CHARACTERIZATION OF *PHYTOCHROME B* AND *CONSTANS*: KEY REGULATORS OF FLOWERING TIME IN SORGHUM

A Dissertation

by

SHANSHAN YANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	John E. Mullet
Committee Members,	Scott A. Finlayson
	Gregory D. Reinhart
	Paul D. Straight
Head of Department,	Gregory D. Reinhart

August 2014

Major Subject: Biochemistry

Copyright 2014 Shanshan Yang

ABSTRACT

Sorghum genotypes used for grain production in temperate regions are photoperiod insensitive, while energy sorghum is highly photoperiod sensitive and flowers late in long days, resulting in enhanced biomass accumulation. The sorghum floral repressors *PSEUDORESPONSE REGULATOR PROTEIN* (*SbPRR37, Ma1*) and *GRAIN NUMBER, PLANT HEIGHT AND HEADING DATE 7 (SbGHD7, Ma6)* contribute to photoperiod sensitivity by delaying flowering in long days with minimal influence in short days. Phytochrome B (PhyB) was found to mediate light signaling required for expression of the floral repressors *SbPRR37* and *SbGHD7* in the evening of long days and for photoperiod regulated flowering in sorghum. In genotypes lacking *PHYB*, SbPRR37 and SbGHD7 did not delay flowering in long days. The floral activators *EARLY HEADING DATE 1* (*SbEHD1*) and *SbCN8* (*ZCN8*) were highly expressed in long and short days in 58M (*phyB-1*) but repressed in 100M (*PHYB*) in long days. Sorghum alleles of *CONSTANS* (*SbCO*) were identified through QTL analysis. SbCO is an activator of flowering that is repressed post-transcriptionally in long days by the floral inhibitor SbPRR37, contributing to photoperiod sensitive flowering in Sorghum. Analysis of the flowering time QTL on SBI10 in Recombinant Inbred Lines (RILs) derived from BTx642 and Tx7000 revealed that BTx642 encodes a recessive *CONSTANS* allele containing a His106Tyr substitution in B-box 2. Genetic analysis characterized *SbCO* as a floral activator that promotes flowering by inducing the

expression of *SbEHD1*, *SbCN8* and *SbCN12* (*ZCN12*). The floral repressor SbPRR37 inhibits SbCO activity and flowering in long days.

These studies further characterize the genetic regulatory pathway that modulates photoperiodic flowering-time in sorghum. The balance between repressors and activators provides the basis for a wide range of flowering times in response to diverse environmental factors contributing to sorghum's wide adaptation.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. John Mullet, and my committee members, Dr. Scott Finlayson, Dr. Paul Straight, and Dr. Gregory Reinhart, for their guidance and support throughout the duration of this research.

Thanks also go to all the colleagues of my lab, especially to Dr. Rebecca Murphy, Dr. Daryl Morishige and Dr. Brock Weers for helping me collect tissues, prepare experiments and manuscripts. I would also like to thank my friends and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother and father for their encouragement and love.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Sorghum and the maturity genes

Sorghum (*Sorghum bicolor* [L.] Moench) is a C4 grass species of the Poaceae tribe (true grasses). Domesticated Poaceous cereal crops, including rice, maize, wheat, barley, sorghum, oat, rye, and millet, are cultivated for their edible grain constituting one of the major sources of food, forage, and biofuel in modern times [1, 2]. Sorghum is the 5th most important grain crop for human consumption and 2nd most important crop for livestock feed around the world [3]. It possesses good drought tolerance and high water-use efficiency, enabling this plant to survive and maintain productivity in arid and semi-arid areas [4-6]. Thus, this species serves as an important food source in Africa, Central America and South Asia [1, 7].

Sorghum originated in Africa, diverging from rice ~50 million years ago [8]. The cultivated sorghum, *Sorghum bicolor*, has been divided into five main races: guinea, bicolor, caudatum, durra, and kafir; each of which is originated in different regions of Africa [9]. The distribution of sorghum races reflects the adaption to distinct climate zones of Africa. The climate of central Africa is wet, receiving more than 180cm of rainfall per year [10]. Guinea is adapted to this high rainfall region [9]. The next climate zone outside of this wet region is the tropical savannah, receiving 50~100cm of rain per year [10]. Bicolor, which is considered as the most primitive race, and caudatum are widely distributed in this area [9]. The transition region from the tropical savannah to the

steppe climate zone receives 25-50cm of rainfall per year [10]. Durra originated from north of this transition region while kafir predominated in the south [9].

Sorghum is now widely adapted and grown as an annual crop from 0 to >40 degrees N/S latitude globally. The U.S. currently produces about one sixth of the sorghum worldwide grain output, approximately 60,000 metric tons per year, second only to Nigeria [1]. During the process of cultivation in North America, sorghum crops have been selected for a range of flowering times depending on growing location and uses, such as grain, sugar, forage, or biomass [2]. Grain sorghum is generally selected for early flowering (60-80 days) to enhance grain yield stability by avoiding drought, adverse temperatures, and insect pressure during the reproductive phase. In contrast, energy sorghum hybrids are designed with high photoperiod sensitivity in order to delay flowering and extend the duration of vegetative growth, resulting in more than 2-fold increases in biomass production [2, 11].

An essential milestone in sorghum cultivation history was the discovery and selection of maturity loci (*Ma*) [12, 13]. Six major maturity loci have been identified in sorghum, regulating the duration of growth or days to flowering and grain maturity. Dominance of these maturity genes results in the delay of floral initiation in long days. *Ma1-Ma4* were identified by Quinby and his colleagues (1940- 1970s) [13]. The discovery of these four maturity genes had a significant impact on commercial sorghum breeding for grain and sugar yield, and became the foundation of the sorghum conversion program established in 1963 [14]. Among the four original maturity loci, *Ma1* has the largest impact on photoperiod regulation of flowering time. The gene

corresponding to *Ma1* has recently been identified at *SbPRR37* [15]. *Ma3* encodes phytochrome B (PhyB) [16]. Based on Quail et al. paper about phytochrome nomenclature published in 1994, PHY was proposed to phytochrome apoprotein, while phytochrome or phy was proposed to represent holoprotein, which is the fully assembled chromoprotein with chromophore covalently attached to the apoprotein [17]. Since all phytochrome proteins referred in this dissertation are holoproteins, Phy is used to represent wild type holoprotein, while phy is used to represent holoprotein with mutations. There are two recessive *Ma3* alleles: $ma3^R$ and *ma3* (weak allele). The $ma3^R$ allele, identified in 58M, contains a deletion of a single base, causing a frame shift mutation in the gene. The frame shift results in a premature stop codon and the prematurely terminated phyB-1 protein that lacks regions of the protein necessary for dimerization and biological activity [16]. The homozygous recessive $ma3^R$ allele can override the effects of all other maturity genes, causing early flowering in long days [12]. 90M has the genotype *Ma1Ma2ma3Ma4*. This line initiates flowering earlier than 100M (*Ma1-Ma4*) and later than 58M.

In 1999, Rooney and Aydin identified two additional maturity loci, designated as *Ma5* and *Ma6* [18]. These loci were identified when it was observed that crossing R.07007 to BTx623, two early flowering genotypes, produced a very photoperiod-sensitive (PS) late-flowering hybrid. R.07007, flowering after about 85 days, was proposed to be *ma5ma5Ma6Ma6*. A/BTx623 and other female lines used in hybrid production that flower in 65~77 days were proposed to have the genotype *Ma5Ma5ma6ma6*. Thus, the very late-flowering photoperiod sensitive F1 hybrids were

heterozygous dominant *Ma5ma5Ma6ma6* and initiated flowering ~175 days after planting in the field when day lengths decreased below 12.2h. The results suggested that extreme lateness of flowering requires both *Ma5* and *Ma6* alleles [18]. Among these six maturity genes, *Ma1*, *Ma3* and *Ma5Ma6* in combination are the main modulators of flowering time in response to day length [12].

Table 1 shows the genotypes and flowering dates of some important sorghum lines used in this research.

Table 1. Genotypes and flowering dates of sorghum lines.

Sorghum Genotype	Maturity Loci	Days to Flowering (LD) *
BTx623	*ma1Ma2Ma3Ma4Ma5ma6*	71
R.07007	*Ma1ma2Ma3Ma4ma5Ma6*	95
(BTx623/R.07007) F1	*Ma1Ma2Ma3Ma4Ma5Ma6*	>160
100M	*Ma1Ma2Ma3Ma4Ma5ma6*	126
90M	*Ma1Ma2ma3Ma4Ma5ma6*	97
58M	*Ma1Ma2ma3^{R}Ma4Ma5ma6*	62

* Days to Flowering (LD) are determined in greenhouse LD (14h light/10h dark) condition.

Photoperiod responsive flowering time pathway in Arabidopsis

The transition from vegetative growth to flowering is a highly regulated process that has a large impact on plant adaption and reproductive success. Signals from many internal and external factors are integrated to control flowering time, including photoperiod, gibberellins, temperature, and age [19-22]. Of these, photoperiod stands as one of the most important external cues that non-equatorial plants interpret to regulate flowering.

The flowering-time regulatory pathway has been extensively characterized in *Arabidopsis thaliana*. The signals from four distinct pathways: photoperiod, vernalization, gibberellins and autonomous are integrated in the shoot apical meristem (SAM) through regulation of the meristem identity genes *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS (SOC1), LEAFY (LFY)* and *APETALA1 (AP1)*, which are activated during transition of the SAM from a vegetative meristem to a floral meristem. SOC1, as a master switch, initiates floral development and triggers expression of *LFY*. Then LFY activates AP1, and these genes together with the floral organ identity genes, control the formation of floral organs [20, 22, 23].

Photoperiodism makes it possible for an event, for example the initiation of flowering, to occur at a particular time. The first research about photoperiodism was published in 1920 by W. W. Garner and H. A. Allard. In their discoveries, they found the length of daylight that was critical for the developmental responses of plants, including flowering time [24]. However, it was later discovered that the length of the night was the controlling factor [25, 26]. There are three main photoperiodic responses:

long-day, short-day and day-neutral [27]. Long day (LD) plants, such as Arabidopsis, show early flowering in LD [28]. In contrast, short day (SD) plants such as rice and sorghum, show delayed floral initiation under LD conditions. The basic model of the photoperiod-response regulatory pathway, regardless of day-length response, is shown in Figure 1 (Fig 1). Light as an external cue is perceived by photoreceptors and together with the output from an endogenous circadian clock these inputs regulate the transition to flowering, consistent with external coincidence models of flowering time regulation, thus allowing the plant to sense and respond to seasonal changes in photoperiod [19, 29].

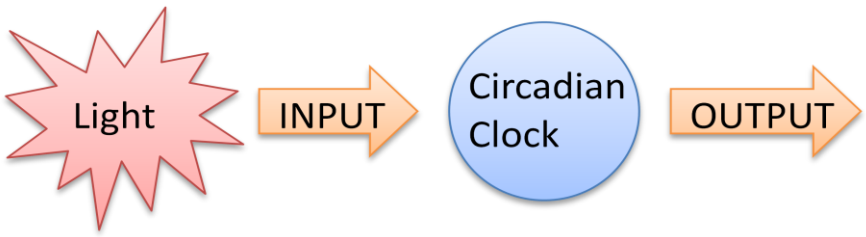


Figure 1. The basic model of photoperiodism. The output from endogenous circadian clock, coincident with the light signal of the day, regulates expression of the genes in the photoperiod flowering time pathway.

Plant photoreceptors include red/far-red light sensing phytochromes (PHY) and blue light/ultraviolet wavelength sensing cryptochromes (CRY), phototropins and Zeitlupes [30] [31]. The Arabidopsis genome encodes five phytochrome genes: *PHYA*-*PHYE*. In the dark, PhyB forms homodimers which are in an inactive form with red light absorbing ability (Pr form). Following the absorption of red light, Pr is converted to the

active far-red light absorbing (Pfr) form. Pfr translocates from the cytoplasm into nuclei [32]. In nuclei, PhyB specifically interacts with PIFs (phytochrome interacting factors) (Fig 2). PIFs are members of the basic helix-loop-helix (bHLH) family of transcription factors [33]. Each phytochrome may interact with one or more PIFs to regulate expression of light-controlled genes [34].

