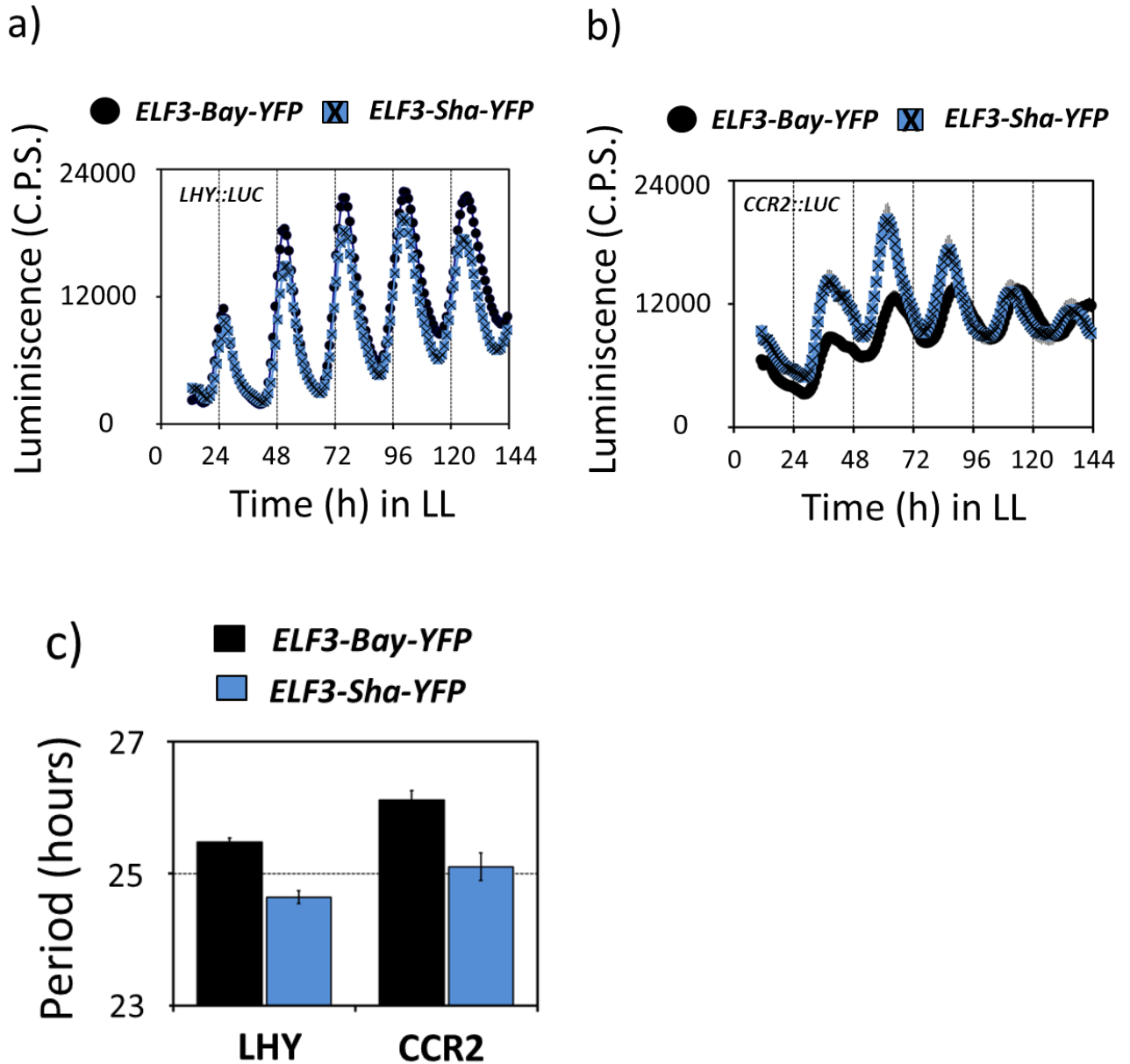
Figure 3.3: *ELF3-Sha* and *ELF3-Bay* complements the null mutant phenotype.

Free-running profile of *LHY::LUC* (a) and *CCR2::LUC* (b) expression in T2 transgenic lines carrying either *ELF3-Bay-YFP* or *ELF3-Sha-YFP* in Ws-2 background. The data is the average of three independent single insert lines displaying similar rhythm profile. (c) Period estimates of the lines shown in (a) and (b). Error bars represent SEM, n = 96. Note that all lines used here were in Ws-2 background.

3.2.4 Sequence analysis Of ELF3 revealed Alanine as an evolutionary conserved residue in Arabidopsis

A change in gene function normally links to underlying DNA polymorphisms (Koornneef et al., 2004; Weigel, 2012; Weigel and Nordborg, 2005). To find such nucleotide differences in *ELF3-Sha*, I compared the genomic DNA sequence of *ELF3* from the Bay and Sha wild types. Consistent with previous reports (Coluccio et al., 2011; Jimenez-Gomez et al., 2010), several polymorphisms were detected (Figure 3.4 a). Notably, I found five SNPs in the promoter region, a single nucleotide transition from 'C' to 'T' in the middle-domain and a 24 base pair deletion in the C-terminal domain (Figure 3.4 a). The middle domain and C-terminal domain polymorphisms resulted in non-synonymous amino-acid changes. Specifically, the middle-domain transition resulted in an encoded amino-acid change from alanine to valine at position 362 (hereafter A362V), whereas the C-terminal deletion caused an eight glutamine deletion (hereafter 8Q558) (Figure 3.4 a). Any of these changes either independently, or in combination with the other could be the reason for *ELF3-Sha* short-period phenotype.

To understand the evolutionary behavior of above mentioned changes, I screened a collection of 62 geographically diverse Arabidopsis accessions using polymorphic markers specific for A362V and 8Q558. Consistent with previous reports (Coluccio et al., 2011; Takeomi et al., 2007), I found in these accessions that the encoded Alanine residue is fully conserved, whereas the length of the glutamine stretch is highly variable (Appendix 2). Further, to detect if variation in glutamine numbers is linked to circadian periodicity, I estimated the extent of their relationship by calculating correlation coefficient between number of glutamines and free-running periodicity after various entrainment conditions, of listed Arabidopsis accession (Appendix 2). Though a weak negative relationship had been reported between these parameters (Takeomi et al., 2007), I could not detect any correlation under any entrainment condition (Figure 3.4 b-e). This led me to conclude that 8Q558 C-terminal domain deletion might not directly be involved in *ELF3-Sha* periodicity acceleration.
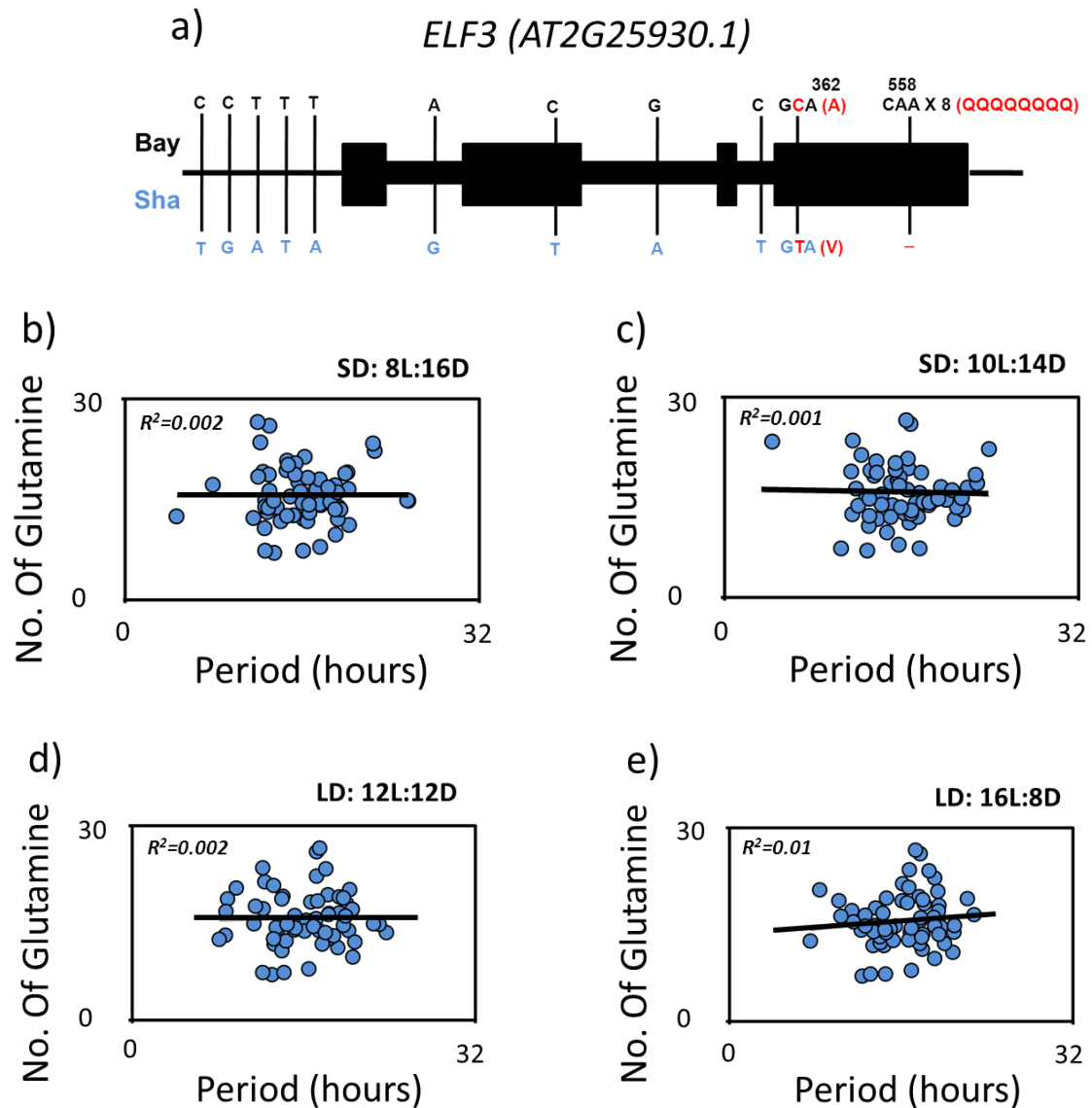
Figure 3.4: No correlation was detected between circadian periodicity and number of glutamine.

A schematic representation of *ELF3 (AT2G25930)* (a). Vertical bars show the location of the nucleic acid transition. The letters above the bars represent the nucleic acid in Bay and letter below represent the correspondent base pair in Sha. The letters in parenthesis show the amino acid with its position shown as numbers above. (b)-(e) Correlation of number of Q-repeats with circadian periodicity estimated after various entrainment conditions (b) SD: 8 hours light and 16 hours dark (c) SD: 10 hours light and 14 hours dark (d) LD: 12 hours light and 12 hours dark (e) LD: 16 hours light and 8 hours dark.

Period estimates and DNA samples for all Arabidopsis accessions used here were kindly provided by Amaury de Montaigu (George Coupland Lab).

3.2.5 Identification of quantitative trait nucleotide (QTN): A362V underlies ELF3-Sha periodicity phenotype

To further discriminate among different *ELF3-Sha* polymorphisms as a causal of period acceleration, I separately cloned the *ELF3* promoter and coding region from Bay and Sha wild types. In this way, four independent constructs were obtained: a) *ELF3* promoter cloned from Sha (hereafter Sp), b) *ELF3* coding cloned from Sha (hereafter Sc), c) *ELF3* promoter cloned from Bay (hereafter Bp), d) *ELF3* coding cloned from Bay (hereafter Bc). A single nucleotide change was then induced in Sc and Bc that resulted in encoded amino-acid changes, from valine to alanine (hereafter Sv2a) and alanine to valine (hereafter Ba2v) in Sc and Bc, respectively. Different coding constructs Sc, Bc, Sv2a and Ba2v were then separately recombined with either Sp or Bp (Figure 3.5 a). These constructs were then fused to YFP and transformed in *elf3-4* null mutant harboring stable *LHY::LUC*. The free-running period of resultant T2 transgenic lines carrying different coding regions driven by Sha promoter was estimated under LL. I found that the lines SpSc and SpBa2v that contained 'Valine' displayed shorter period as compared to SpBc and SpSv2a, lines carrying 'Alanine' (Figure 3.5 b). These differences in period were irrespective to the number of glutamine present. Thus these results demonstrated that a single amino-acid change from alanine to valine associate to period shortening of *ELF3-Sha.*

Finally, to eliminate the minor possibility of any differential promoter regulation of *ELF3-Bay* and *ELF3-Sha*, I further confirmed these results in the transgenic lines expressing *ELF3-Bay* and *ELF3-A362V* under the Bay promoter. The free-running profile of *LHY::LUC* was analyzed under LL and in DD. Both *ELF3-Bay* and *ELF3-A362V* displayed robust rhythm of *LHY::LUC* expression under LL. Here, low amplitude was observed in *ELF3-A362V* (Figure 3.5 c). Furthermore, consistent with previous results in HIFs, NIL-S, and transgenic complementation lines, a ~1.5 hour period difference was observed between *ELF3-Bay* and *ELF3-A362V*. Additionally, no difference in clock speed was detected in these lines in DD (Figure 3.5 d). Collectively, these results provided strong evidence to conclude that a single encoded A362V residual change associates to the *ELF3-Sha* short-period phenotype.
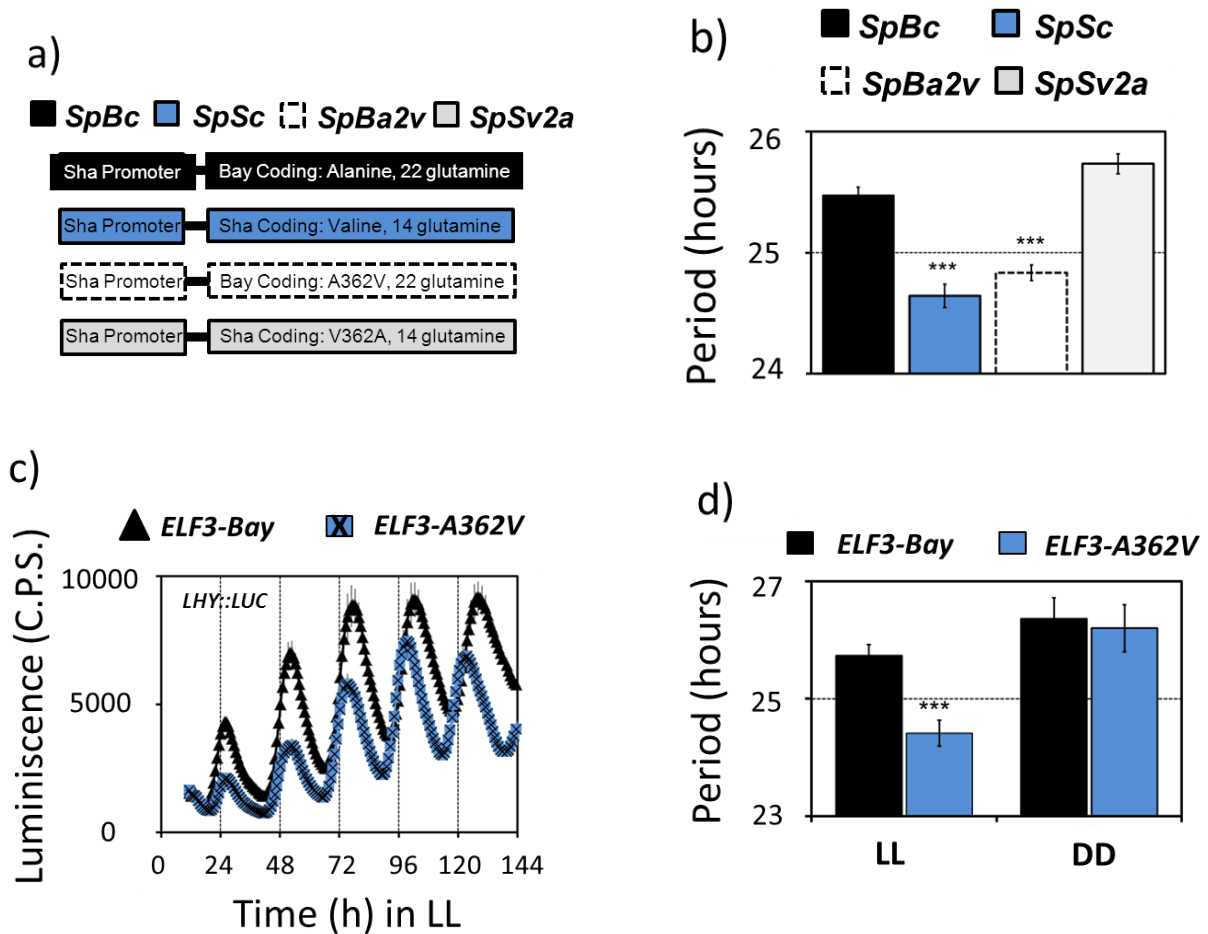
Figure 3.5: Identification of QTN.

a) Schematic diagram representing different promoter-coding combinations used in (b). Sha promoter fused with different coding regions is shown (for details see chapter 2). The amino acid residue at position 362 along with the number of glutamine is shown. (b) Period estimates of *LHY::LUC* expression in lines explained in (a. Note that the lines *SpSc* and *SpBa2v*, carrying valine at residue 362 displayed period acceleration regardless the number of glutamines. Error bars represent SEM, n = 48. Significance as explained in figure 3.1, compared to *SpBc*. (c) Free-running profile of *LHY::LUC* expression in T2 transgenic lines carrying either *ELF3-Bay* or *ELF3-A362V.* A single nucleotide change was induced in *ELF3-Bay* to change the encoded alanine residue at position 362 to valine (*ELF3-A362V*). The data is the average of three independent single-insert lines with similar rhythm profile. (d) Period estimates of the lines *ELF3-Bay* or *ELF3-A362V* under LL and DD. Error bars represent S.E.M., n = 96. Significance as explained in figure 3.1. Note that all lines used in these experiments were in the *Ws-2 elf3-4* background.

3.2.6 Alanine is required for sustained rhythmicity in darkness

ELF3-Sha and hypomorphic allele *elf3-12* share common periodicity phenotypes under light and in DD. Though, *elf3-12* maintained robust rhythms of *LHY::LUC*, it was

reported that in DD, the transcript profile for many clock genes in this allele resembled null alleles (Kolmos et al., 2011). In *ELF3-Sha,* I noticed that the amplitude of expression for *CCR2::LUC* in NIL-S diminished after ~4 days in darkness. Contrarily, both Bay and Sha maintained robust rhythms of *CCR2::LUC* expression for at least 7 days (Figure 3.6 a). This revealed that ELF3-Sha is defective in proper clock function in darkness.

To further investigate if this dampening of *CCR2::LUC* is marker specific and to inquire if A362V residual change underlies this, I monitored the expression of *LHY::LUC* in *ELF3-Bay* and *ELF3-A362V* transgenic lines in DD. Consistent with NIL-S, I found that the *LHY::LUC* expression was significantly decreased in *ELF3-A362V* after ~4 days in darkness, whereas, *ELF3-Bay* maintained robust *LHY::LUC* expression (Figure 3.6 b). These observations led me to hypothesize that light plays an important role in proper functioning of *ELF3-Sha*.
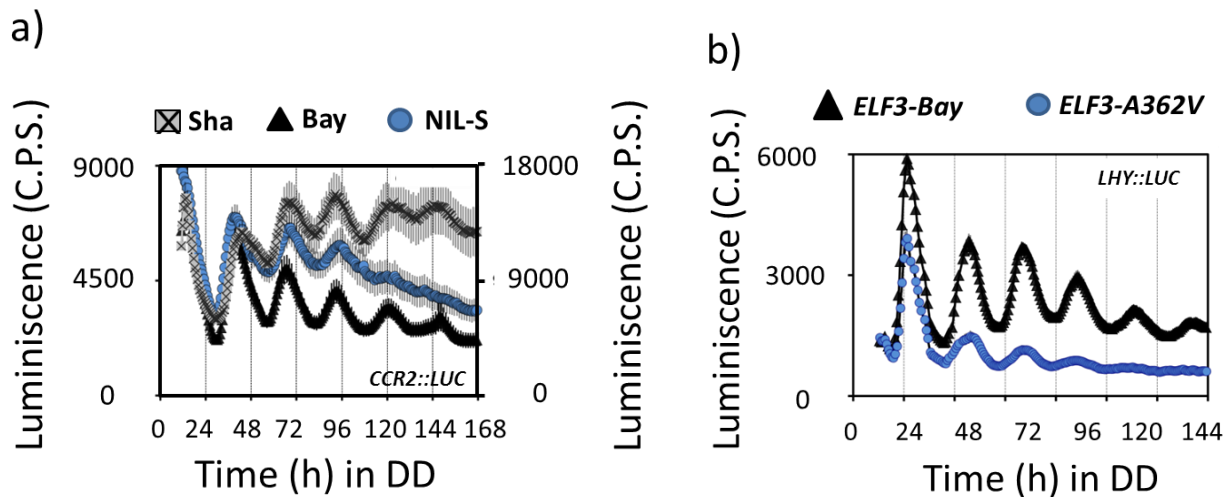
Figure 3.6: *ELF3-Sha* in prolonged darkness.

Free-running profile of *CCR2::LUC* expression in Bay, Sha, and NIL-S in darkness (a), Left y-axis shows the luminescence measures for Bay and NIL-S, and right y-axis shows the luminescence measures for Sha. The NIL-S with introgression of Sha at *ELF3* could not sustain the robust rhythms of *CCR2::LUC* after four days in darkness. Error bars represent S.E.M., n ≥ 24. (b) Shows the free running profile of *LHY::LUC* expression in *ELF3-Bay* and *ELF3-A362V* transgenic lines, in darkness. Note that the *ELF3-A362V* line with valine at position 362 displayed similar loss of robustness of *LHY::LUC* expression as was observed in NIL-S in panel (a). Error bars represent S.E.M., n = 96.

If light is required for sustaining rhythms in *ELF3-Sha* then the plants transferred from darkness to light should restore proper expression of *LHY::LUC*. To investigate this hypothesis, I performed an experiment where the free-running expression profile of *LHY::LUC* was monitored for 15 days continuously in *ELF3-Bay* and *ELF3-A362V* under alternative intervals of light and darkness. I observed that both *ELF3-Bay* and *ELF3-A362V* displayed robust *LHY::LUC* expression as long as they remained under light (Figure 3.7 a). However, as the plants were transferred in darkness, the *LHY::LUC* expression dampened down in both *ELF3-Bay* and *ELF3-A362V*. Although there was an acute reduction of *LHY::LUC* expression in *ELF3-Bay,* it maintained the proper rhythm during the dark period. In contrast there was almost no rhythmic expression of *LHY::LUC* observed in *ELF3-A362* (Figure 3.7 a). To test if the loss of rhythm of *LHY::LUC* in *ELF3-A362V* was light dependent, or was due to a ''permanent'' defect in the clock caused by the continuous darkness, I returned these plants to a light regime after an interval of 4 days in darkness. Both *ELF3-Bay* and *ELF3-A362V* recovered robust rhythms of *LHY::LUC* after 11 days of free running conditions, including an interruption of 4 days in darkness (Figure 3.7 a). This experiment thus confirmed that light is necessary for the oscillating functionality in *ELF3-A362V*.

Robustness of the rhythms depends on precision of the circadian clock and relative amplitude error (R.A.E.) is one measure of precision. R.A.E. defines the extent to which the mathematic model of FFT-NLLs analysis for a perfect curve fits to actual data. Hence, R.A.E estimates the precision of rhythmicity, ranging from 0 (a perfect fit) to 1 (arrhythmic oscillation). To test the precision of the *LHY::LUC* expression rhythms in *ELF3-Bay* and *ELF3-A362V*, I calculated the R.A.E. of these lines under LL and in DD (*LHY::LUC* profile, as shown in Figure 3.7 a). Consistent with previous results, under LL, *ELF3-A362V* displayed a ~1.5h shorter period compared to *ELF3-Bay*, with both lines exhibiting high precision (R.A.E. < 0.20; Figure 3.7 b, e; ZT40-120). However, in darkness, although the circadian periodicity of both lines was similar, *ELF3-A362V* could not maintain precision, as observed by a very high R.A.E. (R.A.E. > 0.75) compared to *ELF3-Bay* (R.A.E. < 0.35; Figure 3.7 c, e; ZT160-240). Interestingly, this decline in precision of *LHY::LUC* rhythm was completely recovered in

ELF3-A362V when the plants were returned to light at ZT250 (Fig 3.7 d, e; ZT260-340). In summary, these results established ''alanine362'' as a required residue in ELF3 for sustained rhythmicity in prolonged darkness.

3.2.7 Clock network in ELF3-Sha

It has been reported that under continuous light and in darkness, the *elf3* loss-of-function mutant displays arrhythmia (Hicks et al., 1996; Thines and Harmon, 2010). Accordingly, expression profiling revealed that *elf3* exhibits reduced expression of the oscillator genes *CCA1* and *LHY*, and has elevated levels of the evening genes *TOC1* and *GI* (Dixon et al., 2011; Fowler et al., 1999; Kikis et al., 2005). All these studies were based on the null *elf3* mutants, where the arrhythmic nature of these mutants made it difficult to place *ELF3* in the clock network. As, *ELF3-Sha* displayed complete and extended rhythmicity under LL and in DD, respectively (Figure 3.3 a, b; Figure 3.6 a, b), expression study of oscillator genes in this allele could provide a better insight into placement of *ELF3* in the clock network. Therefore, I monitored the luciferase reporter expression profile driven by established oscillator gene promoters *PRR7*, *PRR9*, *CCA1*, *LHY*, *TOC1*, and *GI* in *NIL-S* and compare it with the Bay wild type and the null mutants *elf3-1* and *elf3-4*, under both LL and in DD. Consistent with the previous reports (Alabadí et al., 2001; Dixon et al., 2011; Kikis et al., 2005), both null mutants *elf3-1* and *elf3-4* displayed complete arrhythmia for all genes studied (Figure 3.8). Specifically, constant low levels of central oscillator genes *CCA1* and *LHY* were found. In contrast, high levels were detected for morning loop genes *PRR7*, *PRR9* and the evening loop genes *TOC1* and *GI* (Figure 3.8). Moreover, in agreement with the short-period phenotype of *CCR2::LUC* in NIL-S (Figure 3.1 b), all clock genes assayed exhibited similar acceleration of periodicity in NIL-S thus depicting a global effect of *ELF3-Sha* on the circadian clock. This is consistent with a broad role of ELF3 as an oscillator gene (Figure 3.8).

The over expression of ELF3 (*ELF3-OX)* was found to exhibit a longer period (Covington et al., 2001; Herrero et al., 2012). Thus, acceleration of periodicity in *ELF3-Sha*
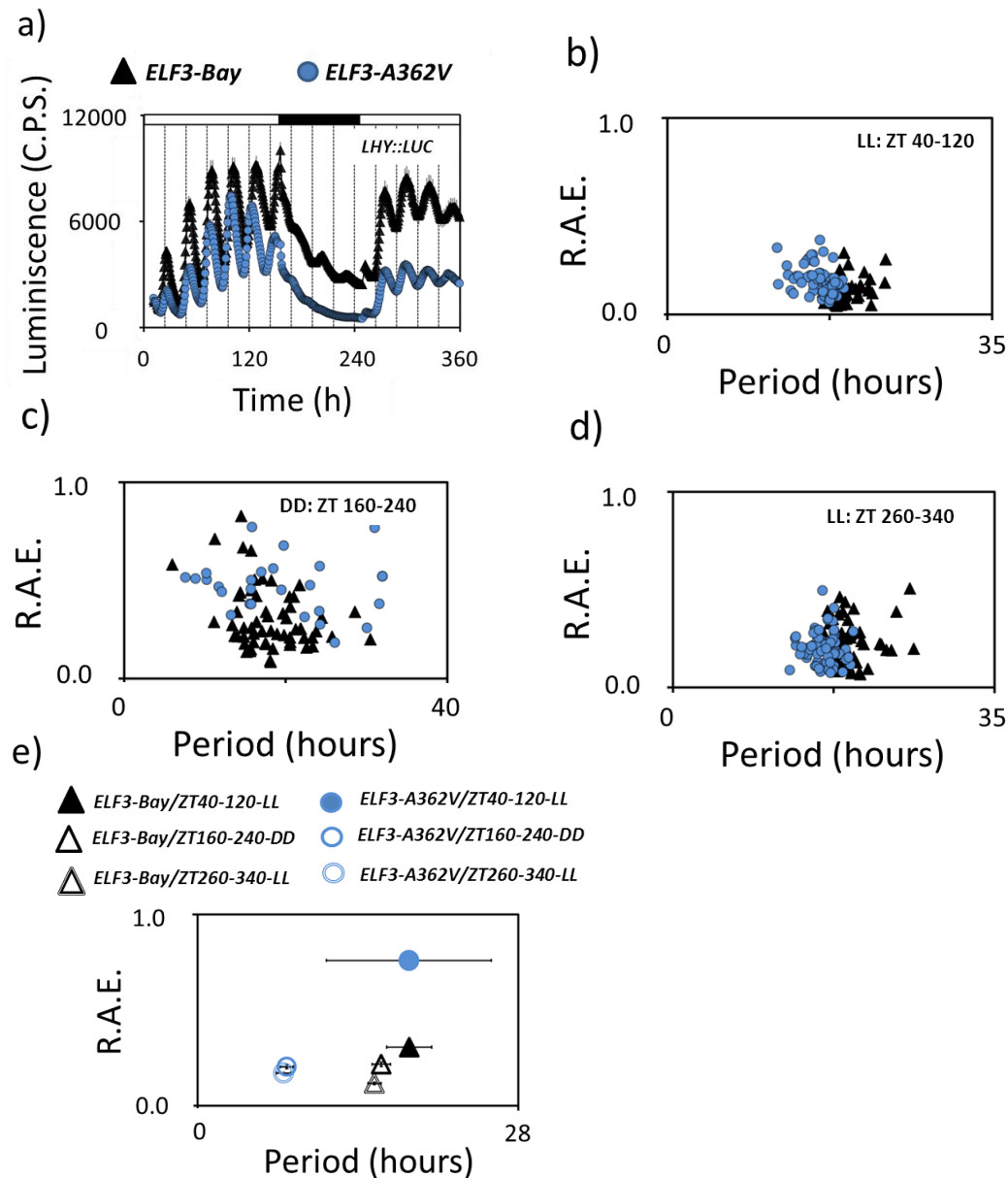
Figure 3.7: Alanine sustains robust oscillations in darkness.

(a) Free-running profile of *LHY::LUC* expression in *ELF3-Bay* and *ELF3-A362V* in a fifteen days continuous experiment, in consecutive light and dark conditions. The plants were entrained for 7 days under 12h:12h light dark cycles and then transferred in LL at ZT0 and LHY::LUC expression was measured for 6 days. On day 7 plants were transferred in darkness and the measurement of luminescence was continued in DD for four more days. On day 11, the plants were transferred under lights again and expression profile of *LHY::LUC* was measured for further four days. Open bars represent time under LL while closed bar represents time in DD. Error bars represent S.E.M. and is shown on every third reading. (b-d) Scatter plot for R.A.E against period showing estimates of individual *ELF3-Bay* and *ELF3-A362V* lines shown in (a). (e) Scatter plot for R.A.E. against average period of ELF3-Bay and ELF3-Sha shown in (a-d). Vertical error bars represent S.E.M., while horizontal error bars represent standard deviation (SD). n=96.

could be associated to reduced functionality. Expectedly, the comparison of expression profiles of luciferase driven by several clock genes in NIL-S, *elf3-1*, *elf3-4* and Bay-0 confirmed the hypomorphic nature of ELF3-Sha, as an intermediate expression of all clock genes were observed in *ELF3-Sha*, compared to null mutants and the Bay wild type (Figure 3.8). Specifically, in NIL-S, the expression of *CCA1::LUC* and *LHY::LUC* was higher than the *elf3-1* and *elf3-4*, but was lower than the Bay-0. Contrarily, the expression of *PRR7::LUC*, *PRR9::LUC*, *TOC1::LUC* and *GI::LUC* in NIL-S was lower than *elf3-1* and *elf3-4*, but higher than Bay-0 (Figure 3.8).

In DD, similar intermediate profiles of luciferase expression for clock genes were observed in NIL-S, as they were detected under LL. Furthermore, an overall decrease in amplitude of all genes assayed was found (Figure 3.9). Additionally, I noticed that *TOC1::LUC* expression in NIL-S was significantly higher than Bay and was comparable to the null mutants (Figure 3.9 e). Furthermore, a continuous increase in *GI::LUC* expression over time was also observed in NIL-S (Figure 3.9 f). Finally, consistent with previous results (Figure 3.1 b), no significant period difference was observed in NIL-S and Bay-0, for any clock gene in DD (Figure 3.9 e). Taken together, these results demonstrated that unlike *elf3-1* and *elf3-4* loss-of-function mutants, *ELF3-Sha* is a hypomorphic allele of intermediate activity.