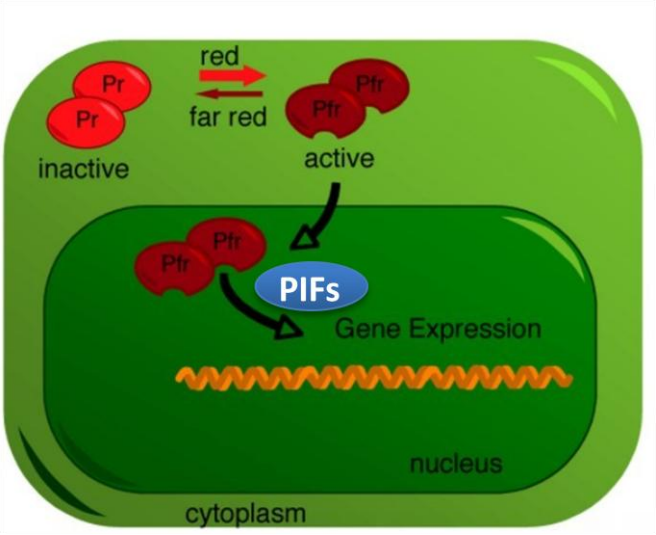


Figure 2. The model of phytochrome B regulating downstream gene expression. In the dark, PhyB are homodimers and in an inactive form with red light absorbing ability (Pr form). Upon absorbing red light, Pr is converted to the active far-red light absorbing (Pfr) form. Pfr translocates from cytoplasm into nuclei and specifically interacts with PIFs (phytochrome interacting factors), controlling downstream gene expression. (Christian Fankhauser © Unil)

The central oscillators of the circadian clock form a negative feedback loop: *TIMING OF CAB EXPRESSION 1* (TOC1) activates expression of the Myb factors *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* and *LATE ELONGATED HYPOCOTYL*

(LHY), which in turn negatively regulate *TOC1* expression [35, 36]. Rhythmic expression of these central oscillators modulates the expression and activity of *GIGANTEA (GI)*, a direct output gene of the circadian clock. In long days, GI activates *CONSTANS* (*CO*) expression in conjunction with *FLAVIN-binding KELCH DOMAIN F BOX PROTEIN1* (FKF1) by inducing degradation of *CYCLING DOF FACTORS 1*(CDF1), repressors of *CONSTANS* transcription [37, 38]. CO responds to photoperiod through changes in protein abundance in Arabidopsis [39]. CO accumulates in LD due to stabilization mediated by cryptochromes (Cry1/2) [40], phytochrome A (PhyA) and *SUPPRESSOR OF PHYA-105* (SPA1) [40-42] that counteract degradation of CO mediated by phytochrome B (PhyB): *CONSTITUTIVE PHOTOMORPHOGENIC 1*(COP1) [41, 43, 44]. CO activates expression of *FLOWERING LOCUS T* (*FT*), referred to as florigen in Arabidopsis [45]. Florigen is a mobile flowering signal that moves from leaves to the SAM where it binds to *FLOWERING LOCUS D* (FD). Together with SOC1, FT promotes expression of meristem identity genes *LFY* and *AP1*, leading to floral transition [41]. (Fig 3)

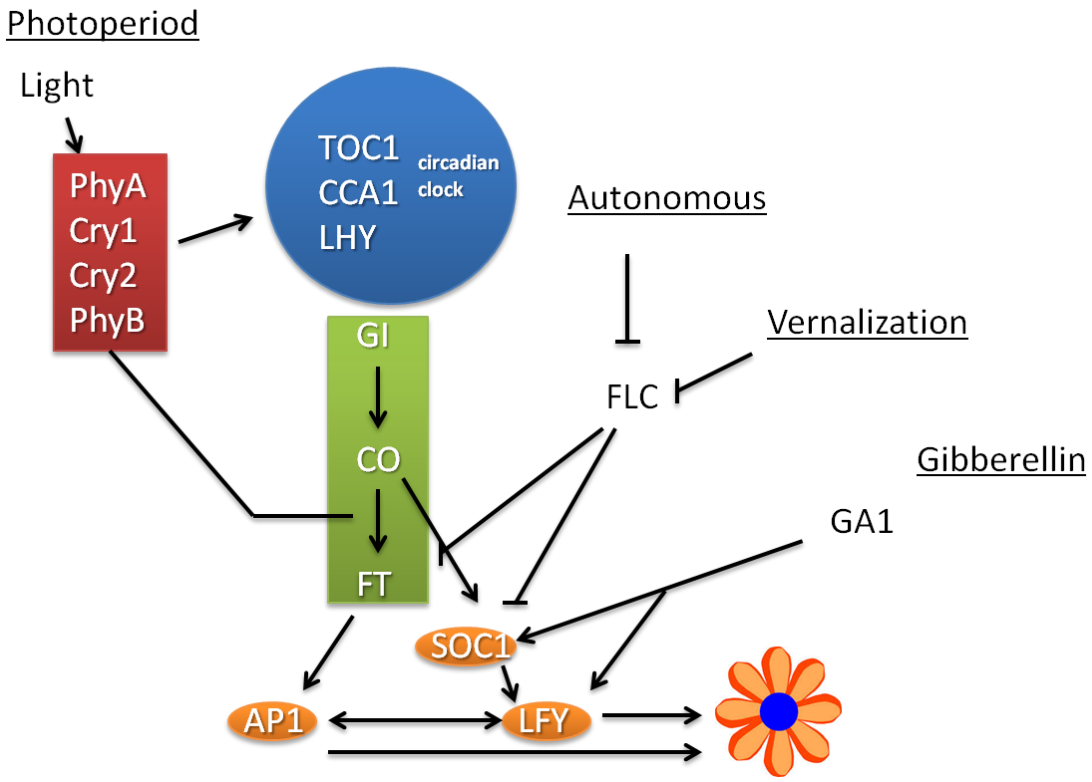


Figure 3. Flowering-time regulatory pathway in Arabidopsis. Four main pathways: photoperiod, vernalization, gibberellins and autonomous are integrated in meristem identity genes *SOC1*, *LFY* and *AP1*, modulating the floral transition. In the photoperiod pathway, light signal is perceived by phytochromes and cryptochromes. Rhythmic expression of central oscillators, TOC1, CCA1 and LHY modulates *GI*, a direct output gene of the circadian clock. GI-CO-FT form the core photoperiod regulatory pathway, coincident with both light and circadian clock signals. GI activates CO, which induces the expression of *FT*, the florigen of Arabidopsis.

Photoperiod flowering time pathway in Rice

The core photoperiod regulatory pathway GI-CO-FT in Arabidopsis is conserved in rice, a short day (SD) plant, although the day-length response is opposite [46]. In inductive SD, *HEADING DATE1* (Hd1) [47], the ortholog of CO, activates *Heading date 3a (Hd3a)*, one of the florigens in rice [48]; while in non-inductive LD, through the

mediation of PhyB, the function of Hd1 is converted to a repressor. Thus, photoperiod sensitivity in rice mainly depends on activity of CO (Hd1) [49, 50] (Fig 4). Rice has two florigen genes: *Hd3a* and its paralog *RICE FLOWERING LOCUS T1* (*RFT1*) [51]. Similar to *Hd3a*, *RFT1* is expressed in leaves and moves through the phloem to the SAM where it induces floral induction. However, RFT1 is a LD specific florigen, making it different and complementary to Hd3a [52].

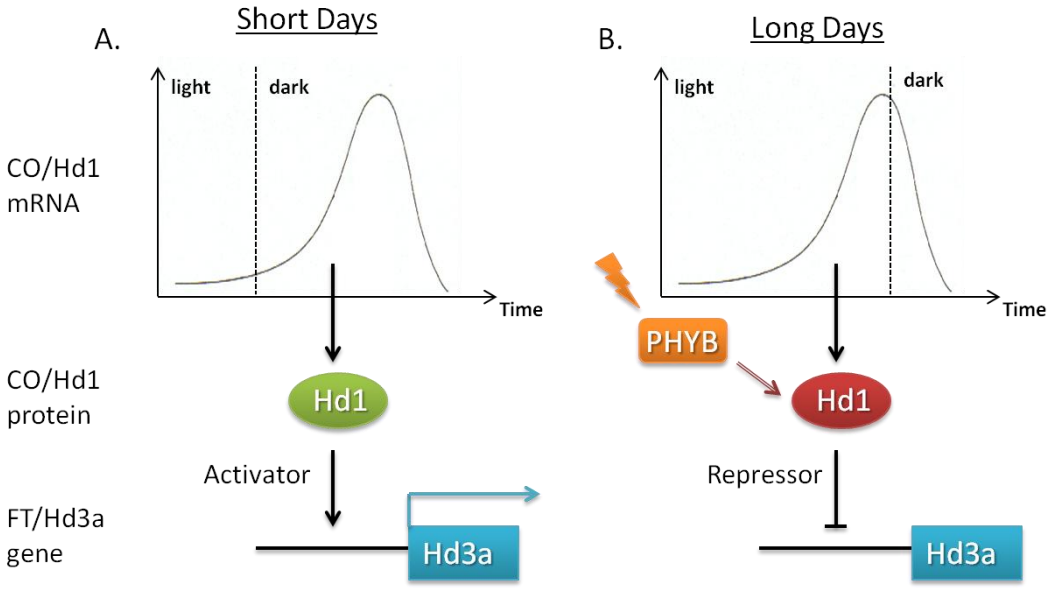


Figure 4. The core photoperiod regulatory pathway OsGI-Hd1-Hd3a in rice. Rice is a short-day plant. The mRNA level and protein abundance of *Hd1*, the ortholog of *CO*, remain similar in both LD and SD. In inductive SD (A), Hd1 activates *Hd3a*, one of the florigens in rice; while in non-inductive LD (B), through the mediation of PhyB (possible through post-transcriptional modification), the function of Hd1 is converted to a repressor.

In addition, two unique grass modulators of flowering were identified in rice: *EARLY HEADING DATE 1* (EHD1) and *GRAIN NUMBER, PLANT HEIGHT AND HEADING DATE 7* (GHD7) [53]. EHD1 is a B-type response regulator, containing a response regulator receiver domain at the N-terminus and a GARP DNA binding motif in the middle of the protein. It activates the expression of *Hd3a* and *RFT1*, in SD and LD respectively and has been shown to promote *FT*-like gene expression and flowering independently of Hd1 in rice [54]. The expression of *EHD1* is controlled by several upstream modulators including the repressors GHD7, *GRAIN NUMBER, PLANT HEIGHT AND HEADING DATE 8* (GHD8) [55], *Os CONSTANS-LIKE4*(OsCOL4) [56], *Os LEC1 AND FUSCA-LIKE1* (OsLFL1)[57], OsMADS56 [58] and the activators GI, *EARLY HEADING DATE 2* (EHD2*)/Os indeterminate1* (OsID1) [21] and OsMADS50/51 [50, 59] (Fig 5). *EHD2/OsID1*, encoding a C2-H2 zinc-finger protein, is an ortholog of maize *indeterminate1* (*id1*). OsMADS50 is an homolog of Arabidopsis SOC1 and might form a protein complex with OsMADS56 [58]. In contrast to the activator activity of EHD1, GHD7, a CCT (CO, CO-LIKE and TIMING OF CAB1) domain protein, represses flowering by down-regulation of *EHD1* and *Hd3a* expression in non-inductive LD conditions [53]. The expression of *GHD7* is photoperiod-regulated, which plays a key role for photoperiod sensitivity in rice, in addition to regulation of Hd1. Under inductive SD condition, the sensitivity of *GHD7* expression to red light is gated at midnight. However, there is no red light at midnight, leading to low *GHD7* mRNA level. Without repression from GHD7, *EHD1* is highly expressed and activates *Hd3a*. In contrast, under non-inductive LD condition, the time period of *GHD7*'s

sensitivity shifts to morning, coincident with high red light intensity. Thus, red light induces the expression of *GHD7*, which inhibits the downstream activator *EHD1* and *Hd3a*, resulting in limited induction of LD-specific florigen *RFT1* [49, 60] (Fig 6). In summary, the phytochrome modulation of Hd1 activity and *GHD7* expression are of central importance for regulating flowering time in rice in response to variation in photoperiod.

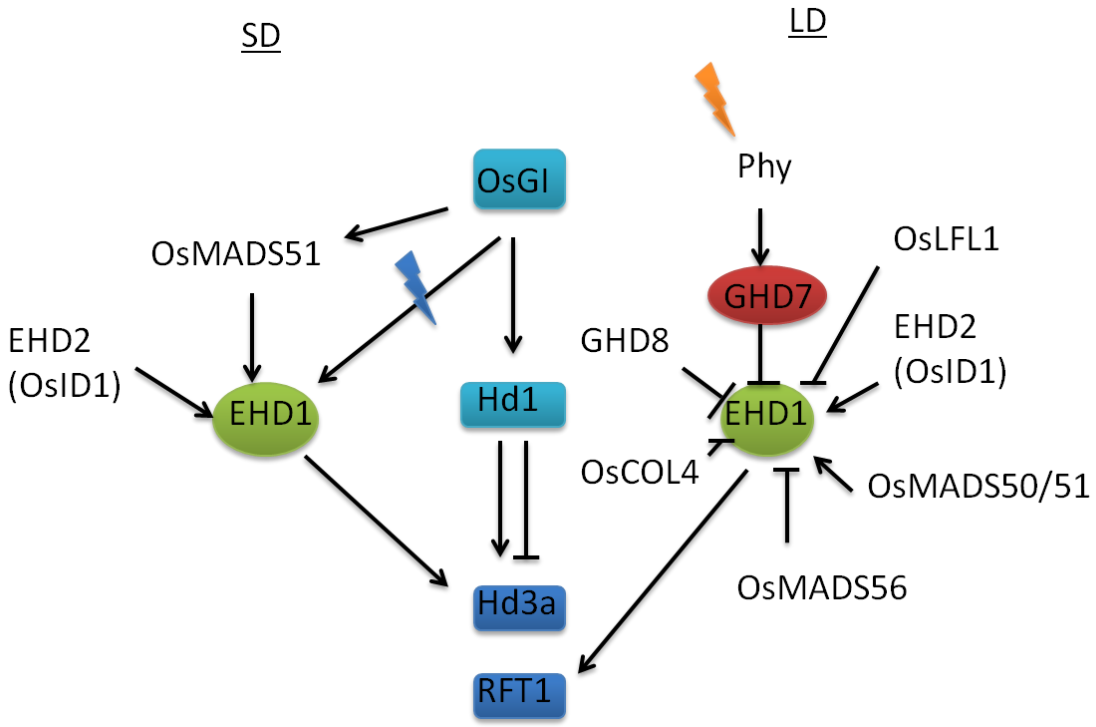


Figure 5. Model of photoperiod flowering-time pathway in rice. In SD, EHD2 (OsID1) and OsMADS51 activate EHD1, a floral activator unique in grasses. EHD1 increases the expression of *Hd3a* and triggers the induction of flowering. In non-inductive LD, EHD1 activates *RFT1*, a LD-specific florigen in rice. GHD7, GHD8, OsCOL4, OsLFL1 and OsMADS56 down-regulate *EHD1* and *Hd3a*, postponing the flowering time.

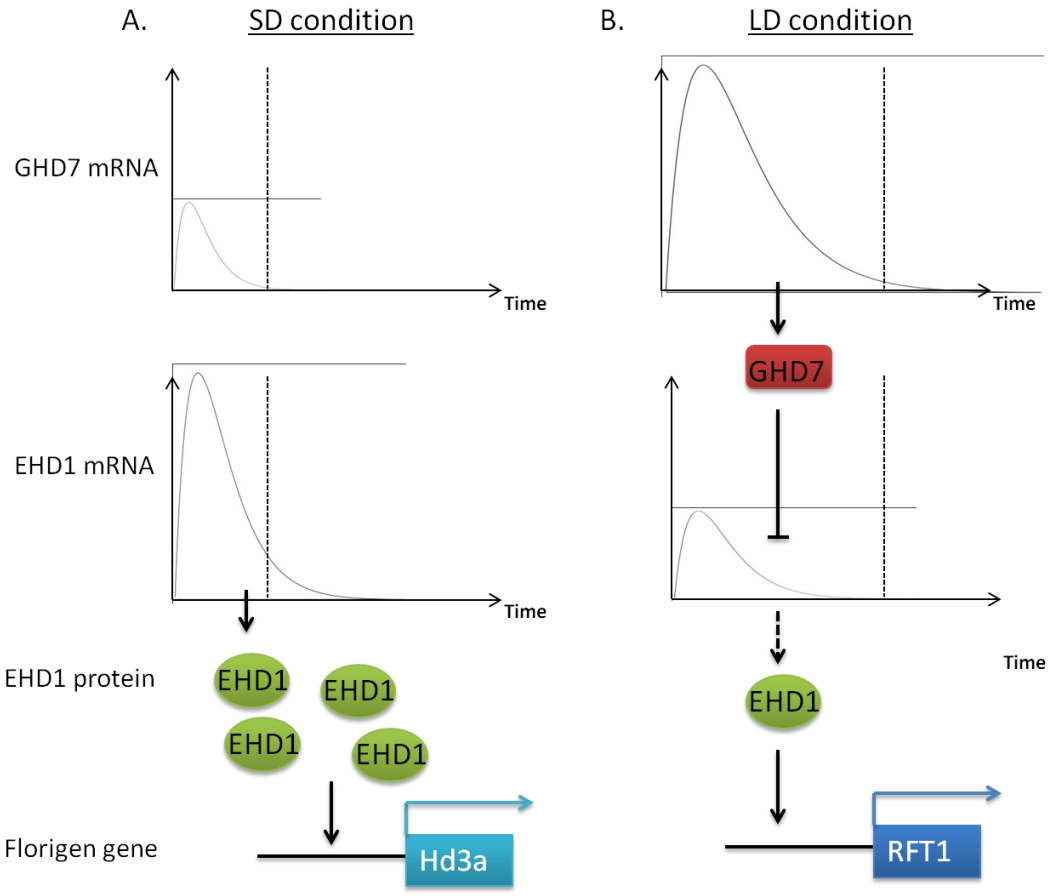


Figure 6. GHD7 is a LD-specific floral repressor in rice. (A) Under inductive SD conditions, the sensitivity of *GHD7* expression to red light is gated at midnight. There is no red light at midnight, resulting in low *GHD7* mRNA levels. Without repression from GHD7, *EHD1* is highly expressed and activates the SD-florigen gene *Hd3a*. (B) Under non-inductive LD condition, the time period of *GHD7*'s sensitivity shifts to morning, coincident with high red light intensity. Red light induces the expression of *GHD7*, which inhibits downstream activator *EHD1* and *Hd3a*, resulting in limited induction of LD-specific florigen *RFT1*.

The analysis of the genetic mechanism controlling the diversity of flowering time in cultivated rice elucidated three essential factors: the functionality of Hd1, the

promoter type of *Hd3a* and the expression level of *EHD1* [61]. The assessment by applying a linear model incorporating these three factors revealed that the independent activity and interactions between them contribute to 26.6% of the entire flowering time variance. Among these three factors, the largest single effect associated with flowering time variation can be attributed to activity of variant *Hd1* alleles (44%) [61]. Moreover, natural variation of heading date in cultivated rice can also be partially explained by the diversity of *GHD7* alleles [62].

Light signal transduction of phytochromes

Light regulates developmental processes throughout the entire plant life cycle, including seed germination, seedling photomorphogenesis, shade avoidance, phototropism and photoperiod regulated flowering [31, 63]. In order to respond to multiple light qualities (wavelengths), plants have evolved at least five distinct families of photoreceptors (Fig 7). Cryptochromes, phototropins and Zeitlupes sense blue light and ultraviolet A (UVA) [30]. Cryptochromes are well characterized in Arabidopsis as Cry1 and Cry2. Typical cryptochromes contain an N-terminal photolysase-related (PHR) domain, which non-covalently binds to two chromophores: a flavin adenine dinucleotide (FAD) and a pterin, and a C-terminal DAS domain [64, 65]. Phototropins constitute two flavin mononucleotide (FMN) chromophore-binding LOV domains (LOV1 and LOV2) at the N-terminus for light-sensing and a serine/threonine kinase domain at the C-terminal end [66]. Phototropin is plant-specific blue-light photoreceptors and PHOT1/PHOT2 are two well-characterized phototropins in Arabidopsis. In contrast,

phytochromes (Phy) perceive red and far-red wavelength light. The photoreceptor absorbing UVB has been identified as UV8 recently [67].

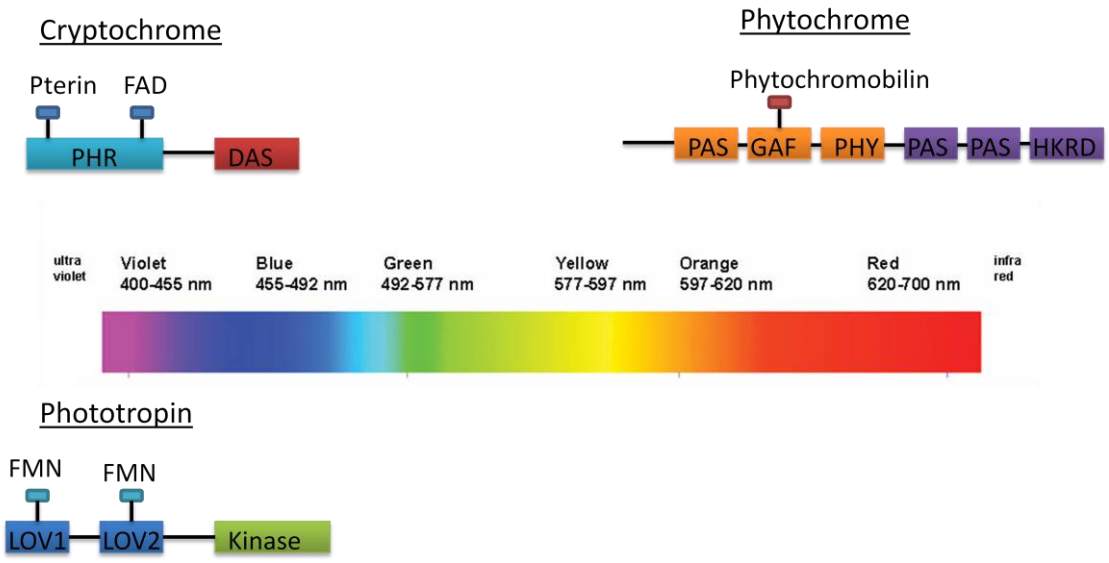


Figure 7. Photoreceptors of higher plants in response to light spectrum. Cryptochromes (Cry), phototropins and Zeitlupes sense blue light and ultraviolet A (UVA); while phytochromes (Phy) perceive red and far-red wavelength. Cryptochromes contain an N-terminal photolysase-related (PHR) domain, which non-covalently binding to two chromophores: a flavin adenine dinucleotide (FAD) and a pterin, and a C-terminal DAS domain. Phototropins constitute two flavin mononucleotide (FMN) chromophore-binding LOV domains (LOV1 and LOV2) at the N-terminus for light-sensing and a serine/threonine kinase domain at the C-terminus of the protein. Phytochromes contain an N-terminal photosensory moiety (PAS, GAF and PHY) and a C-terminal dimerization moiety (PAS and HKRD).

It has been suggested that plant phytochromes originally evolved from bacterial two-component receptors [68]. Bacteria, such as cyanobacterium *Fremyella*, use the

two-component signaling system to sense and respond to environmental cues. The first component is a photoreversible sensor protein, which is weakly similar to the chromophore-binding domain of plant phytochrome, transmitting the received light signal to the second component. The second component is the response protein, accepting the signal from the sensor protein, and linked to a histidine kinase domain as an output domain to initiate the cellular response cascade [69].

The divergence of the phytochrome family occurred very early in the evolution of seed plants. *PHYA*, *PHYB* and *PHYC* are present widely in angiosperms, suggesting duplication from a single progenitor into these three members occurred prior to the radiation of the angiosperm lineage. Dicots, such as Arabidopsis, contain five members of phytochrome genes (*PHYA*-*PHYE*), while monocots, including rice and sorghum, encode only three universal phytochromes members *PHYA*, *PHYB* and *PHYC*. The absence of *PHYD* and *PHYE* in monocotyledonous plants indicates the divergence of *PHYD/E* from other phytochromes has occurred after the divergence of monocots from dicots or that *PHYD/E* have been lost in monocots [70] (Fig 8). Base on their stability in the light, phytochromes have been classified into two types. PhyA belongs to Type I phytochrome (photo-labile), which accumulate in the dark and degrades rapidly upon light exposure. All others are Type II phytochromes (photo-stable), which are relatively stable in the light [71].

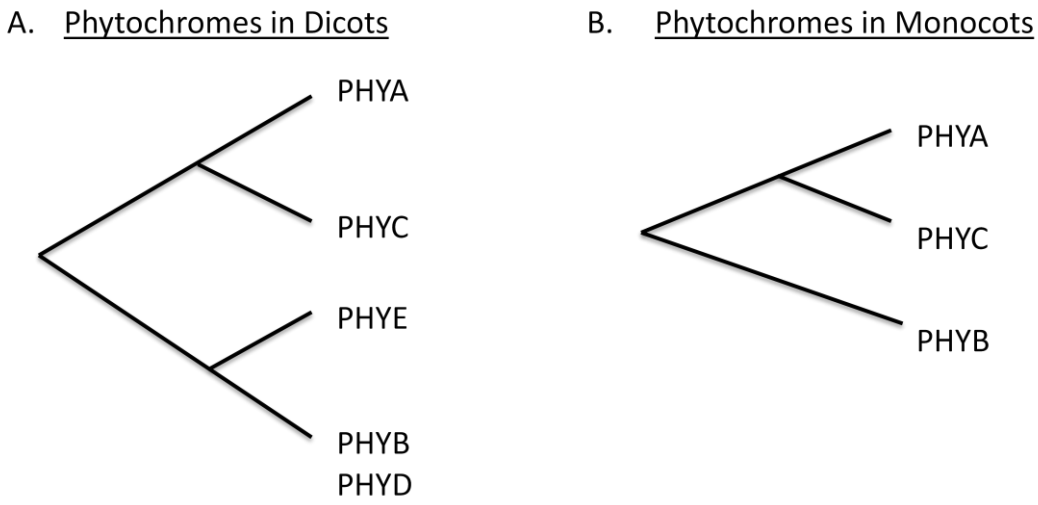


Figure 8. The phytochromes family in angiosperms. (A) Dicots contain five members of phytochrome genes: *PHYA*-*PHYE*. *PHYD* and *PHYE* are more closely related to *PHYB* than *PHYA* and *PHYC*. (B) Monocots include three phytochromes: *PHYA*, *PHYB* and *PHYC*. The absence of *PHYD* and *PHYE* in monocots indicates the divergence of *PHYD/E* from other phytochromes has occurred after the divergence of monocots from dicots or *PHYD/E* have been lost in monocots.

Phytochromes are soluble chromoproteins that contain an N-terminal photosensory domain and a C-terminal dimerization moiety. There are three sub-domains in the N-terminal moiety: PAS (PER, ARNT and SIM), GAF (cGMP phosphodiesterase, adenylate cyclase, Fh1A) and PHY (phytochrome-specific GAF-related), which form a unique structure, the "light-sensing knot" [72] (Fig 9). Since the "light-sensing knot" is involved in the direct interaction with PIFs and PAS/GAF domains can still transduce light signals without other parts of phytochrome [73], it contributes directly to light signaling and transduction. Phytochromobilin, a linear tetrapyrrole chromophore is covalently attached to the cysteine of GAF domain. The C-

terminal moiety, consisting of two PAS and HKRD (histidine-kinase-related domain), is responsible for dimerization and nuclear localization.

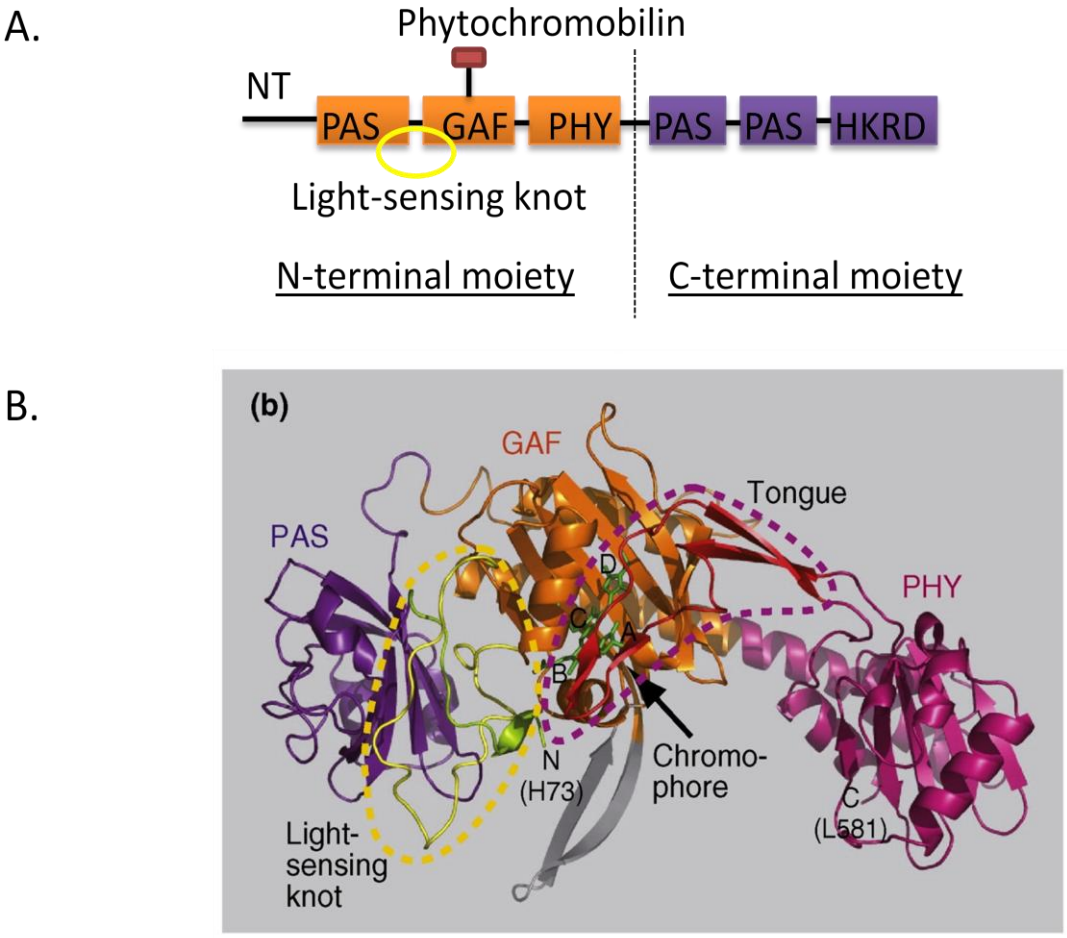


Figure 9. Domain structure of phytochromes. (A) Phytochromes contain three sub-domains in the N-terminal photosensory moiety: PAS (PER, ARNT and SIM), GAF (cGMP phosphodiesterase, adenylate cyclase, Fh1A) and PHY (phytochrome-specific GAF-related), which form a unique structure, the "light-sensing knot". Phytochromobilin, a linear tetrapyrrole chromophore is covalently attached to the cysteine of GAF domain. The C-terminal moiety, consisting of two PAS and HKRD (histidine-kinase-related domain), is responsible for dimerization and nuclear localization. (B) A 3-dimensional (3D) model of the N-terminal moiety of Arabidopsis PHYA (residues 73-581) [72].