3.2.8 The *ELF3-Sha* periodicity effect was neutralized in a Sha genomic context

A stark difference in circadian periodicity of two contrasting lines carrying same *ELF3-Sha* allele was observed. These lines include NIL-S, an introgression line carrying *ELF3-Sha* allele in Bay genetic background and Sha wild type, which donated the *ELF3* introgression to NIL-S. If *ELF3-Sha* causes short-period phenotype, than the speed of circadian oscillation for these two lines should be similar, as both carries same *ELF3* allele. However, I found that where NIL-S exhibited short-period phenotype under LL, the speed of circadian oscillation remained essentially similar in both Sha and Bay, as was detected by *CCR2::LUC* (Figure 3.1 b).
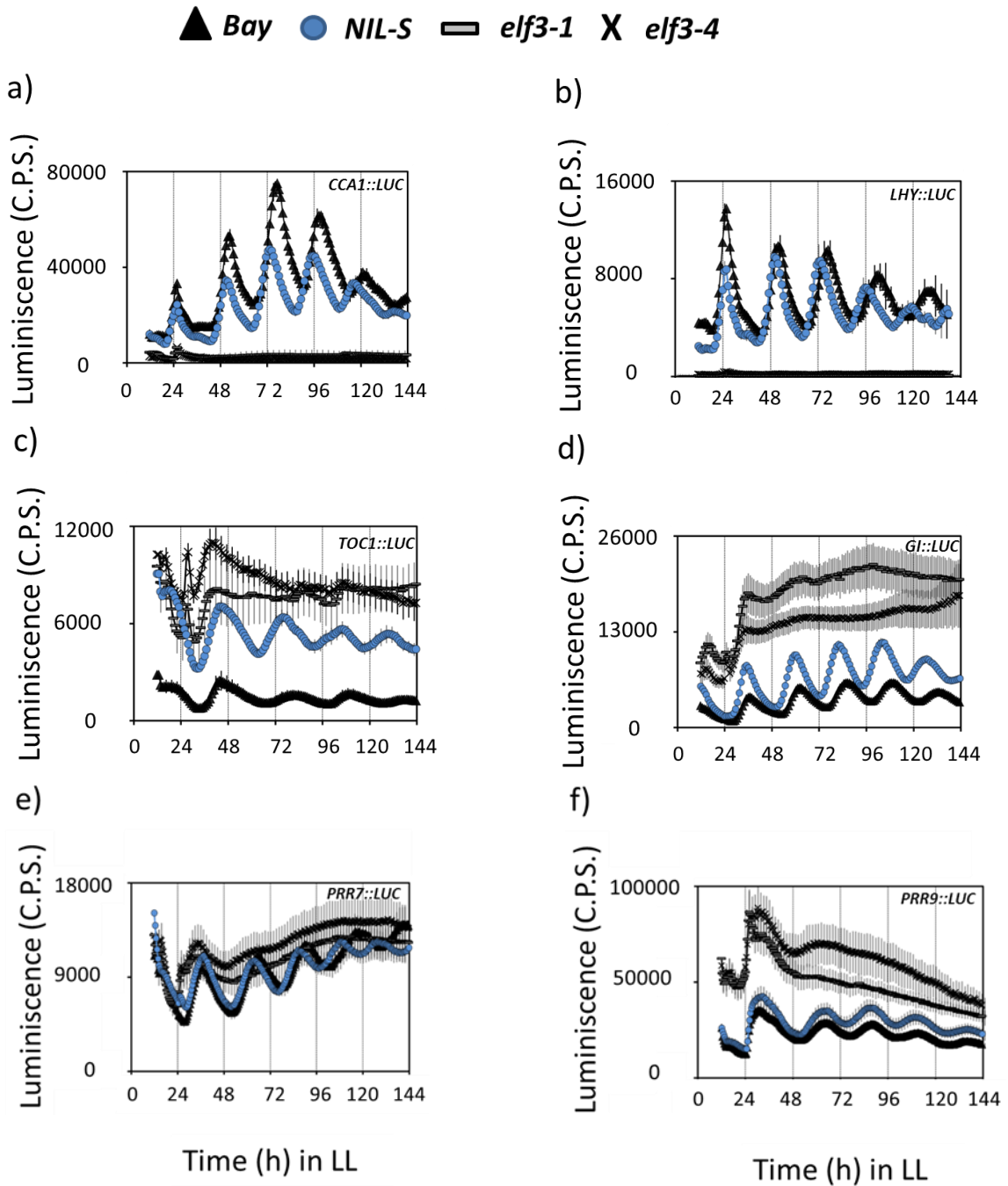
Figure 3.8: Clock network in *ELF3-Sha* under LL.

(a-f) Luciferase expression profile of different clock genes in NIL-S, *elf3-1*, *elf3-4* and Bay-0, under LL. (a) *CCA1::LUC* (b) *LHY::LUC* (c) *TOC1::LUC*, (d) *GI::LUC,* (e) *PRR7::LUC,* (f) *PRR9::LUC*. Error bars represent S.E.M. and is shown on every third reading. Note that the NIL-S exhibited an intermediate expression of all clock genes relative to null mutants *elf3-1* and *elf3-4* and Bay-0.

▲ *Bay* ● *NIL-S* ▭ *elf3-1* ✕ *elf3-4*

a)

Luminiscence (C.P.S.)

CCA1::LUC

b)

Luminiscence (C.P.S.)

LHY::LUC

c)

Luminiscence (C.P.S.)

TOC1::LUC

d)

Luminiscence (C.P.S.)

GI::LUC

e)

Luminiscence (C.P.S.)

PRR7::LUC

f)

Luminiscence (C.P.S.)

PRR9::LUC

Time (h) in DD

Time (h) in DD
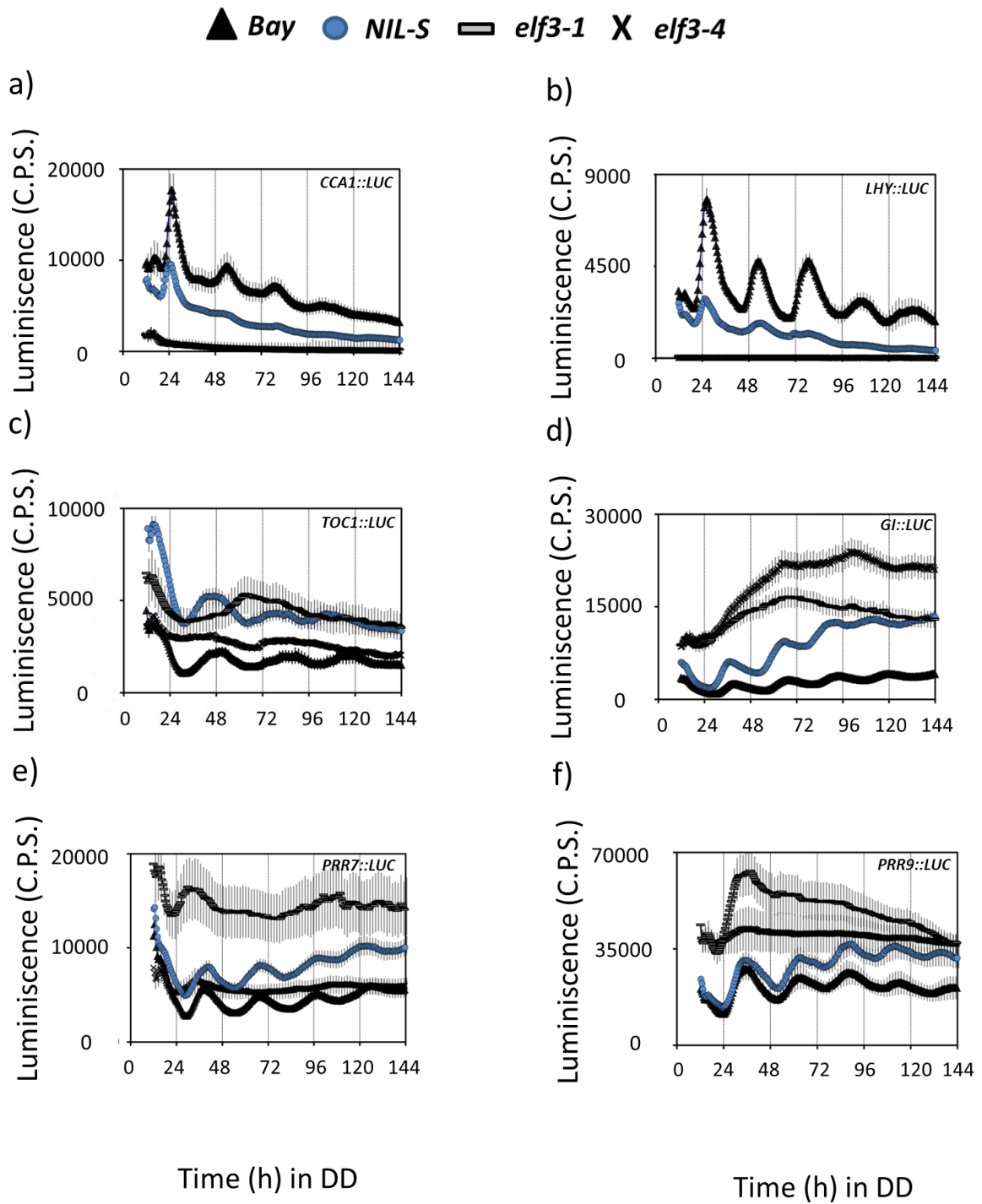
Figure 3.9: Clock network in *ELF3-Sha* in DD.

(a-f) Luciferase expression profile of different clock genes in NIL-S, *elf3-1*, *elf3-4* and Bay-0, in DD. (a) *CCA1::LUC* (b) *LHY::LUC* (c) *TOC1::LUC*, (d) *GI::LUC*, (e) *PRR7::LUC*, (f) *PRR9::LUC*. Error bars represent S.E.M. and is shown on every third reading. Note that *TOC1::LUC* expression is significantly higher than Bay and is similar to *elf3-1*.

This prompted me to examine the transcript profiles of major oscillator genes in NIL-S and compare this to Sha. This was also compared with the profiles of Bay and *elf3-4*. As *promoter::luciferase* lines for all major clock genes were not available in Sha, I monitored the mRNA abundance of *PRR7*, *PRR9*, *CCA1*, *LHY*, *TOC1* and *GI* in these lines under LL. The mRNA-accumulation pattern of these genes in NIL-S, Bay and *elf3-4* followed their respective luciferase expression profile under LL (Figure 3.10 a-f; for luciferase data see figure 3.8). Specifically, in addition to an early peak of expression of all genes studied, a de-repression for *PRR7*, *PRR9*, *TOC1* and *GI* was observed in NIL-S compare to Bay. Moreover, consistent with previous reports (Dixon et al., 2011; Herrero et al., 2012), arrhythmic low levels of *CCA1*, *LHY* and arrhythmic higher levels of *PRR7*, *PRR9*, *TOC1* and *GI* was detected in *elf3-4* (Figure 3.10 a-f). Interestingly, except TOC1, which displayed a de-repression, no difference in transcript profile of any other genes tested was observed in Sha, compared to Bay (Figure 3.10 a-f). Specially, the advancement of expression of peaks detected in NIL-S was absent in Sha (Figure 3.10 a-f), establishing the fact that some genetic components in this background buffers *ELF3-Sha* periodicity effect.

3.2.9 Cellular basis of ELF3-Sha phenotype

A higher expression of morning genes *PRR7* and *PRR9* and evening genes *TOC1* and *GI* suggests a repressive role of ELF3 (Kolmos et al., 2011). Thus, the short-period phenotype of *ELF3-Sha* could be due to reduced level of ELF3. To test this hypothesis, I assessed mRNA abundance of *ELF3* in NIL-S and Bay-0. I found that the mean level of *ELF3* transcript was slightly higher in NIL-S compare to Bay-0 (Fig 3.10 g). Thus transcript levels of ELF3 in Sha do not explain the short-period phenotype. I then estimated the amount of protein generated in fluorescence tagged ELF3-Bay-YFP and ELF3-Sha-YFP lines. As measured by YFP fluorescence signals, the overall amount of ELF3-Sha was higher compared to ELF3-Bay, (Fig 3.11 a-d). This elevated level of ELF3 protein thus also did not explain the short-period phenotype of ELF3-Sha, as ELF3-OX exhibits long-period (Covington et al., 2001; Herrero et al., 2012). A more detailed comparison of ELF3-Bay-YFP
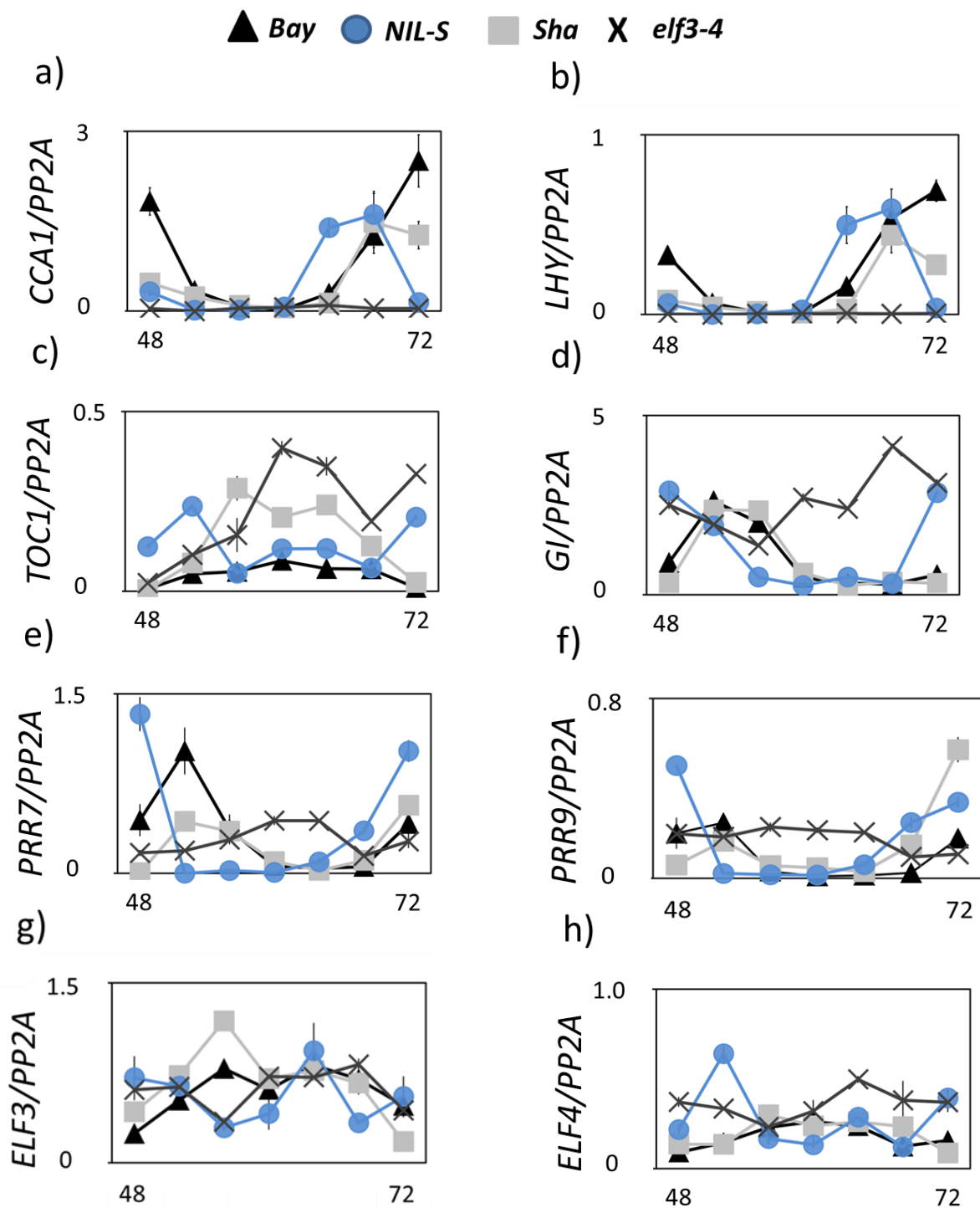
Figure 3.10: Clock network remains unaltered in Sha.

Transcript accumulation of different clock genes in Bay-0, Sha, NIL-S and *elf3-4* under LL. (a) *CCA1::LUC* (b) *LHY::LUC* (C) *TOC1::LUC* (d) *GI::LUC* (e) *PRR7::LUC* (f) *PRR9::LUC*. Error bars represent standard deviation of three technical repeats. Expression levels are normalized for *PROTEIN 19 PHOSPHATASE 2a subunit A3* (*PP2A*). For this experiment, plants were entrained for 7 days under 12L:12D cycles. The following day at ZT0, plants were put under LL for 2 days. Samples were then taken every 4 hours starting at ZT48.

and ELF3-Sha-YFP revealed marked localization differences. I observed that the formation of distinct nuclear bodies, a characteristic of ELF3, (Herrero et al., 2012; Yu et al., 2008) was markedly attenuated in ELF3-Sha. Whereas, these nuclear bodies were clearly observable in ELF3-Bay (Figure 3.11 a, b). Moreover, the preferential nuclear localization of ELF3 was compromised in ELF3-Sha. It also displayed a considerably higher amount of cytoplasmic content compared to ELF3-Bay (Figure 3.11 a-d). Quantitatively, the nucleus to cytoplasmic ratio of ELF3-Sha was four times less than ELF3-Bay (Figure 3.11 e). Based on these observations, I concluded that it is not the level but the inappropriate sub-cellular distribution of ELF3-Sha that associates to its short-period phenotype.

3.2.10 ELF3-Sha displayed unaffected developmental morphology

elf3 null mutants display photoperiod independent early flowering and unconstrained elongated hypocotyl (Liu et al., 2001; Nozue et al., 2007; Nusinow et al., 2011; Zagotta et al., 1996). To inquire if *ELF3-Sha* is also defective in these developmental processes, I examined the flowering time of the HIFs (89-S and 539-B), NIL-S and the Bay-0 wild type under long day and short day. I did not observe any significant difference in the flowering time under long days. However, under short days, the HIF 89-S line showed slightly advanced flowering time compared to HIF 539-B, as was measured by number of days to bolt (Figure 3.12 a). Extending on flowering time, I then measured the hypocotyl length of HIFs, NIL-S, Bay, Sha and *elf3-1*, under short days (SD), constant red light (RR) and constant blue light (BB). Similar to lack of flowering-time phenotype, I did not observe any significant difference in the hypocotyl length of HIFs, NIL0, and Bay-0 in RR or BB. In short days, HIF 89-S displayed a marginally elongated hypocotyl compared to HIF 539-B, but four times shorter than the *elf3-1* (Figure 3.12 b). Interestingly, contrarily to lack of periodicity phenotype, a significant difference in hypocotyl length was detected between Sha and Bay under all conditions tested, with Sha being shorter in hypocotyl length (Figure 3.12 b).
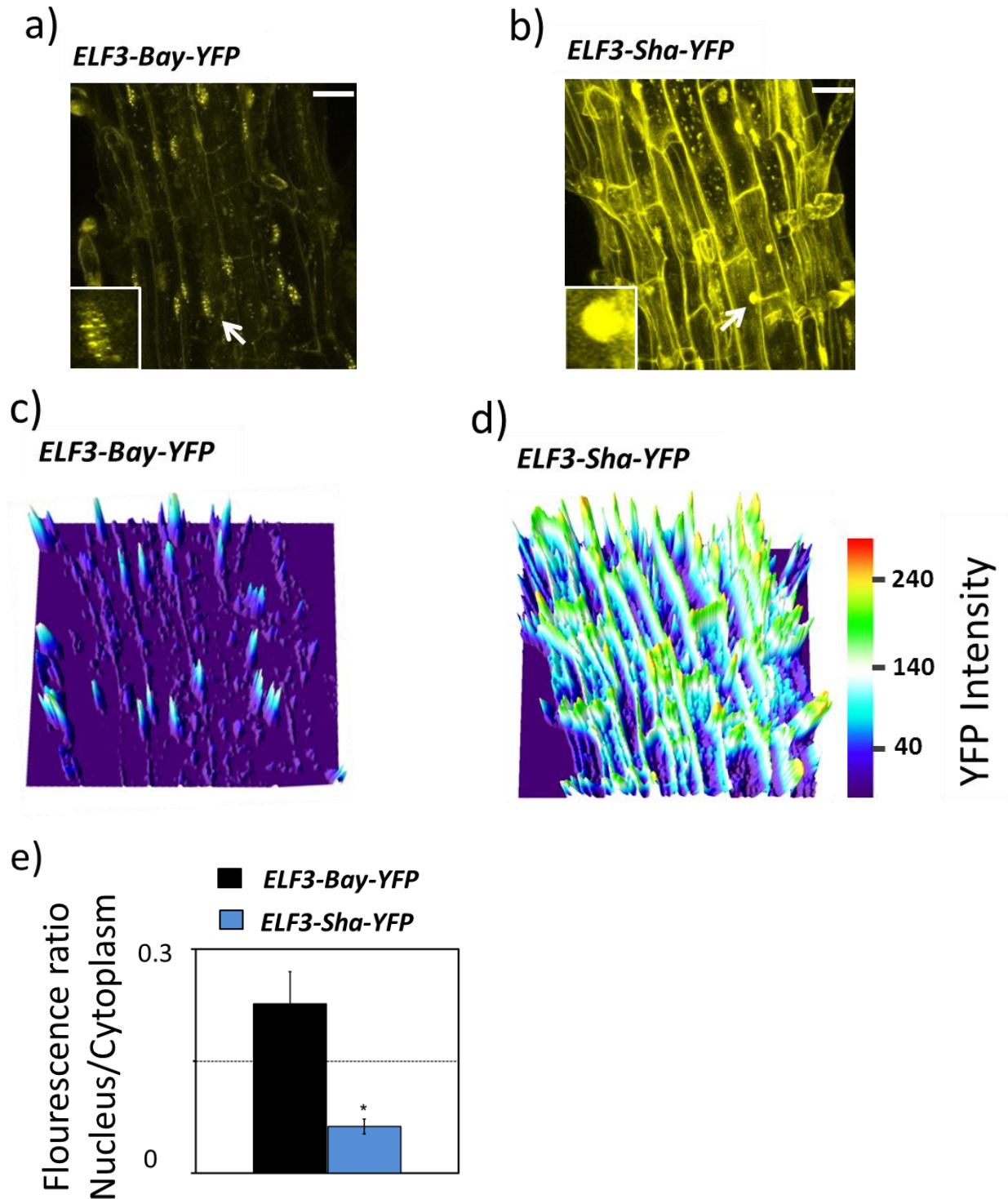
Figure 3.11: ELF3 sub-cellular localization is defective in *ELF3-Sha*.

(a) and (b) show maximum intensity projection of ELF3-YFP localization in root cells of *ELF3-Bay-YFP* (a) and *ELF3-Sha-YFP* (b). Arrows indicate nuclei that are four times magnified and shown in small insets at the bottom of (a) and (b). Note that the ELF3-Bay-YFP forms distinct nuclear (continue)

foci whereas, ELF3-Sha-YFP is diffused in the nucleus. Scale bar is 20µm. (c) and (d) display the YFP intensity distribution of (a) and (b) in thermal units, respectively. Same settings of the 'interactive 3D image' plugin of 'ImageJ' were used to plot the graphs. Note that The ELF3 cytoplasmic contents were higher in ELF3-Sha-YFP as compared to ELF3-Bay-YFP. (e) shows the nucleus to cytoplasmic fluorescence ratio of ELF3-Bay-YFP and ELF3-Sha-YFP, as calculated by Image J. Error bars represent S.E.M., n=3. Significance as described in Figure 3.1. The representative data of three independent experiments and three independent transgenic lines is shown.
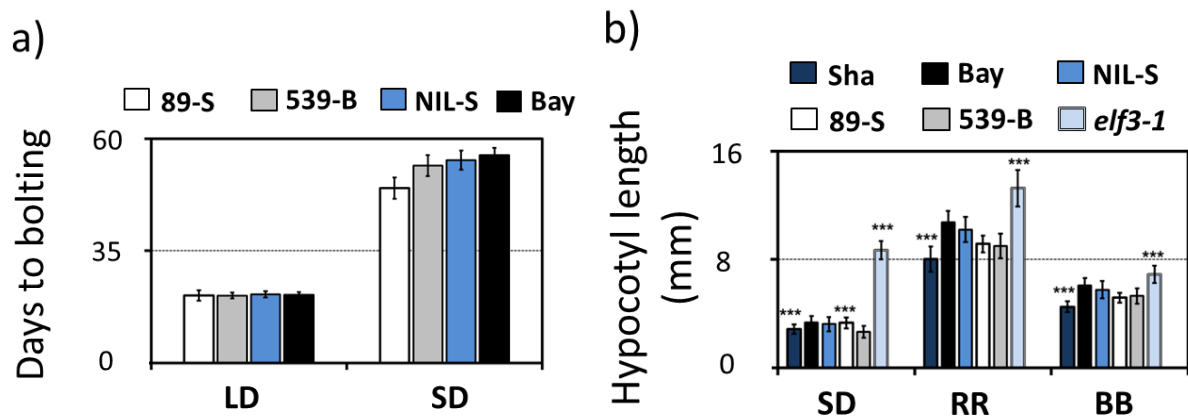
Figure 3.12: Flowering time and hypocotyl growth remain unaffected in *ELF3-Sha*.

(a) Flowering time of HIF 89-S, HIF 539-B, NIL-S, and Bay under long day (16L:8D) and short day (8L:16D). Flowering time was counted as number of days at the appearance of 1cm bolt. (b) Hypocotyl length of HIF 89-S, HIF 539-B, NIL-S, *elf3-1,* Sha and Bay under short day (8L:16D), in RR (15 µmol m-2s-1), and in BB (15 µmol m-2s-1). Error bars represent the standard deviation. Significance as described in figure 3.1 compared to Bay-0, except HIF 89-S, which was compared to HIF 539-B.

3.3 Discussion

Previous studies had established ELF3 as a required component for the generation of circadian rhythms and modulation of light and temperature inputs to the clock (Covington et al., 2001; Kolmos et al., 2011; Thines and Harmon, 2010). However, most *elf3* mutants used in these studies were identified in mutagenesis screens and many of them display arrythmicity under free-running conditions of LL and DD (Dixon et al., 2011; Herrero et al., 2012; Liu et al., 2001; Zagotta et al., 1996). In this chapter, I reported the cloning and characterization of *ELF3-Sha* as a natural allele of *ELF3* that displayed a light dependent short-period-phenotype. Unlike loss-of-function mutants, this allele displayed robust circadian rhythms under continuous light. However, in prolonged darkness the expression

profiles of many clock genes were almost arrhythmic. Connection of these clock phenotypes of *ELF3-Sha* to its cellular basis led me to hypothesize that sub-cellular distribution of ELF3 play pivotal role in its normal function in control of circadian clock.

The identification and validation of QTLs by genetic mapping is a well-established procedure. To understand the molecular basis of the locus by characterizing a quantitative trait locus (QTL) to a quantitative trait nucleotide (QTN) level is still an arduous task (Koornneef et al., 2004). Using positional isolation and transgenic complementation approaches, I was able to successfully clone *ELF3-Sha* to a single nucleotide resolution. My results revealed that the acceleration of *ELF3-Sha* periodicity in light and loss of robustness in darkness was associated to a 'C' to 'T' nucleotide transition in middle domain of *ELF3* that results in an encoded Alanine to Valine residual change (Figure 3.5). Two proteins, ELF4 and GI, have been reported to interact with this domain of ELF3 (Herrero et al., 2012; Yu et al., 2008). Thus, a defect in the binding affinity of any of these proteins could explain *ELF3-Sha* light-modulated periodicity acceleration.

Unlike arrhythmic nature of null mutants, I found that *ELF3-Sha* displays differential behavior under LL and in DD. Under LL, besides periodicity acceleration, this allele exhibited robust rhythms. In contrast, in prolonged darkness, the precision and robustness was compromised and that was promptly recovered on re-exposure to light (Figure 3.7, Figure 3.9). Similar phenotype in darkness has also been reported for clock and light regulated gene (*ZGT*) that showed robust rhythms in light and arrhythmia in prolonged darkness (Xu and Johnson, 2001). Thus I speculate that like *ZGT*, regulation of *ELF3* could also be controlled by both the circadian clock and by light (directly or indirectly). This idea is further strengthened with previous observations that demonstrated that the stability of ELF3 in nucleus largely depends on light (Liu et al., 2001). Based on these results, I propose that in darkness, a minimal amount of *ELF3-Sha* that is required for its proper function might get degraded rapidly, compare to *ELF3-Bay.* A COP1 mediated proteasomal decay could account for this process (Yu et al., 2008), which is discussed in detail in Chapter 5.

Chapter Four: **Quantitative genetics and genomics study of circadian clock and flowering-time variation in *Arabidopsis thaliana***

Whole genome re-sequencing of Ws-2, Strand and Tnz-1 genomes was performed by Jun Cao at the Max Plant Institute for Developmental Biology, Tübingen. Genome assembly and SNP calling was performed by Xue Dong and Karl Nordstorm at Max Plank Institute for Plant Breeding Research Cologne.

4.1 Introduction

Natural genetic variations in Arabidopsis accessions provide a rich resource of allelic variants that has been exposed to natural selection in contrasting ecological habitat, and thus, hold promise to answer difficult evolutionary questions which could not be properly addressed using induced mutants (Alonso-Blanco and Koornneef, 2000; Mitchell-Olds and Schmitt, 2006). How plants evolved the mechanism of integration of environmental signals to control complex developmental processes is one such question that could be interrogated by exploiting natural-genetic resources. As one example, I described in chapter 3, the identification and isolation of a natural allele of *ELF3 (ELF3-Sha)* that was detected in Bay X Sha RIL population (Chapter 3). The characterization of this allele provided an insight, how a single nucleotide variant can affect a broad spectrum mechanism like circadian clock. This success encouraged me to further exploit the ''natural variation-treasure'' to investigate the role of circadian clock in mediating the complex developmental process of flowering-time.

In Arabidopsis, QTL mapping using RILs has been successfully employed to investigate the genetic architecture of the circadian clock (Boikoglou et al., 2011; Darrah et al., 2006; Michael et al., 2003a; Swarup et al., 1999). In many of these studies, leaf-movement rhythms were detected and used to estimate the periodicity of RIL lines under free-running conditions (Michael et al., 2003a; Swarup et al., 1999). However, this approach was not suitable for large RIL sets. This limitation was overcome in recent studies where pre-existing RIL populations were modified with a *CCG::LUC* and used to estimate various circadian parameters (Boikoglou et al., 2011; Darrah et al., 2006). The luciferase-based modification enabled high-throughput screening of large RIL sets; however, the positional effect of the *CCG::LUC* transgene on phenotype cannot be ignored in these RILs. To eliminate this drawback, RILs can be generated by crossing one of the parental accessions harboring stable *CCG::LUC* transgene to another wild type parental accession, and then selecting for lines in the resultant progeny that are luciferase positive. These lines can then be selfed using SSD. In this way, all of the resultant lines will have same fixed

position of the *CCG::LUC*, thus the possibility of phenotypic variation conferred by positional effect of the transgene will be eliminated.

Based on the above strategy, I generated three independent RIL sets by making pair-wise crosses among three geographically distant Arabidopsis accessions: Strand, Ws-2 and Tnz-1. Complete genomes of all three parental accessions were re-sequenced using next generation Solexa platform. This revealed extensive polymorphisms which allowed for dense linkage maps in the three RIL populations. Further, quantitative genetics and genomics approaches were applied to characterize these diverse RIL sets for circadian periodicity and flowering-time traits. Several QTLs were detected contributing to the phenotypic variation for these traits. Based on DNA-sequence differences, possible candidate genes underlying the QTLs were predicted. Furthermore, large effect QTLs were validated in NILs. Collectively, these results provided an insight about the possible genetic architecture of the circadian clock and flowering-time variation in Arabidopsis.

4.2 Results

4.2.1 Selection of the parental accessions

Three Arabidopsis accessions Strand, Ws-2 and Tnz-1 were selected for quantitative-genetic studies of circadian-clock mediated flowering time pathway. Geographical location and climatic differences at local habitat were the primary criteria for the selection of these accessions. The detail information about the location and environment differences of these accessions is provided in Chapter 2.2.1. Briefly here, Strand and Ws-2 are two northern accessions that grow at latitude 68°N and 51°N, respectively. Strand is a coastal accession and Ws-2 is from continental climate. Contrarily, Tnz-1 grows on high mountains of Kitumbeine forest near the equator (-3°N). The local habitat for these accessions varies tremendously and this includes wide differences in their latitudinal positions. Notably, light and temperature are two primary factors that follow distinctly different yearly pattern of variation on these locations (Table 2.13). Specifically,

both light and temperature varies tremendously on northern altitudes (where Strand and Ws-2 grow), whereas, these elements remain almost same throughout the year on equator (where Tnz-1 grow). As these factors adjust the dynamics of circadian clock and flowering-time, I expected that different gene pool for these pathways would have been selected in these accessions due to their adaptation to different conditions. Based on this assumption, I selected these three accessions and transformed them with *CCR2::LUC* promoter-reporter constructs. This makes them useful for high-throughput circadian clock assays. Next, I measured both the free-running period of these accessions as a measure of circadian-clock activity and the number of days to bolting as a flowering-time trait. Circadian periodicity was measured under constant light, after entraining the plants under light dark (LD) condition (for detail experimental conditions, see Chapter 2.2.) and flowering-time was measured under long day (LgD) and short day (StD).
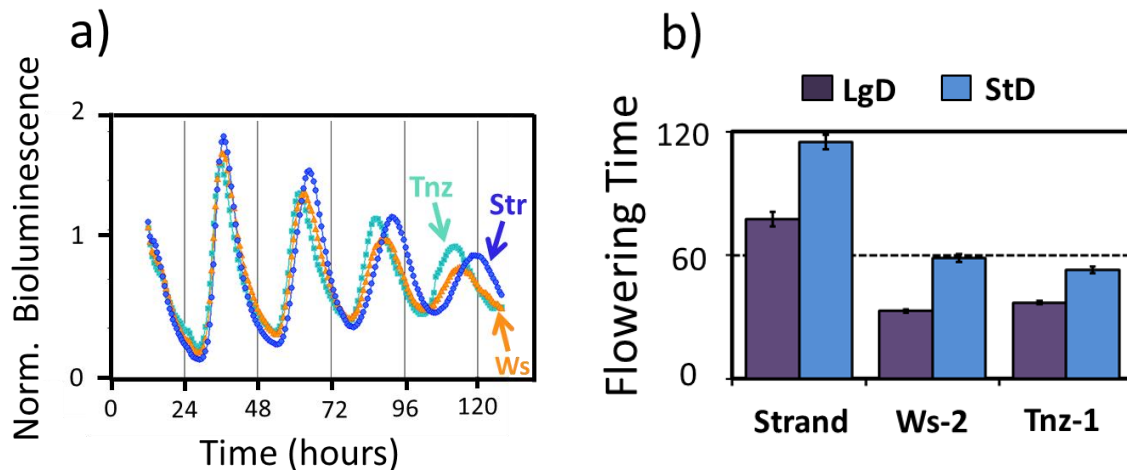
Figure 4.1: Circadian clock and flowering-time assay of parental accessions.