Phytochromes have two interconvertable conformations: a red-light (R, λ=660nm) absorbing Pr form and a far-red (FR, λ=730nm) absorbing Pfr form. Upon light absorption, the Pr form converts to the Pfr form, which is considered as an active form since many physiological responses of plants are promoted by R light. This photoconversion is caused by the isomerization around the C15-C16 double bond between the C and D rings of the tetrapyrrol [71] (Fig 10). The light triggered conformational change of the chromophore leads to the change of phytochrome's structure, kinase activity and localization, initiating its interaction with transcription factors of the PIF family and downstream signal transduction.

PIF3 was the first identified member of the PIF family and was isolated in a yeast two-hybrid screen by interaction with the C-terminal domain of PhyB [33]. It was shown later that PIF3 interacts more strongly with the PhyB N-terminal domain. Based on the identification of several mutations that impair the interaction between PhyB and PIF3, it was suggested that the "light-sensing knot" of phytochrome may be directly involved in the interaction with PIFs. Other members of the PIF family, such as *PIF4*, *PIF5* and *PIF6* were identified by either double mutants or BLAST analysis [74]. The Arabidopsis PIF-subfamily contains 15 members and belongs to the bHLH transcription factor superfamily [34]. Only some of these members are involved in phytochrome signaling. The bHLH domain is in the middle of the C-terminal portion of PIFs and is responsible for dimerization and specific DNA binding. The target cis-element for PIFs is the G-box (5'-CACGTG-3'). The N-terminus of all PIF members contain a conserved APB (Active

Phytochrome B-binding) motif, which is required and sufficient to specifically bind the Pfr (active) form of PhyB [34].

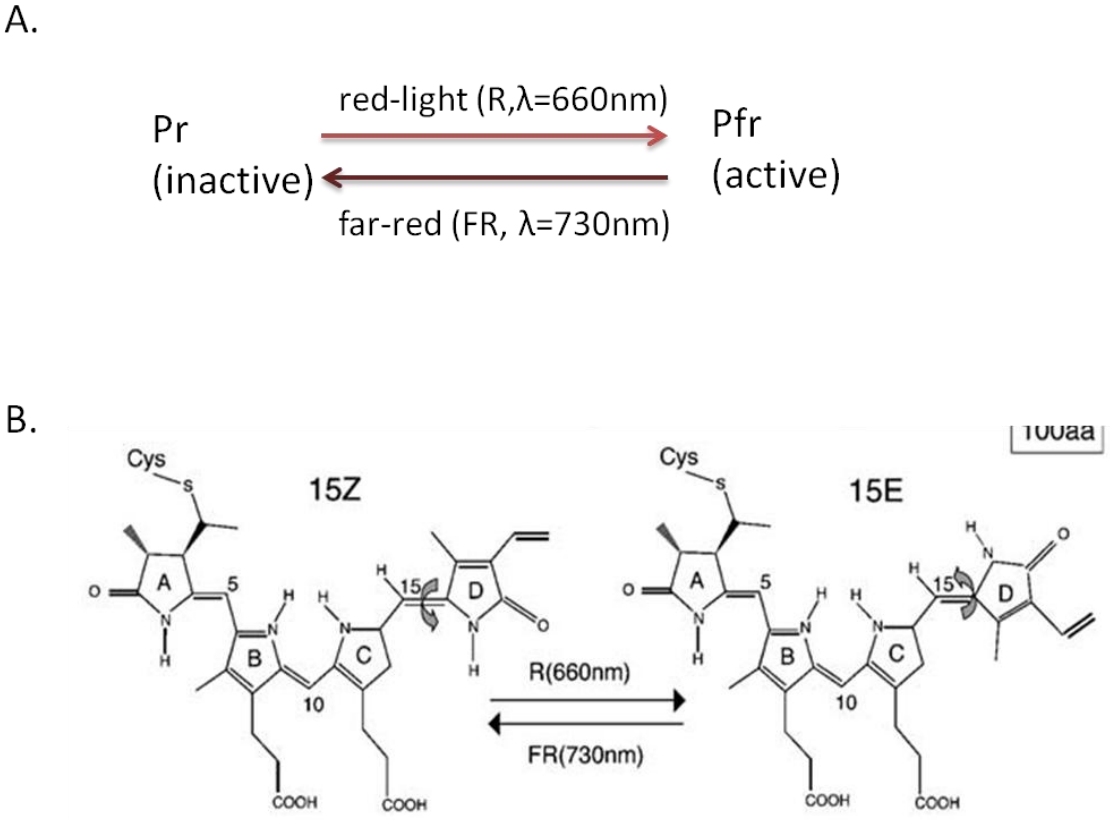


Figure 10. The photoconversion of phytochromes. (A) Phytochromes have two interconverable conformers: a red-light (R, λ=660nm) absorbing Pr form and a far-red (FR, λ=730nm) absorbing Pfr form. Upon red-light exposure, Pr form converts to Pfr form, an active form of phytochrome. (B) Light triggers the configuration change of the phytochrome chromophore, phytochromobilin. The isomerizaion around C15-C16 double bond between the C and D rings of the tetrapyrrol (15Z to 15E) explains the mechanism of phytochromes' photoconversion [71].

A model for PIF function in the phytochrome signaling pathway was suggested by Castillon et al. in 2007 [33]. In the dark, phytochromes are presumed to be inactive and localized in the cytoplasm. PIFs are bound to G-box motifs in nuclei and modulate the expression of genes that repress photomorphogenesis. Upon illumination, phytochromes are converted to the active Pfr form and translocate into nuclei, where they interact with PIFs. The physical interaction triggers phosphorylation of PIFs, resulting in their degradation. Thus, light induced phytochromes remove PIFs, thereby derepressing photomorphogenesis [75, 76]. (Fig 11)

In addition to flowering time determination, phytochromes-PIFs play critical roles throughout the entire life cycle of Arabidopsis. They modulate multiple developmental responses, including seed germination, seedling de-etiolation, shade avoidance and stomatal development [34, 77, 78]. Recent studies indicate that other pathways also intersect with this phytochrome-PIF pathway in response to light, such as the gibberellin pathway [79, 80], the circadian clock and high temperature [78], making the PIFs a cellular signaling hub that integrates complex environmental cues. The table published by Franklin and Quail in 2010 [81] summarized the diversity of phytochrome functions based on analysis of Aradidopsis mutants. (Table 2)

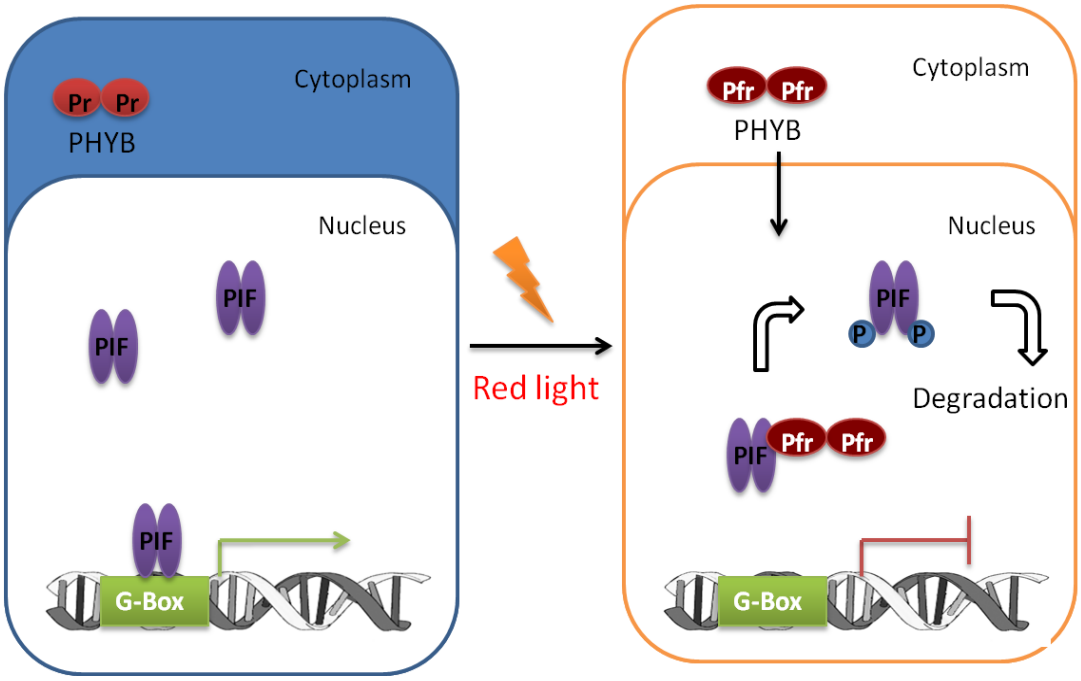


Figure 11. The model of PIF function in the phytochrome signaling pathway. In the dark, phytochromes are presumed to be inactive and localized in the cytoplasm. PIFs are bound to G-box motifs in the nucleus and modulate the expression of genes that repress photomorphogenesis. Upon illumination, phytochromes are converted to the active Pfr form and translocate into the nucleus, where they interact with PIFs. The physical interaction triggers phosphorylation of PIF, resulting in degradation. Light induced phytochromes remove PIFs from G-box motifs, thereby de-repressing photomorphogenesis [33].

Table 2. Summary of phytochrome functions through analysis of Arabidopsis mutants.

Function	Phytochromes
Promotion of seed germination	PhyA, PhyB, PhyE
Regulation of seedling de-etiolation	PhyA, PhyB, PhyC, PhyD, PhyE
Regulation of root gravitropic curvature	PhyB
Suppression of root hair growth	PhyB
Regulation of leaf architecture	PhyA, PhyB, PhyC, PhyD, PhyE
Suppression of internode elongation	PhyA, PhyB, PhyE
Suppression of shade avoidance	PhyB, PhyD, PhyE
Regulation of stomatal index	PhyB
Entrainment of the circadian clock	PhyA, PhyB, PhyD, PhyE
Photoperiodic perception	PhyA, PhyC
Repression of flowering	PhyB, PhyC, PhyD, PhyE

Function of the CO-FT regulon in Arabidopsis and grasses

CONSTANS was initially isolated by Putterill et al. in Arabidopsis [82] as an important transcriptional regulator in control of flowering. It belongs to a family of transcription factors unique to plants with three characteristic domains: either one or two N-terminal zinc finger B-box domains and a CCT domain at the C-terminus [83] (Fig

12). Two conserved cysteine and histidine amino acids in the B-box domain coordinate binding to Zn atoms and thus are essential for CO activity [43]. Arabidopsis mutants with amino acid substitutions in these positions have late flowering phenotypes. The C-terminal CCT domain is named from its presence in CO, CO-like genes and some circadian oscillators, such as TOC1 and Pseudo Response Regulators (PRRs). The CCT domain contains a nuclear import signal and was proposed to mediate DNA binding by forming a complex with HEME ACTIVATOR PROTEIN (HAP) [84, 85]. Yeast two-hybrid screens showed that CCT domain proteins can strongly interact with HAP3 and HAP5, but not HAP2, which are members of the heterotrimeric CCAAT-box-binding factor (CBF) complex. It suggested that CO may substitute for HAP2 and associate with HAP3 and HAP5 in a CBF complex, which is recruited to the CCAAT motif in the promoter [84, 86, 87]. Recently, it was reported that TOC1 can bind to its cognate motif in the promoter of *LHY/CCA1* directly through its CCT-domain, indicating CCT domain proteins may also access DNA sequence directly [88]. Further biochemical analysis will be required to distinguish between these possibilities. Extensive gene duplication events occurred in this gene family, leading to ~17 members present in Arabidopsis, ~16 members in rice and ~9 members in barley [89]. The major role of CCT domain genes is to modulate photoperiod responses. However, there is little information about the individual function of most *CO*-like genes.

Figure 12. Domain structure of CONSTANS. CONSTANS consists of three characteristic domains: either one or two zinc finger B-box domains at the N-terminus and a CCT domain at C-terminus. The B-box may involve in protein-protein interaction. Two conserved cysteine and histidine amino acids in the B-box domain coordinate the binding to the Zn atoms. The CCT domain contains a nuclear import signal and mediates DNA or/and protein binding.