(a) Free-running profile of *CCR2::LUC* in Strand, Ws-2 and Tnz-1 under LL. (b) Flowering-time measured as number of days to bolting in Strand, Ws-2 and Tnz-1, under LgD and StD. Error bars represent standard deviation. n = 24.

Expectedly, all three accessions displayed considerable phenotypic variation for both traits analyzed. Specifically, for circadian periodicity, I found that the speed of oscillations in Tnz-1 is faster, compare to both Strand and Ws-2. Further, in Ws-2, the clock runs faster compared to Strand (Figure 4.1a). The comparison of flowering time in these accessions showed that under both LgD and StD, Strand was later flowering than both Tnz-1 and Ws-2, whereas, flowering time for Tnz-1 and Ws-2 was almost similar under both conditions (Figure 4.1b). From these results, I concluded that the northern accessions Strand and Ws-2 are physiologically different than the equatorial accession Tnz-1, and thus, these accessions could be utilized as a suitable genetic resource for quantitative studies, including QTL mapping.

4.2.2 Re-sequencing of the parental accessions

4.2.2.1 Genome assembly, read mapping and SNP calling

Rapid advances in next generation sequencing (NGS) technologies have made it possible for the individual labs to re-sequence the whole genomes. As an assistant for QTL mapping using RILs, the whole genome sequences (WGS) of the parental accessions can greatly facilitate the processes of linkage-map construction, and further, positional isolation of the candidate genes. To exploit the full benefits of these technological advances, I re-sequenced the complete genomes of all three parental accessions.

The re-sequencing was carried out on Solexa\Illumina platform that provided the sequence output in the form of short pair-end reads. For the Ws-2 genome, a 76bp pair-end reads with 15X genome-wide coverage was obtained, whereas, longer pair-end reads of 120bp with genome coverage of 26.6x and 23.4x was obtained for Strand and Tnz-1 genomes, respectively (Table 4.1). The coverage was sufficient to assemble these genomes by aligning against the reference genome (Columbia, TAIR9).

The genome assemblies and alignments were followed by detection of polymorphism between the three target genomes and the Columbia reference genome. Enormous polymorphisms in the form of SNPs and InDels (insertions and deletions) were

detected in all genomes. Specifically, the overall SNPs detected in the three genomes after aligning raw sequence reads to the reference genome were ~0.58 million, 1.10 million and 0.83 million, for Ws-2, Tnz-1 and Strand genomes, respectively (Table 4.1). These raw SNPs were then further filtered and only SNPs with a Phred quality score of above 25 were kept. Here, a Phred score of 20 corresponds to 1% error in SNP calling. Thus, the final number of high-quality SNPs obtained had well above 99% probability to be accurate. With this filtering, ~45% of SNPs were discarded as low-quality. Using the same quality parameters, InDels were also filtered. Eventually, I used only those SNPS and InDels detected with high quality.

Table 4.1: Details of NGS data output and variant calling.

				Number of SNPs		No. of InDels
Accession	NGS Plateform	Read Length (bp)	Coverage (X)	Raw	Filtered (q=25)*	Filtered (q=25)*
WS-2	Solexa	2 X 76	15	581175	377764	53327
TNZ-1	Solexa	2 X 120	23.4	1109792	726412	87989
Strand	Solexa	2 X 120	26.6	838671	550381	69590

* Phred quality score (q), where q, ranges from 1-40.

The final number of SNPs and InDels varied considerably among three parental accessions. Precisely, ~0.72 million SNPs were detected in Tnz-1, whereas, the SNPs detected in Ws-2 and Strand were ~0.37 million and ~0.55 million, respectively. Further, compared to SNPs, the presence of InDels was detected less frequently for all three genomes. However, their relative difference pattern compared to Columbia genome was same, with Tnz-1 being more polymorphic compared to Ws-2 and Strand. Accurately, 0.087 million InDels were detected in Tnz-1 genome, compared to 0.053 million and 0.069 million in Ws-2 and Tnz-1 genomes, respectively (Table 4.1, Figure 4.2a).

Further, I calculated the approximate average distance between these polymorphisms, as SNPs/Kb and as InDels/Kb. As abundance of polymorphisms was considerably higher in Tnz-1 than Ws-2 and Strand, accordingly, a more dense distribution of SNPs and InDels were observed in Tnz-1 genome. On average 5.81 SNPs/kb and 0.70 InDels/Kb were detected in Tnz-1, whereas, these numbers were relatively lower in Ws-2 (3.02 SNPs/Kb, 0.42 InDels/Kb) and Strand (4.4 SNPs/Kb, 0.55 InDels/Kb) (Figure 4.2). Collectively, the comparison of these three genomes to the 'Columbia' genome revealed that Tnz-1 is more polymorphic than Ws-2 and Strand.
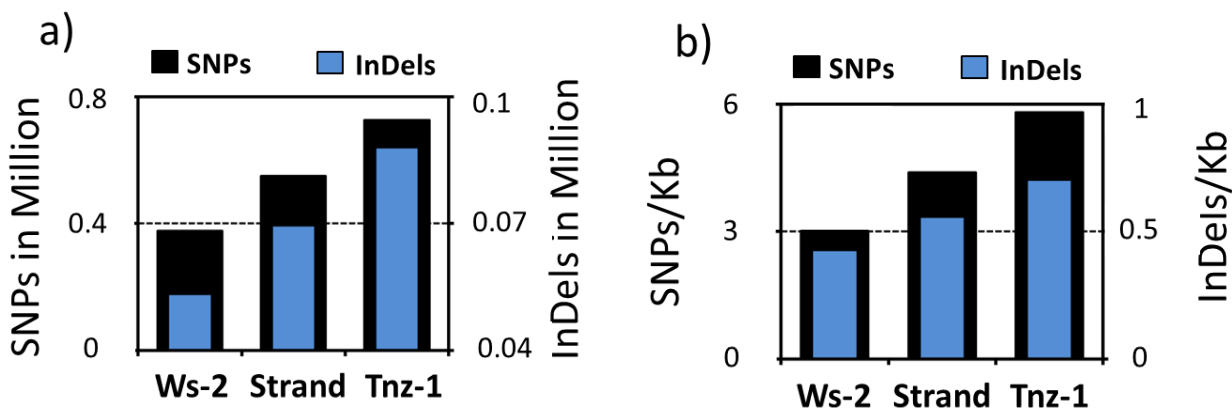
Figure 4.2: Distribution of SNPs and InDels in Tnz-1, Ws-2 and Strand genomes.

(a) SNPs and InDels detected in Ws-2, Strand, and Tnz-1. Left Y-axis represents SNPs in million, right Y-axis represents InDels in million. (b) SNPs and InDels per Kb in Ws-2, Strand, and Tnz-1. Left Y-axis represents SNPs/Kb, right Y-axis represents InDels/Kb.

4.2.2.2 Comparative polymorphism analysis of Tnz-1, Ws-2 and Strand genomes.

I further analyzed the SNPs data to estimate the extent of polymorphism shared among these three genomes. For that, I calculated the number of common and unique SNPs in all three genomes by making pair-wise comparison among them. I found that Tnz-1 and Ws-2 shares the least number of SNPs between them, whereas, northern accession Strand and equatorial accession Tnz-1 shared the highest. Moreover, a relatively low number of shared SNPs were detected in northern accessions Ws-2 and Strand (Figure 4.3).
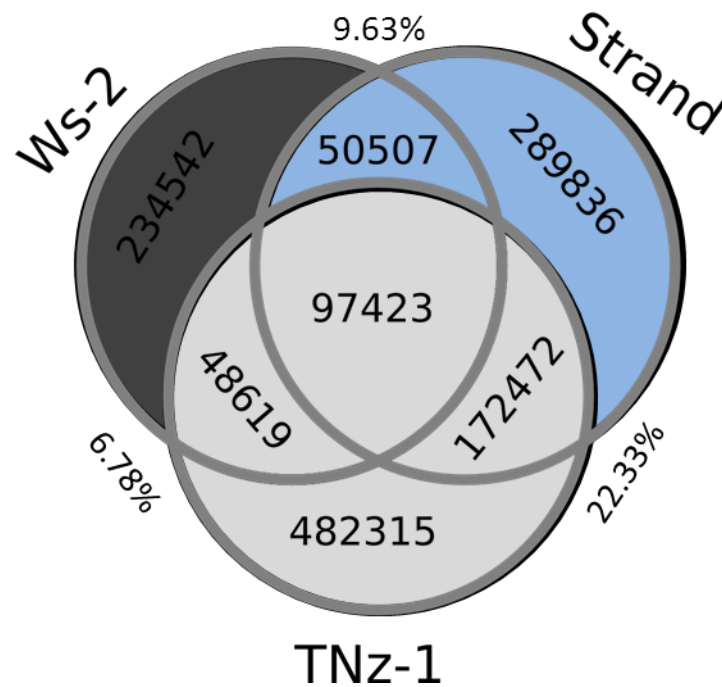
Figure 4.3: Comparative analysis of SNPs in Tnz-1, Ws-2 and Strand genomes.

Numbers of shared and unique SNPs are shown inside the circles for all three accessions. Numbers outside the circle represent the percentage of shared SNPs between respective genomes. This percentage was calculated by dividing the number of shared SNPs to the total number of SNPs detected in those genomes (% shared SNPs = No. of shared SNPs / (total no. of SNPs in genome A + total no. of SNPs in genome B). Total number of SNPs for each genome is given in Table 4.1.

Further analysis revealed that only 6.78% of the total number of Ws-2 and Tnz-1 SNPs were shared, whereas, in the case of Tnz-1 and Strand, this number was 22.33%. The percentage of shared SNPs between Ws-2 and Strand was 9.63. Finally, I calculated the number of unique SNPs in all three genomes and found that Tnz-1 contained the highest number of unique SNPs, which were almost twice as many as were detected in Ws-2 and Strand. Taken together, these results provided an indication that on genome-wide scale Ws-2 and Strand are genomically distinct from each other, even though they are geographically less isolated from each other than they are to Tnz-1. Furthermore, Tnz-1 and Strand are more different from each other than any other comparison.

4.2.2.3 Chromosomal distribution of SNPs and InDels

Distribution of SNPs and InDels over individual chromosomes was calculated to estimate the relative contribution of each chromosome to the total number of polymorphisms. A direct relationship between the length of the chromosome and number of unique SNPs was observed, with longer chromosome supplying more variation to the genome (Figure 4.4a). Specifically, for all three genomes, chromosome 1 and chromosome 5 contributed ~50% of the detected polymorphisms, whereas, the rest of the ~50% was contributed by the other three chromosomes (2, 3, and 4). Furthermore, the relative participation of each chromosome in overall number of SNPs followed a similar trend for all three genomes. Chromosome 1 contributed ~25%, chromosome 2 ~17%, chromosome 3 ~19%, chromosome 4 ~16% and chromosome 5 ~23% of the total variation. However, a small genomic aberration of polymorphism density was detected in chromosome 3 and chromosome 4 in Ws-2 genome, where 2% more and 2% less participation of SNPs were detected for chromosome 3 and chromosome 4, respectively (Table 4.2).
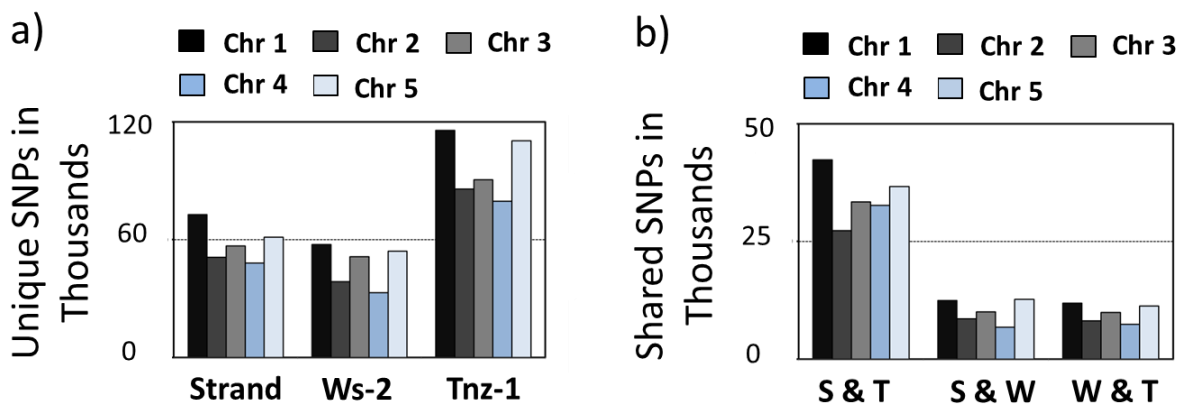
Figure 4.4: Distribution of unique and common SNPs over individual chromosomes.

(a) Number of unique SNPs contributed by each chromosome for all three genomes. (b) Number of common SNPs contributed by each chromosome all pair-wise genome combinations. 'S & T' denote SNPs shared between Strand and Tnz-1 genomes, 'S & W' between Strand and Ws-2 genomes and 'W & T' between Ws-2 and Tnz-1 genome. Note that the numbers of both unique and shared SNPs are directly related to the length of the chromosome.

The estimation of unique SNPs/Kb revealed a uniform distribution of SNPs over all chromosomes for each genome. I noticed that, on average, ~4.0 SNPs/Kb were detected on each chromosome of Tnz-1 genome, whereas, unique SNPs/Kb detected for Ws-2 and Strand were 2.0 and 2.5, respectively (Table 4.2). A similar pattern of chromosome-wide distribution was also observed for the shared SNPs for all genomic combinations (Figure 4.4b, Table 4.3).

I further examined the chromosome-wide distribution of InDels in all genomes and found that they also followed the same uniform distribution pattern as was observed for SNPs. Consistent results were obtained as the InDels distribution pattern followed the SNPs distribution pattern and covered complete length of the chromosomes with uniform gaps.

Table 4.2: Chromosome-wide distribution of unique SNPs in all genomes studied.

Chr	Total SNPs			% contribution			SNPs/Kb		
	Strand	Ws-2	Tnz-1	Strand	Ws-2	Tnz-1	Strand	Ws-2	Tnz-1
1	72810	57555	115757	25.12	24.53	24.00	2.427	1.91	3.85
2	50978	38607	85877	17.58	16.46	17.80	2.60	1.96	4.38
3	56772	51269	90625	19.58	21.85	18.78	2.46	2.22	3.94
4	48028	32949	79602	16.57	14.04	16.50	2.66	1.83	4.42
5	61250	54162	110445	21.13	23.09	22.89	2.35	2.08	4.24

Table 4.3: Chromosome-wide distribution of shared SNPs in all genomes studied.

Chr	Total SNPs			% contribution			SNPs/Kb		
	S & T	S & W	W & T	S & T	S & W	W & T	S & T	S & W	W & T
1	42387	12466	11876	24.58	24.68	24.43	1.41	0.42	0.40
2	27315	8561	8142	15.84	16.95	16.75	1.39	0.44	0.42
3	33414	10043	9913	19.37	19.88	20.39	1.45	0.44	0.43
4	32692	6742	7401	18.95	13.35	15.22	1.82	0.37	0.41
5	36666	12697	11288	21.26	25.14	23.22	1.41	0.49	0.43

Centromere regions were the only locations where relatively large gaps of polymorphism detection were observed. The largest centromere gap of 10Kb was detected for chromosome 1 in the Ws-2 genome (Figure 4.5). Taken together, I concluded that for all three genomes, both SNPs and InDels were essentially uniformly distributed and covered the complete length of all chromosomes. This thus ensured high quality sequence polymorphisms data of parental accessions. I extensively utilized this data for my further research. Specially, it was used to design targeted markers for genotyping the RILs and to promptly identifying the sequence differences in potential candidate genes underlying interesting QTLs. I will discuss these processes in detail in next subsections of this chapter.

a)

b)

c)
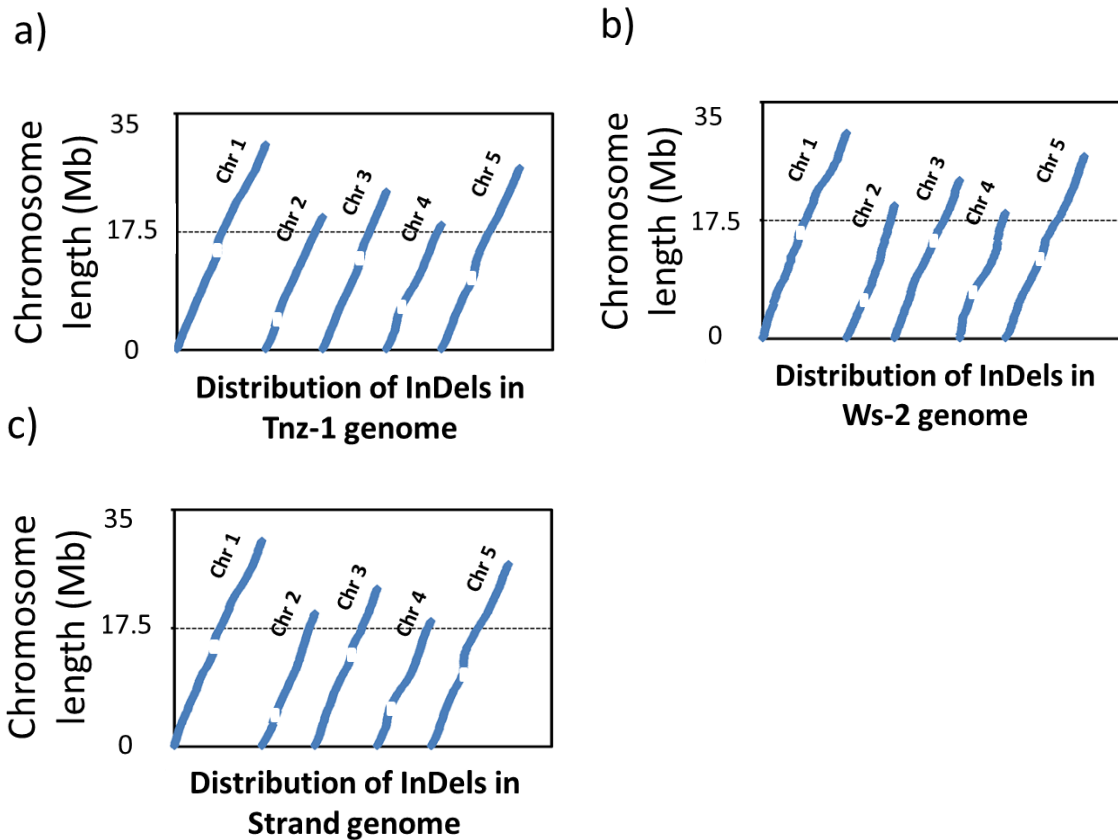
Figure 4.5: Distribution of InDels over Tnz-1, Ws-2 and Strand genomes.

Distribution of InDels in (a) Tnz-1 genome (b) Ws-2 genome and (c) Strand genome. The small white boxes indicate the predicted centromere position.

4.2.3 Generation of Recombinant Inbred Lines (RIL) populations

To generate appropriate RILs, geographical divergence and phenotypic variation in the traits of interest are two important criterions for the selection of the parental accessions (Koornneef et al., 2004; Loudet et al., 2002; Weigel, 2012). As I discussed in the previous section (4.2.1) that Strand, Ws-2 and Tnz-1 exhibits extensive variation in clock periodicity and flowering time. Their genomes also display extensive polymorphisms. Furthermore, they belong to geographically distant locations. Thus, these accessions were selected as suitable parental accessions for RILs generation and QTL mapping.

I generated three different RIL sets by making pair-wise crosses among Strand, Ws-2 and Tnz-1, in all possible combinations. Crossing was then followed by self-fertilization for several generations. This resulted in three independent RIL sets, (1) Ws-2 X Tnz-1, (2) Ws-2 X Strand and (3) Strand X Tnz-1. Each of these RIL set comprised different sub-populations. Specifically, the sub-populations in Ws-2 X Tnz-1 RIL set include WT1W and T1W. In RIL set Ws-2 X Strand, they were WS, WSW, and 35WS. Finally, in RIL set Strand X Tnz-1, these include ST1, ST1T1, and ST1S. All of the individual plants in each RIL population were harboring stable *CCR2::LUC* transgene. The detail process for generation of these RIL sets and their sub-populations is described in chapter 2.2.2.

4.2.4 Characterization of Ws-2 X Tnz-1 RIL set

4.2.4.1 Generation of Linkage map

In total, Ws-2 X Tnz-1 comprised 78 lines. All of these lines were genotyped using 55 evenly distributed markers that covered the complete lengths of all chromosomes (Appendix). Three different kinds of markers were used. These include SSLP, InDels and SNPs markers. All SSLP markers were designed against the simple sequence repeats using Columbia genome as a reference, whereas, InDels and SNP markers were specially designed on the target regions of choice, which were polymorphic in Ws-2 and Tnz-1 genomes (Figure 4.6).
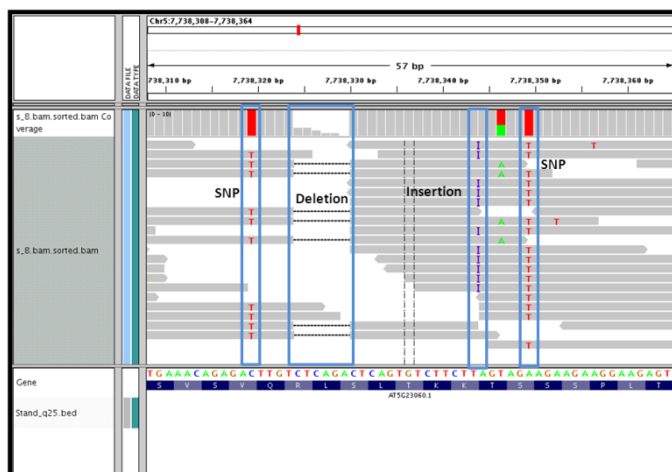
Figure 4.6: IGV genome browser showing polymorphic sites in Ws-2 and Tnz-1 genomes.
Red letter shows SNPs, dotted line represent Deletion, whereas purple letter is site of insertion.

Calculations of the contribution of parental genomes in RILs revealed that overall the contribution of Ws-2 genome was 63.1%, which was expected for a mixed population containing progenies of both backcross and F_2 lines (see chapter 2.2.2). A similar pattern of the contribution of parental genome was also observed in four chromosomes, with the exception being chromosome 3. A considerably high bias for selection of Ws-2 was detected on chromosome 3, where the overall contribution of Ws-2 genome was 79.1% (Table 4.4). Moreover, 6.7 % overall heterozygosity was detected, which was a little more than the expected 6.3%. One of the potential reasons for this increase could be the mixture of DNA that was extracted by bulking tissue samples from nine plants of the same line. Any line segregating for any locus will lead towards bias for heterozygosity.

Table 4.4: Genome participation of parental accessions.

Genotype	Genome participation (%)					
	Global	Chr. 1	Chr. 2	Chr. 3	Chr. 4	Chr. 5
Ws-2	63.1	59.7	59	79.1	58.6	58.7
Tnz-1	28.8	31.2	29.4	13.4	32.9	34.2
Ws-2/Tnz-1	6.7	8	7.9	7.2	6.2	6.3
Missing	1.4	1.1	3.6	0.3	2.3	0.8

A clear bias for selection of Ws-2 in chromosome 3 prompted me to check all markers for allelic segregation distortion. I found an overall increased in proportion of Ws-2 allele on all chromosomes with a Tnz-1/Ws-2 ratio of 1:2.3. However, this kind of segregation distortion is common in Arabidopsis and has been reported in several RILs (Alonso-Blanco et al., 1998; Lister and Dean, 1993; Loudet et al., 2002; Singer et al., 2006). A more dramatic aberration in segregation was found on chromosome 3, where the top arm of the chromosome displayed an extreme bias for selection of Ws-2 allele (Figure 4.7). An unavoidable selection for *CCR2::LUC* for several generations could be one explanation for this deviation on a genomes.
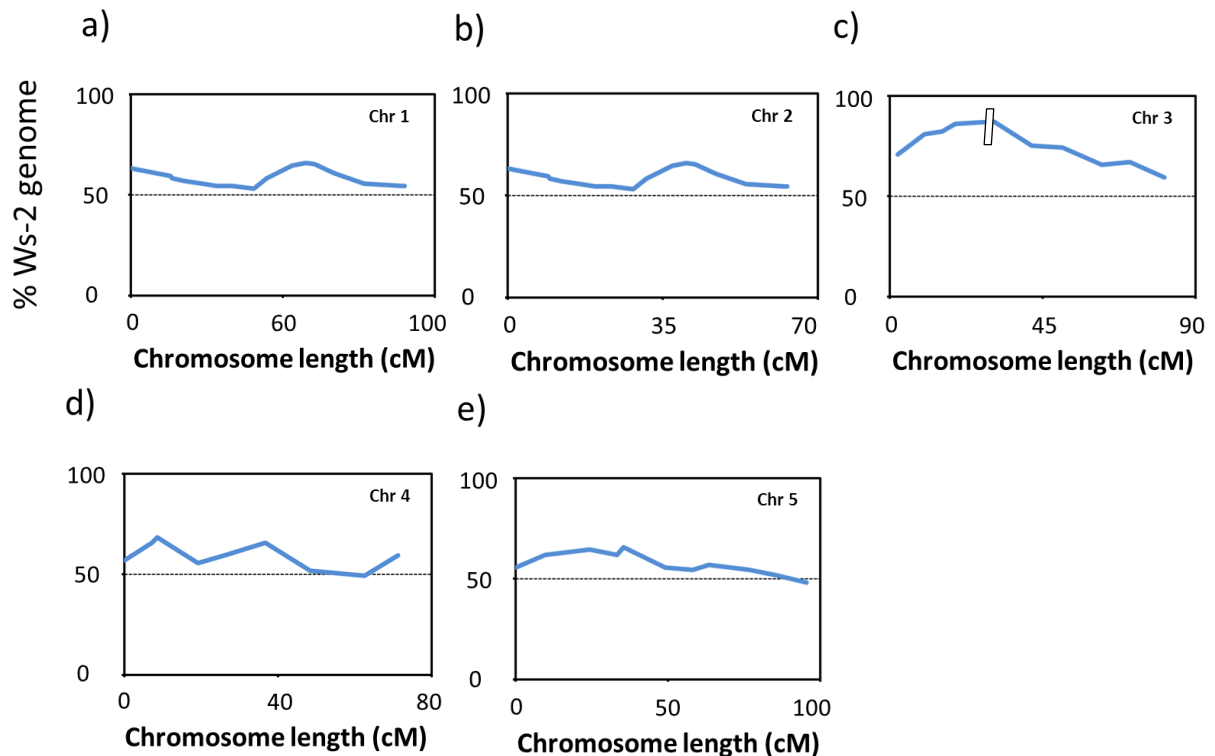
Figure 4.7 Segregation between Ws-2 and Tnz-1 alleles along the five chromosomes.

The percentage of Ws-2 allele is represented along Y-axis. W-axis represents chromosome length in centimorgans. Note that the top of chromosome 3 is highly distorted towards Ws-2 allele. Vertical box in (c) represents possible position of the *CCR2:LUC* transgene.

A genetic map was calculated using JoinMap 4 and is presented in Figure 4.8. All markers were assigned to the desired linkage group, without any ambiguity, corresponding to the order of their physical positions in the genomes. However, an overall ''shrinkage'' of chromosome 3 was observed, where the length of the chromosome was severely reduced. This problem was ''fixed'' by allocating positions to marker by taking information from other maps where they had already been assigned or through estimations based on their physical positions (Singer et al., 2006).

The five linkage groups of the genetic map represent a total of 411.9 cM, with an average marker to marker distance of 7.48 cM and a maximum distance of 15.8 cM on the top of chromosome 1. Most chromosomes showed a linear relationship between their physical and genetic length, with the exception at chromosome 4, which was longer than the chromosome 2, despite being physically the shortest one. This increase in the genetic length of chromosome 4 relative to chromosome 2 has been reported previously (Alonso-Blanco et al., 1998; Loudet et al., 2002).
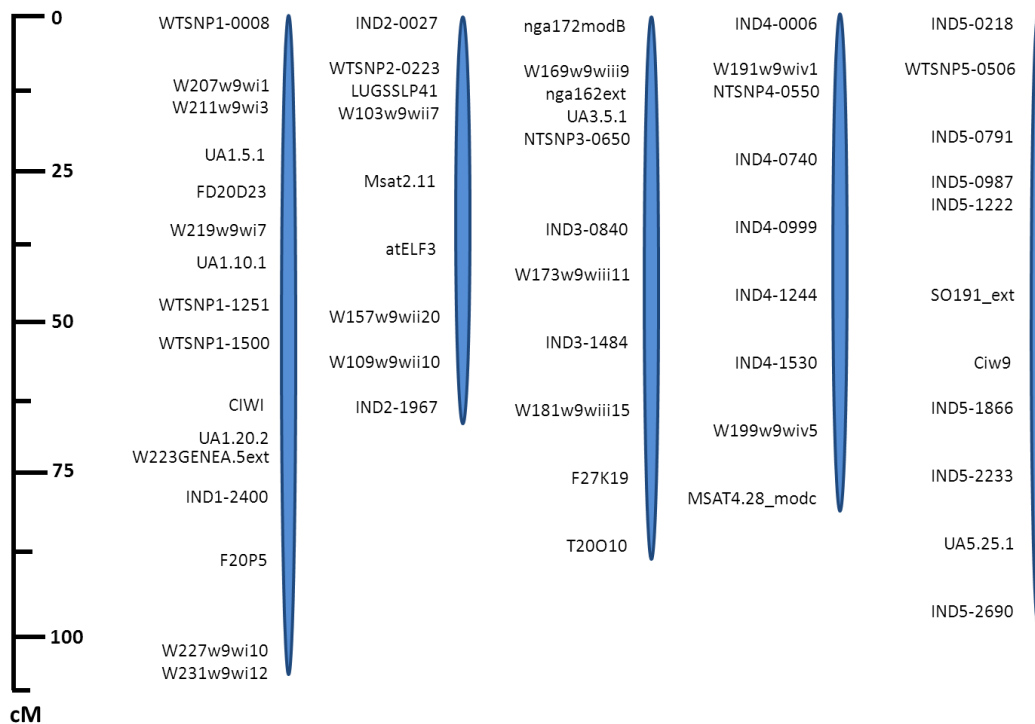
Figure 4.8: Genetic map of Ws-2 X Tnz-1 RIL population.

4.2.4.2 Phenotypic characterization of Ws-2 x Tnz-1 RIL set

Identifying novel components that contributes to light and temperature mediated circadian-clock variation was the main objective of the creation of these RIL sets. Additionally, I was also interested to know how variation in clock components modulates the induction of flowering time. To achieve these goals, I first scrutinized the Ws-2 X Tnz-1 RIL population for two parameters: speed of circadian oscillations, and flowering time. To further dissect the genetic components of light and temperature regulation of the clock, the speed of oscillation was measured after two preceding environmental conditions. These were, light-dark entrainment (LD) and warm-cool entrainment (WC). Detail experimental procedures and conditions are described in chapter 2.2.6.2. Briefly here, for LD entrainment, RILs were entrained for 7 days under 12L:12D cycles under constant temperature (22°C), whereas, for WC entrainment, the same RILs were entrained for 7 days under 12 hour 22°: 12 hour 16°C under constant light. The free-running profiles of *CCR2::LUC* oscillations for all RILs (LD entrain and WC entrain) were then measured under constant light and temperature, importantly at same time, in the same experiment. Thus, any difference observed in *CCR2::LUC* oscillation properties in a particular RIL after the two entrainment conditions must reflect the preceding environmental effect.

The *CCR2::LUC* profiles were analyzed to estimate the period length of each RIL after LD and WC. Enormous variation was detected after both entrainment conditions. I observed that the mean period of Ws X Tnz after WC entrainment was shorter than after LD entrainment (Figure 4.9 a,b). This acceleration in periodicity after WC entrainment was also observed in the individual RIL lines, where above 90% of the lines displayed shorter period after thermal cycles compared to light cycles. Further, I found that after both entrainment conditions, Tnz-1 exhibited a shorter period compare to Ws-2. Specifically, the period difference between two parental accessions was 0.63 hours after LD entrainment, whereas, a larger difference of 2.60 hours was observed after WC entrainment. Interestingly, the main contributor for this period difference after WC was Tnz-1 that displayed ample increase in speed of oscillations after WC cycles (Period-LD: 26.73, Period-

WC: 24.14) while Ws-2 maintained almost similar periodicity after both entrainments (Period-LD: 26.60, Period-WC: 26.69) (Figure 4.9 a,b). These results depicted that Tnz-1 is more responsive to temperature changes than Ws-2.

Frequency distribution analysis of periodicity revealed a transgressive effect in RILs where individual lines had phenotypic values both above and below the means values of the parents (Figure 4.9 a,b). Transgression in RILs is a common phenomenon and has been observed in many QTL studies involving clock parameters (Boikoglou et al., 2011; Darrah et al., 2006; Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999). Further, the period values for RILs were continuously distributed for both LD and WC, with the lowest and highest values at 24.5 and 29, respectively (Figure 4.9 a,b). This continuous variation confirmed circadian periodicity as a quantitative trait in this RIL population.
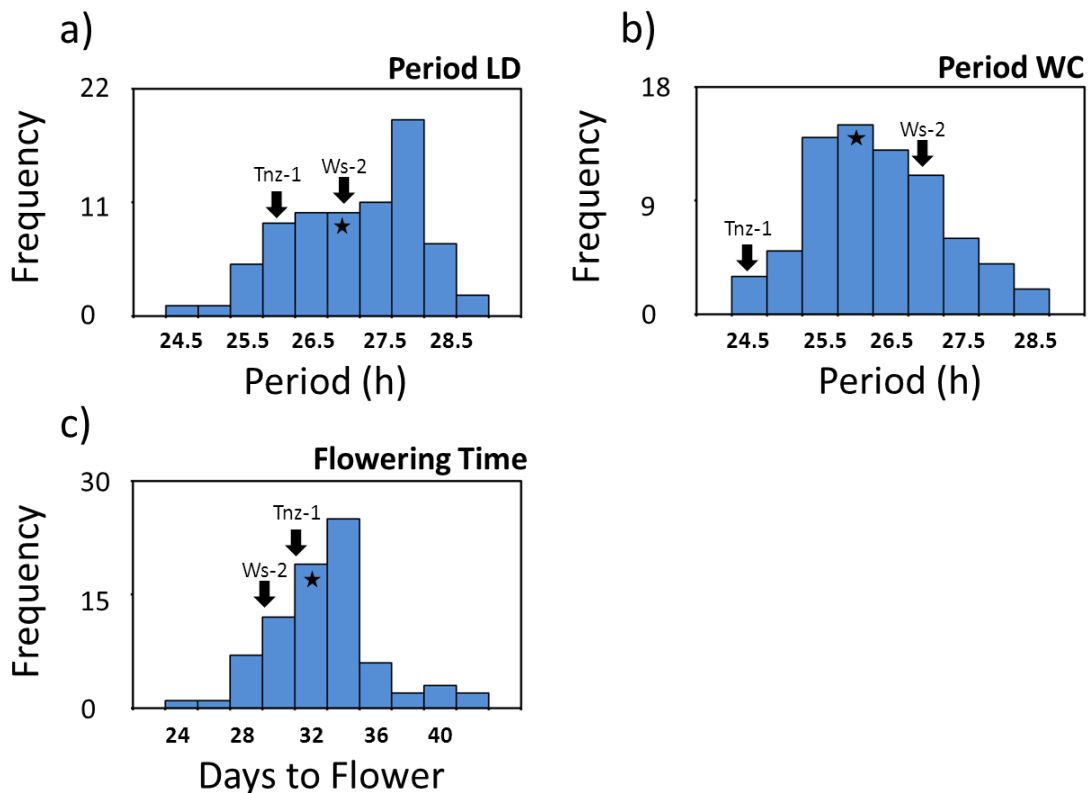
Figure 4.9: Distribution of circadian period and flowering time in Ws-2 X Tnz-1 RILs.

Distribution of mean period (a) after LD entrainment (b) after WC entrainment and (c) flowering-time. Mean values of parental accessions are marked with black arrows. Black stars represent population means. Bins are labeled with the upper bounds.

Flowering time was measured in as the number of days to bolting in a temperature-control greenhouse under conductive long days (see chapter 2.2.7). Though, no large difference in mean flowering time of both accessions was detected, the variation in flowering time of RILs was considerable. This defined transgression in the flowering-time trait. Similar to circadian periodicity, flowering time in the RILs varied continuously. This was between 24 to 42 days. This thus defines flowering time as quantitative trait in this RIL population (Figure 4.9 c).

Variations in the circadian clock can result in changes in flowering time, and thus, these two components seems to be linked to some extent (de Montaigu et al., 2010). To examine such a relationship in Ws-2 X Tnz-1 RILs, I performed correlation analysis between circadian periodicity and flowering time in the individual RILs. I could not detect any significant relationship between flowering time and circadian period (both after LD or after WC) (Figure 4.10). However, a strong positive relationship was observed between LD-period and WC-period, with a Pearson correlation coefficient value of 0.568 and P-Value <0.001 (Figure 4.10, Table 4.5). These results thus indicate that no statistically significant correlation exists between speed of clock and flowering time but components of periodicity traits are shared.

4.2.4.3 Genetic dissection of phenotypic variation

The three basic requirements of QTL analysis are (i) availability of a linkage map (ii) phenotypic values of a trait that varies quantitatively and (iii) a statistical method that can find significant associations between genotypic and phenotypic data. After fulfilling these requirements, one can perform QTL mapping to find genetic components of the trait under study. As circadian periodicity and flowering time varies quantitatively in Ws-2 X Tnz-1, and as I had genotyped this population to generate linkage map, I could perform statistical-analysis for QTL mapping. However, most statistical analysis methods assume that the phenotypic trait data is normally distributed. A deviation from normality affects the accuracy of the results.
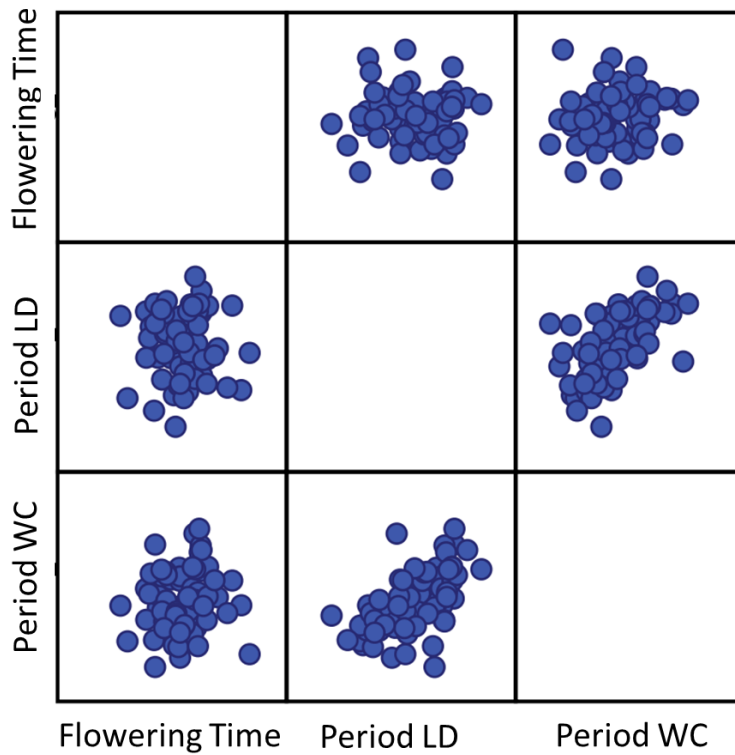
Figure 4.10: Estimation of correlations among Period-LD, Period-WC, and Flowering-time.

Table 4.5: Statistics of correlation analysis.

		Flowering Time	Period LD	Period WC
Flowering Time	Pearson Correlation	1	-.022	.085
	Sig. (2-tailed)		.851	.476
	N	78	75	73
Period LD	Pearson Correlation	-.022	1	.568[**]
	Sig. (2-tailed)	.851		<.001
	N	75	75	73
Period WC	Pearson Correlation	.085	.568[**]	1
	Sig. (2-tailed)	.476	<.001	
	N	73	73	73

** Correlation coefficient is significant at the level 0.01 (2-tailed)

To check if the periodicity and flowering time data in Ws-2 X Tnz-1 population is normally distributed, I performed Q-Q (quantile-quantile) plots. Q-Q graphs plot the observed values against the expected normal values and thus show the relationship between them. A linear relationship suggests normality in data. I found that all traits, Period-LD, Period-WC and flowering time were, normally distributed, as indicated by a strong linear relationship in their respective Q-Q graphs (Figure 4.11)
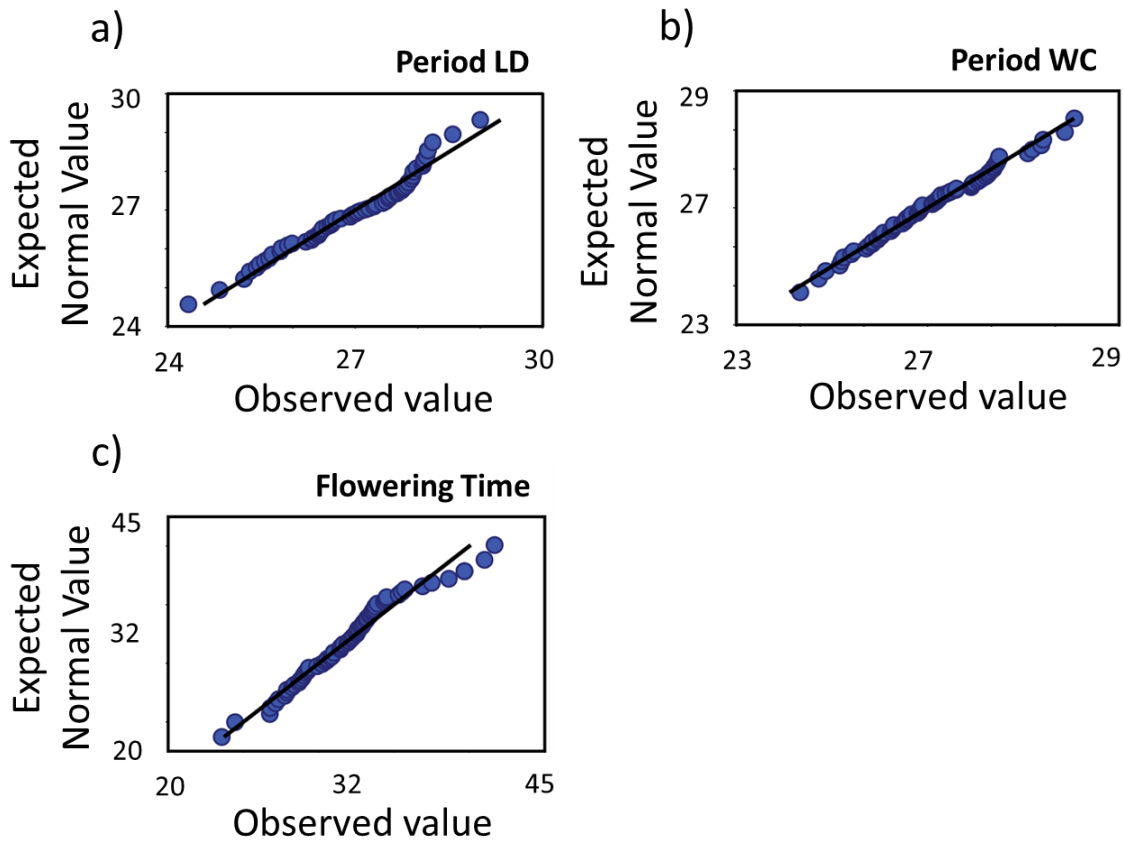
Figure 4.11: Q-Q plot showing distribution of phenotypic trait values.

(a) Period-LD (b) Period-WC and (c) Flowering-time. Note that all traits were normally distributed.

As the trait values for Period-LD and Period-WC were normally distributed, I next performed an ANOVA to examine if the mean difference in the periods of two entrainment conditions is significantly different, and thus, that both Period-LD and Period-WC can be taken as independent traits in the QTL analysis. The analysis of variance revealed that at a

95% confidence interval, the lower bound of the mean of Period-LD (26.71) did not overlap with the upper bound of the mean of Period-WC (Table 4.6). In simple words this means that 95% of the period estimates of the RILs after LD entrainment are above 26.71 (lower bound), whereas, 95% of the period estimates of the RILs after WC entrainment are below 26.28 (upper bound). These differences in means were found to be statistically significant as an F-ratio of 31.07 was obtained between groups (Period-LD and Period-WC) by ANOVA, which was highly significant with a $P < 0.001$ (Table 4.7). These results thus confirmed that the difference in the means observed for Period-LD and Period-WC is not just by chance, but had a strong genetic basis. Hence, both Period-LD and Period-WC can be analyzed as independent traits and different genetic architecture could be expected.

Table 4.6: Statistical analysis of mean differences in Period-LD and Period-WC.

Ent.	N	Mean	Std. Dev.	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
LD	75	26.9476	.99241	.11459	26.7193	27.1759
WC	73	26.0636	.93495	.10943	25.8454	26.2817
Total	148	26.5116	1.05859	.08702	26.3396	26.6835

Ent. = entrainment condition, N = total number of observations, Mean = population mean, Std. Dev. = standard deviation, Std. Error = standard error , lower and upper limits of mean are given at 95% confidence interval.

Table 4.7: Analysis of Variance.

Period	SS	df	MS	F	Sig.
Between Groups	28.911	1	28.911	31.078	<.001
Within Groups	135.818	146	.930		
Total	164.729	147			

Groups indicate Period-LD and Period-WC, SS = sum of squares
df = degree of freedom, MS = mean square, F = F-ratio, Sig. = significance

4.2.4.4 QTL mapping

QTL mapping was performed using MapQTL 6.0 statistical software. Interval mapping (IM) followed by multiple-QTL mapping (MQM) analysis, using markers as co-factors was employed. Use of markers as co-factors in MQM analysis reduces the residual variance and thus the power to detect QTL is increased. An automatic cofactor selection (ACS) procedure that uses backward elimination method was applied to detect potential markers that were then used as cofactors in the subsequent MQM analysis. The detail procedure for this analysis is described in chapter 2.2.8). Similar results were obtained for both IM and MQM. Here I will only show the results obtained by MQM analysis, as overall QTL resolution achieved by MQM analysis was better than IM.

Three QTL were detected for circadian periodicity. One was for Period-WC and two for Period-LD. A period-WC QTL was detected at chromosome 1 at 20-40 cM. This was a main effect QTL with a corresponding LOD score of 7.78 that explained ~40% of the total phenotypic variation observed in the RILs. A QTL that completely co-localizes with this region was also detected for Period-LD. The maximum LOD score obtained for this QTL was 7.83 and the variance explained by this QTL was 40.7%. A known clock gene *GI* localizes at this position of the genome, thus it is a potential candidate for this QTL. A second QTL for Period-LD was detected at the bottom of chromosome 4 at 60-65 cM with a LOD score of 2.69. This was a minor effect QTL that explained 9.1% of the total phenotypic variation. No known clock gene has been identified on this genomic region. Further, this QTL was not detected in Period-WC, and thus assigns it as a light-specific locus (Figure 4.12).

Four QTLs, one each on chromosome 1, chromosome 2, chromosome 3 and chromosome 5 were detected for flowering time (Figure 4.12). Two of these QTLs, at chromosome 3 and chromosome 5, were main effect QTL, with a corresponding LOD score of 6.5 and 5.2 respectively. The QTLs on chromosome 1 and chromosome 2 were minor effect QTLs, with a LOD score of 2.38 and 2.59, respectively. These QTLs explained 6.1% and 7.4% of the variation, respectively. Chromosome 2 QTL, detected at 15-26 cM, co-localizes with the photoreceptor *PHYB* (Michael et al., 2003b). Chromosome 3 QTL that

explained 21.5% of the variation, with a LOD value of 6.45, was detected between intervals of 10-15 cM. *SPY*, a known flowering-time gene underlies this region. Finally, a QTL at the bottom of chromosome 5, at 80-93 cM, explained 16.1% of the variation, with a LOD score of 5.05. This region of the chromosome contains many known circadian clock and flowering-time genes. These include *ZTL*, *PRR3*, *TOC1*, *ELF5*, *CDF1* and *LFY* (Darrah et al., 2006; Michael et al., 2003b). Any of these genes can be a possible candidate for this QTL. The presence of more than one QTL at this region cannot be ignored because the density of the markers here is low that could have resulted in low QTL resolution.

4.2.4.5 Statistical confirmation of QTLs and allelic interaction analysis

In MapQTL6.0, the statistical model (MQM) used to detect QTLs only implements the additive and dominant gene actions. Thus other uses of markers as co-factors, for instance, to study gene X environment or gene X gene interactions, is not possible with this software. Both circadian clock and flowering time are complex mechanisms. Therefore, it is plausible that there are interactions among genes in these pathways.

To inspect the epistatic interactions between QTLs identified for circadian periodicity and flowering-time, a univariate GLM analysis was performed using SPSS 17.0 statistical package (For details see 2.2.9). Period-LD, Period-WC or Flowering-time was used as dependent variable, and markers that were selected as co-factors in MQM analysis were used as fixed factors. The main effects and interactions were calculated under a custom model using 'type III' sum of squares.

All three QTLs identified for circadian periodicity were significant. Specifically, both QTLs at chromosome 1, for Period-LD and Period-WC, were highly significant for P-values <0.001 (Table 4.8, 4.9). Similarly, the chromosome 4 QTL for Period-LD was also significant for P=0.009 (Table 4.8). For Period-WC, this QTL did not cross the significance threshold in MQM analysis, therefore was not designated as a *bona fide* QTL. Further, I examined if there was any interaction between the chromosome 1 and chromosome 4 QTL for Period-LD. No statistically significant evidence for an epistatic interaction was detected (Table 4.8, 4.9). Thus, all three periodicity QTLs were main effect QTLs that independently participate
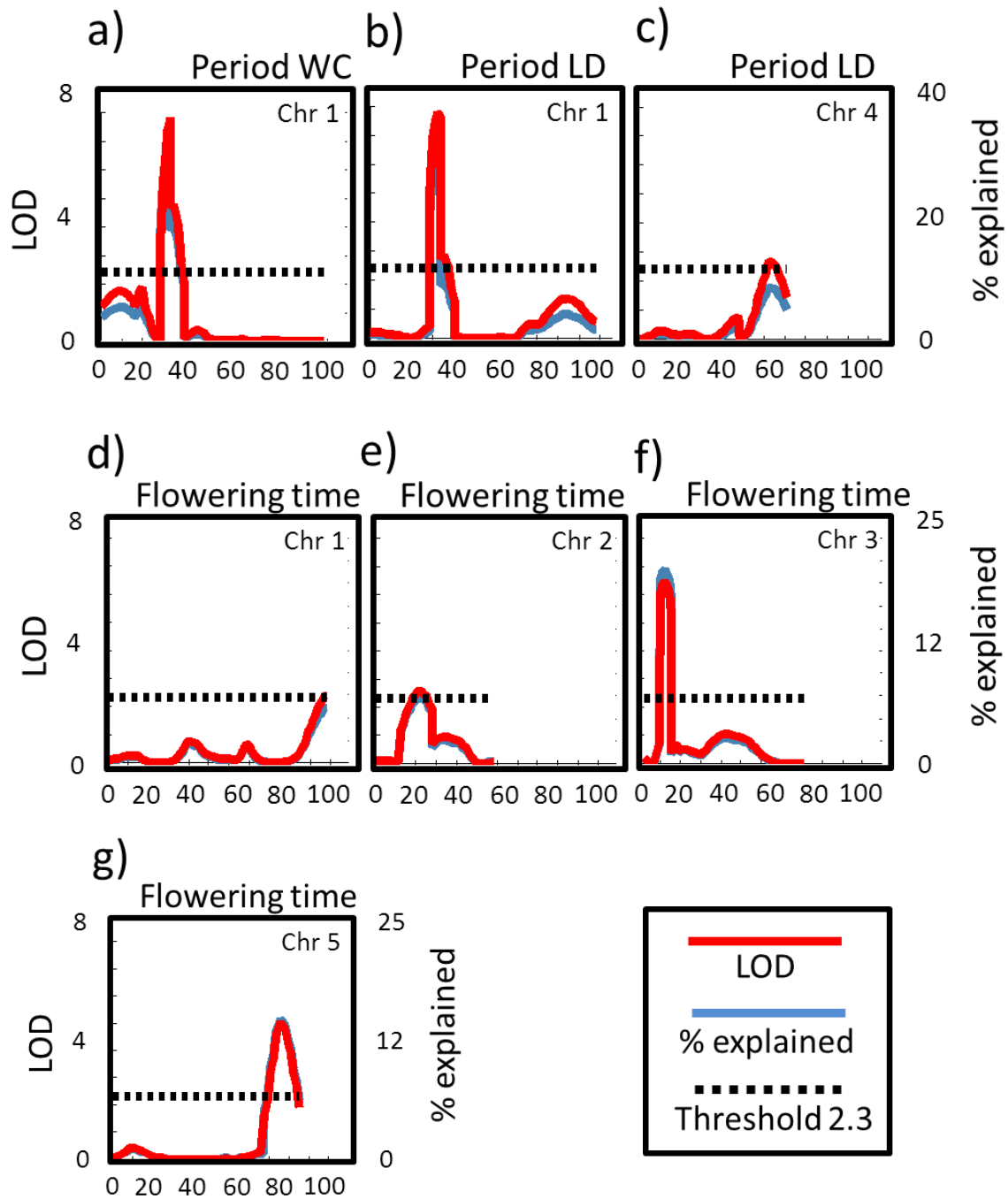
Figure 4.12: QTL analysis of circadian periodicity and flowering-time.

QTLs detected on (a) chromosome 1 for Period-WC (b) chromosome 1 for Period-LD (c) chromosome 4 for Period-LD (d) chromosome 1 for flowering-time (e) chromosome 2 for flowering-time (f) chromosome 3 for flowering-time (g) chromosome 5 for flowering-time. X-axis represents distance in cM. Left y-axis and right y-axis represent LOD score and percentage of variance explained, respectively.

Table 4.8: Statistical interaction analysis for Period-LD QTLs.

Source	SS	df	MS	F	Sig.
Corrected Model	31.443	4	7.861	15.278	<.001
Intercept	23882.867	1	23882.867	46418.715	<.001
Chr. 1	14.761	1	14.761	28.689	<.001
Chr. 4	5.184	2	2.592	5.038	.009
Chr. 1 * Chr. 4	.023	1	.023	.045	.833
Error	30.871	60	.515		
Total	47280.601	65			
Corrected Total	62.314	64			

Table 4.9: Statistical interaction analysis for Period-WC QTLs.

Source	SS	df	MS	F	Sig.
Corrected Model	18.852	3	6.284	11.089	<.001
Intercept	32209.866	1	32209.866	56836.528	<.001
Chr. 1	12.474	1	12.474	22.010	<.001
Chr. 4	.618	1	.618	1.091	.301
Chr. 1 * Chr. 4	.118	1	.118	.208	.650
Error	31.736	56	.567		
Total	40803.216	60			
Corrected Total	50.588	59			

Table 4.10: Statistical interaction analysis for Flowering-time QTLs.

Source	SS	df	MS	F	Sig.
Corrected Model	483.654	10	48.365	11.236	<.001
Intercept	6371.519	1	6371.519	1480.248	<.001
Chr. 1	1.417	2	.709	.165	.849
Chr. 2	41.726	1	41.726	9.694	.003
Chr. 3	119.958	1	119.958	27.869	<.001
Chr. 5	109.580	1	109.580	25.458	<.001
Error	176.479	41	4.304		
Total	51438.903	52			
Corrected Total	660.133	51			

in the control of clock speed.

For flowering time, chromosome 2, chromosome 3 and chromosome 5 QTLs were significant and the chromosome 1 QTL was not. In MQM analysis, this latter QTL was appeared as a minor QTL with a LOD score of 2.38, which was just above the threshold 2.3. Statistical analysis did not confirm any association of this locus with flowering-time variation. All other QTLs, at chromosome 2, chromosome 3 and chromosome 5 were highly significant with P-values of 0.003, <0.001, <0.001, respectively (Table 4.10). No epistatic interactions among any of these QTLs were detected. Thus all QTLs identified in Ws-2 X Tnz-1 RIL population were main-effect QTLs that directly contributed to the detected flowering-time variation.

4.2.4.6   Allelic effect analysis

One of the important outputs of these statistical analyses is defining the extent and direction of the effect that a particular parental allele delivers to the phenotype. For each QTL, markers selected as cofactors in the MQM analysis were used to calculate the allelic effect of the parents. These markers were selected because the QTL LOD score and percentage of variation associated to these markers were highest during MQM analysis. FD20D23 was selected as cofactor for chromosome 1 periodicity QTL. For chromosome 4 periodicity QTL it was W199. For flowering time, W103, W169 and UA5.25.1 were selected for the chromosome 2, 3 and 5 QTLs, respectively.

The presence of Tnz-1 allele at FD20D23 locus resulted in one hour period acceleration after either LD or WC entrainment. The effect at locus W199 was opposite in Tnz-1, where this allele reduced the speed of oscillation by ~45 minutes (Figure 4.13 a). For flowering time, the presence of Tnz-1 allele delayed flowering for all loci. The largest effect was observed at locus W169, where the bolting date was delayed by ~7 days. The late flowering effect of Tnz-1 allele was three and four days, for locus W103 and UA5.25.1, respectively (Figure 4.13 b). Thus, overall, Tnz-1 alleles slowed down the reproductive transition process, in the Ws-2 X Tnz-1 RIL population.
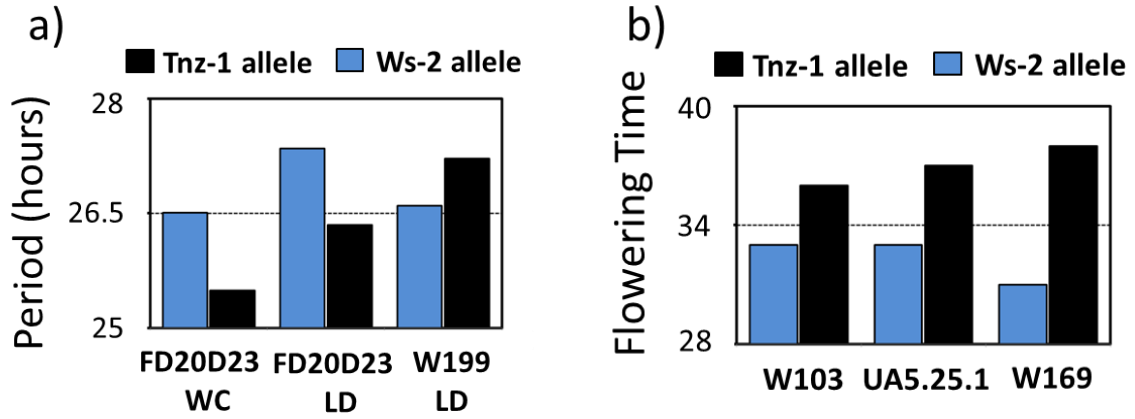
Figure 4.13: Allelic contribution of the parental alleles.
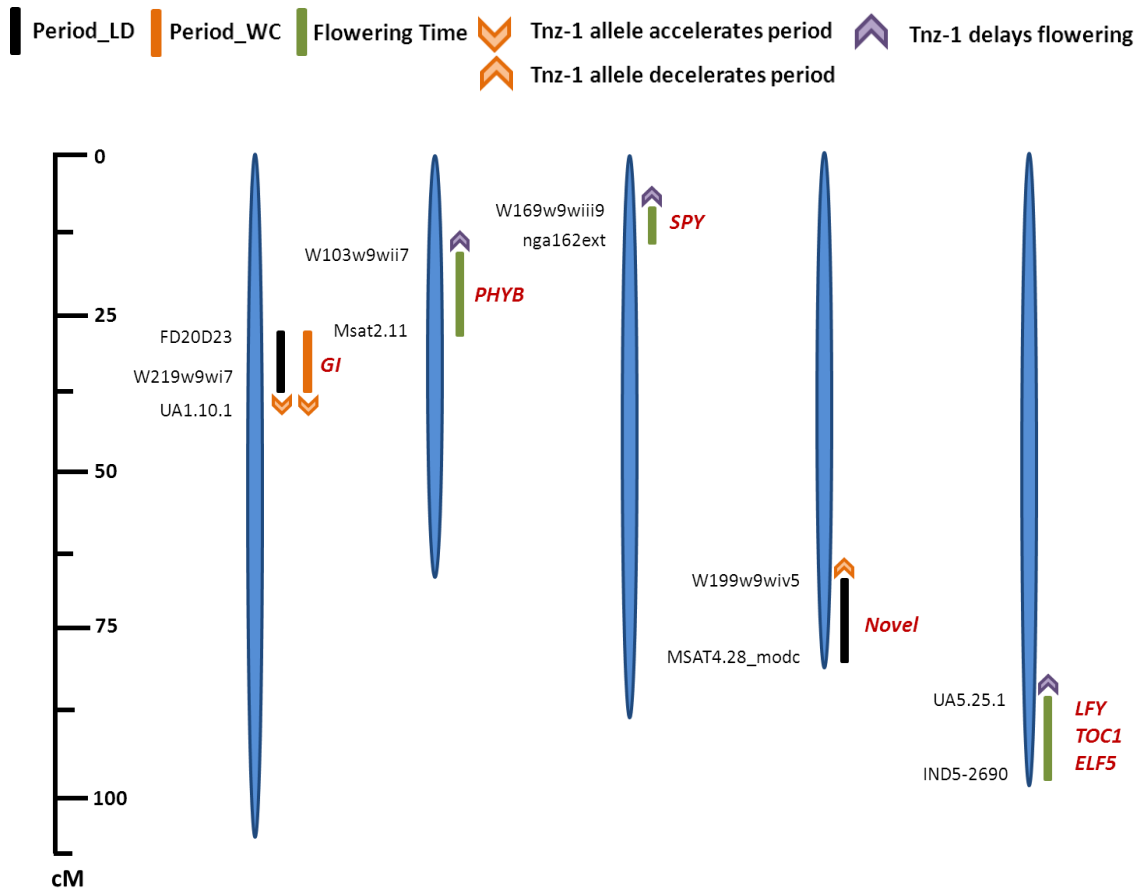
Figure 4.14: Summary of QTL mapping.

Scale bar represents distance in centimorgans (cM). Vertical bars represent QTLs location on the genome. Markers flanking QTL confidence interval are given on the left side of the respective bar. Candidate genes are shown on the right side of the bars. Arrows denote the effect of the Tnz-1 allele on the Ws-2 phenotype.

4.2.4.1 Validation of Chromosome 1 periodicity QTL: *GI* as a candidate

Validation of QTL authenticates the accuracy of QTL mapping. One way to validate a QTL is to confirm its allelic effect in almost isogenic lines, such as NILs. NILs can be generated by successive backcrossing a particular RIL to one of the parent, until an introgression of one parent is obtained in otherwise homogenous background of the second parent. I adopted this sort of a strategy to validate the chromosome 1 periodicity QTL.

NIL29 was obtained in the Ws-2 genetic background containing a Tnz-1 introgression at the QTL confidence interval. NIL29 was not completely homogeneous for the Ws-2 background as the bottom of the chromosome 5 was still segregating (Figure 4.15 a). No QTL for periodicity was detected at this genomic region (Figure 4.14). Further backcrosses would be required to homogenize this part to Ws-2 for a ''perfect'' NIL. Nevertheless, as NIL29 was > 95% homogenous in Ws-2 background with a Tnz-1 introgression at the QTL, this line provided a good estimation of the periodicity effect of Tnz-1 allele. Therefore, I estimated the free-running period of *CCR2::LUC* in NIL29 and Ws-2 after both LD and WC entrainment. Consistent with the allelic affect detected in QTL mapping (Figure 4.13a), NIL29 exhibited a ~1h shorter period compared to Ws-2. This difference in the speed of *CCR2::LUC* oscillations in NIL29 and Ws-2 was further increased after WC entrainment, as I detected 2h period difference between NIL29 and Ws-2 after WC (Figure 4.15 b). Taken together, these results validated the chromosome 1 QTL, and also ensured the accuracy of the mapping.