FLOWERING LOCUS T –like (FTL) genes belong to one of three major classes of PEBP (phosphatidylethanolamine-binding proteins) domain gene families and function in floral promotion and inflorescence architecture determination in both dicots and monocots [90]. The phylogenetic analysis of all PEBP sequences in cereals hypothesized independent evolution by duplication or gene loss in each taxon, leading to obscured orthology relationship structure and functional diversification [91, 92]. There are at least 13 *FTL* genes in rice, including two florigens: *Hd3a* (*OsFTL2*) and *RFT1* (*OsFTL3*). In maize, 25 *FT* homologs were identified and designated as *Zea mays CENTRORADIALIS* (*ZCN*) genes [93]. *ZCN15*, the most closely related homolog of rice *Hd3a*, is expressed predominantly in kernels and not in leaves, ruling out its potential as a maize florigen. Instead, *ZCN8* and *ZCN12* mRNA accumulate in leaves and can be induced by a transient SD treatment, implying their roles in floral promotion [93]. It was shown that *ZCN8* transcripts were strongly up-regulated in a diurnal manner under floral

inductive SD in a SD tropical maize variety [94]. Moreover, ectopic expression of *ZCN8* in vegetative shoot apex can trigger early flowering in transgenic plants, confirming its role as the primary floral activator in maize [94]. Information on the role of other *FT*-like family genes in Arabidopsis and cereals is lacking.

The most characterized member of the *CO*-like gene family is *CONSTANS* in Arabidopsis. The circadian pattern of *CO* expression is activated and regulated by the direct clock output gene *GI* together with FKF1 [37], which can release *CO* expression by blue-light mediated degradation of DOF (DNA binding with One Finger) proteins [38]. The circadian pattern of *CO* expression is characterized as low expression early in the day, followed by a rapid increase after dawn and peaking at ~15h [83]. CO responds to photoperiod through changes in protein abundance. CO only accumulates under LD, in which stabilization from cryptochrome 2 (Cry2), cryptochrome 1 (Cry1) and phytochrome A (PhyA) counters the degradation from phytochrome B (PhyB) and other factors [83] (Fig 13). Thus, CO becomes a LD-specific activator for *FT* expression and floral initiation in Arabidopsis.

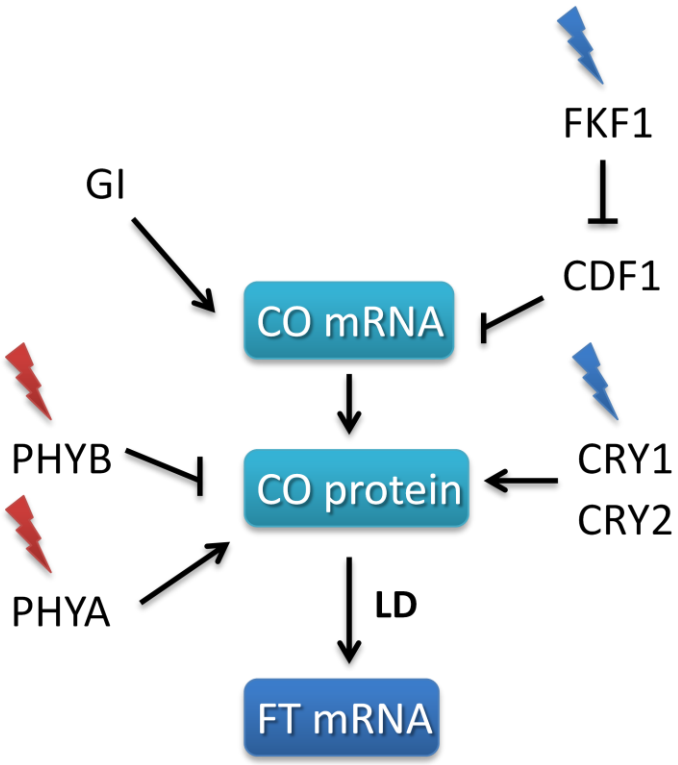


Figure 13. CO protein abundance is altered by photoperiod in Arabidopsis. *CO* expression is activated and regulated by the clock output gene *GI* together with FKF1, which induces *CO* expression by blue-light mediated degradation of CDF1. CO accumulates under inductive LD through stabilization from Cry1, Cry2 and PhyA, countering the degradation mediated by PhyB and other factors. High levels of CO protein can increase the expression of *FT* and initiate floral transition in LD.

CO-like genes in the grass family, including rice, maize and barley, have been characterized [95, 96]. Rice and sorghum are classified as SD plants, barley and wheat as LD plants, and maize is classified as a day-neutral plant. It has been reported that *Hd1*, the rice homolog of *AtCO*, also plays a critical role in photoperiod sensitivity and flowering time modulation [47]; however, the mechanism is distinct from Arabidopsis. In rice, experiments showed that *Hd1* transcript levels remained the same and Hd1

protein levels were not altered under SD and LD, suggesting that post-translational modification modulates activity [97]. Hd1 protein could function as either an activator or suppressor of flowering depending on light signals mediated by PhyB. Ishikawa et. al [97] discovered that the suppression of *Hd3a* and a delay of flowering caused by over expression of *Hd1* was PhyB-dependent and the effect was not associated with the degradation of Hd1 protein. In addition, day length extension decreased *Hd3a* expression. However, the expression pattern of *Hd1* mRNA was altered slightly in wild type and *phyB* mutant plants, leading to the explanation of *Hd3a* repression by extending day length, focusing on direct effects from PhyB-mediated light signaling rather than indirect changes of *Hd1* expression level. Thus, the conclusion was that PhyB plays an essential role in determining Hd1's activity as an activator or repressor in response to day-length. Natural variation in *OsPRR37* has also been shown to regulate heading date in rice [98]. Together with the fact that the largest single effect associated with flowering time variation in cultivated rice comes from the variation of *Hd1* alleles, all of these results indicate that rather than protein abundance and expression level, variation in the activity of Hd1 in rice is determined by PhyB and/or PRR37-mediated post-translational regulation [97] and *Hd1* allele variation [47, 61].

A similar characterization of *CO* homologs was also done in barley. *HvCO1*, showing the highest similarity to Arabidopsis *CO* and collinearity with rice *Hd1*, was reported to control flowering time in barley [99]. The study demonstrated that over-expression of *HvCO1* increased the mRNA level of *HvFT1* in LD and accelerated flowering time under both LD and SD conditions. Moreover, the transgenic plants with

over-expressed *HvCO1* remained photoperiod sensitivity, which is mediated by *PHOTOPERIOD-H1* (*PPD-H1*) [100]. The studies in barley showed there are two additional CCT domain proteins: PPD-H1 and *VERNALIZATION 2* (VRN2). Both of them are involved in flowering time regulation in response to photoperiod via controlling the expression of *HvFT1*. *PPD1*, a pseudo-response regulator and an ortholog of *AtPRR7* [101], has been identified as a central photoperiod modulator that acts in LD in barley and wheat. In barley, a *ppd-H1* mutant showed reduced photoperiod responsiveness [100]; while in wheat, miss expression of *Ppd-D1a* is also associated with reduced day-length sensitivity [102]. Thus, *PPD1* provides adaptation to photoperiod and modulates flowering time in LD- grasses. *VRN2* is the barley homolog of rice *GHD7*, and has a similar function as GHD7, which represses *HvFT1* expression in LD. It is involved in the crosstalk between photoperiod and vernalization. The process of vernalization, can down-regulate *VRN2*, releasing the expression of *HvFT1* and floral initiation [103, 104]. The induction of *FT*-like genes is regulated through interactions between various CCT domain proteins, including PPD1, CO and VRN2, and the HAP complex [85, 95].

In contrast, there is no evidence indicating that the maize *CO*-like gene, *constans of Zea mays1* (*conz1*), is involved in flowering time regulation. Maize was initially domesticated from teosinte, a photoperiod sensitive SD plant that originated from tropical areas of Mexico (Central America). To grow maize in a broad range of latitudes, post-domestication breeding developed modern maize that is insensitive to day-length. Thus, modern temperate maize was mostly characterized as a day-neutral plant with SD

tropical landraces still retaining the photoperiod sensitivity [105]. Maize *conz1* is the ortholog of rice *Hd1* [106] and its expression in short days exhibits a diurnal fluctuation, peaking at ~15h post-dawn, similar to expression pattern of *Hd1*. However, the expression of *conz1* in long days peaks twice, one at evening (~15h) and the other one at night (~21h) [107].

In summary, although the CO-FT regulon is conserved in angiosperms, including Arabidopsis and grasses, the functions of *CO*-like and *FT*-like gene families in one taxa or *CO* and *FT* homologs in various taxa are diverse [21, 108]. The mechanism of floral regulation is taxa specific, dependent on the expression pattern, protein abundance, day-length response and other regulators in the flowering time pathway. Thus, the diverse functions of *CO* homologs endow the flexibility needed by plants to respond to various environmental cues and become one of the reasons for their wide adaption and reproductive success.

Sorghum development and the role of two floral repressors

Sorghum goes through four development stages after germination: juvenile, vegetative, booting and grain filling [12] (Fig14). The juvenile phase is characterized by vegetative growth and floral induction cannot be induced by external signals, including day length and temperature [109]. The juvenile phase varies from 14-30 days and is regulated by several microRNAs [110, 111]. The transition from juvenile phase to adult growth is controlled by genes in the age-related pathway [109]. The length of adult growth as a vegetative plant varies from 5-180 days, dependent on the photoperiod sensitivity and hormone level, such as gibberellins (GA) [80, 112]. LD specific floral

repressors in sorghum mainly contribute during this stage [12]. Once flowering is induced, the plants go through the boot stage, which is approximately 30 days in length, from floral induction to anthesis. During this phase, all leaves become fully expanded and heads are swelling inside the flag leaf sheath. This stage is critical since the reproductive structures of panicles are forming. The final stage is grain filling, beginning with floral anthesis and continuing until grain maturity. Sorghum usually takes about 40 days since floral anthesis for complete grain maturity [12].

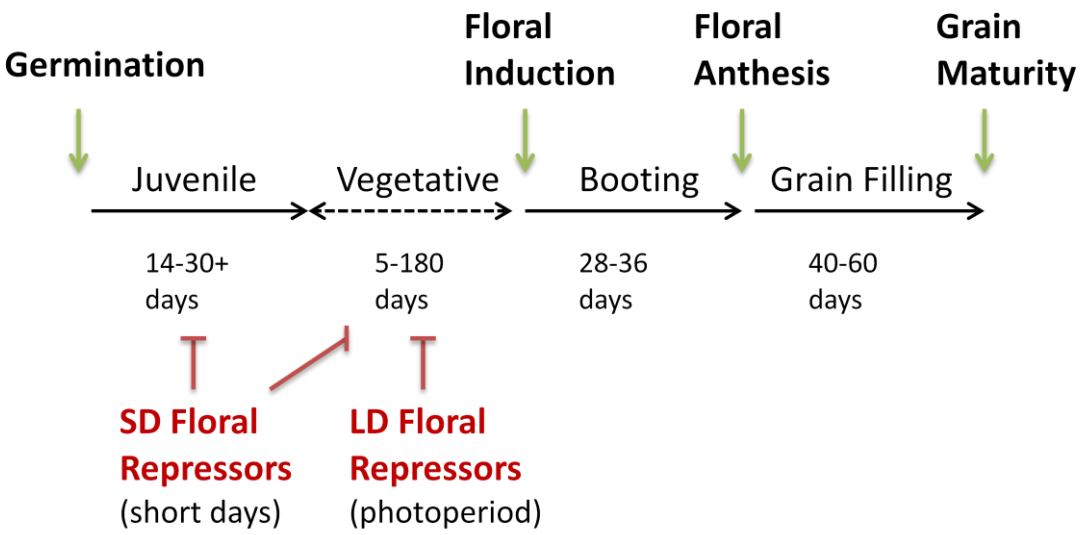


Figure 14. Sorghum developmental stages. Sorghum goes through four development stages after germination: juvenile, vegetative, booting and grain filling. Juvenile phase is characterized by vegetative growth and usually takes about 14-30 days. The transition from juvenile phase to adult growth is controlled by genes in age-related pathway. The length of adult growth varies from 5-180 days, dependent on the photoperiod sensitivity and hormone level. LD specific floral repressors in sorghum mainly contribute during this stage. Once floral induced, the plants go to the boot stage, which is about 6-8 days just prior to heading. The final stage is grain filling, beginning with floral anthesis and continuing till grain maturity. It usually takes about 20 days for complete grain maturity.

SbPRR37 and *SbGHD7* were identified as two major LD specific floral repressors in sorghum, and corresponding to *Ma1* and *Ma6* respectively [15, 113]. In sorghum, *Ma1*, which has the largest effect on photoperiod regulation, has been positionally cloned and identified as *SbPRR37* [15]. Sequence variation of *SbPRR37* alleles showed early flowering grain sorghum contain loss-of-function mutant alleles. The *Sbprr37-1* allele contains a single nucleotide deletion, resulting in the premature termination upstream of a pseudo-receiver domain of the SbPRR37 protein. In long days, SbPRR37 inhibits expression of floral activators *EHD1*, *SbCN8*, the sorghum ortholog of *ZCN8*, and *SbCN15*, the ortholog of rice *Hd3a*, and delays floral initiation. *SbPRR37* transcript levels are regulated by the circadian clock and light. When 100M (*Ma1-Ma5ma6*) plants are planted in LD or exposed to continuous light, *SbPRR37* RNA levels peak in the morning (~3h) and the evening (~15h) [15]. In contrast, in SD, *SbPRR37* RNA levels peak in the morning but the evening peak of expression is not observed since plants are in darkness during the evening phase. In addition, plants transferred to continuous dark show neither peak of *SbPRR37* expression, demonstrating a light requirement for *SbPRR37* expression [15]. This result is consistent with the proposed External Coincidence Model of flowering time regulation, which involves coincidence between light signaling and output from the circadian clock [29].

In LD, the light-dependent expression peaks of *SbPRR37* in morning and evening provide a sufficient level of SbPRR37, leading to suppression of sorghum florigen gene expression and floral transition. In contrast, in SD, when plants are in darkness during the evening, *SbPRR37* expression is reduced, resulting in de-repression of the floral activators *EHD1*, *SbCN8* and *SbCN15*, and floral induction. The model demonstrating SbPRR37 plays an essential role in inhibiting floral initiation in response to photoperiod is shown in Fig15.