GI is an established circadian clock and flowering-time regulator and co-localized at the QTL confidence interval. As a strong possible candidate, I promptly decided to compare the nucleotide sequence of *GI* in Ws-2 and Tnz-1 using the available re-sequencing data of the parental accessions. If *GI*, at FD20D23 locus is accelerating periodicity in Tnz-1, then there should be sequence polymorphism(s) between *GI-Ws-2* and *GI-Tnz-1.* Several polymorphisms were detected in the promoter and coding region of both accessions (Figure 4.15 c). However, only one of those, a 'T' to 'C' nucleotide change, resulted in an

non-synonymous encoded amino-acid change. This was Alanine to Valine at translated position 343 (A343V). If *GI* is the underlying gene for chromosome 1 periodicity QTL, then this residue change could be a plausible reason for the acceleration of periodicity in Tnz-1.
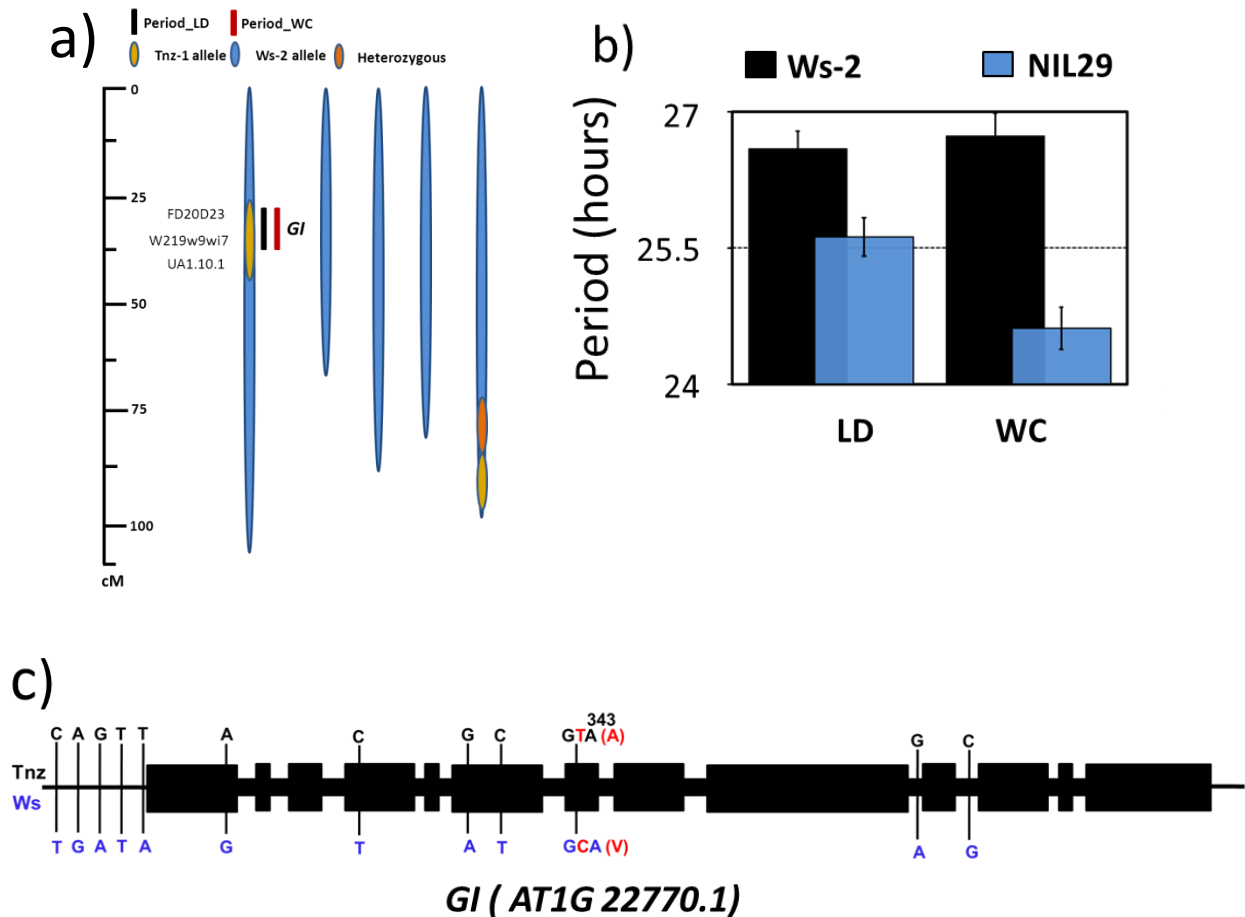
Figure 4.15: Validation of chromosome 1 QTL.

(a) Shows the genetic structure of the NIL29. Note that the NIL29 contains an introgression of Tnz-1 genome at QTL confidence interval in otherwise almost homogenous Ws-2 background, except for at the bottom of chromosome 5. (b) Period estimates of *CCR2::LUC* in NIL29 and Ws-2 after LD and WC entrainment. Note the acceleration of periodicity in NIL29 after both entrainments. (c) Schematic diagram showing gene structure of *GI* along with sequence comparison in Tnz-1 and Ws-2. Vertical bars show the position of the nucleic-acid transition. The letters above the bars represent the nucleic acid in Tnz-1 and letter below represent the nucleic acid in Ws-2 relative to Columbia reference. The numbers above the letters represent the position of encoded amino-acid change. The letters in parenthesis show the amino-acid change. Note that a nucleotide substitution from 'C' to 'T' in Tnz-1 results in an encoded amino-acid change from Alanine to Valine at position 343.

4.2.5 Characterization of Ws-2 X Strand RIL sets

4.2.5.1 Generation of linkage map

Two sub-populations of Ws-2 X Strand RIL set, WS and WSW comprising 192 lines were genotype using 54 uniformly distributed SSLP and InDels markers. The analysis of genome participation from parental alleles revealed that, overall, 61.4% of the genome is contributed by Ws-2. Four chromosomes except chromosome 3 followed the same pattern of parental allele contribution. In chromosome 3, the overall representation of Ws-2 was 80% whereas, Strand contributed 12.9%. Furthermore, 6% of the total heterozygosity was observed in this RIL set, which was similar to the expected 5.8% for this kind of mixed population (Table 4.11).

Similar to Ws-2 X Tnz-1 RIL population, an overall segregation distortion in favor of Ws-2 allele was observed at all chromosomes. Specially, the top of chromosome 3 displayed an extreme aberration for selection of Ws-2 (Figure 4.16). Interestingly, the comparison of distortion pattern over the complete lengths of chromosome in Ws-2 X Tnz-1 and Ws-2 X Strand RIL sets was quite similar. For instance, a steep increase in Ws-2 genome at around 60 cM was observed for both RIL sets. A similar increase was also evident at the bottom of chromosome 4 (Figure 4.7, Figure 4.16). As discussed in the previous section, one potential reason for the former segregation distortion is the selection for the luciferase transgene, which is located at the top of chromosome 3.

Table 4.11: Genome participation of parental accessions.

Genotype	Genome participation (%)					
	Global	Chr. 1	Chr. 2	Chr. 3	Chr. 4	Chr. 5
Ws-2	61.4	58.2	63.4	80	56.1	56.6
Strand	30	31.9	26.6	12.9	36.6	34.1
Ws-2/Strand	6	6.2	6.8	6	4.9	6.9
Missing	2.6	3.6	3.8	1.1	2.4	2.4
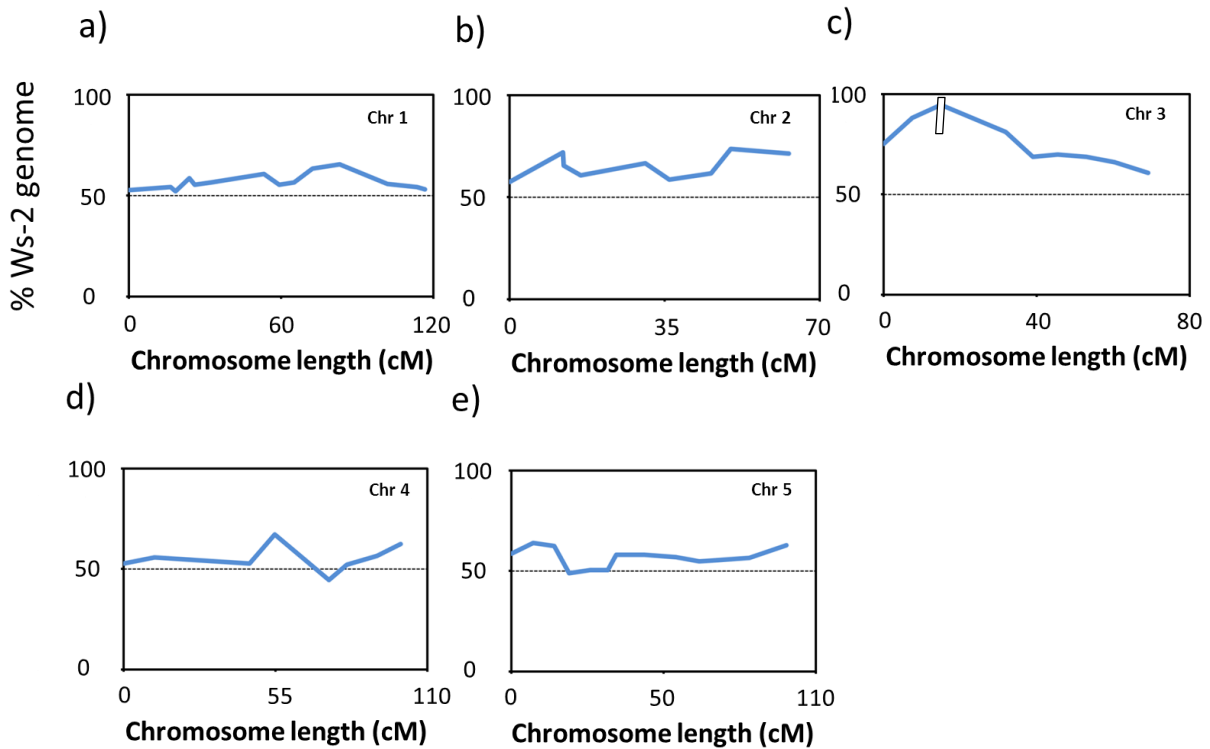
Figure 4.16: Segregation between Ws-2 and Strand alleles along the five chromosomes.

The percentage of Ws-2 allele is represented along Y-axis. W-axis represents chromosome length in centimorgans. Note that the top of chromosome 3 is highly distorted towards Ws-2 allele.

The linkage map was then calculated using Joinmap 4.0 (Figure 4.17). All markers were assigned to the desired linkage group, without any ambiguity, corresponding to the order of their physical positions in the genomes. However, similar to Ws-2 X Tnz-1 linkage map, an overall ''shrinkage'' of chromosome 3 was observed in Ws-2 X Strand map, which was ''fixed'' using the marker location information available from other maps. A potential reason for this ''shrinkage'' could be the high segregation distortion at the top of chromosome 3, which resulted in a decrease in the recombination events in that region. This ultimately affected the total length of the chromosome.

The total length of the Ws-2 X Strand population was estimated to be 448 cM, which was 37 cM longer than the Ws-2 X Tnz-1 RIL population. The average marker to

marker distance was 8.29 cM. Similar to Ws-2 X Tnz-1 map, all chromosomes in Ws-2 X Strand RIL set showed a linear relationship between their physical and genetic length. Notably, the length of the chromosome 4 was even further extended in this RIL population, which was 99 cM in Ws-2 X Strand, compare to 72 cM in Ws-2 X Tnz-1.
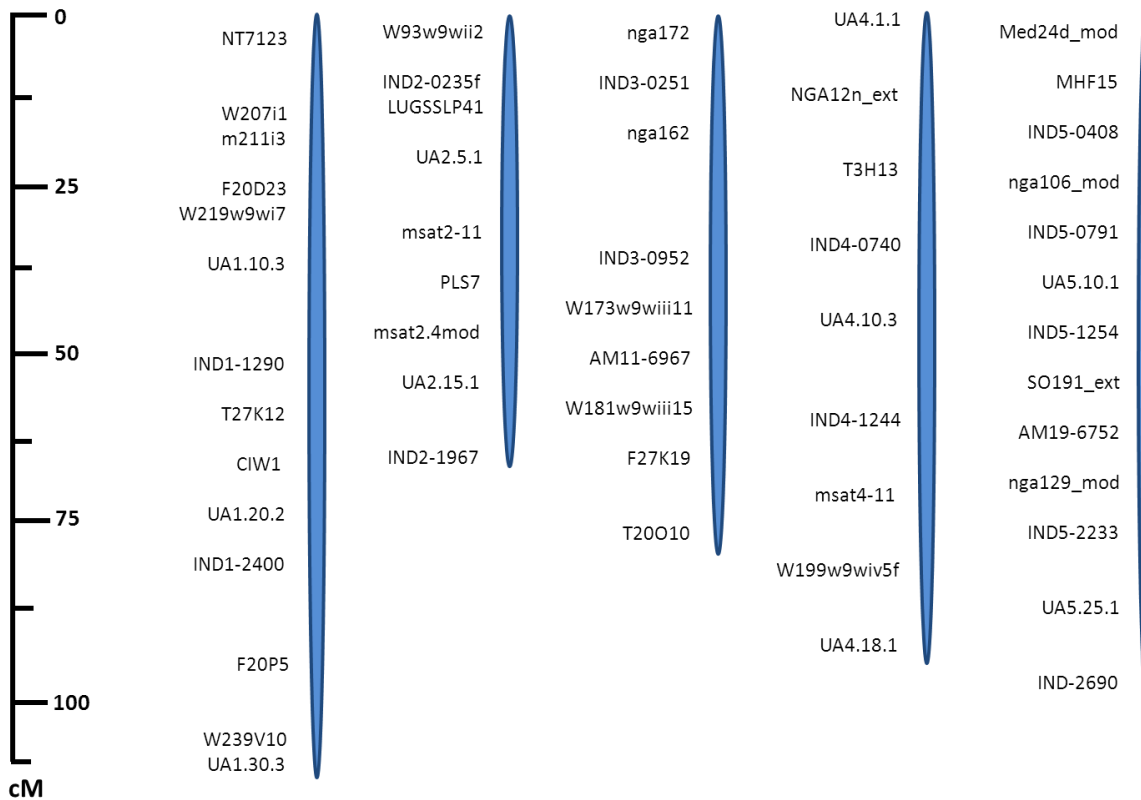
Figure 4.17: Genetic map of Ws-2 X Strand RIL population.

4.2.5.2 Phenotypic characterization of Ws-2 X Strand RIL set

Similar to Ws-2 X Tnz-1, this RIL set was phenotyped for two traits: circadian periodicity and flowering time. Free-running period of *CCR2::LUC* was both measured under continuous white light after LD and WC. Flowering time was measured as number of days to bolting. Considerable variation in all of these traits was observed. The frequency distribution revealed that all three traits in the Ws-2 X Strand RIL set varied continuously

with a population mean of 27.5 h, 27.0 h and 32 days for Period-LD, Period-WC and flowering time, respectively. Moreover, a transgressive behavior of these RILs was observed in Ws-2 X Strand population with individual RIL lines exhibiting phenotypes both below and above the parent's phenotypes (Figure 4.18). Next, I performed bivariate correlation analysis to check if the circadian periodicity in Ws-2 X Strand RIL set is linked to flowering time. Similar to WS-2 X Tnz-1 RIL population (Figure 4.10), I could not detect any significant relationship between circadian periodicity and flowering time. However, a strong positive relationship was observed for Period-LD and Period-WC (Figure 4.19).
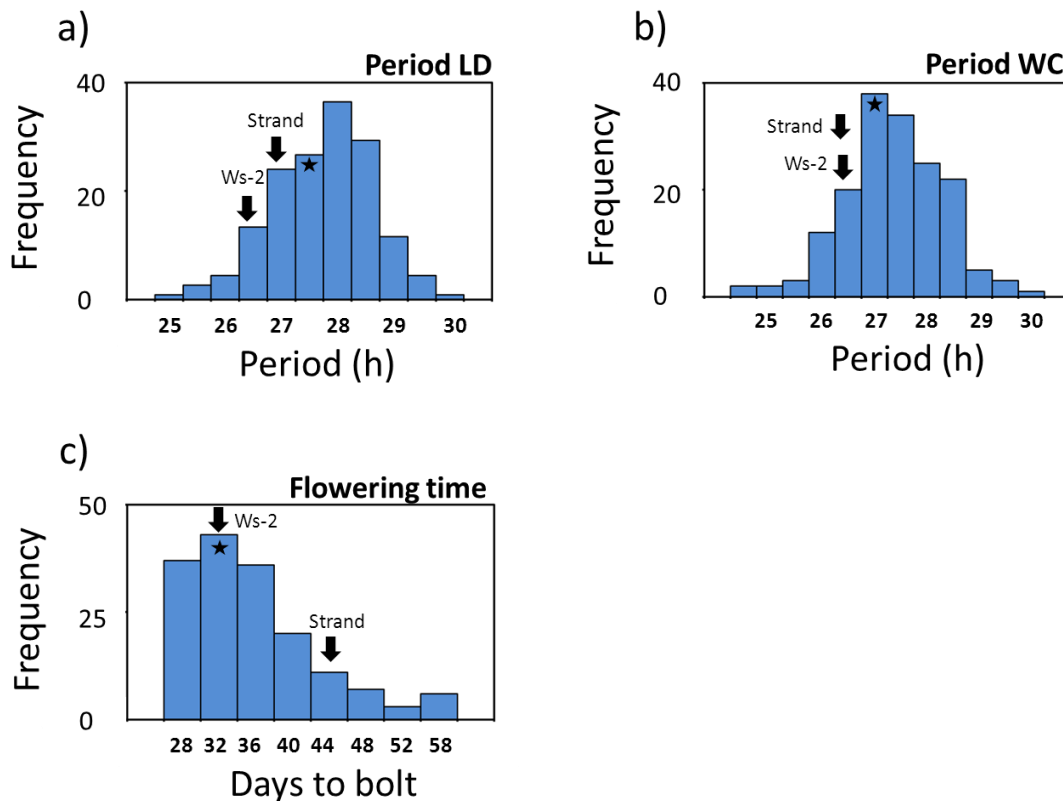
Figure 4.18: Distribution of circadian period and flowering-time in Ws-2 X Tnz-1 RILs.

Distribution of mean (a) period after LD entrainment (b) period after WC entrainment and (c) flowering-time. Mean values of parental accessions are marked with black arrows. Black stars represent population means. Bins are labeled with the upper bounds.

4.2.5.3 Genetic dissection of phenotypic variation: QTL mapping

Q-Q plots were performed to check the normality of the phenotypic data. I observed a strong linearity between the expected normal values and the observed values for both Period-LD and Period-WC, depicting a completely normal distribution of the values (Figure 4.20 a,b). However, for the flowering-time data, a considerable deviation from normality was observed, as I found that some trait values diverged from the normal expected value (Figure 4.20 c).

Deviation from normality is common in developmental traits and if absolutely required, statistical approaches for data transformation can be used to mimic normality (Jimenez-Gomez et al., 2010). However, QTL mapping methods, for instance, IM, MQM and ACS are quite robust against deviation from normality. In case of MapQTL 6.0, using the maximum likelihood method, it can be often seen from the number of iteration if the deviation from normality becomes problematic. In case of acceptable, normal distribution, the maximum likelihood algorithm in MQM analysis converges in less than 10 iterations. If the number of iterations increases this number than the accuracy of the detected LOD score could have been compromised and the data might need to be transformed. In QTL analysis that I discuss below, the number of iterations never exceeded from two. This ensured that the extent of deviation from normality observed in flowering-time data did not significantly affect the MQM analysis.

MQM analysis using MapQTL 6.0 was performed selecting markers as cofactors. Markers were selected using ACS. Four QTLs were detected for Period-LD, two on chromosome 1 and two on chromosome 5. Chromosome 1 QTLs were detected at 1-18 cM and 105-116 cM with corresponding LOD scores of 3.0 and 3.14, respectively. The total variation explained by these QTLs was 9.1% and 9.4%, for top and bottom QTLs, respectively. Both QTLs were significant with P=0.019 and 0.025, respectively. The additive affect associated to the QTL at the top of chromosome 1 was 0.65 hours, with the Strand allele delayed periodicity. For the bottom QTL, the additive affect was 0.72 hours, and again, the Strand allele delayed periodicity (Figure 4.21 a; Table 4.12).
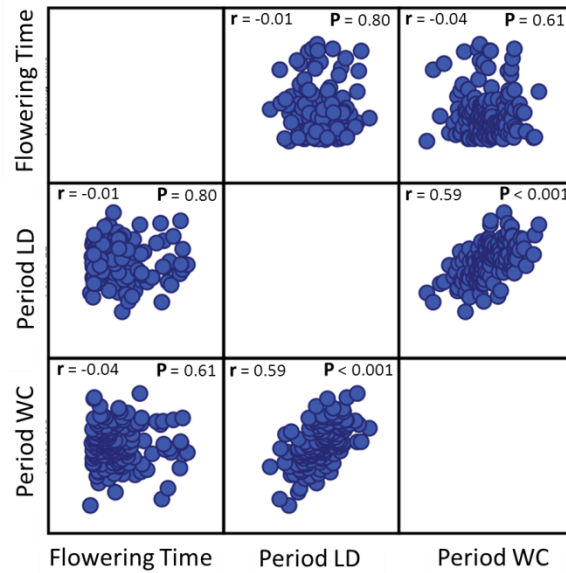
Figure 4.19: Estimation of correlations among Period-LD, Period-WC, and Flowering-time.
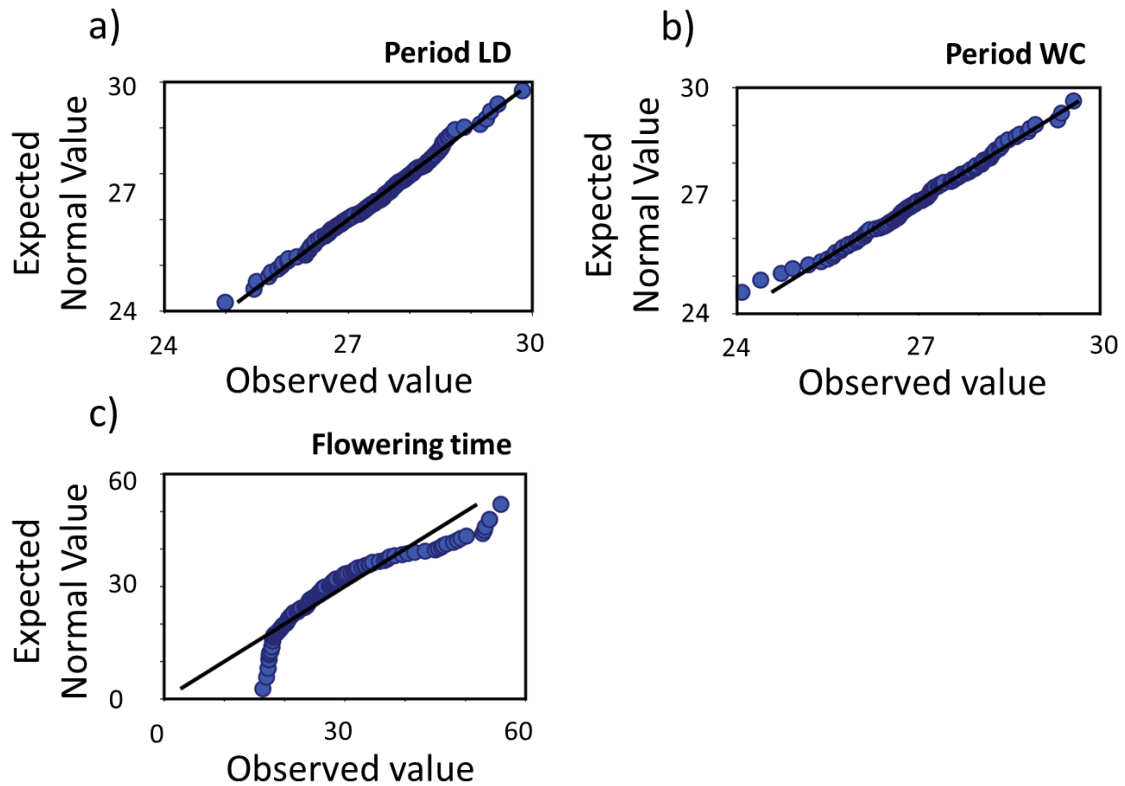
Figure 4.20: Q-Q plot showing distribution of phenotypic trait values.

(a) Period-LD (b) Period-WC and (c) Flowering-time. Note that Flowering-time values are not completely normally distributed.
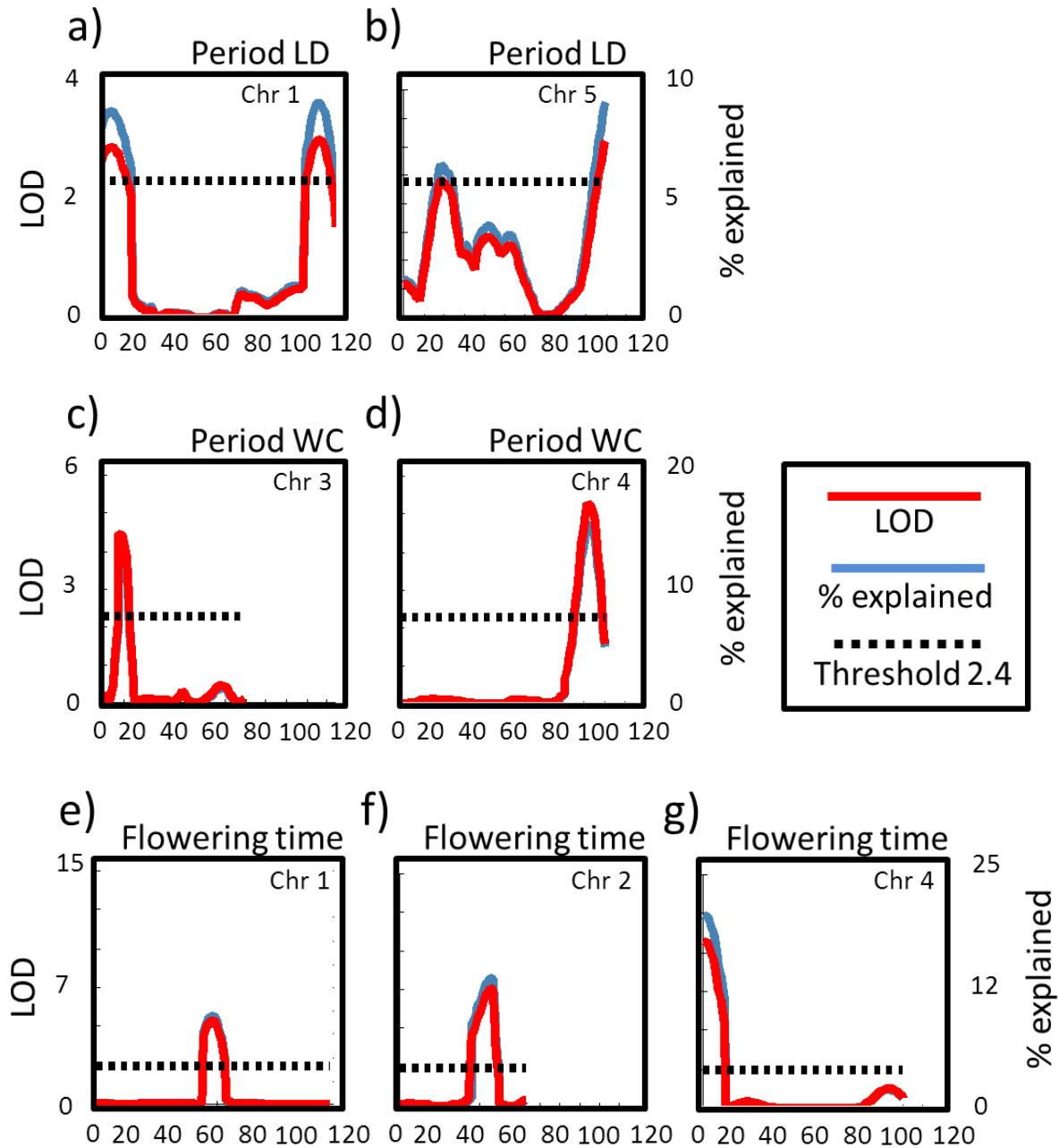
Figure 4.21: QTL analysis of circadian periodicity and flowering-time.

QTLs detected on (a) chromosome 1 for Period-LD (b) chromosome 5 for Period-LD (c) chromosome 3 for Period-WC (d) chromosome 4 for Period-WC (e) chromosome 1 for flowering-time (f) chromosome 2 for flowering-time (g) chromosome 4 for flowering-time. X-axis represents distance in cM. Left y-axis and right y-axis represent LOD score and percentage of variance explained, respectively.

Photoreceptors *CRY2* and *PHYA* were the candidate genes for the top chromosome 1 QTL, whereas, *FKF1* co-localize with the bottom QTL (Michael et al., 2003b). Further, QTLs identified on chromosome 5 were located at 20-32 cM and and 95-99 cM with corresponding LOD scores of 2.42 and 3.10 and explained variance of 6.7% and 9.5%, respectively. Additionally, both QTLs were significant with P=0.007 and 0.028, respectively. The additive affect associated to the QTL at the top of chromosome 1 was 0.53 hours, with the Strand allele delaying periodicity. For the bottom QTL, the additive affect was -0.72 hours. In this case, the Strand allele accelerated periodicity (Figure 4.21 b; Table 4.12). *COL1* co-localizes with the QTL at the top of chromosome 5 (Ledger et al., 2001), whereas, *PRR3*, *ZTL*, and *TOC1* are the possible candidates for the bottom QTL (Darrah et al., 2006; Michael et al., 2003b).

Table 4.12: Statistical analysis of QTL mapping.

Trait	Ent.	Chr.	Marker	Position (cM)	F	P	D
Period	LD	1	NT7123	1	5.57	0.019	0.65
		1	W239V10	113	5.16	0.025	0.72
		5	nga106	20.9	7.59	0.007	0.53
		5	Ind5-2690	99	4.91	0.028	-0.58
Period	WC	3	Ind3-0251	7.5	8.83	0.004	-1.01
		4	W199iv5	92	13.57	<0.001	0.56
Flowering -time	-	1	T27K12	59	10.09	0.002	7.05
		2	UA2.15.1	50	23.42	<0.001	8.97
		4	UA4.1.1	1	56.46	<0.001	15
		1*4	UA2.15.1*UA4.1.1	-	6.16	0.014	-

Ent. denotes entrainment, Chr. represent chromosome, Marker shows the exact marker used as cofactor in MQM analysis, in position column the position of the same marker is given, F is F-ratio, P represent P-value, D denotes the additive effect of the QTL and is calculated as (mean value Strand allele – mean value Ws-2 allele). D is measured in hours for Period and days for flowering-time. * represents interaction between QTLs.

One QTL each on chromosome 3 and chromosome 4 were identified for Period-WC. A LOD value of 4.46 and 5.30 corresponded to the chromosome 3 and chromosome 4 QTLs, respectively. The percentage of variation explained was 13% and 15.8%, respectively (Figure 4.21 c,d). Moreover, the additive effect contributed by Strand allele at chromosome 3 locus was -1.01 hour, whereas, the Strand allele delayed periodicity by 0.56 hours at the chromosome 4 locus. Although, the allelic affect associated to the chromosome 4 locus was not large, this QTL was highly significant with a P < 0.001 (Table 4.12). As of yet, no clock gene has been reported at this locus, whereas, *SPY* localizes at chromosome 3 QTL.

Analysis of flowering-time data revealed three potential QTL, each on chromosome 1, chromosome 2 and chromosome 4. These QTLs explained 9.5%, 13.9% and 20.7% of the variation in the phenotypic trait, respectively. The corresponding LOD scores were 5.43, 7.61 and 10.73, for chromosome 1, chromosome 2 and chromosome 4 QTLs, respectively (Figure 4.21 e-g). Chromosome 2 and chromosome 4 QTLs were highly significant for P < 0.001. Similarly chromosome 1 QTL was also significant with a P = 0.002. The presence of Strand allele at these loci all delayed flowering. Specifically, the Strand allele at chromosome 1 locus delayed flowering by 7.05 days, whereas, the additive effect associated to chromosome 2 QTL was a delay of 8.97 days. The highest effect of a Strand allele was observed for the chromosome 4 QTL, where it delayed bolting time by ~15 days (Table 4.12). *FRI*, a strong repressor of flowering, is a potential candidate for this QTL (Johanson et al., 2000).

Univariate GLM analysis revealed a strong interaction between the chromosome 2 and chromosome 4 QTLs with a P-value of 0.014 (Table 4.12. The markers used as fixed factors for this analysis were UA2.15.1 (for chromosome 2 QTL) and UA4.1.1 (for chromosome 4 QTL). These were the same markers that were used as cofactors in the MQM analysis. I found that mean flowering time was 23.29 days when both Ws-2 alleles were present on UA2.15.1 and UA4.1.1. The flowering time was increased by 4.5 days when the Strand allele substituted the Ws-2 allele at UA2.15.1. Flowering was delayed by 10.34 days when Strand allele at UA4.1.1 interacted with Ws-2 allele at UA2.15.1. A more

significant delay of 24 days was detected when both alleles at UA2.15.1 and UA4.1.1 were Strand (Figure 4.22). Taking all these flowering-time results together, I concluded that the overall, Strand alleles at all identified loci slowed down the vegetative to floral transition. The extremely late flowering of Strand wild type (Figure 4.1 b) further supports this statement.
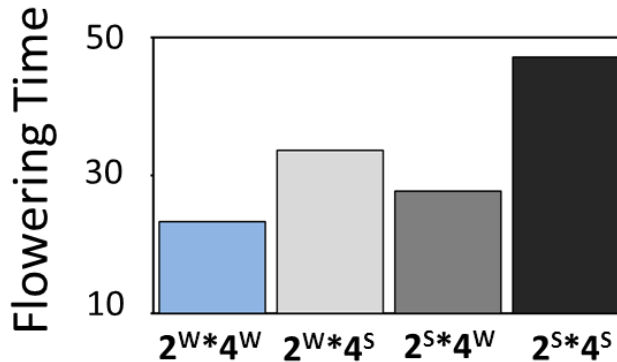
Figure 4.22: Flowering-time variation due to interaction of two QTLs at chromosome 2 and chromosome 4.

In X-axis the number represents the chromosome number and the superscripted W and S are designated for the Ws-2 and Strand allele respectively.

Taken together, I found that many of the flowering-time QTLs detected in both Ws-2 X Tnz-1 and Strand X Ws-2 RIL populations co-localized with the known genomic regions assigned to circadian-clock genes. Further, QTLs detected for circadian periodicity in one population were detected for the flowering time in the second population. However, co-localizing QTLs for both circadian periodicity and flowering time in a single RIL set were not detected. This shows that the link of the circadian clock to flowering time is not direct or is greatly influenced by crosstalk between several other pathways, for instance, hormone signaling. Probably, this is the reason that no strong evidence for a correlation between circadian rhythms and flowering time has yet been reported (de Montaigu et al., 2010).

Chapter Five: **Final Discussion**

5.1 Characterization of *ELF3-Sha*

In chapter 3 of my thesis, I described the positional isolation and functional characterization of *ELF3-Sha*. I found that this allele accelerates circadian oscillations in a light-dependent manner, but does not significantly affect developmental processes under normal growth conditions. Further, I confirmed that the periodicity phenotype of *ELF3-Sha* is a result of a single residual change from an encoded Alanine to Valine. Moreover, unlike *elf3* loss-of-function mutants, *ELF3-Sha* maintains robust circadian rhythms. Accordingly, a highly altered, but rhythmic expression of clock genes was observed under free-running conditions. This supported a reduced activity of ELF3-Sha. The cellular studies revealed that ELF3-Sha is defective in its proper localization in the nucleus, which explained the cellular basis of *ELF3-Sha* circadian phenotypes.

5.1.1 Importance of NILs for validation of ELF3-Sha QTL

A periodicity QTL previously detected at chromosome 2 in BXS RIL population was the starting point of this project. The presence of Sha allele on the locus was shown to result in the acceleration of clock speed. I confirmed this periodicity effect in HIFs and a NIL. The presence of a Sha allele caused short-period in both HIFs and NIL-S, thus validated the QTL identity (Figure 3.1). Co-localizing QTLs for shade-avoidance response (SAR) has been reported in two separate studies (Coluccio et al., 2011; Jimenez-Gomez et al., 2010). Both studies validated the QTLs using HIFs. One key factor that determines the SAR is hypocotyl length, which is known to be diurnally controlled (Nozue et al., 2007). Thus, a common candidate gene *ELF3* was proposed for both the periodicity and SAR QTLs. One SAR study reported that the SAR of HIFs carrying Sha allele was completely opposite than the transgenic lines expressing *ELF3-Sh*a in Columbia *elf3-1* null background. They argued that epistatic interactions of *ELF3* with Columbia genetic background underlie this differential response (Jimenez-Gomez et al., 2010). Similar epistatic interactions were also observed for *ELF3-Sha* periodicity response in Bay and Sha genetic backgrounds. I found that the *ELF3-Sha* exhibited short period in NIL-S (Bay-0 background), whereas, in Sha, it

conferred no periodicity difference (Figure 3.1 b). Thus the genetic background defines the response of *ELF3-Sha* and highlights the importance of studying this allele in a homogenous genetic background, such as NILs.

5.1.2 ELF3-Sha as a QTG and a possible role of the Q-Stretch

Transgenic lines expressing *ELF3-Bay* and *ELF3-Sha* complemented the arrhythmic phenotype of *elf3* loss-of-function mutants in such a way that *ELF3-Sha* lines exhibited periodicity acceleration compared to *ELF3-Bay* (Figure 3.3). These result confirmed *ELF3* as a quantitative trait gene (QTG) underlying chromosome 2 QTL. Further, the sequence comparison of *ELF3-Bay* and *ELF3-Sha* revealed two non-synonymous changes (Figure 3.4 a). A single encoded residual replacement 'A362V' in the middle domain, and an 8 glutamine deletion in the Q-stretch (glutamine stretch) in the C-terminal domain. I found that 'A362V' residue was fully conserved in 62 accessions. This was consistent with the previous finding where alanine was found conserved in ELF3 sequences from 20 accessions (Coluccio et al., 2011). It has been reported that the Q-stretch is highly polymorphic between Arabidopsis accessions, which was consistent with my finding in a different set of accessions. A weak correlation between the circadian period and the number of glutamine was reported (Takeomi et al., 2007). However, I could not find such a relationship between glutamine number and period length (Figure 3.4 b-e). Different sets of accessions used in both studies might be one reason for this inconsistency. Further, the SAR observed in 22 Arabidopsis accessions did not correlate with glutamine number (Coluccio et al., 2011). Collectively, it seems that the variation in number of glutamine residues in C-terminal domain of ELF3 is, at most only marginally linked to the clock speed. Considering the fact that ELF3 is a multifunctional protein, and different domains of ELF3 modulate interactions with different components, the involvement of the Q-stretch in some other physiological processes is plausible. Especially, considering that the Q-stretch in the mouse CLOCK protein is required for its histone acetyltransferase (HAT) activity (Doi et al., 2006), the functional involvement of ELF3 Q-stretch in other physiological process cannot be ignored.

5.1.3 ELF3-Sha: the first natural allele of a clock component

After eliminating the possible role of glutamine deletion in *ELF3-Sha* period acceleration, I demonstrated that a single residual change A362V in the middle domain of ELF3 underlies the *ELF3-Sha* short-period phenotype. Proof was obtained from the observation that transgenic lines harboring 'Valine' (*SpSc* and *SpSa2v*) displayed period acceleration compare to transgenic lines harboring 'Alanine' (*SpBc* and *SpSv2a*) (Figure 3.5 a,b). Further using *ELF3-A362V* lines that were driven under Bay promoter, I excluded the possibility of any differential promoter regulation in *ELF3-Sha* (Figure 3.5 c,d). This was the first report where a QTL detected in a dedicated clock study was resolved to a resolution of quantitative trait nucleotide (QTN). In previous clock studies, the QTLs detected for periodicity, phase or temperature compensation experiments were only validated in NILs and based on sequence comparison possible candidate genes were proposed (Boikoglou et al., 2011; Darrah et al., 2006; Edwards et al., 2005; Michael et al., 2003b; Swarup et al., 1999). None of these studies fully confirmed the underlying QTN or QTG. Probably this was the reason that most of the known clock genes were detected in mutagenesis screens. The isolation of *ELF3-Sha* to a QTN resolution and further its characterization to the sub-cellular level provides a reminder that a rich resource of allelic variants is available as natural accessions.

5.1.4 ELF3: a light and clock regulated gene

The differential response of *ELF3-Sha* to light and darkness was intriguing. Under LL, the *ELF3-A362V* lines displayed robust rhythms with enhanced speed of oscillation (Figure 3.5 c), whereas, in prolonged darkness, the robustness of the rhythms was gradually dampened and the increase in the oscillator speed disappeared (Figure 3.6 b). This provided a clue that the defect in *ELF3-Sha* was augmented in DD. This was confirmed in an experiment where free-running profile of *LHY::LUC* expression was measured in *ELF3-Bay* and *ELF3-A362V* for 15 continuous days under different regimes of light and darkness. As long as the plants remained under light, the expression of *LHY::LUC* was rhythmic for both

ELF3-Bay and *ELF3-A362V*. However, in darkness a rapid decrease in *LHY::LUC* expression was detected for both lines. Importantly, *ELF3-Bay* maintained proper rhythms with low precision, whereas, in *ELF3-A362V* the *LHY::LUC* expression was almost arrhythmic. Further, both lines quickly restored normal rhythms when transferred back into the light (Figure 3.7).

One of the reason for this lack of precision of *LHY::LUC* in *ELF3-A362V* in DD could be the stability of ELF3 itself. It is known that ELF3 protein rapidly degrades in darkness and it accumulates under LL (Liu et al., 2001). Thus a relatively rapid decay of ELF3-A362V compared to ELF3-Bay in DD could explain its arrhythmic phenotype. Support for this hypothesis was obtained in an experiment that mimics the standard gating experiment. It has been established that ELF3 represses light regulated gene expression, so that the gate remains closed when the expression of ELF3 is highest in the early subjective night (Covington et al., 2001). Conversely, in *elf3* null mutants gate remains open all the time points depicting no repression of light. In such an experiment, I applied light pulses to the *ELF3-Bay* and *ELF3-A362V* plants that were adapted in DD for 4 days (an approximate time when *ELF3-A362V* becomes arrhythmic) and monitored the *LHY::LUC* expression. I observed a phase advance in *ELF3-A362V* compare to ELF3-Bay depicting an open gate (data not shown). This kind of hyper-response to light has been observed for *ZTL*, whose expression is regulated by both circadian clock and light (Xu and Johnson, 2001). Thus I propose that *ELF3* is also regulated by both light and circadian clock.

5.1.5 ELF3-Sha: supports dual repressor function of ELF3

The *elf3* loss-of-function mutants under free-running conditions were previously shown to display major defects in the expression of core oscillator genes *PRR7*, *PRR9*, *CCA1*, *LHY*, *TOC*1 and *GI* (Dixon et al., 2011; Kikis et al., 2005; Kolmos et al., 2011). To monitor the expression of these clock genes in *ELF3-Sha*, I used the robust luciferase reporter-expression system and measured their expression profiles in Bay-0, NIL-S and the nulls *elf3-1* and *elf3-4*. My results were consistent with the previous reports showing the

arrhythmic low levels of *CCA1* and *LHY* and high levels of *TOC1* and *GI* in *elf3* loss-of-function mutants. In contrast to the null phenotype, under LL, *ELF3-Sha* was rhythmic for all genes studied, and it displayed an intermediate level of expression compared to wild type and loss-of- function *elf3* (Figure 3.8 a-f). Specifically, NIL-S displayed higher expression of *TOC1* and *GI* compared to Bay-0, but lower expression than *elf3-1* and *elf3-4*. In contrast, expression of *CCA1* and *LHY* was lower in NIL-S, compared to Bay-0, but it was higher than *elf3-4*. Consistent with the previous finding that CCA1 and LHY activates *PRR7* and *PRR9* (Farre et al., 2005; Nakamichi et al., 2010), a higher transcript abundance of *PRR7* and *PRR9* was observed in NIL-S compare to null mutants (Figure 3.8 e,f). These data and recent findings that the evening complex (ELF3, ELF4, and LUX) directly binds to the *PRR9* promoter to repress transcription support the repressive action of ELF3 in the core oscillator [reviewed in (Herrero et al., 2012)]. Higher expression levels of the evening genes *TOC1* and *GI*, as well as dramatically dampened oscillations of *ELF3-Sha* in darkness, cannot be simply explained by only considering ELF3 at *PRR9*. Especially, considering the observation that double loss-of-function mutant *cca1-11 lhy-21* is rhythmic, an additional repressive role of ELF3 by targeted degradation of GI is conceptually plausible (Clark et al., 2007; Yu et al., 2008). Taken together, my data support a hypothesis which holds that the EC-containing ELF3 has more than one entry point in the circadian clock (Herrero et al., 2012; Kolmos et al., 2011; Kolmos et al., 2009).

5.1.6 ELF3-Sha*: sub-cellular misplacement associates to its functional defect*

ELF3 is a reported nuclear protein that has several binding partners. These include the red-light photoreceptor phyB, the ubiquitin E3-ligase COP1 and the clock-related proteins GI and ELF4 (Covington et al., 2001; Herrero et al., 2012; Nusinow et al., 2011; Yu et al., 2008). While the increased level of *ELF3-Sha* transcript does not support its short-period phenotype, I propose that ELF3-Sha localization defects, which include a decrease in its nuclear pool and formation of fewer nuclear speckles, could be causal for its short-period phenotype (Figure 3.11). Nuclear speckles could be suspected as interaction points

of binding proteins (Herrero et al., 2012; Yu et al., 2008). Thus, formation of fewer speckles in ELF3-Sha might be the result of a defect in the binding of one of its interacting proteins. This could explain why *ELF3-Sha* displayed a shorter period despite the generation of more total protein. Since the A362V variant is located in the middle domain of ELF3, the defect in the binding of ELF3-Sha with either GI and/or ELF4 is plausible. Similarly, ELF3 has been shown to be an adaptor protein required for the proper function of both GI and ELF4 (Herrero et al., 2012; Nusinow et al., 2011; Yu et al., 2008). Since overexpression of ELF3 only mildly affects the circadian clock network, as compared to the severe effect of null mutants, it has been proposed that a minimal level of fully functional ELF3 would be enough for its normal function (Covington et al., 2001; Dixon et al., 2011). Thus I concluded that a defect in cellular distribution, not the level of ELF3-Sha, explains its short-period phenotype.

5.1.7 ELF3-Sha and the possibly of a quantitative-interacting modifier

The introgressions line (NIL-S) carrying Sha allele at *ELF3* locus displayed shorter-period compare to Bay-0, whereas, no periodicity difference was observed in Sha and Bay-0 (Figure 3.1 b). Further, the *ELF3-Sha-YFP* and *ELF3-A362V* transgenic lines in Ws-2 background were also short-period (Figure 3.3; Figure 3.5 b). Moreover, the effect of *ELF3-Sha* on hypocotyl length in NIL-S and Sha was contrasting. Under all light qualities tested, the hypocotyl length of NIL-S was similar to Bay-0, whereas, it was significantly different than Sha (Figure 3.12 b). These result provided indications that *ELF3-Sha* behaves differentially in different genetic backgrounds. Specifically, the periodicity effect of ELF3-Sha appeared to be neutralized in Sha background. I further confirmed this in transcript abundance assays, where mRNA accumulation profiles of several clock genes in NIL-S, Sha, Bay and *elf3-4* were monitored under LL. Transcript abundance of all clock genes tested were essentially similar in Bay-0 and Sha. Conversely, NIL-S displayed an early peak of accumulation for all genes along with de-repression of *PPR7*, *PRR9*, *TOC1* and *GI* and repression of *CCA1* and *LHY* (Figure 3.10 a-h). This leads me to hypothesize that Sha wild

type must have a modifier to offset its *ELF3-Sha*, which would attenuate the hypomorphic defects that were seen in the context of the Bay-0 and Ws-2 genomic backgrounds. Such trans modifiers of defective allelic state have similarly been reported for seed longevity in the Arabidopsis mutants *abscisic acid insensitive3* and *leafy cotyledon1 (Sugliani et al., 2009)*, yield related traits in Tomato (Eshed and Zamir, 1996) and odor-guided behavior in Drosophila (Fedorowicz et al., 1998). The explanation I propose is that the encoded A362V residual change underlying *ELF3-Sha* was a probable deleterious event. To compensate, the preceding accession state of Sha overcame this with a second-site modifier. Although the identity of such a hypothetical modifier has not yet been established, such a hypothesis can be confirmed in further studies using realistic environmental conditions in which the Sha accession is derived.

5.1.8 ELF3-Sha and elf3-12: a comparative analysis

ELF3-Sha and recently characterized *elf3-12 (Kolmos et al., 2011)* alleles have many phenotypic features in common. Both alleles carry a single encoded amino-acid change in the middle domain of ELF3 that is proposed to be necessary for ELF3 function. The residual change in *ELF3-Sha* is A362V whereas in *elf3-12* it is G326D. Both these residues are fully conserved among *ELF3* sequences from number of Arabidopsis accessions and from a variety of plants (Kolmos et al., 2011). Both *ELF3-Sha* and *elf3-12* displayed light-dependent short-period for *CCR2::LUC*, whereas, in DD, they both exhibited no period difference. However, in prolonged darkness *ELF3-Sha* became arrhythmic. Contrarily, no such aberration was reported for *elf3-12 (Kolmos et al., 2011)*. Interestingly, in DD, the expression rhythms for many clock genes in *elf3-12* that include *CCA1*, *LHY*, *TOC1*, *GI*, *PRR7* and *PRR9* were essentially similar to the loss-of-function mutants depicting a movement towards arrythmicity (Kolmos et al., 2011). Further, the comparison of expression patterns of all these genes in *ELF3-Sha* and *elf3-12* revealed that they behave similar. Thus it seems that *elf3-12* might also has tendency to be arrythmic in DD.

Developmental processes that include hypocotyl length and flowering time

remained unaltered in both alleles showing that the ELF3 domain required to control these processes is functional. Moreover, similar defects in the subcellular localization were observed for both alleles. Specifically, compare to wild types, the cytoplasmic to nucleus ratio of ELF3 protein was higher for *ELF3-Sha* as well as *elf3-12* [(Figure 3.11; (Kolmos et al., 2011)]. Collectively, based on these observations I assume that the functional mode-of-action of both alleles is similar.

5.1.9 Mechanistic explanation of ELF3-Sha clock defects

ELF3 is a nuclear localized, multifunctional protein that has several binding partners. These include PHYB, COP1, ELF4 and GI (Herrero et al., 2012; Yu et al., 2008). Different domains of ELF3 specifically interact with different proteins. Both PHYB and COP1 interact with the N-terminal domain (Kolmos et al., 2011; Liu et al., 2001; Yu et al., 2008), whereas, ELF4 and GI interact with the middle domain (Herrero et al., 2012; Yu et al., 2008). No protein has yet been shown to interact with the C-terminal domain. However, this domain carries nuclear localization signal (NLS) and shown to be important for the nuclear localization of ELF3 (Herrero et al., 2012; Liu et al., 2001). Further, all these proteins co-localize in the nucleus where they form distinct nuclear bodies (Chen et al., 2010; Herrero et al., 2012; Mas et al., 2000; Yu et al., 2008). Thus, it is plausible that these proteins functions together in control of complex physiological mechanisms.

It could be expected that the residual change in the middle domain of ELF3-Sha has induced a binding defect with its interacting proteins that resulted in attenuated function. Both PHYB and COP1 could be excluded for any such possibility because their region of interaction lies outside the point of residual change. Both ELF4 and GI bind to the ELF3 middle domain, the region where A362V is localized. Thus a binding defect with one or both of these proteins could be a potential reason for ELF3-Sha attenuated function.

It was proposed that ELF4 act as an effector protein that interacts with ELF3 and activates it in the nucleus to initiate its repressive action (Kolmos et al., 2009). Such an activation defect in *elf3-12* has been proposed to result in its attenuated function (Kolmos

et al., 2011). As both *elf3-12* and ELF3-Sha carries a mutation in the same ELF4-binding domain, this kind of activation defect can also be assigned to ELF3-Sha. However, it was notable that in yeast two hybrid (Y-2H) assay, ELF4 was capable to interact with ELF3-12 protein (Kolmos et al., 2011). Thus, other possibility of ELF3-Sha functional defect, such as a compromised binding affinity with GI, cannot be ignored.

In a Y-2H assay GI and ELF3 were found to physically interact (Yu et al., 2008). Further, ELF3 modulates GI expression by mediating its targeted decay that involves its physical binding with COP1. Notable here is that ELF3 is essential for this process. In absence of ELF3, COP1 could not interact with GI and thus cannot initiate its decay (Yu et al., 2008). Contrarily, both ELF3 and COP1 can interact even in the absence of GI. This interaction leads to ELF3 degradation (Yu et al., 2008). However, COP1 itself is also a target of proteasomal decay. Important to note here is that all these processes starts in the darkness, when the level of both ELF3 and GI reach to maximum in the evening (Yu et al., 2008). Thus these processes could be taken as an active mechanism to regulate the proper amount of ELF3 and GI.

Based on the observation that *ELF3-Sha* could not sustain rhythmicity under prolonged darkness, I propose that *ELF3-Sha* encodes a protein that is defective in binding with GI. This binding defect results in a weak physical association of ELF3-Sha and GI that disrupts the balance of COP1 action on these proteins, favoring ELF3-Sha degradation more than GI. In this way, in prolonged darkness, ELF3-Sha reaches to the minimum level required for its function, more rapidly compare to normal ELF3-Bay, which results in arrythmicity.

The above hypothesis can be further supported by phase response curve (PRC) experiments results performed in ELF3-Sha functionally similar allele *elf3-12*. For these experiments, the *elf3-12* plants were entrained for 7 days under LL and then transferred for one day in DD. The next day light pulses of one hour was given every 3 hour and phase changes were monitored (Kolmos et al., 2011). During the initial subjective evening and early night (CT12 to CT18), light pulses delayed the phase similarly in *elf3-12* and the wild

type. By contrast, *elf3-12* had increased phase advances (up to 4 h) during the late subjective night and the following early subjective day (CT18 to CT27) (Kolmos et al., 2011). As the degree of clock resetting was suggested to be inversely correlated with the level of ELF3 (Covington et al., 2001), a comparatively low level of ELF3-12 protein during late subjective night could account for these phase differences.

      ELF3-Sha exhibited short period under LL with higher levels of *PRR7*, *PRR9*, *TOC1* and *GI*, and lower levels of *CCA1* and *LHY* (Figure 3.8 a-f). As it is known that ELF3 directly represses *PRR9* by binding to its promoter (Herrero et al., 2012; Nusinow et al., 2011), a lower amount of ELF3-Sha protein sustained after initial light dark cycles could account for the higher expression of *PRR9/PRR7*. Further, higher expression of *PRR7/PRR9* resulted in increased repression of *CCA1/LHY* which ended in higher *TOC1* expression. This higher *TOC1* expression could also contribute to the low levels of *CCA1/LHY*, as TOC1 is shown to repress *CCA1/LHY* by binding to their promoters (Gendron et al., 2012). Higher *GI* expression could be explained by derepression of CCA/LHY. The higher level of *GI* is consistent with previous findings that exclude the direct link from TOC1 to GI (Pokhilko et al., 2010). Collectively, the initial derepression of *PRR9/PRR7* defines the pace of the oscillation in *ELF3-Sha* that is maintained through the LL and results in a short-period phenotype.

      Though this model can explain all phenotypes of *ELF3-Sha*, a defect in ELF3-SHA to ELF4 binding is still plausible. Especially, considering the recent finding that ELF4 recruits ELF3 in the nucleus (Herrero et al., 2012), the higher cytoplasmic pool of ELF3-Sha could be a result of ELF3-Sha to ELF4 binding defect. Moreover, as the peak of expression of *ELF3*, *ELF4* and *GI* follows same diurnal pattern, and all have been reported to diminish under darkness (Doyle et al., 2002; Liu et al., 2001; McWatters et al., 2007; Yu et al., 2008), it is possible that ELF3 also recruit ELF4 in the COP1-GI complex and mediate its degradation: especially when all of these proteins localize in same nuclear bodies (Herrero et al., 2012; Yu et al., 2008). Collectively, I propose that the kinetics of ELF3 degradation in darkness is an important mechanism that modulates the pace of the circadian oscillation. Here, proper

binding of ELF3 with GI provides ELF3 stability in darkness.

5.2 Characterization of RILs

In chapter 4 of my thesis, I described the generation of three RIL populations that include Strand X Ws-2 and Ws-2 X Tnz-1 and Strand X Tnz-1. Two of these populations were then further characterized for flowering time and circadian-periodicity traits. Several QTLs were detected and potential candidates were proposed. One periodicity locus was also validated in a NIL. All these processes were supported by the whole genome re-sequencing of the parental accessions. Collectively, my results described the possible genetic connection between flowering time and clock speed.

5.2.1 Re-sequencing of the parental accessions

The re-sequencing of the three parental accessions: Strand, Ws-2 and Tnz-1 was performed using next-generation Illumina/Solexa platform. After aligning the short-sequence reads obtained from Solexa, the average genome coverage achieved for these accessions was 15 to 27X (Table 4.1). This genome coverage proved sufficient for the subsequent variant calling and was consistent with the coverage reported for the re-sequencing of the 80 Arabidopsis accessions (Cao et al., 2011). Enormous polymorphisms as SNPs and InDels were detected in the variant calling for all three genomes. I found that, on whole-genome scale, the equatorial accession Tnz-1 was highly diverged compared to the northern accessions Strand and Ws-2. This conclusion was based on the observation that highest numbers of total polymorphisms were detected in Tnz-1 genome. In contrast, Ws-2 was the least polymorphic among the three accessions (Table 4.1; Figure 4.2). The total numbers of SNPs detected in Ws-2 genome were 377764, which were higher than the previously reported SNPs for the re-sequencing of this accession (Ashelford et al., 2011). Several factors including difference in NGS platform, coverage depth, and statistical approach used of the analysis in both studies could contribute to these dissimilarities.

In all three genomes, both SNPs and InDels were uniformly distributed over all five

chromosomes, and the contribution of variation supplied by a chromosome was directly proportional to its length. This confirms the accuracy of the sequencing as well as the subsequent data analysis (Figure 4.4; Figure 4.5). The Ws-2 results were consistent with the previous report for this genome using ABI SOLiD sequencing platform (Ashelford et al., 2011).

Though the SNPs and InDels were uniformly distributed over the complete genome, small gaps around centromeres were detected (Figure 4.5). The centromeres are known to comprise repetitive DNA sequences and short reads obtained from NGS platforms are difficult to align on such kind of regions which often results in gaps (Treangen and Salzberg, 2012). Collectively, the re-sequencing of the Strand, Ws-2 and Tnz-1 genomes greatly facilitated the later steps of my quantitative genetic analysis of RILs that were generated using these accessions as parent.

5.2.2 Generation of RILs

Three sets of RILs that include Strand X Ws-2, Ws-2 X Tnz-1 and Strand X Tnz-1, were generated by making pair-wise crosses in such a way that each RIL set had one parent in common among them (Chapter 2.2.2; Figure 2.1). This sort of connected-RIL populations have been shown to increase the power to detect QTL, which is achieved by making a single integrated genetic map of all RILs followed by joint QTL analysis, using all RILs as one set (Klasen et al., 2012; Kover et al., 2009). In maize, such joint analysis resulted in as much as twice the number of QTL identified, as compare to the single population analysis of six connected F2 populations (Blanc et al., 2006). Additional power was provided by generating RIL populations in such a way that each RIL set was further comprised of sub-populations, which included F2, backcross (BC) and intercross (F2X) lines (Chapter 2.2.2; Table 2.14). The backcross and intercross lines had been shown to increase the QTL identification power by increase the recombination events in the resultant lines (Klasen et al., 2012). Collectively, a comprehensive genetic resource in the form of systematic RILs was generated for further quantitative analysis.

Finally, it was ensured during RILs generation that each RIL must contain the *CCR2::LUC* (see chapter 2.2.2). Thus, these lines provide an outstanding resource for the genetic analysis of circadian assays under high-throughput platforms. Though, such kind of RILs harboring *CAB2::LUC* (Darrah et al., 2006) or *CCR2::LUC* (Boikoglou et al., 2011) already exist, those RILs were generated by incorporating the luciferase transgene in already existing RIL populations. Hence, the phenotypic effects delivered by the positional insertion of the transgene cannot be ignored. Importantly, these phenotypic effects do not have genetic basis. Therefore, the subsequent mapping analysis could be compromised. I eliminated this problem by generating RILs in such a way that only one of the two parents in the initial cross had the luciferase transgene, while in subsequent generation luciferase positive lines were selected. In this way, all the RILs obtained in the particular cross had a fixed position of the *CCR2::LUC* donated by a single parent. Thus, the possibility of phenotypic variation delivered by positional effect of the luciferase transgene was excluded. However, the selection of the *CCR2::LUC* positive lines for several generations resulted in a higher segregation distortion in the RILs at this genomic position, which I have discussed in the next section.

5.2.3 Linkage mapping

Two RIL populations, Strand X Ws-2 and Ws-2 X Tnz-1, were genotyped using uniformly distributed markers. The even distribution of the markers is an important factor, since Zeng et al. (1999) showed that, the QTL mapping statistic test can be affected by an uneven distribution of markers over the genome (Zeng et al. 1999). For both RIL populations, an overall bias for the selection of Ws-2 was observed, as expected base on the crossing scheme employed (Chapter 2.2.2). Except at chromosome 3, the relative contribution of Ws-2 allele never exceeded 63% in the genome (Table 4.4; Figure 4.7). This level of segregation is common in Arabidopsis and has been reported for several RILs (Alonso-Blanco et al., 1998; Lister and Dean, 1993; Loudet et al., 2002; Singer et al., 2006). Such distortions could simply be explained by the existence of unintentional selection

pressure. For instance, an epistatic relation between two loci might lead to a decreased germination rate of the seeds containing a certain combination of alleles. These alleles would be, at least partly, counter-selected and, by linkage, the markers around the loci would reflect this distortion. However, on chromosome 3 an extreme segregation distortion was observed. I found that the representation of Ws-2 allele on the top of chromosome 3 was ~80% (Table 4.4; Figure 4.7). This was the region where the *CCR2::LUC* transgene was inserted in the Ws-2 genome. As Ws-2 (*CCR2::LUC*) was the common pollen donor parent for both RIL sets, a repetitive selection of *CCR2::LUC* positive lines for several generations could be the reason for this high segregation distortion.

The higher segregation distortion on chromosome 3 resulted as a reduced chromosomal length in the linkage map. This problem was ''fixed'' using marker location information available from the other maps, and by calculating the approximate relationship between physical and genetic distance, based on the genome-wide recombination rate of that population. However, it was obvious that the overall accuracy of the marker positions on chromosome 3 had been compromised.

A direct relationship between the physical and the genetic length of the chromosomes 1, 2, 4 and 5 was observed. Further all marker assigned to these chromosomes fully corresponded to their physical location on the genome (Figure 4.8). This ensured that the segregation distortion on chromosome 3 did not affect the mapping accuracy of the other chromosomes, which could have arisen because of spurious linkages.

The genetic map of Strand X Ws-2 (448 cM) was longer than the Ws-2 X Tnz-1 RIL population (412 cM) (Figure 4.8; Figure 4.17). However, map differences in RILs are normal and have been reported for several populations (Alonso-Blanco et al., 1998; Balasubramanian et al., 2009; Lister and Dean, 1993; Loudet et al., 2002). Usually, it arises because overall recombination frequencies varies when different genomes combines. Another reason for this difference could be the size of the populations. As Strand X Ws-2 comprised 192 lines compare to 78 lines of Ws-2 X Tnz-1, it is likely that the few recombination events have been missed In Ws-2 X Tnz-1 RIL set due to lack of sufficient

lines.

For both RIL populations, the average genetic distance between two adjacent markers was ~8.0 cM. The marker density was sufficient to achieve optimum efficiency of QTL mapping. Indeed, simulation studies have shown that the advantages of increasing marker density beyond one marker every 10 cM are less significant than those obtained when increasing the size of the population (Darvasi and Soller, 1994; Gilles, 2000). Collectively, I generated good quality genetic maps of Strand X Ws-2 and Ws-2 X Tnz-1 RIL sets that could be used as an invaluable resource for further genetic studies.

5.2.4 Genetic analysis of circadian periodicity and flowering time

The RIL populations Strand X Ws-2 and Ws-2 X Tnz-1 were phenotypically characterized for circadian periodicity and flowering time. Further, circadian periodicity was measured after two entrainment conditions: LD and WC. The statistical analysis revealed that RILs periodicity after both the entrainment conditions was significantly different; hence, each can be taken as an independent trait (Table 4.6; Table 4.7). The frequency distribution of all three traits revealed a transgressive variation in the phenotypes in both directions from the parents (Figure 4.9; Figure 4.18). This suggested that parents had obtained a balance of short-period/long-period alleles or early-flowering/late-flowering alleles through evolution, and this unique balance was broken in the RILs because of recombination, which appeared as a transgression in the RILs. This transgression is a common mechanism and has been reported in several RILs for both traits (Darrah et al., 2006; Edwards et al., 2005; Giakountis et al., 2010; Loudet et al., 2002; Michael et al., 2003b).

In total, for both RIL populations, 15 QTLs were detected: 9 for circadian periodicity and 6 for flowering time (Figure 5.2). Many of these QTLs overlapped with the already reported loci for these traits. As such, potential candidate genes can be proposed. Two loci were novel, one at the bottom of chromosome 4 for circadian periodicity and other at the middle of chromosome 1 for flowering time. No clock or flowering-time gene has yet been

identified on these positions (Figure 5.1). Thus, these regions could be a potential targets for the isolation of novel components.
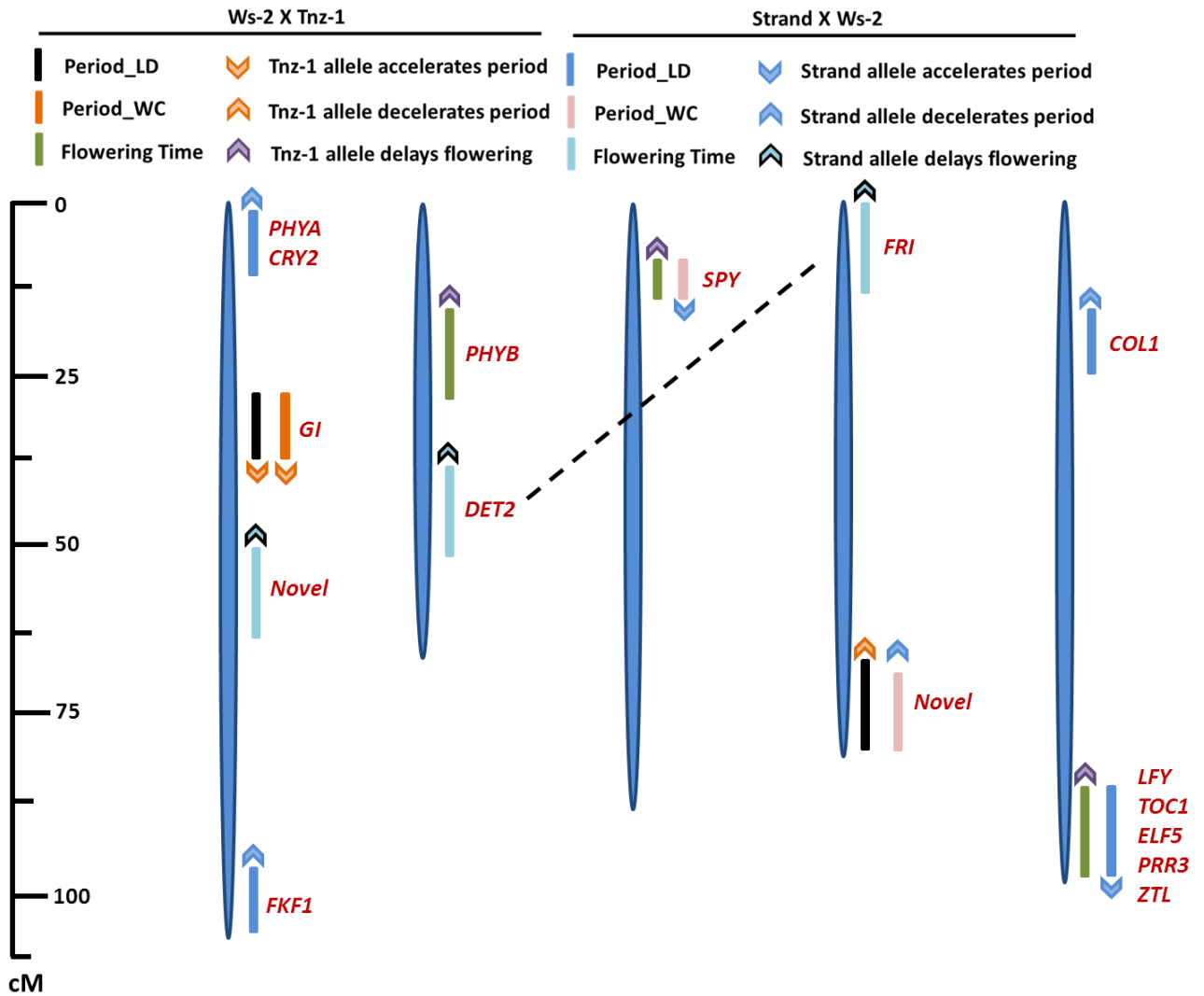
Figure 5.1: Concentrated summary of the QTL mapping.