SbGHD7, a homolog of wheat *VRN2,* modulates photoperiod sensitivity and floral repression in an additive fashion with *SbPRR37* [113]. It increases photoperiod sensitivity in a similar manner of *SbPRR37*. *SbGHD7* mRNA levels only peak in the evening under long day conditions. The similar expression pattern of *SbPRR37* and *SbGHD*7, both regulated by coincidence of circadian clock and light, suggested common upstream regulation. SbGHD7 also inhibits expression of the floral activators *EHD1* and *SbCN8,* resulting in repression of flowering in LD but not in SD. *SbGHD7* was identified as *Ma6* through map based cloning [113]. The model representing the function of SbGHD7 in the photoperiod flowering time pathway is shown in Fig16.

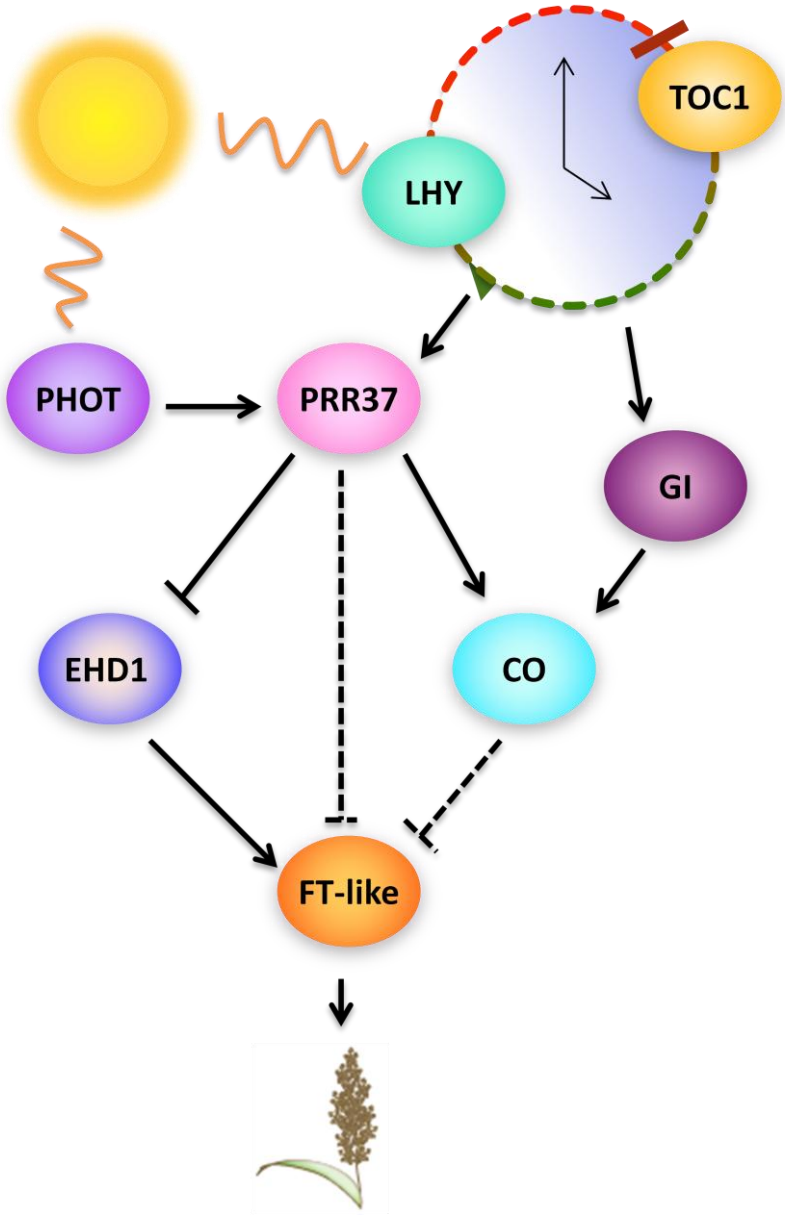


Figure 15. SbPRR37 inhibits floral initiation in response to photoperiod in sorghum. In LD, the light-dependent expression of *SbPRR37* in morning and evening provide a sufficient level of SbPRR37, leading to suppression of floral activators *EHD1*, *FT*-like genes *SbCN8* and *SbCN15*, and postpone floral initiation. In inductive SD, when plants are in darkness during the evening, *SbPRR37* expression is reduced, resulting in de-repression of the floral activators and floral induction.

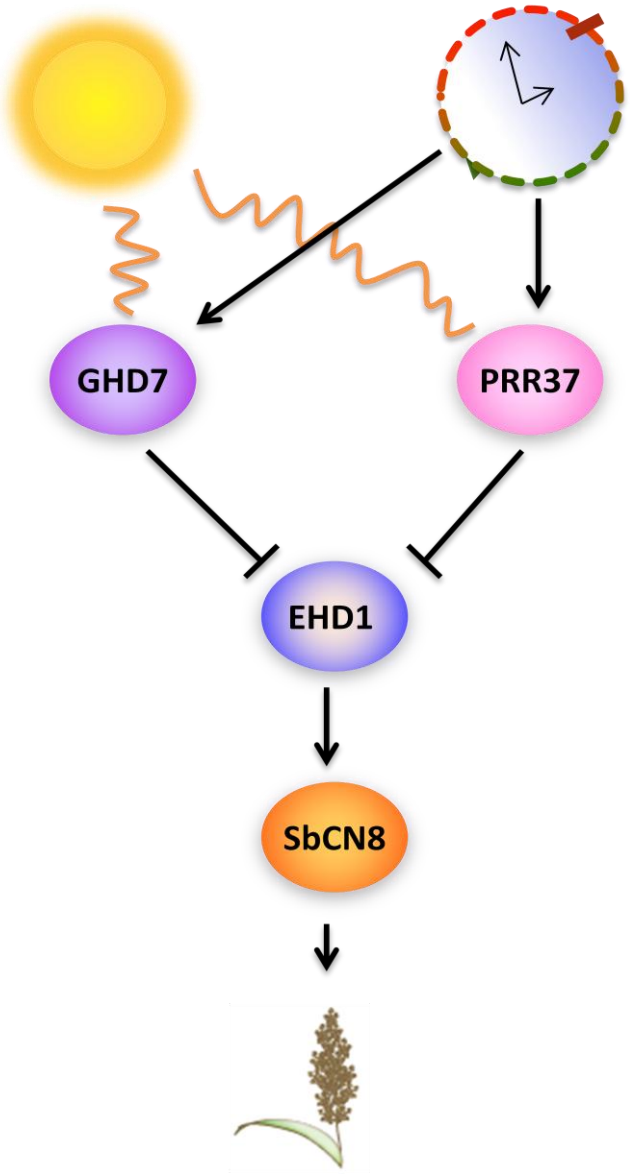


Figure 16. The model of SbGHD7 in photoperiod flowering time pathway in sorghum. *GHD7* modulates photoperiod sensitivity and floral repression in an additive fashion with *SbPRR37*. GHD7 inhibits expression of the floral activators *EHD1* and *SbCN8*, resulting in repression of flowering in LD but not in SD.

Dissertation Overview

In order to further characterize the genetic regulatory pathway that modulates photoperiod flowering time in sorghum, this research was focus on two activators:

In Chapter II, the study of phytochrome B (PHYB) elucidated that light signaling mediated by PHYB is required for expression of the floral repressors *SbPRR37* and *SbGHD7* in the evening of long days and for photoperiod regulated flowering in sorghum. In long days, the phyB deficient genotype 58M (*phyB-1*, *ma3R*) flowered ~60 days earlier than 100M (*PHYB, Ma3*) and ~11 days earlier in short days. Populations derived from 58M (*Ma1*, *ma3R, ma6*) and R.07007 (*Ma1, Ma3, Ma6*) segregated for flowering time QTL aligned to *PHYB/phyB-1* (*Ma3*), *GHD7/ghd7-1* (*Ma6*) and *Ma5*. In genotypes lacking PHYB, *SbPRR37* (*Ma1*) and *SbGHD7* (*Ma6*) did not delay flowering in long days. Light signaling mediated by PHYB was required for high expression of the floral repressors *SbPRR37* and *SbGHD7*, especially during the evening of long days. The floral activators EARLY HEADING DATE 1 (*SbEHD1*) and *SbCN8* (*ZCN8*) were highly expressed in long and short days in 58M (*phyB-1*) but repressed in 100M (*PHYB*) in long days.

In Chapter III, the study of sorghum CONSTANS (SbCO) demonstrated that SbCO is an activator of flowering that is repressed post-transcriptionally in long days by the floral inhibitor PRR37, contributing to photoperiod sensitive flowering in *Sorghum*, a short day plant. Quantitative trait loci (QTL) that modify flowering time in sorghum were identified by screening Recombinant Inbred Lines (RILs) derived from BTx642 and Tx7000 in long days, short days, and under field conditions. Analysis of the

flowering time QTL on SBI10 revealed that BTx642 encodes a recessive *CONSTANS* allele containing a His106Tyr substitution in B-box 2 known to inactivate CONSTANS in *Arabidopsis thaliana*. Genetic analysis characterized sorghum CONSTANS as a floral activator that promotes flowering by inducing the expression of *EARLY HEADING DATE 1* (*SbEHD1*) and sorghum orthologs of the maize FT genes *ZCN8* (*SbCN8*) and *ZCN12* (*SbCN12*). The floral repressor *PSEUDORESPONSE REGULATOR PROTEIN 37* (PRR37) inhibits sorghum CONSTANS activity and flowering in long days.

Various alleles of flowering time modulatorcan be used by sorghum breeders to control flowering time more precisely and produce sorghum varieties for different purposes, including food, forage and biofuel.

CHAPTER II

SORGHUM PHYTOCHROME B INHIBITS FLOWERING IN LONG DAYS

Background

Flowering time has a significant impact on plant adaptation to agro-ecological environments, biomass accumulation and grain yield [2]. Floral initiation is regulated by plant development, photoperiod, shading, temperature, nutrient status, and many other factors [19-22]. Signals from many input pathways are integrated in the shoot apical meristem (SAM) through regulation of the meristem identity genes *LFY* and *AP1*, which are activated during transition of the SAM from a vegetative meristem to a floral meristem. Long day (LD) plants, such as *Arabidopsis*, flower earlier in LD compared to short days (SD). In contrast, SD plants, such as rice and sorghum, show delayed floral initiation under LD conditions. Photoperiod regulated flowering is mediated by light signaling from photoreceptors and output from the endogenous circadian clock consistent with external coincidence models of flowering time regulation [29]. Photoperiod sensitive *Sorghum bicolor* genotypes delay floral initiation when grown under LD conditions. Sorghum genotypes with reduced photoperiod sensitivity have been identified and used by breeders because they flower early and at similar times in both long and short days, enhancing grain production [12]. In contrast, bioenergy sorghum is highly photoperiod sensitive, flowering in long day environments only after an extended phase of vegetative growth, thereby increasing biomass accumulation and nitrogen use efficiency [2, 11].

Photoperiod regulated flowering requires perception of light and signaling by plant photoreceptors such as the red/far-red light sensing phytochromes (Phy), blue light/ultraviolet wavelength sensing cryptochromes (Cry), phototropins, and Zeitlupes [30, 31]. Phytochromes play an important role in flowering time regulation in most plants including rice [114] , barley [115], and sorghum [16]. The sorghum genome encodes three phytochrome genes, *PHYA*, *PHYB* and *PHYC*. Inactivation of PhyB results in early flowering in long days [16]. Phytochromes are soluble chromoproteins that contain an N-terminal photosensory domain and a C-terminal dimerization moiety. There are three sub-domains in the N-terminal moiety: PAS (PER, ARNT and SIM), GAF (cGMP phosphodiesterase, adenylate cyclase, Fh1A) and PHY (phytochrome-specific GAF-related), which form a unique structure, the "light-sensing knot"[72]. The PAS/GAF domains transduce light signals and the C-terminal domain, consisting of two PAS and HKRD (histidine-kinase-related domain), is responsible for dimerization and nuclear localization.

The central oscillators of the plant circadian clock are encoded by *TOC1, CCA1* and *LHY* [35]. Rhythmic expression of these central oscillators modulates the expression of *GI*, an output gene of the circadian clock. GI, in concert with other factors, activates expression of *CO*, a zinc-finger transcription factor that plays an essential role in photoperiod regulation of flowering time in Arabidopsis [82], rice [47] and sorghum [116]. In *Arabidopsis*, CO is stabilized and accumulates during the evening of long days through the action of Cry1, Cry2 and PhyA, where it activates expression of *FT* and flowering. In SD, CO is not stabilized during the evening because *CO* expression occurs

in darkness [41]. FT is produced in leaves and translocated to the SAM where it binds to FD. In Arabidopsis, FT together with SOC1, promote expression of meristem identity genes *LFY* and *AP1*, leading to floral transition [41].

The core of photoperiod regulatory pathway *GI-CO-FT* is present in *Arabidopsis,* a LD plant, and the SD plants rice and sorghum. In rice, *OsGI*, *Hd1*, and *Hd3a* are orthologs of *GI*, *CO*, and *FT*, respectively [46]. Hd1 (*OsCO*) delays flowering time in LD in rice and activates flowering in SD. In addition, Itoh et al. [53] identified a pair of genes, *EHD1* and *GHD7* that regulate flowering in response to day length by modifying expression of *Hd3a* (florigen) in rice. *EHD1* activates *Hd3a* expression and induces the floral transition in SD. In contrast, *GHD7*, a homolog of wheat *VRN2* [117], represses flowering in LD by down-regulating *EHD1* and *Hd3a*. In sorghum CO activates flowering in SD by inducing expression of *SbEHD1*, *SbCN8* and *SbCN12,* whereas in LD, CO activity is inhibited by SbPRR37 [116].