Scale bar represents distance in centimorgans (cM). Vertical bars represent QTLs location on the genome. Candidate genes are shown on the right side of the bars. Arrows denote the effect of the Strand or Tnz-1 allele on the Ws-2 phenotype.

5.2.4.1 *GI* as a major-effect contributor of periodicity variation

A major periodicity QTL was detected at chromosome 1 in Ws-2 X Tnz-1 RIL population (Figure 4.12). Co-localizing QTLs for temperature compensation, circadian phase and clock speed have been reported (Boikoglou et al., 2011; Darrah et al., 2006; Michael et al., 2003b; Swarup et al., 1999). *GI* a known regulator of flowering time and circadian clock was the potential candidate. I validated this QTL using a NIL, and further, by sequence comparison of the parental accessions, identified that the GI allele of Tnz-1 carries a single nucleotide change from 'C' to 'T' in Tnz-1 that resulted in an encoded amino-acid change from 'V343A' (Figure 4.15). I found by sequence comparison of 20 accessions that 'Valine' is a conserved residue (data not shown). Thus, I proposed that 'V343A' in Tnz-1 *GI* variant underlies the chromosome 1 QTL.

GI regulates both light and temperature inputs to the clock. Further, *gi* mutants exhibit short-period phenotype under free-running conditions (Gould et al., 2006; Mizoguchi et al., 2005). Consistently, the NIL-29 carrying Tnz-1 allele at *GI* locus displayed an increase in the clock speed after both LD and WC entrainment. Moreover, the periodicity effect of Tnz-1 was larger after WC entrainment compare to LD entrainment (Figure 4.15). This was consistent with the response of Tnz-1 wild type to the temperature cycles, as I observed that speed of oscillation after WC was increased by ~2h compare to LD (Figure 4.9). This showed that *GI* largely contributes to the hyper-temperature response of Tnz-1.

GI is a positive regulator of flowering time and mutations in *GI* normally cause late flowering. However, mutants such as *gi-611* have been reported that exhibited short-period, but no flowering-time phenotype (Gould et al., 2006). As, no flowering-time QTL at *GI* locus was identified, I assumed that *GI* in Tnz-1 is functional in flowering-time control.

5.2.4.2 *FRIGIDA* as a major-effect contributor to flowering-time variation

In the Strand X Ws-2 RIL population, a major flowering-time QTL was detected at the top chromosome 4 (Figure 4.21). The presence of Strand allele at this locus delayed the flowering time by 15 days, which was consistent with the action of the candidate gene,

FRIGIDA (FRI). *FRI* had been established as a major contributor in flowering-time variation in natural accessions (Clark et al., 2007). A functional *FRI* allele delays flowering by promoting the expression of the flowering repressor *FLC (Geraldo et al., 2009)*. However, under prolonged coldness, a process called vernalization, *FLC* becomes epigenetically silenced and flowering is induced (Shindo et al., 2006). In this way, *FRI* ensures that the Arabidopsis accessions growing under extreme cold conditions can grow vegetatively until the inductive and less stressed flowering-time conditions have seasonally arrived. Consistently, many northern accessions carry a functional copy of *FRI*, allowing them to cope with the extreme cold winter (Johanson et al., 2000). Thus, it was not surprising that I found in sequence analysis that Strand carries a functional copy of *FRI*. Conversely, as reported previously that Ws contained a non-functional *FRI*, a 16bp deletion in the coding region of *FRI* allele in Ws-2 was detected (data not shown). This explained the genetic basis of the chromosome 4 top QTL.

It has been reported that *FRI* and *FLC* epistatically interact to confer flowering-time variation. In Strand X Ws-2 RIL population, I identified another locus at chromosome 2 that interacted with *FRI* locus to modulate flowering time. *DET2* was the candidate gene here. *det-2* mutants are known to increase flowering time in Arabidopsis (Chory et al., 1991). The presence of Strand allele at this locus delayed flowering by 9 days. However, an epistatic interaction of Strand allele at both *DET2* interval and *FRI* loci act additively and cause further delay in flowering by 24 days (Table 4.12). Further study to reveal the molecular mechanism of this genetic interaction would be interesting.

5.2.4.3 *PHYB* and the possibility of a hyperactive allele

The allelic effect of the most of the QTLs identified for clock periodicity and flowering time, in both Strand X Ws-2 and Ws-2 X Tnz-1 RIL population, followed the effect of the mutant alleles of their proposed candidate genes (Figure 5.2). As one example, presence of Strand allele at chromosome 1 periodicity QL, delayed speed of clock by 0.65 h (Table 4.12). *PHYA* was proposed as a candidate gene for this locus. *phyA* mutants delay speed of oscillation (Somers, 1998). Thus it is plausible that Strand allele of *PHYA* is a

mutant variant. Conversely, *PHYB* was the candidate gene for the chromosome 2 flowering-time QTL, detected in Ws-2 X Tnz-1 RIL population (Figure 5.2). *phyB* mutants exhibit early flowering (Liu et al., 2001). Here, the allelic effect of the candidate gene, was opposite to its known mutant phenotype. The sequence analysis of Ws-2 and Tnz-1 *PHYB* gene revealed two non-synonymous changes in the coding region, depicting that the *PHYB* allele in Tnz-1 is a plausible mutant variant. However, presence of Tnz-1 at this locus delayed flowering. This raised two possibilities. Either there is a novel flowering-time gene locates at this region or Tnz-1 allele of *PHYB* is a hyperactive allele. Such kind of hyperactive allele affecting Arabidopsis growth vigor has been reported (Beuchat et al., 2010). The further positional isolation could discriminate these possibilities.

5.2.4.4 Conclusion and perspective

The genetic dissection of the two RIL populations Strand X Ws-2 and Ws-2 X Tnz-1 revealed several QTLs underlying phenotypic variation in circadian periodicity and flowering time. The majority of these QTLs were co-localized with the already reported regions underlying these traits. For instance, the top and middle of chromosome 1, and bottom of chromosome 5 have been identified in many clock studies for various clock related parameters. Further, for flowering time, a recent study reported that as many as 5 regions of the Arabidopsis genome can be attributed to majority of the natural variation in flowering time. These regions include bottom of chromosome 1, middle part of chromosome 2, top of chromosome 4, and top and bottom of chromosome 5 (Salomé et al., 2011). My analysis of flowering time revealed that at least three of these regions, middle part of chromosome 2, top of chromosome and bottom of chromosome 5, are associated to the flowering-time variation in the Strand X Ws-2 and Ws-2 X Tnz-1 mapping populations. Collectively, my results add new genetic components to these traits as well as support the existing ones.

The unique connected design of my RILs population raised interesting questions. As one example, the chromosome 4 QTL with *FRI* as candidate was only detected in Strand X Ws-2 RIL population, whereas, this QTL was absent in the Ws-2 X Tnz-1. As Ws-2 was a

common parent in both these populations and further Ws-2 carried a non-functional *FRI* allele, it is possible that Tnz-1 also contain a non-functional allele. Thus the presence of two alleles with similar effects resulted in the absence of this QTL in Ws-2 X Tnz-1 RIL population. Indeed, the sequence analysis revealed 3 non-synonymous residual change along with a frame shift deletion in the first exon of Tnz-1 FRI allele, depicting that Tnz-1 also carries non-functional *FRI* (data not shown). If this statement is true then the analysis of the third population Strand X Tnz-1 should reveal the FRI as a flowering-time QTL. This will be soon tested.

References:

Alabadí D., Oyama T., Yanovsky M.J., Harmon F.G., Más P., Kay S.A. (2001) Reciprocal Regulation Between TOC1 and LHY/CCA1 Within the Arabidopsis Circadian Clock. Science 293:880-883.

Alonso-Blanco C., Bentsink L., Hanhart C.J., Blankestijn-de Vries H., Koornneef M. (2003) Analysis of Natural Allelic Variation at Seed Dormancy Loci of *Arabidopsis thaliana*. Genetics 164:711-729.

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

Alonso-Blanco C., Peeters A.J.M., Koornneef M., Lister C., Dean C., Van Den Bosch N., Pot J., Kuiper M.T.R. (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. The Plant Journal 14:259-271.

Aschoff J. (1979) Circadian Rhythms: Influences of Internal and External Factors on the Period Measured in Constant Conditions. Zeitschrift für Tierpsychologie 49:225-249.

Ashelford K., Eriksson M., Allen C., D'Amore R., Johansson M., Gould P., Kay S., Millar A., Hall N., Hall A. (2011) Full genome re-sequencing reveals a novel circadian clock mutation in Arabidopsis. Genome Biology 12:R28.

Aukerman M.J., Hirschfeld M., Wester L., Weaver M., Clack T., Amasino R.M., Sharrock R.A. (1997) A Deletion in the PHYD Gene of the Arabidopsis Wassilewskija Ecotype Defines a Role for Phytochrome D in Red/Far-Red Light Sensing. The Plant Cell Online 9:1317-1326.

Balasubramanian S., Schwartz C., Singh A., Warthmann N., Kim M.C., Maloof J.N., Loudet O., Trainer G.T., Dabi T., Borevitz J.O., Chory J., Weigel D. (2009) QTL Mapping in New *Arabidopsis thaliana* Advanced Intercross-Recombinant Inbred Lines. PLoS ONE 4:e4318.

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

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

Beuchat J., Li S., Ragni L., Shindo C., Kohn M.H., Hardtke C.S. (2010) A hyperactive quantitative trait locus allele of Arabidopsis BRX contributes to natural variation in root growth vigor. Proc Natl Acad Sci U S A.

Blanc G., Charcosset A., Mangin B., Gallais A., Moreau L. (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. TAG Theoretical and Applied Genetics 113:206-224.

Boikoglou E. (2008) Quantitative genetic analysis of temperature entrainment in the *Arabidopsis thaliana* circadian clock. PhD thesis. University of Cologne.

Boikoglou E., Ma Z., von Korff M., Davis A.M., Nagy F., Davis S.J. (2011) Environmental Memory From a Circadian Oscillator: The *Arabidopsis thaliana* Clock Differentially Integrates Perception of Photic Versus Thermal Entrainment. Genetics.

Caicedo A.L., Richards C., Ehrenreich I.M., Purugganan M.D. (2009) Complex Rearrangements Lead to Novel Chimeric Gene Fusion Polymorphisms at the *Arabidopsis thaliana* MAF2-5 Flowering Time Gene Cluster. Molecular Biology and Evolution 26:699-711.

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

Chen M., Galvão R.M., Li M., Burger B., Bugea J., Bolado J., Chory J. (2010) Arabidopsis HEMERA/pTAC12 Initiates Photomorphogenesis by Phytochromes. Cell 141:1230-1240.

Chory J., Nagpal P., Peto C.A. (1991) Phenotypic and Genetic Analysis of det2, a New Mutant That Affects Light-Regulated Seedling Development in Arabidopsis. The Plant Cell Online 3:445-459.

Clark R.M., Schweikert G., Toomajian C., Ossowski S., Zeller G., Shinn P., Warthmann N., Hu T.T., Fu G., Hinds D.A., Chen H., Frazer K.A., Huson D.H., Schölkopf B., Nordborg M., Rätsch G., Ecker J.R., Weigel D. (2007) Common Sequence Polymorphisms Shaping Genetic Diversity in *Arabidopsis thaliana*. Science 317:338-342.

Clarke J., Mithen R., Brown J., Dean C. (1995) QTL analysis of flowering time *Arabidopsis thaliana*. Molecular and General Genetics MGG 248:278-286.

Coluccio M.P., Sanchez S.E., Kasulin L., Yanovsky M.J., Botto J.F. (2011) Genetic mapping of natural variation in a shade avoidance response: ELF3 is the candidate gene for a QTL in hypocotyl growth regulation. J Exp Bot 62:167-76.

Covington M.F., Panda S., Liu X.L., Strayer C.A., Wagner D.R., Kay S.A. (2001) ELF3 Modulates Resetting of the Circadian Clock in Arabidopsis. The Plant Cell Online 13:1305-1316.

Dai S., Wei X., Pei L., Thompson R.L., Liu Y., Heard J.E., Ruff T.G., Beachy R.N. (2011) BROTHER OF LUX ARRHYTHMO Is a Component of the Arabidopsis Circadian Clock. The Plant Cell Online 23:961-972.

Darrah C., Taylor B.L., Edwards K.D., Brown P.E., Hall A., McWatters H.G. (2006) Analysis of phase of LUCIFERASE expression reveals novel circadian quantitative trait loci in Arabidopsis. Plant Physiol 140:1464-74.

Darvasi A., Soller M. (1994) Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. TAG Theoretical and Applied Genetics 89:351-357.

Davis A., Hall A., Millar A., Darrah C., Davis S. (2009) Protocol: Streamlined sub-protocols for floral-dip transformation and selection of transformants in *Arabidopsis thaliana*. Plant Methods 5:3-3.

de Montaigu A., Toth R., Coupland G. (2010) Plant development goes like clockwork. Trends Genet 26:296-306.

Devlin P.F., Kay S.A. (2000) Cryptochromes Are Required for Phytochrome Signaling to the Circadian Clock but Not for Rhythmicity. The Plant Cell Online 12:2499-2510.

Ding Z., Millar A.J., Davis A.M., Davis S.J. (2007) TIME FOR COFFEE Encodes a Nuclear Regulator in the *Arabidopsis thaliana* Circadian Clock. The Plant Cell Online 19:1522-1536.

Dixon L.E., Knox K., Kozma-Bognar L., Southern M.M., Pokhilko A., Millar A.J. (2011) Temporal repression of core circadian genes is mediated through EARLY FLOWERING 3 in Arabidopsis. Curr Biol 21:120-5.

Dodd A.N., Salathia N., Hall A., Kévei E., Tóth R., Nagy F., Hibberd J.M., Millar A.J., Webb A.A.R. (2005) Plant Circadian Clocks Increase Photosynthesis, Growth, Survival, and Competitive Advantage. Science 309:630-633.

Doi M., Hirayama J., Sassone-Corsi P. (2006) Circadian Regulator CLOCK Is a Histone Acetyltransferase. Cell 125:497-508.

Dong G., Golden S.S. (2008) How a cyanobacterium tells time. Current Opinion in Microbiology 11:541-546.

Doyle M.R., Davis S.J., Bastow R.M., McWatters H.G., Kozma-Bognar L., Nagy F., Millar A.J., Amasino R.M. (2002) The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. Nature 419:74-77.

Eckardt N.A. (2005) Temperature Entrainment of the Arabidopsis Circadian Clock. The Plant Cell Online 17:645-647.

Edwards K.D., Lynn J.R., Gyula P., Nagy F., Millar A.J. (2005) Natural Allelic Variation in the Temperature-Compensation Mechanisms of the *Arabidopsis thaliana* Circadian Clock. Genetics 170:387-400.

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

Eshed Y., Zamir D. (1995) An Introgression Line Population of Lycopersicon pennellii in the Cultivated Tomato Enables the Identification and Fine Mapping of Yield-Associated QTL. Genetics 141:1147-1162.

Eshed Y., Zamir D. (1996) Less-Than-Additive Epistatic Interactions of Quantitative Trait Loci in Tomato. Genetics 143:1807-1817.

Farre E.M., Harmer S.L., Harmon F.G., Yanovsky M.J., Kay S.A. (2005) Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock. Curr Biol 15:47-54.

Fedorowicz G.M., Fry J.D., Anholt R.R.H., Mackay T.F.C. (1998) Epistatic Interactions Between smell-impaired Loci in *Drosophila melanogaster*. Genetics 148:1885-1891.

Fornara F., Panigrahi K.C.S., Gissot L., Sauerbrunn N., Rühl M., Jarillo J.A., Coupland G. (2009) Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. Developmental Cell 17:75-86.

Fowler S., Lee K., Onouchi H., Samach A., Richardson K., Morris B., Coupland G., Putterill J. (1999) GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. EMBO J 18:4679-4688.

Gazzani S., Gendall A.R., Lister C., Dean C. (2003) Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. Plant Physiol 132:1107-1114.

Gendron J.M., Pruneda-Paz J.L., Doherty C.J., Gross A.M., Kang S.E., Kay S.A. (2012) Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. Proc Natl Acad Sci U S A.

Geraldo N., Bäurle I., Kidou S.-i., Hu X., Dean C. (2009) FRIGIDA Delays Flowering in Arabidopsis via a Cotranscriptional Mechanism Involving Direct Interaction with the Nuclear Cap-Binding Complex. Plant Physiol 150:1611-1618.

Giakountis A., Cremer F., Sim S., Reymond M., Schmitt J., Coupland G. (2010) Distinct Patterns of Genetic Variation Alter Flowering Responses of Arabidopsis Accessions to Different Day lengths. Plant Physiol 152:177-191.

Gilles C. (2000) Power and accuracy of QTL detection: simulation studies of one-QTL models. Agronomie 20:309-323.

Gould P.D., Locke J.C.W., Larue C., Southern M.M., Davis S.J., Hanano S., Moyle R., Milich R., Putterill J., Millar A.J., Hall A. (2006) The Molecular Basis of Temperature Compensation in the Arabidopsis Circadian Clock. The Plant Cell Online 18:1177-1187.

Graf A., Schlereth A., Stitt M., Smith A.M. (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. Proceedings of the National Academy of Sciences.

Hall A., Bastow R.M., Davis S.J., Hanano S., McWatters H.G., Hibberd V., Doyle M.R., Sung S., Halliday K.J., Amasino R.M., Millar A.J. (2003) The TIME FOR COFFEE Gene Maintains the Amplitude and Timing of Arabidopsis Circadian Clocks. The Plant Cell Online 15:2719-2729.

Harmer S.L., Kay S.A. (2005) Positive and Negative Factors Confer Phase-Specific Circadian Regulation of Transcription in Arabidopsis. The Plant Cell Online 17:1926-1940.

Hazen S.P., Schultz T.F., Pruneda-Paz J.L., Borevitz J.O., Ecker J.R., Kay S.A. (2005) LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc Natl Acad Sci U S A 102:10387-10392.

Helfer A., Nusinow D.A., Chow B.Y., Gehrke A.R., Bulyk M.L., Kay S.A. (2011) LUX ARRHYTHMO Encodes a Nighttime Repressor of Circadian Gene Expression in the Arabidopsis Core Clock. Current Biology 21:126-133.

Herrero E., Kolmos E., Bujdoso N., Yuan Y., Wang M., Berns M.C., Uhlworm H., Coupland G., Saini R., Jaskolski M., Webb A., Gonçalves J., Davis S.J. (2012) EARLY FLOWERING4 Recruitment of EARLY FLOWERING3 in the Nucleus Sustains the Arabidopsis Circadian Clock. The Plant Cell Online.

Hicks K.A., Albertson T.M., Wagner D.R. (2001) EARLY FLOWERING3 Encodes a Novel Protein That Regulates Circadian Clock Function and Flowering in Arabidopsis. Plant Cell 13:1281-1292.

Hicks K.A., Millar A.J., Carré I.A., Somers D.E., Straume M., Meeks-Wagner D.R., Kay S.A. (1996) Conditional Circadian Dysfunction of the Arabidopsis early-flowering 3 Mutant. Science 274:790-792.

Hilscher J., Schlötterer C., Hauser M.-T. (2009) A Single Amino Acid Replacement in ETC2 Shapes Trichome Patterning in Natural Arabidopsis Populations. Curr Biol 19:1747-1751.

Huang X., Feng Q., Qian Q., Zhao Q., Wang L., Wang A., Guan J., Fan D., Weng Q., Huang T., Dong G., Sang T., Han B. (2009) High-throughput genotyping by whole-genome resequencing. Genome Research 19:1068-1076.

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

Imaizumi T. (2010) Arabidopsis circadian clock and photoperiodism: time to think about location. Curr Opin Plant Biol 13:83-9.

Ito S., Niwa Y., Nakamichi N., Kawamura H., Yamashino T., Mizuno T. (2008) Insight into Missing Genetic Links Between Two Evening-Expressed Pseudo-Response Regulator Genes TOC1 and PRR5 in the Circadian Clock-Controlled Circuitry in *Arabidopsis thaliana*. Plant and Cell Physiology 49:201-213.

Ito S., Song Y.H., Josephson-Day A.R., Miller R.J., Breton G., Olmstead R.G., Imaizumi T. (2012) FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. Proc Natl Acad Sci U S A. 109:3582-3587.

Jimenez-Gomez J.M., Wallace A.D., Maloof J.N. (2010) Network analysis identifies ELF3 as a QTL for the shade avoidance response in Arabidopsis. PLoS Genet 6.

Johanson U., West J., Lister C., Michaels S., Amasino R., Dean C. (2000) Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. Science 290:344-347.

Jones M.A., Covington M.F., DiTacchio L., Vollmers C., Panda S., Harmer S.L. (2010) Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. Proc Natl Acad Sci U S A 107:21623-8.

Kiba T., Henriques R., Sakakibara H., Chua N.-H. (2007) Targeted Degradation of PSEUDO-RESPONSE REGULATOR5 by an SCFZTL Complex Regulates Clock Function and Photomorphogenesis in *Arabidopsis thaliana*. The Plant Cell Online 19:2516-2530.

Kikis E.A., Khanna R., Quail P.H. (2005) ELF4 is a phytochrome-regulated component of a negative-feedback loop involving the central oscillator components CCA1 and LHY. Plant J 44:300-13.

Kim J.-Y., Song H.-R., Taylor B.L., Carre I.A. (2003) Light-regulated translation mediates gated induction of the Arabidopsis clock protein LHY. EMBO J 22:935-944.

Kim J., Kim Y., Yeom M., Kim J.-H., Nam H.G. (2008) FIONA1 Is Essential for Regulating Period Length in the Arabidopsis Circadian Clock. The Plant Cell Online 20:307-319.

Kim W.-Y., Fujiwara S., Suh S.-S., Kim J., Kim Y., Han L., David K., Putterill J., Nam H.G., Somers D.E. (2007) ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature 449:356-360.

Kim W.Y., Hicks K.A., Somers D.E. (2005) Independent roles for EARLY FLOWERING 3 and ZEITLUPE in the control of circadian timing, hypocotyl length, and flowering time. Plant Physiol 139:1557-69.

Klasen J.R., Piepho H.P., Stich B. (2012) QTL detection power of multi-parental RIL populations in *Arabidopsis thaliana*. Heredity.

Kobayashi Y., Weigel D. (2007) Move on up, it's time for change—mobile signals controlling photoperiod-dependent flowering. Genes & Development 21:2371-2384.

Kolmos E., Herrero E., Bujdoso N., Millar A.J., Toth R., Gyula P., Nagy F., Davis S.J. (2011) A Reduced-Function Allele Reveals That EARLY FLOWERING3 Repressive Action on the Circadian Clock Is Modulated by Phytochrome Signals in Arabidopsis. Plant Cell.

Kolmos E., Nowak M., Werner M., Fischer K., Schwarz G., Mathews S., Schoof H., Nagy F., Bujnicki J.M., Davis S.J. (2009) Integrating ELF4 into the circadian system through combined structural and functional studies. HFSP J 3:350-66.

Koornneef M. (1983) Linkage map of *Arabidopsis thaliana*. Journal of Heredity 74:265-272.

Koornneef M., Alonso-Blanco C., Vreugdenhil D. (2004) Naturally Occurring Genetic Variation In *Arabidopsis thaliana*. Annual Review of Plant Biology 55:141-172.

Kover P.X., Valdar W., Trakalo J., Scarcelli N., Ehrenreich I.M., Purugganan M.D., Durrant C., Mott R. (2009) A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in *Arabidopsis thaliana*. PLoS Genet 5:e1000551.

Laubinger S., Marchal V., Gentilhomme J., Wenkel S., Adrian J., Jang S., Kulajta C., Braun H., Coupland G., Hoecker U. (2006) Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. Development 133:3213-3222.

Ledger S., Strayer C., Ashton F., Kay S.A., Putterill J. (2001) Analysis of the function of two circadian-regulated CONSTANS-LIKE genes. The Plant Journal 26:15-22.

Lee M., Sharopova N., Beavis W.D., Grant D., Katt M., Blair D., Hallauer A. (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 population. Plant Molecular Biology 48:453-461.

Li D., Liu C., Shen L., Wu Y., Chen H., Robertson M., Helliwell C.A., Ito T., Meyerowitz E., Yu H. (2008) A Repressor Complex Governs the Integration of Flowering Signals in Arabidopsis. Developmental Cell 15:110-120.

Lister C., Dean C. (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. The Plant Journal 4:745-750.

Liu X.L., Covington M.F., Fankhauser C., Chory J., Wagner D.R. (2001) ELF3 Encodes a Circadian Clock–Regulated Nuclear Protein That Functions in an Arabidopsis PHYB Signal Transduction Pathway. The Plant Cell Online 13:1293-1304.

Locke J.C., Kozma-Bognar L., Gould P.D., Feher B., Kevei E., Nagy F., Turner M.S., Hall A., Millar A.J. (2006) Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. Mol Syst Biol 2:59.

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

Loudet O., Michael T.P., Burger B.T., Le Metté C., Mockler T.C., Weigel D., Chory J. (2008) A zinc knuckle protein that negatively controls morning-specific growth in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A. 105:17193-17198.

Lu S.X., Webb C.J., Knowles S.M., Kim S.H.J., Wang Z., Tobin E.M. (2012) CCA1 and ELF3 Interact in the Control of Hypocotyl Length and Flowering Time in Arabidopsis. Plant Physiol 158:1079-1088.

Mackay T.F.C., Stone E.A., Ayroles J.F. (2009) The genetics of quantitative traits: challenges and prospects. Nat Rev Genet 10:565-577.

Makino S., Matsushika A., Kojima M., Yamashino T., Mizuno T. (2002) The APRR1/TOC1 Quintet Implicated in Circadian Rhythms of *Arabidopsis thaliana*: I. Characterization with APRR1-Overexpressing Plants. Plant and Cell Physiology 43:58-69.

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

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

Mas P., Devlin P.F., Panda S., Kay S.A. (2000) Functional interaction of phytochrome B and cryptochrome 2. Nature 408:207-211.

McClung C.R. (2006) Plant Circadian Rhythms. The Plant Cell Online 18:792-803.

McClung C.R. (2011) The genetics of plant clocks, Advances in genetics. pp. 105-139.

McWatters H.G., Bastow R.M., Hall A., Millar A.J. (2000) The ELF3zeitnehmer regulates light signalling to the circadian clock. Nature 408:716-720.

McWatters H.G., Kolmos E., Hall A., Doyle M.R., Amasino R.M., Gyula P., Nagy F., Millar A.J., Davis S.J. (2007) ELF4 Is Required for Oscillatory Properties of the Circadian Clock. Plant Physiol 144:391-401.

Méndez-Vigo B., Picó F.X., Ramiro M., Martínez-Zapater J.M., Alonso-Blanco C. (2011) Altitudinal and Climatic Adaptation Is Mediated by Flowering Traits and FRI, FLC, and PHYC Genes in Arabidopsis. Plant Physiol 157:1942-1955.

Michael T.P., Salomé P.A., McClung C.R. (2003a) Two Arabidopsis circadian oscillators can be distinguished by differential temperature sensitivity. Proceedings of the National Academy of Sciences 100:6878-6883.

Michael T.P., Salomé P.A., Yu H.J., Spencer T.R., Sharp E.L., McPeek M.A., Alonso J.M., Ecker J.R., McClung C.R. (2003b) Enhanced Fitness Conferred by Naturally Occurring Variation in the Circadian Clock. Science 302:1049-1053.

Michaels S.D., Amasino R.M. (1999) FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering. The Plant Cell Online 11:949-956.

Millar A.J. (2004) Input signals to the plant circadian clock. J Exp Bot 55:277-283.

Mitchell-Olds T., Schmitt J. (2006) Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. Nature 441:947-952.

Mizoguchi T., Wheatley K., Hanzawa Y., Wright L., Mizoguchi M., Song H.-R., Carré I.A., Coupland G. (2002) LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis. Developmental Cell 2:629-641.

Mizoguchi T., Wright L., Fujiwara S., Cremer F., Lee K., Onouchi H., Mouradov A., Fowler S., Kamada H., Putterill J., Coupland G. (2005) Distinct Roles of GIGANTEA in Promoting Flowering and Regulating Circadian Rhythms in Arabidopsis. The Plant Cell Online 17:2255-2270.

Nakamichi N., Kiba T., Henriques R., Mizuno T., Chua N.-H., Sakakibara H. (2010) PSEUDO-RESPONSE REGULATORS 9, 7, and 5 Are Transcriptional Repressors in the Arabidopsis Circadian Clock. The Plant Cell Online 22:594-605.

Nozue K., Covington M.F., Duek P.D., Lorrain S., Fankhauser C., Harmer S.L., Maloof J.N. (2007) Rhythmic growth explained by coincidence between internal and external cues. Nature 448:358-361.

Nusinow D.A., Helfer A., Hamilton E.E., King J.J., Imaizumi T., Schultz T.F., Farre E.M., Kay S.A. (2011) The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. Nature 475:398-402.

Onai K., Ishiura M. (2005) PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. Genes to Cells 10:963-972.

Păcurar D.I., Păcurar M.L., Street N., Bussell J.D., Pop T.I., Gutierrez L., Bellini C. (2012) A collection of INDEL markers for map-based cloning in seven Arabidopsis accessions. J Exp Bot.

Park D.H., Somers D.E., Kim Y.S., Choy Y.H., Lim H.K., Soh M.S., Kim H.J., Kay S.A., Nam H.G. (1999) Control of Circadian Rhythms and Photoperiodic Flowering by the Arabidopsis GIGANTEA Gene. Science 285:1579-1582.

Perales M., Más P. (2007) A Functional Link between Rhythmic Changes in Chromatin Structure and the Arabidopsis Biological Clock. The Plant Cell Online 19:2111-2123.

Platt A., Horton M., Huang Y.S., Li Y., Anastasio A.E., Mulyati N.W., Ågren J., Bossdorf O., Byers D., Donohue K., Dunning M., Holub E.B., Hudson A., Le Corre V., Loudet O., Roux F., Warthmann N., Weigel D., Rivero L., Scholl R., Nordborg M., Bergelson J., Borevitz J.O. (2010) The Scale of Population Structure in *Arabidopsis thaliana*. PLoS Genet 6:e1000843.

Plautz J.D., Straume M., Stanewsky R., Jamison C.F., Brandes C., Dowse H.B., Hall J.C., Kay S.A. (1997) Quantitative Analysis of Drosophila period Gene Transcription in Living Animals. Journal of Biological Rhythms 12:204-217.

Pokhilko A., Hodge S.K., Stratford K., Knox K., Edwards K.D., Thomson A.W., Mizuno T., Millar A.J. (2010) Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model. Mol Syst Biol 6.

Ravi M., Chan S.W.L. (2010) Haploid plants produced by centromere-mediated genome elimination. Nature 464:615-618.

Rebai A., Goffinet B. (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. TAG Theoretical and Applied Genetics 86:1014-1022.

Reed J.W., Nagpal P., Bastow R.M., Solomon K.S., Dowson-Day M.J., Elumalai R.P., Millar A.J. (2000) Independent action of ELF3 and phyB to control hypocotyl elongation and flowering time. Plant Physiol 122:1149-1160.

Resco V., Hartwell J., Hall A. (2009) Ecological implications of plants' ability to tell the time. Ecology Letters 12:583-592.

Salomé P.A., Bomblies K., Laitinen R.A.E., Yant L., Mott R., Weigel D. (2011) Genetic Architecture of Flowering-Time Variation in *Arabidopsis thaliana*. Genetics 188:421-433.

Salomé P.A., McClung C.R. (2005) PSEUDO-RESPONSE REGULATOR 7 and 9 Are Partially Redundant Genes Essential for the Temperature Responsiveness of the Arabidopsis Circadian Clock. The Plant Cell Online 17:791-803.

Salome P.A., Weigel D., McClung C.R. (2010) The role of the Arabidopsis morning loop components CCA1, LHY, PRR7, and PRR9 in temperature compensation. Plant Cell 22:3650-61.

Schaffer R., Ramsay N., Samach A., Corden S., Putterill J., Carré I.A., Coupland G. (1998) The late elongated hypocotyl Mutation of Arabidopsis Disrupts Circadian Rhythms and the Photoperiodic Control of Flowering. Cell 93:1219-1229.

Schomburg F.M., Patton D.A., Meinke D.W., Amasino R.M. (2001) FPA, a Gene Involved in Floral Induction in Arabidopsis, Encodes a Protein Containing RNA-Recognition Motifs. The Plant Cell Online 13:1427-1436.

Scortecci K., Michaels S.D., Amasino R.M. (2003) Genetic interactions between *FLM* and other flowering-time genes in *Arabidopsis thaliana*. Plant Molecular Biology 52:915-922.

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

Seymour D.K., Filiault D.L., Henry I.M., Monson-Miller J., Ravi M., Pang A., Comai L., Chan S.W.L., Maloof J.N. (2012) Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. Proc Natl Acad Sci U S A. 109:4227-4232.

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

Singer T., Fan Y., Chang H.-S., Zhu T., Hazen S.P., Briggs S.P. (2006) A High-Resolution Map of Arabidopsis Recombinant Inbred Lines by Whole-Genome Exon Array Hybridization. PLoS Genet 2:e144.

Somers D.E. (1998) Phytochromes and Cryptochromes in the Entrainment of the Arabidopsis Circadian Clock. Science 282:1488-1490.

Somers D.E., Schultz T.F., Milnamow M., Kay S.A. (2000) ZEITLUPE Encodes a Novel Clock-Associated PAS Protein from Arabidopsis. Cell 101:319-329.

Southern M.M., Millar A.J. (2005) Circadian Genetics in the Model Higher Plant, *Arabidopsis thaliana*, in: W. Y. Michael (Ed.), Methods in Enzymology, Academic Press. pp. 23-35.

Srikanth A., Schmid M. (2011) Regulation of flowering time: all roads lead to Rome. Cellular and Molecular Life Sciences 68:2013-2037.

Suarez-Lopez P., Wheatley K., Robson F., Onouchi H., Valverde F., Coupland G. (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature 410:1116-1120.

Sugano S., Andronis C., Green R.M., Wang Z.-Y., Tobin E.M. (1998) Protein kinase CK2 interacts with and phosphorylates the Arabidopsis circadian clock-associated 1 protein. Proc Natl Acad Sci U S A 95:11020-11025.

Sugliani M., Rajjou L., Clerkx E.J., Koornneef M., Soppe W.J. (2009) Natural modifiers of seed longevity in the Arabidopsis mutants abscisic acid insensitive3-5 (abi3-5) and leafy cotyledon1-3 (lec1-3). New Phytol 184:898-908.

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

Takase T., Ishikawa H., Murakami H., Kikuchi J., Sato-Nara K., Suzuki H. (2011) The Circadian Clock Modulates Water Dynamics and Aquaporin Expression in Arabidopsis Roots. Plant and Cell Physiology 52:373-383.

Takeomi T., Atsushi O.D.A., Mayu N., Hiroshi K., Tsuyoshi M. (2007) Natural variation of polyglutamine repeats of a circadian clock gene ELF3 in Arabidopsis. Plant biotechnology 24:237-240.

Thines B., Harmon F.G. (2010) Ambient temperature response establishes ELF3 as a required component of the core Arabidopsis circadian clock. Proc Natl Acad Sci U S A 107:3257-62.

Tóth R., Kevei É., Hall A., Millar A.J., Nagy F., Kozma-Bognár L. (2001) Circadian Clock-Regulated Expression of Phytochrome and Cryptochrome Genes in Arabidopsis. Plant Physiol 127:1607-1616.

Treangen T.J., Salzberg S.L. (2012) Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet 13:36-46.

Troein C., Corellou F., Dixon L.E., van Ooijen G., O'Neill J.S., Bouget F.-Y., Millar A.J. (2011) Multiple light inputs to a simple clock circuit allow complex biological rhythms. The Plant Journal 66:375-385.

Tuinstra M.R., Ejeta G., Goldsbrough P.B. (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. TAG Theoretical and Applied Genetics 95:1005-1011.

Valverde F., Mouradov A., Soppe W., Ravenscroft D., Samach A., Coupland G. (2004) Photoreceptor Regulation of CONSTANS Protein in Photoperiodic Flowering. Science 303:1003-1006.

Wang L., Fujiwara S., Somers D.E. (2010) PRR5 regulates phosphorylation, nuclear import and subnuclear localization of TOC1 in the Arabidopsis circadian clock. EMBO J 29:1903-1915.

Wang Z.-Y., Tobin E.M. (1998) Constitutive Expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) Gene Disrupts Circadian Rhythms and Suppresses Its Own Expression. Cell 93:1207-1217.

Weigel D. (2012) Natural Variation in Arabidopsis: From Molecular Genetics to Ecological Genomics. Plant Physiol 158:2-22.

Weigel D., Nordborg M. (2005) Natural Variation in Arabidopsis. How Do We Find the Causal Genes? Plant Physiol 138:567-568.

Weinig C., Ungerer M.C., Dorn L.A., Kane N.C., Toyonaga Y., Halldorsdottir S.S., Mackay T.F.C., Purugganan M.D., Schmitt J. (2002) Novel Loci Control Variation in Reproductive Timing in *Arabidopsis thaliana* in Natural Environments. Genetics 162:1875-1884.

Wigge P.A., Kim M.C., Jaeger K.E., Busch W., Schmid M., Lohmann J.U., Weigel D. (2005) Integration of Spatial and Temporal Information During Floral Induction in Arabidopsis. Science 309:1056-1059.

Wu J.-F., Wang Y., Wu S.-H. (2008) Two New Clock Proteins, LWD1 and LWD2, Regulate Arabidopsis Photoperiodic Flowering. Plant Physiol 148:948-959.

Xu Y., Johnson C.H. (2001) A clock- and light-regulated gene that links the circadian oscillator to LHCB gene expression. Plant Cell 13:1411-1425.

Yakir E., Hilman D., Harir Y., Green R.M. (2007a) Regulation of output from the plant circadian clock. FEBS Journal 274:335-345.

Yakir E., Hilman D., Hassidim M., Green R.M. (2007b) CIRCADIAN CLOCK ASSOCIATED1 Transcript Stability and the Entrainment of the Circadian Clock in Arabidopsis. Plant Physiol 145:925-932.

Yerushalmi S., Green R.M. (2009) Evidence for the adaptive significance of circadian rhythms. Ecology Letters 12:970-981.

Yerushalmi S., Yakir E., Green R.M. (2011) Circadian clocks and adaptation in Arabidopsis. Molecular Ecology 20:1155-1165.

Yoshida R., Fekih R., Fujiwara S., Oda A., Miyata K., Tomozoe Y., Nakagawa M., Niinuma K., Hayashi K., Ezura H., Coupland G., Mizoguchi T. (2009) Possible role of early flowering 3 (ELF3) in clock-dependent floral regulation by short vegetative phase (SVP) in *Arabidopsis thaliana*. New Phytol 182:838-50.

Yu J.W., Rubio V., Lee N.Y., Bai S., Lee S.Y., Kim S.S., Liu L., Zhang Y., Irigoyen M.L., Sullivan J.A., Lee I., Xie Q., Paek N.C., Deng X.W. (2008) COP1 and ELF3 control circadian function and photoperiodic flowering by regulating GI stability. Mol Cell 32:617-30.

Zagotta M.T., Hicks K.A., Jacobs C.I., Young J.C., Hangarter R.P., Meeks-Wagner D.R. (1996) The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. The Plant Journal: For Cell and Molecular Biology 10:691-702.

Zagotta M.T., Shannon S., Jacobs C., Meeks-Wagner D.R. (1992) Early-Flowering Mutants of *Arabidopsis thaliana*. Functional Plant Biol. 19:411-418.

Zeilinger M.N., Farre E.M., Taylor S.R., Kay S.A., Doyle F.J. (2006) A novel computational model of the circadian clock in Arabidopsis that incorporates PRR7 and PRR9. Mol Syst Biol 2.

Zhang E.E., Kay S.A. (2010) Clocks not winding down: unravelling circadian networks. Nat Rev Mol Cell Biol 11:764-776.

Zuo Z., Liu H., Liu B., Liu X., Lin C. (2011) Blue light-dependent interaction of CRY2 with SPA1 regulates COP1 activity and floral initiation in Arabidopsis. Curr Biol 21:841-7.

Appendecis

Appendix 1: List of the markers used for genotyping of RILs.

Chr	Marker Name	Sequence 5´->3´ f	Sequence 5´->3´ r	Position (Mb)
I	WTSNP1-0008	CCGACAGCAGAATCTCCAAT	GAGTCGCCGGAGAATAACAA	0.008
I	W207W9WI1	TTATTAAACTGAACTAGTCTCGCCAAAACC	AAAAGGGGAGGAGAAAGAAAACATGAATA	3.1
I	W211W9WI3	ATTCCAAATGACTGTAGAATTTTCACACAA	TTTCCCAGTGTACTGAGTTTCTCAACTTTT	3.6
I	F3F19	CCACAAAACAATTTGGTTCACTC	TCCCGTTGGGGATATTAAAG	4.4
I	UA1.5.1	TATCCATCTCCTTTTTCTTTTTGGTCGT	ACATCTCTCTCGTTTCTGGTTGTAATCC	4.8
I	F20D23	TTATGCCAACTCATGTGGAAAG	TGTCAAAGCGTCTGGTTCTG	6
I	W219W9WI7	CTATTTTGCTTGGCGTATGTGAATATTTTT	AATGGTGAGATACTGAGATTATCCTTGATT	7.29
I	UA1.10.3	CAGGCTTCTAAAATGAAGGACAAGAAAA	AGAGTTTTAATCCGGGCTACCTAAGAGA	10
I	UA1.10.1	TGGACGACTTGTTTGAAGGAAGAAGTAT	ATTGCCAAGACTAAAACCAAATAGAGCA	10.34
I	WTSNP1-1251	AAGGCCATGTTTCCTTTTTC	CCACCAGTAATTAACCCCTGT	12.51
I	IND1-1290	CCCACATCACACATCTCTCA	TCTTATTGGCGCCTATTCTG	12.9
I	WTSNP1-1500	ACTCGATGAGAATCGCAGGA	CGATACAAATGAATCGGAGGA	15
I	T27K12	GGAGGCTATACGAATCTTGACA	GGACAACGTCTCAAACGGTT	15.93
I	CIW1	ACATTTTCTCAATCCTTACTC	GAGAGCTTCTTTATTTGTGAT	18
I	UA1.20.2	TATATCTTCTCGAAGTTGTGTTCTCATT	TCGTATAGACACTAGTTAGAAGCTGACC	20.04
I	NGA128	ATCTTGAAACCTTTAGGGAGGG	GGTCTGTTGATGTCGTAAGTCG	20.63
I	W223GENEA.5	ACATAACCACAAATAGGGGTGCTGAGT	ACCATGCATAGCTTAAACTTCTTGTTA	22
I	NTSNP1-2200	CACACCCTCTTCCATAAGGTT	CTCTTCCAATGCCTTCAATG	22
I	IND1-2400	CTGCTAAATGGCCTCAGGTA	CCACAGGAGTAGCAGGAGAA	24
I	IND1-2411	GATGAAGCGGAAACAAAAGA	AAACAACCGTCCAACGATTA	24.11
I	F20P5	GATACGTTCAAAATTAGGGACTTC	TGTATTTTGCTAATTGAGGTTATGG	26
I	W239V10	TAGGCTGAGCTAAGTTTAGAGAGAGAGAGC	TAACACTATCGAGAAAGAAAACGGAGAAAA	28.83
I	W227W9WI10	AATTAAGATCAAAAATTCTATCCGCTCACC	CAAAACACAATCTATTATTAGTTGGCACGA	29
I	W231W9WI12	CGAATAATCCTCAAATCTCAAGGTTAAGTG	AAAGGGAGTAGATCAAAGCTCAAAGGTTAG	30.01
I	UA1.30.3	TTAATGTGACACTTTTAATTGCTAATCC	AAGACGAAAGAAGAGACTTTTATGAATC	30.27
I	UA1.30.2	GAGAGAAAAAGATTAGGACTCACAGATT	CTAGCCGTAATAATAAGGTCCAAAAC	31
II	IND2-0027	GCGGTGCGATTCAAATAGT	TGACCATGTGATGCAATATCA	0.2
II	W93W9WII2	TGCTAGCTTTAGGCTATAATGTTGTTGTGA	AAAAGGGATACTGACAAAGCTAAACTCCTC	0.25
II	WTSNP2-0223	AGGGTATTTTGGGAGATTGA	GGAATTTTTGCTTCTGTGTGA	2.23
II	IND2-0235	GTTTCAAAAAGGAGGCGAAT	CTGTGACACCAGCAACTTGA	2.35
II	W103W9WII7	CCTCAAGGAACAAAATTTATCTTTCATAATCA	AAGTTCGTATTTTCGGTTTTAAAAAGTTTG	5.2
II	UA2.5.1	ATCCTAGATTCCAAACCAAATCTCACAA	ACTGAACCTCTCACACATTCAGCATAAA	5.12
II	MSAT2-11	GATTTAAAAGTCCGACCTA	CCAAAGAGTTGTGCAA	8.22
II	PLS7	GATGAATCTTCTCGTCCAAAAT	GACAAACTAAACAACATCCTTCTT	9.81
II	ATELF3	ATGATGCCCACCATAATGAACCCATATTG	AAAGGACTTGCTACCAGAGATTCCCTGTG	11.06
II	W157W9WII20	ATACTATAAACCCCAAACTCAACACGATAC	CCAGTGATGTTTTAGGAGTTACAGAGTTAG	12.13

Appendecis

II	**MSAT2.4MOD**	GAGGATCACCTAACCAACTCATGGAC	TCTTTCCTGTAATCTGGGTTTTTGTG	**13.8**
II	**UA2.15.1**	TTTTGTCCCATTAATTAAACTAAGACCT	ATTAAAACCCTGAAATACACAAAACATA	**14.79**
II	**W109W9WII10**	GATCAAATCAATCATAGCTACAGTCACCAC	CTCCTGGAACCTAAATCTAAGAATATGACT	**16.23**
II	**NGA168 EXT**	GAGGACATGTATAGGAGCCTCGAATG	TCGTCTACTGCACTGCCGGCATCTC	**16.3**
II	**IND2-1967**	TGAGGAGAACAAGTCCTTCG	AGAGAGGCAAGGCCAGTTAT	**19.67**
II	**LUGSSLP41**	TGCATCAGTTTTGGTTGTGTGATCT	GCTGTATTTTCCATAGGGGGCA	
III	**NGA172**	AGCTGCTTCCTTATAGCGTCC	CCATCCGAATGCCATTGTTC	**0.8**
III	**UA 3.1.1**	ATCTTATTACATCCGAATGCCATTGTTC	AATTCGAGCTAAAGCAGAAACACAAAAG	**1.1**
III	**W169W9WIII9**	CTCGTTGAAAATCAGTTACCAATTAACC	GTTACCTTTTGGTTTCTCCAAGCAATTT	**2.5**
III	**IND3-0251**	TCGACGACAACAACAACAAC	GACCCAGATTCAACGACAAC	**2.51**
III	**NGA162**	CATGCAATTTGCATCTGAGG	CTCTGTCACTCTTTTCCTCTGG	**4.6**
III	**UA3.5.2**	GTATCTTTTTCCTTTGTCACCATCGTTC	ATTTCCTCCCAAAACAAAACAAATTTTC	**5**
III	**NTSNP3-0650**	GGAAGACTTGTCTCTGGGAGT	CTTGCATCTGATGACCCTTC	**6.5**
III	**NTSNP3-0655**	TGGATCAAGATCCTCCATTG	ATCCATCCAAAGTCCAAACA	**6.55**
III	**W117W9WIII2**	TGGATGTTTTAAATTATCGAATAACAAAGA	TAAAATTTACATGTGGAGTATGTCCTAAAG	**7**
III	**IND3-0840**	CATTGTTTTAGCGGATTCAAA	TAATGGGACAAGGAGAGTGG	**8.4**
III	**UA3.10.1**	ATTACCGATAATAAGTTTTTCACTTCGT	TAGTGTTATGTGTGTCTCAGTGTGTGTA	**9.32**
III	**IND3-0952**	TGCCAGGGAGACAGACTTAC	AAGCCAACTCCCTTCTCATT	**9.52**
III	**W173W9WIII11**	TCTACGACATAACTATCTCTCTCGAACTC	ATCCACTTAACGTTCTTCATCTCATCATAC	**13**
III	**UA3.15.2**	ATCTTTCTTATCCCTCGTATATTTTGTG	GTTTTCTTGATCTAGAGACTTGATCGTA	**14.76**
III	**IND3-1484**	CCGAACTAGTCCCATCTTCA	GACGGGTCAGTTTATTGCAG	**14.84**
III	**AM11-6967**	CAAAAGAAATGCAACGAGAC	TTTGATCATGAATGGTAGTG	**16.34**
III	**W181W9WIII15**	ATGTATGTATACGAAGTGAAAGTGAAGTGG	ATTCTCTTTGGTGCAAGAGTAGTTATTGT	**19**
III	**NTSNP3-1998**	ACGAGGCTCTCTCTTCGTCT	TCCAAAATCGTACAAAATCAGA	**19.98**
III	**F27K19**	TGCTTTTGAAGAGATGGTTATTAGG	CCCCATTTCACTTATCATTGG	**21**
III	**T20O10**	GTTGCACGATCATGCGTTTAC	CCCCCTTCATTTACGCTGTAG	**23.4**
IV	**IND4-0006**	GGGCTCATTGGGAGAATTTA	CAGGGATTCGTCTGCTGATA	**0.01**
IV	**NTSNP5-0008**	AGCAACCTCCTCCGTTAGAT	GCTAGCGAGGCAGTCATTA	**0.01**
IV	**UA4.1.1**	AGAGGCTTCTCTATTAAGCTAAAACTTTCC	TATTGACATACAAAAGAAAACAAAGATGGT	**1.01**
IV	**W191W9WIV1**	ACAAGTCTCTATTTTAGCGATTTTAGATTT	GTATGAAGTTATCTCATTATCCATAGGTTC	**2.3**
IV	**NTSNP4-0257**	TGACACTTTCTCAACGTCACA	GCAACCTCAAGGTGTGTCA	**2.57**
IV	**NGA12N_EXT**	TAATGTTGTTTCCCCTCCTCGACAAG	CACTGAACCTGAATCGGCGATGTTG	**2.8**
IV	**T3H13**	TTTGGTGGGTCAAGAGTCAAG	GCAAAAGTCATTACGGACAATAC	**4.7**
IV	**NTSNP4-0550**	ATGCAACTGCCTCCTGAAC	TCATGTGAACTCTTCTCTATCCTTT	**5.5**
IV	**NTSNP5-0573**	TTTGCAGGGAAAACATTCTT	GCAAGCATCCTCTCAGACAT	**5.73**
IV	**IND4-0740**	AAATAAATATGATTTGTGCGTTCA	TTCTGGGACCAGTTTGTGAT	**7.4**
IV	**NTSNP4-0746**	AACAACACCTGAAATTGCTTT	TCTTTTCCTCTGTCAAAATCG	**7.46**
IV	**UA4.10.3**	ATGAATCTTTGTGAATTTAGGGTAATTTTC	GACACAAACTCTTTATTTCATTCACATTTT	**9.87**

Appendecis

IV	IND4-0999	TAACCACCCCTCACACTCAC	CAAAGACGTCCCAAATGTTC	9.99
IV	NTSNP4-1210	TTGGTCGAGGTTCAAAACAT	GCATGGAGCCGTAAGTAGTTT	12.1
IV	IND4-1244	GAAGAAATTGAGGAAAATTCTTGT	GGTTTTGAGTTATGTGCGTGA	12.44
IV	NTSNP4-1447	AAATTGCTTGACGGTGAATG	AGAAACAGCAGCCAGGTATG	14.47
IV	MSAT4-11	AAAAATCCGGTAGAGCATCC	CCAATTCCGAGCCAGTAA	14.5
IV	IND4-1530	TCTCTTTCCACAAAGGTTGG	TGATTTTCCGTTTCGTTTGT	15.3
IV	W199W9WIV5	AGTTTCTTATGTAATTTCTCGAAATCTGTT	GTTGTGTATAAAACATAGTGGCACATATAA	17.43
IV	UA4.18.1	TTAGGTGCAAATCTATTACTCTGTACCTGT	TTGTTGTAATTAGAAACGTGTGATTAACCT	17.79
IV	MSAT4.28_MODC	AATTGATGTCGCAATCAGATGAAG	CAACAATGCATGTAGAGAGTGTTTTG	18.5
IV	NTSNP4-1858	TGCTTTCTCTCGTCTTCTCC	GGGATGTGGGGTAGAGTTTT	18.58
V	MED24D_MOD	GGGGGACCTTTTTCTTGATTACC	GCAGAGTCTCACTCTCATCTCC	1
V	MHF15	CTCCTCCTTTAATTTTCTCTCTGTG	AGTTCCAGCTTTGGACTTCTTC	1.9
V	IND5-0218	CCACTTTCGTCTTCTCTCCA	TCTCCGCTCAATTTTGTTTC	2.18
V	IND5-0233	CGATGGTCAGCTTTAATTGC	TAGAACAAGTTGCGGGGATA	2.33
V	IND5-0408	GCCACATTTTTCTGGTGTTT	CTGGAGCTCTACTGGCTGAG	4.08
V	WTSNP5-0506	GAGCTTTGCTTCGTTTCAAG	TGCGTCAATACTTTCGCTCT	5.06
V	NGA106_MOD	GAGAGAGCTTCGAAATGAAGAAAC	GGTTCTTTAAAGCTGGTCCAAATA	5.4
V	IND5 0774	TGTCATGAAACTAGCTTGAATTG	AGAAGGGAGCAAGAGATGGT	7.74
V	IND5-0791	CGGCTAGTCACCACAACTTT	GGGAGCAACTTGCAAGAGTA	7.91
V	IND5-0987	CACTGAGACAAAAACGCAAA	ATCGGAGCACTTGATTTCAC	9.87
V	IND5-0990	GAGCCGGTCATCTGAGTCTA	GATCGGTCTGTTGTTTACCG	9.9
V	UA5.10.1	AAATCACAAATTTTAGAAAACAAATCATCC	ATTTCTATCAGTTTCCCTACTATGGTTGTT	10.12
V	IND5-1222	TTCAAAGATGCGATTCCATT	CGAGGTGGATGTTAAGAACG	12.22
V	IND5-1254	GGGTTGAGATACGGAGGTTT	CCCCAAATTTGAAAAAGGAG	12.54
V	SO191_EXT	CTCCACCAATCATGCAAATGTTTTTG	TGATGTTGATGGAGATGGTCAGAT	15.02
V	CIW9	CAGACGTATCAAATGACAAATG	GACTACTGCTCAAACTATTCGG	16.8
V	AM19-6752	ATTTAGATCGCTTCATTCAC	ACCATTTTCACATTTTATTA	18.21
V	IND5-1866	CGAAAAACGAAACCAGAGAA	CGCGAAACAGAGGAATAAAA	18.66
V	NGA129_MOD	GTCGACTACAACACTGAAGATGGTCT	GTTCTTCAGGAGGAACTAAAGTGAGG	20.13
V	IND5 2228	AACCAAACAAATGCACCACT	TGTGCTTGTGAAGAGCTTTG	22.28
V	IND5-2233	CGATGGTCAGCTTTAATTGC	TAGAACAAGTTGCGGGGATA	22.33
V	UA 5.25.2	GAAATTTCTAGAGAGCAGAGTAAGTCG	TAAGAAATGGATGGTCTGAAGAAGAAGT	24.95
V	5.25.1	CTTCAAACTCTGTATAAGAAAACATTCG	CACACTAGCTTAACCAACTATGGTAGAC	25.16
V	UA 5.25.3	GTACGTTGAGGGTACTTTAGTCATTTAG	GAATATTCTCAAAAACCCTAATTCAGAT	25.25
V	IND5-2690	AAGGGGCTCTCTCTCTGTGT	TTGGACGACTCCTACCACAT	26.9

Appendecis

Appendix 2: Estimated number of glutamines in 62 Arabidopsis accessions.

Serial	Accessions	Estimated Q		Serial	Accessions	Estimated Q
1	Aa-0	16		32	Edi-0	18
2	An-2	17		33	Ei-2	15
3	Ba-1	15		34	Eil-0	17
4	Bay0	22		35	El-0	19
5	Bch-1	18		36	En-2	21
6	Bd0	12		37	Enk-D	13
7	Bla3	11		38	Est	19
8	Br-0	21		39	Et-0	26
9	Bs-1	14		40	Fr-4	14
10	Bs-2	14		41	Gr	20
11	Bs-5	14		42	H55	8
12	Bsch-0	12		43	Hi	19
13	Bsch-2	14		44	Je54	13
14	Bu-0	19		45	Lip-0	15
15	Bu-2	13		46	Lm	19
16	Bur-0	23		47	Lu-2	14
17	C24	7		48	Nd	16
18	Chi-0	15		49	Oy	16
19	Cl-0	16		50	Petergof	14
20	Co	10		51	RLD1	18
21	Col-0	7		52	Rsch	27
22	Col-2	7		53	Rubez	16
23	Col-3	7		54	S96	17
24	Ct-1	12		55	Sha	14
25	Da-0	12		56	Sn(5)-1	23
26	Da1-12	13		57	Sol-0	19
27	Db-0	18		58	Ta	15
28	Dij-G	17		59	Wil	13
29	Dij-M	12		60	Ws-0	16
30	Dr-0	13		61	Yo	15
31	Dra-0	12		62	Ze-0	15

Acknowledgments

This thesis will not be completed without acknowledging few people. First, I would like to thank my parents, **Anwer Waseem** and **Nargis Bano** for their love, prayers, and efforts to provide better education to their children. Thanks to my brothers and sisters, **Sadia**, **Saima**, **Bilal, Sidra** and **Ali**, for their love and support. **I miss you.**

I would like to thanks **Prof. Dr. George Coupland** for being my doctor father and signing all the bureaucratic stuff through all these years. I would like to thanks **Prof. Dr. Ute Höcker** and **Prof. Dr. Martin Hülskamp** for being the reviewers of my thesis.

Special thanks to **Marc Hallstein** for his help in the research work, having lot of lunches together, and sharing funny conversations and noises.

Thanks to **Mandi** for organizing all lab stuff, and helping me in the early days.

Thanks to **Nora** for 'Zussamenfassung' and **Jieun** for her little 'tips and tricks'.

Thanks to Davis lab members, **Elsebeth**, **Eleni**, **Chiarina**, **Ela**, **Alfredo**, **Sarah**, **Eva**, **Ma, Du** and **Lukas** for providing nice working atmosphere in the lab.

Thanks to my friends **Ramzan**, **Ali**, **Hasnain**, **Ilyas**, **Alvi**, **Ghanzanfar**, **Sajid**, **Selva** and **Reena** for nice 'gupshup' and sharing the moments of joy and grief. **You people are great.**

Thanks to my sisters **Ruqia**, **Raheel**, **Shaista** and **Saira** for saving me from starvation in Germany.

Thanks **Aisha**, **Aida**, **Falah**, **Filza**, **Minahal**, **Sawal**, **Liv**, **Luca** and **Zoha** for your sweet smiles. **You kids are cute.**

Finally, from the ''Title'' of my thesis to ''This'' word, I would like to thanks **Dr. Seth Jon Davis** for his efforts, guidance, help and support. **You are really a good teacher**.

Thank You Seth

Erklärung

Köln, 2014

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzen Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken mit Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder universität zur Prüfung vorgelegen hat; dass sie noch nicht veröffentlicht worden ist, sowie dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden.

Max-Planck Institut für Pflanzenzüchtungsforschung,

Köln, 14 Juni 2014 _____

Muhammad Usman Anwer

Lebenslauf

Personal Data:

Name:	Muhammad Usman
Family Name:	Anwer
Date of Birth:	10 December 1983
Address:	Max Planck Institute for Plant Breeding Research Carl-von-Linne Weg 10, D-50829 Cologne, Germany.
E-Mail:	anwer@mpipz.mpg.de

Educational and Scientific Career:

April 2008 - June 2014 **Ph.D. (*magna cum laude*) (HEC-DAAD Fellowship)**
Max Planck Institute for Plant Breeding Research
Cologne Germany
Title: "*Natural variation in the Arabidopsis thaliana circadian clock as a determinant of flowering time: a quantitative genetics and genomics study*"

Aug 2005 - Aug 2007 **M.Sc. (Hons)** Agricultural Biotechnology **(Silver Medal)**
University of Agriculture Faisalabad, Pakistan.
Title: *"Molecular characterization of different tomato genotypes using RAPD markers"*.

Jul 2001 - July 2005 **B.Sc. (Hons)** Plant Breeding and Genetics
 (PG merit-based Internship award)
University of Agriculture Faisalabad, Pakistan.

Work Skills:

<u>Quantitative and population genetics</u>
Construction of recombinant inbred lines (RILs), heterogeneous inbred families (HIFs) and near isogenic lines (NILs), genotyping and phenotyping of RILs, construction of linkage maps, QTL mapping, Genome wide association studies (GWAS) analysis, Validation and positional isolation of QTL, Phylogenetic analysis, Detection of selection pattern
Software command
JoinMap4.1, MapQTL6, QTL network2.0, QTL cartographer, Tassel3.0, MEGA4.0, DnaSP5.0,

Luciferase Imaging
Complete command over several luciferase imaging platforms including TopCount scintillation counter and leaf movement imaging counter.
Software command
TopTempII, BRASS2.0 and BRASS3.0

Microscopic Imaging
Highly skilled in handling light as well as confocal microscopes including Leica TCS SP2, Zeiss LSM 700 and Zeiss LSM 780. Expert in localization, co-localization, FRET and FLIM techniques.
Software command
ImageJ, Adobe Photoshop, Zen 2011, Zen 2012, Leica LAS

High-throughput sequencing
Experience in handling and analysis of large-scale sequence data generated by next generation sequencing platforms such as Illumina/Solexa Genome Analyzer.
Software command
DNASTAR Lasergene 10, IGV2.0, CLC genomics workbench

Molecular techniques
Highly skilled in wide range of techniques such as high-throughput genotyping, gel electrophoresis, qRT-PCR, gene cloning using a number of vector backbones and methods including classical and gateway system, site directed mutagenesis, Arabidopsis transformation, bacterial transformation, RNAi, yeast two-hybrid, protein isolation, western blotting, vector construction and modification etc.

Language skills

Language(s)	UNDERSTANDING				
	LISTENING	READING	SPEAKING		WRITING
English	Proficient	Proficient	Excellent		Excellent
Urdu	Proficient	Proficient	Proficient		Proficient
German	Satisfactory	Satisfactory	Satisfactory		Satisfactory

Publications:

1. **Anwer, M.U.,** Boikoglou, E., Herrero, E., Hallstein, M., Davis, A. M., James, G.V., Nagy, F., and Davis, S.J., **2014**. Natural variation reveals that intracellular distribution of ELF3 protein is associated with function in the circadian clock. **eLife.** http://dx.doi.org/10.7554/eLife.02206.

2. **Anwer, M.U.,** and Davis, S.J., **2013**. An overview of natural variation studies in the Arabidopsis thaliana circadian clock. **Seminars in Cell & Developmental Biology.**

3. Shin, J. **Anwer, M.U.,** and Davis, S.J., **2013**. Phytochrome-interacting factors (PIFs) as bridges between environmental signals and the circadian clock: diurnal regulation of growth and development. **Molecular Plant.**

Scientific Presentations:

1. Floral transition in Arabidopsis, 16-18 September 2013, Tübingen, Germany. **(Oral presentation)**

2. 24th International Conference on Arabidopsis Research (ICAR), 24-28 June 2013, Sydney, Australia. **(Poster presentation)**

3. DFG 2nd Annual Meeting ''Flowering time control: from natural variation to crop improvement'' 22-24 January 2012, Cologne, Germany. **(Oral presentation)**

4. 23rd International Conference on Arabidopsis Research (ICAR), 3-7 July 2012, Vienna, Austria. **(Poster presentation)**

5. DFG First Kickoff Meeting ''Flowering time control: from natural variation to crop improvement'' 22-24 February 2012, Kiel, Germany. **(Oral presentation)**

6. International Conference on Applied Genetics and Biotechnology, 5-8 December 2011, Faisalabad, Pakistan. **(Oral presentation)**

7. Botaniker Tagung Conference, 18-23 September 2011, Berlin, Germany **(Poster presentation)**

8. Next Generation Sequencing symposium at Max Planck Institute for Plant Breeding Research, 3- 5 November 2010, Cologne, Germany.

9. 2nd PhD Retreat in Plant Sciences at Max Planck Institute for Plant Breeding Research, 15-17 April 2010, Cologne, Germany. **(Oral presentation)**