More than 40 flowering time QTL have been identified in sorghum [118]. At least six of the flowering time QTL, termed maturity loci *Ma1-Ma6*, modify photoperiod sensitivity [12, 13, 18]. Dominance at *Ma1-Ma6* delays floral initiation in long days. *Ma3* encodes phytochrome B, indicating that light signaling through this photoreceptor is required for photoperiod sensitive variation in flowering time [16]. *Ma6* was identified as *SbGHD7*, a repressor of flowering in long days [113]. In LD, SbGhd7 increases photoperiod sensitivity by inhibiting expression of the floral activators *SbEHD1*, *SbCN12* and *SbCN8*. *Ma1* was identified as *SbPRR37*, a floral repressor that acts in LD [15]. The orthologs of *SbPRR37* in wheat and barley, *PHOTOPERIOD 1*

(*Ppd1, Ppd-H1, Ppd-D1a*) [100, 102] and rice *OsPRR37* [98], also modulate flowering time in response to photoperiod. In LD, SbPRR37 inhibits expression of *SbEHD1*, *SbCN12*, and *SbCN8,* resulting in repression of flowering [15]. Moreover, *SbPRR37* modulates photoperiod sensitivity and floral repression in an additive fashion together with *SbGHD7*. Expression of *SbPRR37* and *SbGHD7* are regulated by the circadian clock and light, suggesting common upstream regulation [113].

The current study focuses on elucidating how phytochrome B regulates flowering time in response to day-length in sorghum. We report here that *PHYB* is required for light activation of *SbPRR37* and *SbGHD7* expression in the evening of long days, resulting in repression of *SbEHD1*, *SbCN12*, *SbCN8* and floral initiation.

Results

PHYB alleles in diverse sorghum lines

58M, a photoperiod insensitive early flowering sorghum line has the genotype $ma3^{R}ma3^{R}$ corresponding to the *phyB-1* allele [16]. This allele contains a frame shift mutation that results in a prematurely terminated PhyB, lacking regions of the protein necessary for dimerization and biological activity. To confirm and extend prior analysis of *SbPHYB*, alleles from several diverse sorghum lines that vary in photoperiod sensitivity were sequenced and compared. The full length genomic sequence of *PHYB* from BTx623 and 100M (both *Ma3*) was 7285bp in length consisting of four exons encoding a protein with 1178 amino acid residues. *PHYB* sequences from R.07007, Hegari, Tx7000, BTx642, SC56, Shallu and BTx3197 were identical to BTx623 and

100M (*Ma3*). The *PHYB* sequence from 58M ($ma3^R$), referred to as *phyB-1* (Table 3), contains a mutation that renders the gene inactive (3). No coding mutations were identified in 90M, a line that encodes the weak allele *ma3* [13]. IS3620C encodes a third allele, designated *phyB-2*, which differs from *PHYB* by one INDEL and two SNPs resulting in one amino acid deletion and two amino acid substitutions (Table 3). The first substitution could alter function because it produces an $Asp^{308}Gly$ change in the GAF domain of PhyB.

Table 3. Sequencing analysis of *PHYB* coding alleles in different sorghum lines.

	Exon 1	Exon 1	Exon 3	Exon 4	Sorghum Genotypes
Nucleotide Variation	CAC >…	A>G	A>.	C>G	
Protein Modification	His>…	Asp>Gly	Premature stop codon	Leu>Val	
Mutation Position (AA #)	31	308	1023	1113	
Alignment with PHYB in *Arabidopsis* (AA #)	32	293	1007	1096	
Phytochrome Domain		GAF(N)			
***PHYB* (Ma_3 or ma_3)**	-	-	-	-	**BTx623, 100M, 90M, R.07007, Hegari, Tx7000, BTx642, SC56, Shallu, BTx3197**
***phyB-1* (ma_3^R)**	-	-	+	-	**58M**
phyB-2	+	+	-	+	**IS3620C**

PhyB affects flowering time in LD and SD

The sorghum maturity standards, 100M, 90M, and 58M were constructed from Milo genotypes that contain alleles of *Ma1* and *Ma3* that modify flowering time [13]. The sorghum maturity standard 100M is photoperiod sensitive with a maturity genotype *Ma1Ma2Ma3Ma4Ma5ma6* [113]. The genotype 58M is photoperiod insensitive, flowers early in LD and SD, and has the genotype *Ma1Ma2ma3^{R}Ma4Ma5ma6* [113]. Genotype 58M contains null alleles of *Ma3* (*ma3^{R}*, *phyB-1*) and *Ma6* (*ghd7-1)*. When grown in a greenhouse under 14 h LD during the summer, 58M plants were spindly and flowered in ~62 days, whereas 100M flowered in ~126 days due to the repressing action of SbPRR37 (*Ma1*) (Fig 17A). This result confirmed that loss of PhyB activity in 58M reduces the ability of *Ma1* to inhibit flowering in LD [16]. When grown in a greenhouse in 10 h SD during December-February at lower light intensity, 100M flowered in ~59 days while 58M flowered in ~48 days (Fig 17B). Therefore in sorghum, PhyB has a smaller but still significant effect on flowering time in SD.

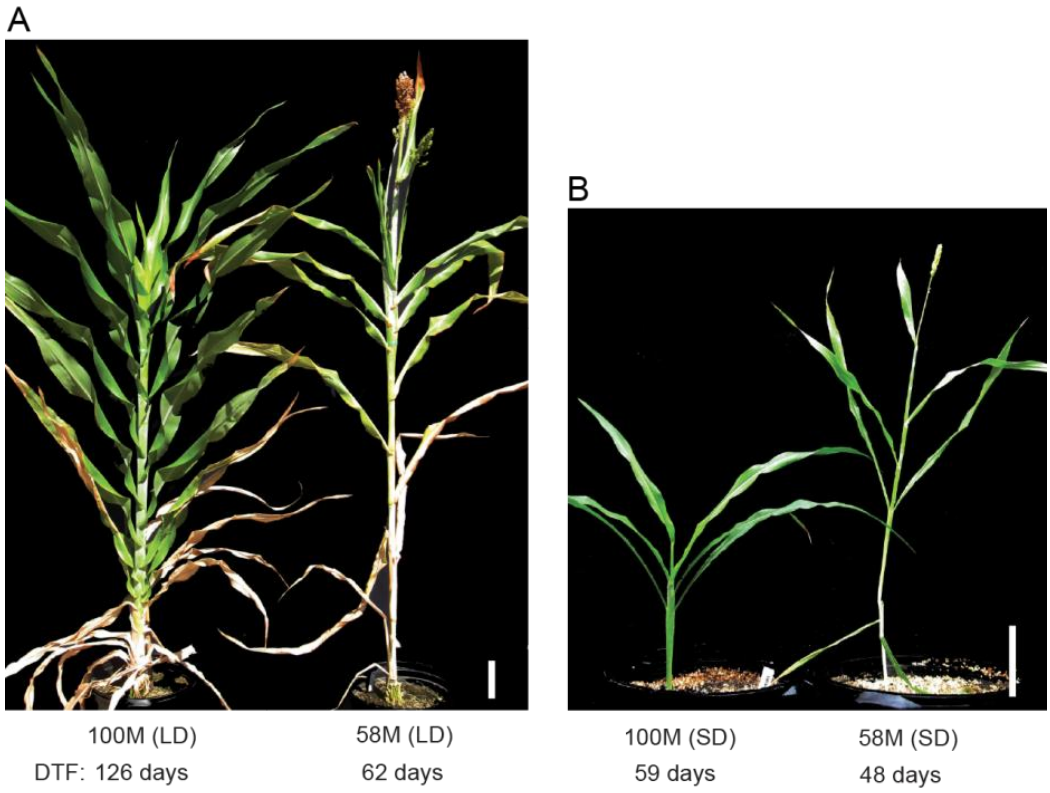


Figure 17. Flowering phenotypes of sorghum lines. (A) Phenotype of 100M (left) and 58M (right) under long day condition (14h light/10h dark). 100M and 58M flowered at 126 days and 62 days respectively. (B) Phenotype of 100M (left) and 58M (right) under short day condition (10h light/14h dark). 100M flowered at 59 days and 58M flowered at 48 days. LD: long day condition. SD: short day condition. DTF: days towards flowering time. Scale bar is 8.6cm.

***PHYB* is epistatic to *Ma1* (*SbPRR37*) and *Ma6* (*SbGHD7*)**

In sorghum, *SbPRR37* (*Ma1*) and *SbGHD7* (*Ma6*) act in additive fashion to inhibit flowering in LD. Expression of both genes is induced by light, although the photoreceptor or photoreceptors that mediate the response were not known prior to the current study [15, 113]. To examine how *PHYB (Ma3)*, *SbPRR37* (*Ma1*), and *SbGHD7*

(*Ma6*) co-regulate the timing of floral initiation, F2 and F3 populations were derived from a cross of R.07007 (*Ma1Ma3ma5Ma6*) and 58M (*Ma1ma3^R Ma5ma6*). These populations segregated for a wide range of flowering times (~85 days) when planted in July and grown in a greenhouse in 14 h LD. Digital genotyping [119] was employed to generate DNA markers for genetic map construction. The genetic map spanned all of the ten sorghum chromosomes, although the long arm of SBI02 and SBI09 were deficient in DNA markers. Analysis carried out using QTL Cartographer identified three significant QTL (LOD score >3.7) for days to anthesis in LD using the F2 population (n=86), which together explained ~50% of the phenotypic variance for flowering time (Fig 18A). The QTL on chromosome 1 aligned with *PHYB* had the highest LOD score (LOD =24.2) (Table 4). This QTL (1 LOD interval) spanned DNA on chromosome 1 from 60,402,909-61,604,749 that encodes *PHYB* (chromosome_1:60,915,677-60,917,553). Recessive *ma$_3^R$* alleles from 58M associated with this QTL contributed early flowering time phenotypes explaining 14.1% of the phenotypic variance. The flowering time QTL on chromosome 6 spanned a physical interval from 203,707-1,716,581 (1 LOD interval) aligned with *SbGHD7* [113]. The recessive *ghd7-1* null allele from 58M was associated with early flowering in LD explaining 15.5% of the variance in flowering time. The third flowering time QTL near the proximal end of chromosome 1 (chromosome 1: 6,139,583-9,077,991) had a LOD score of 8.7 and explained 19.6% percent of phenotype variance. This QTL was tentatively identified as *Ma5* because R.07007 was reported to be recessive for *Ma5*, a rare allele in sorghum [18]. No QTL aligned with *Ma1* as expected because both 58M and R.07007 contain

dominant alleles of *Ma1 (SbPRR37)*. These three flowering time QTL were also identified in the F3 population (Fig 19, 20).

Table 4. Parameters of the QTLs for Days to Flowering in F2 58MxR.07007 population.

QTL	Ma Locus	Chr #	Position (cM) [a]	LOD score	Physical Interval [b]	Additive Effect [c]	Dominant Effect [d]	R2 [e]
1	*Ma5*	Ch_1	1.8	8.6602	6139583-9077991	-17.085	18.1857	0.1964
2	*Ma3*	Ch_1	99.4	24.2142	60402909-61604749	12.5452	16.0913	0.1408
3	*Ma6*	Ch_6	7.2	12.0869	203707-1716581	12.8255	5.8149	0.1549
Total								49.21%

- **a.** Position of likelihood peak (highest LOD score).
- **b.** Physical Interval: physical coordinate interval spanning 1 LOD interval across the likelihood peak.
- **c.** Additive Effect: A positive value means the delay of flowering time due to R.07007 allele. A negative value means the delay of flowering time due to 58M allele.
- **d.** Dominant Effect: A positive value means dominance for the delay of flowering time.
- **e.** R^2 (coefficient of determination): percentage of phenotypic variance explained by the QTL.

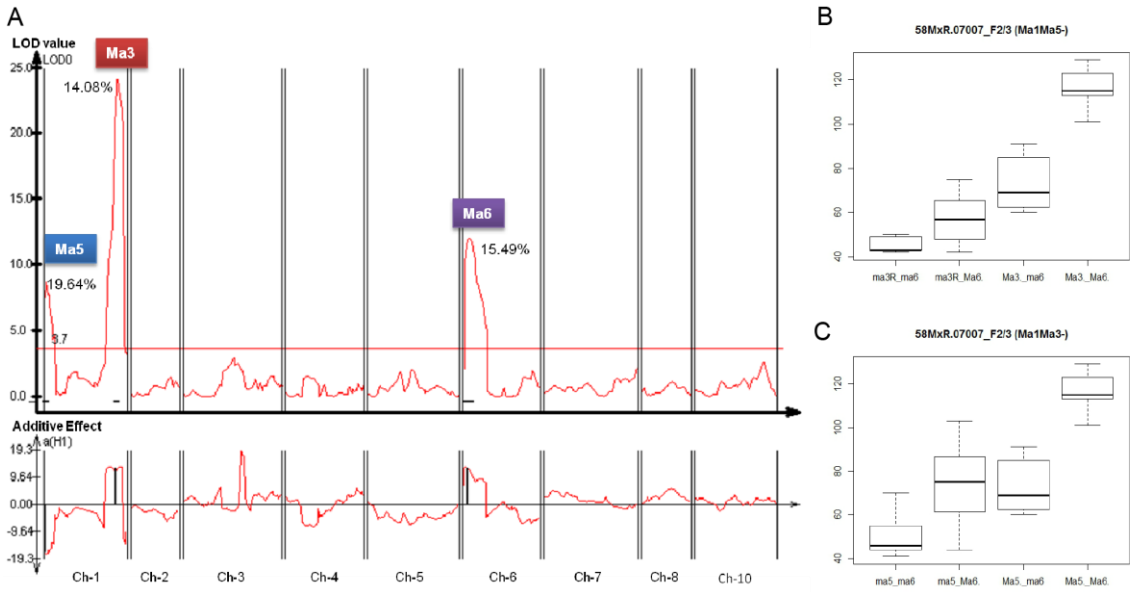


Figure 18. QTL and epistasis analysis of *Maturity* loci. (A) Three major QTLs associated with days to flowering in F2 population of 58MxR.07007 were labeled underscore. The LOD values of all of them passed the significant threshold 3.7 (horizontal red line), calculated based on permutations with 1000 repeats and α=0.05. The QTLs correspond to *Ma5*, *Ma3* and *Ma6* with R^2 value 19.64, 14.08 and 15.49 respectively. R^2 value (the coefficient of determination) correlates with the phenotype variance explained by the specific QTL. For additive effect, positive value indicates allele from R.07007 contribute to the delay of flowering time, while negative value represents R.07007 allele can hasten the floral transition. Chromosome numbers are given below. The number of markers for the QTL detection is 285. The F2 population size (N) is 86. The phenotype variance is 572.3. (B) Boxplot of flowering time distribution in *Ma1Ma5*- background of 58MxR.07007 F2 population. (C) Boxplot of flowering time distribution in *Ma1Ma3*- background of 58MxR.07007 F3 population. Median values are represented by horizontal lines within the boxes. Boxes represent the first and third quartiles. Bars represent the lowest datum still within 1.5 interquartile range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile.

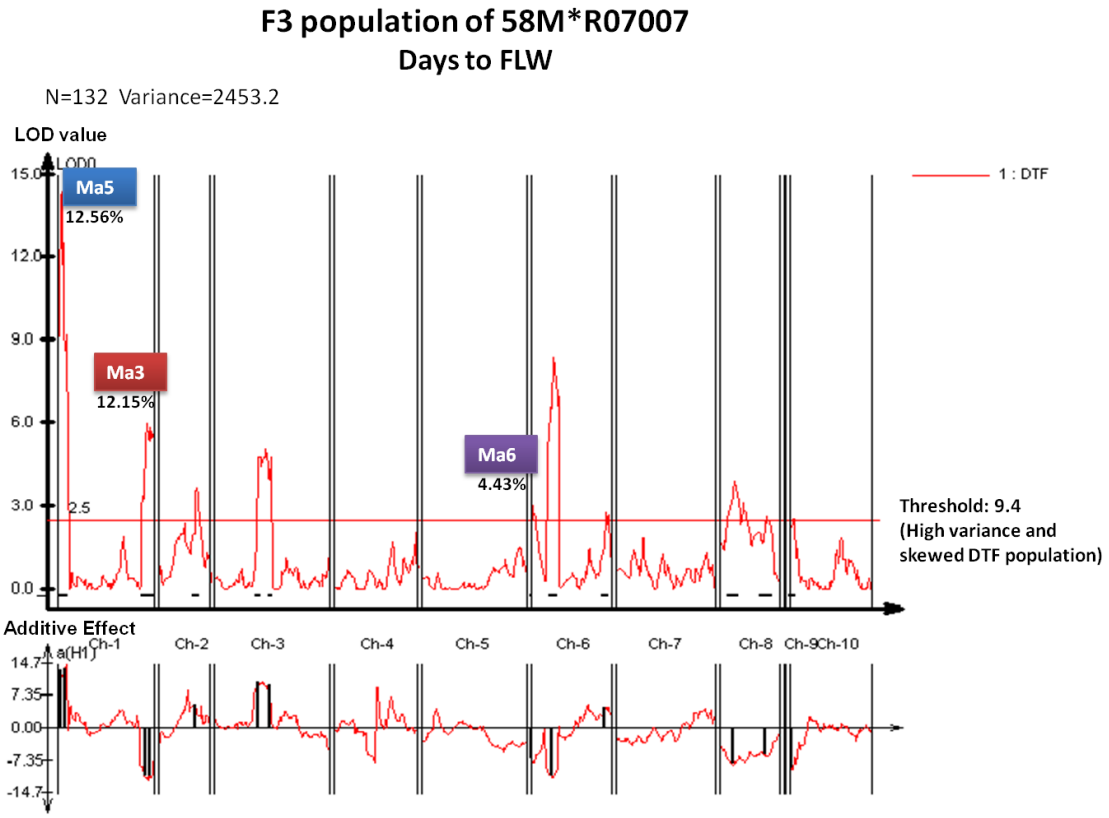


Figure 19. Flowering time QTLs in 58MxR.07007 F3 population under LD condition. Candidate gene with R^2 value (proportion of phenotypic variance explained by the QTL) was labeled above each QTL. Significant threshold of LOD score was noted on the right side. Positive value of additive effect represents allele from R.07007 contributes to delayed flowering time.

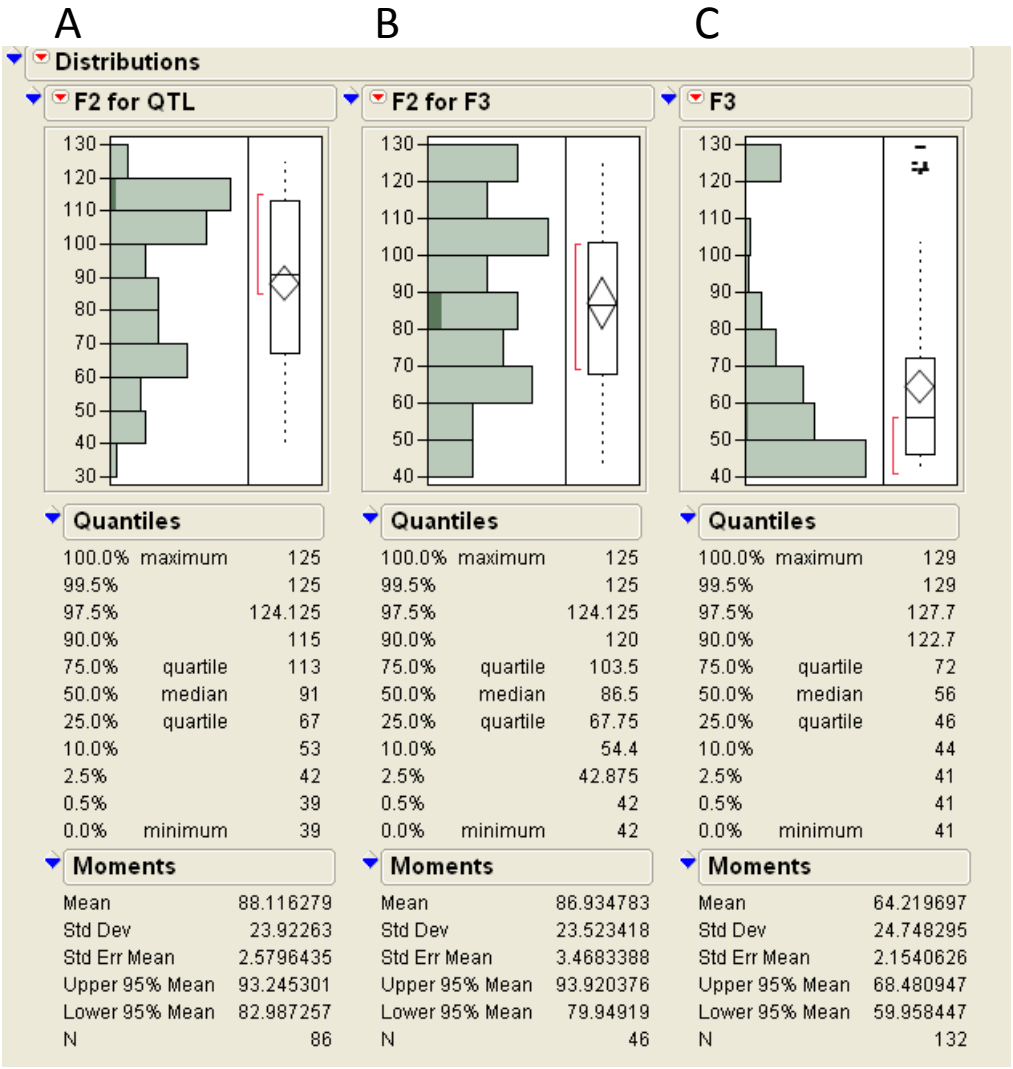


Figure 20. Distribution of flowering time in 58MxR.07007 population. (A) F2 population for QTL analysis. (B) F2 population seeds used for F3 population planting. (C) F3 population for QTL analysis.

Plants from the F2/3 population are homozygous for *Ma1*, a repressor of flowering in LD, but varied in alleles of *Ma3, Ma5* and *Ma6*. Progeny with the genotype *Ma3_Ma5_Ma6_* flowered later (101-129 days) than plants with the genotype *Ma3_Ma5_ma6ma6* (60-91 days) consistent with increased floral repression due to *Ma6* in *Ma1* dominant backgrounds. Progeny with the genotypes *Ma3_Ma5_Ma6_* and *Ma3_Ma5_ma6ma6* flowered later than genotypes that were homozygous recessive for $ma3^R$ showing that *PHYB* is epistatic to the floral repressors encoded by *Ma1* and/or *Ma6* (Fig 18B/C; Table 5). Progeny that were $ma3^Rma3^RMa5\_Ma6\_$ flowered ~14 days later than plants with the genotype $ma3^Rma3^RMa5\_ma6ma6$. It was noted that progeny lacking PhyB with *Ma6* genotypes showed a significant range of flowering times (42-75 days) suggesting that additional genes and/or environmental factors affect *Ma6* action in this genetic background (Fig 18B). A similar wide range of flowering time was observed among plants with the genotype *Ma3_ma5ma5Ma6_* (Fig 18C; Table 5). In addition, plants with the genotype *Ma3_Ma5_Ma6_* flowered later in LD than plants with the genotypes *Ma3ma5ma5Ma6_ or Ma3ma5ma5ma6ma6.* This shows that Ma_5 is also required for late flowering in LD in *Ma1Ma3* backgrounds and that *Ma5* is epistatic to *Ma1* and *Ma6* (Fig 18C). Plants with the genotype $ma3^Rma3^RMa5\_ma6ma6$ and *Ma3_ma5ma5ma6ma6* flowered early and in a similar number of days as genotypes that are homozygous recessive for both $ma3^R$ and *ma5* ($ma3^Rma3^Rma5ma5ma6ma6$) indicating that the products of both *Ma3* and *Ma5* are required in LD for delayed flowering mediated by *Ma1* (*SbPRR37*).

Table 5. Flowering time of F2/F3 progeny from 58MxR.07007 in LD.

Genotype (All plants = Ma1Ma1)			Days to Flowering: median (range)
*Ma3_*	*Ma5_*	*Ma6_*	115 (101-129)
*Ma3_*	*Ma5_*	*ma6ma6*	69 (60-91)
$ma3^{R}ma3^{R}$	*Ma5_*	*Ma6_*	57 (42-75)
$ma3^{R}ma3^{R}$	*Ma5_*	*ma6ma6*	43 (42-50)
*Ma3_*	*ma5ma5*	*Ma6_*	75 (44-103)
*Ma3_*	*ma5ma5*	*ma6ma6*	46 (41-70)
$ma3^{R}ma3^{R}$	*ma5ma5*	*ma6ma6*	44 (39-68)

The need for both PhyB and the product of *Ma5* to observe delayed flowering in LD led us to examine the *Ma5* locus for candidate genes that might explain this interaction. The *Ma5* locus is located on SBI-01 and spans several genes known to affect flowering time in other plants including *AP1, CK2*, and *PHYC*. *PHYC* appeared to be the best candidate gene for *Ma5* because PhyC modifies flowering time in rice specifically in LD similar to *Ma5* in sorghum [120], PhyB stabilizes PhyC , and PhyB:PhyC act as heterodimers in both Arabidopsis [121, 122] and rice [120], consistent with the co-dependence observed between *PHYB* and *Ma5* in this study. Phylogenetic tree of PhyC homologs from sorghum, maize, rice and Arabidopsis is showed in Figure 21. Comparison of *PHYC* sequences from BTx623 (*Ma5*) (reference genome), 100M (*Ma5*), and R.07007 (*ma5*) revealed four differences in PhyC amino acid sequence between BTx623 and R.07007, and two differences between 58M/100M and R.07007

(Table 6) (Fig 22). The latter amino acid variants occur in the PAS domain (Gly:Val) and HKRD domain (Glu:Asp) and SIFT analysis [123] indicated these changes could affect the function of PhyC. These results are consistent with *SbPHYC* as a candidate gene for *Ma5*. Further analysis is underway to test this assignment.

Table 6. Sequencing analysis of *PHYC* coding alleles in different sorghum lines.

	Exon 1	Exon 1	Exon 1	Exon 2	Sorghum Genotypes
Nucleotide Variation	G >T	G>A	T>C	G>T	
Protein Modification	Gly>Val	Gly>Arg	Val>Ala	Glu>Asp	
Mutation Position (AA #)	124	162	190	922	
Alignment with PHYB in *Arabidopsis* (AA #)	160	198	226	954	
Phytochrome Domain	PAS(N)	PAS(N)	PAS-GAF Loop	HKRD(C)	
***PHYC-1* (*Ma_5*)**	**-**	**-**	**-**	**-**	**BTx623**
PHYC-2	**-**	**+**	**+**	**-**	**100M, 90M**
***phyC-1* (*ma_5*)**	**+**	**+**	**+**	**+**	**R.07007**

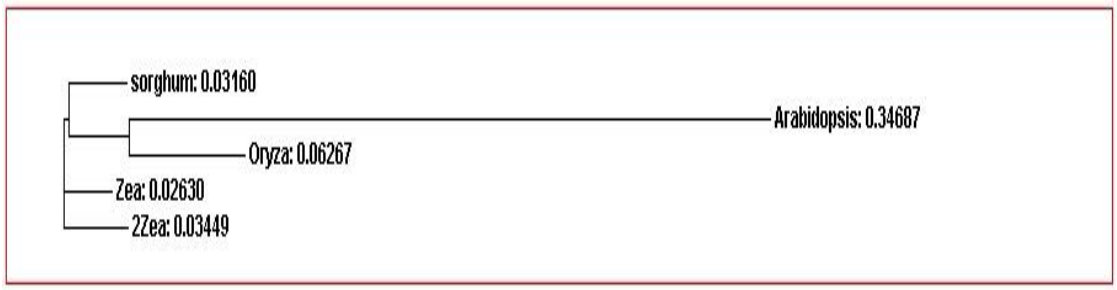


Figure 21. Phylogenetic tree of phytochrome C (PhyC) homologs from sorghum, maize, rice and Arabidopsis. Distances are shown following each species.

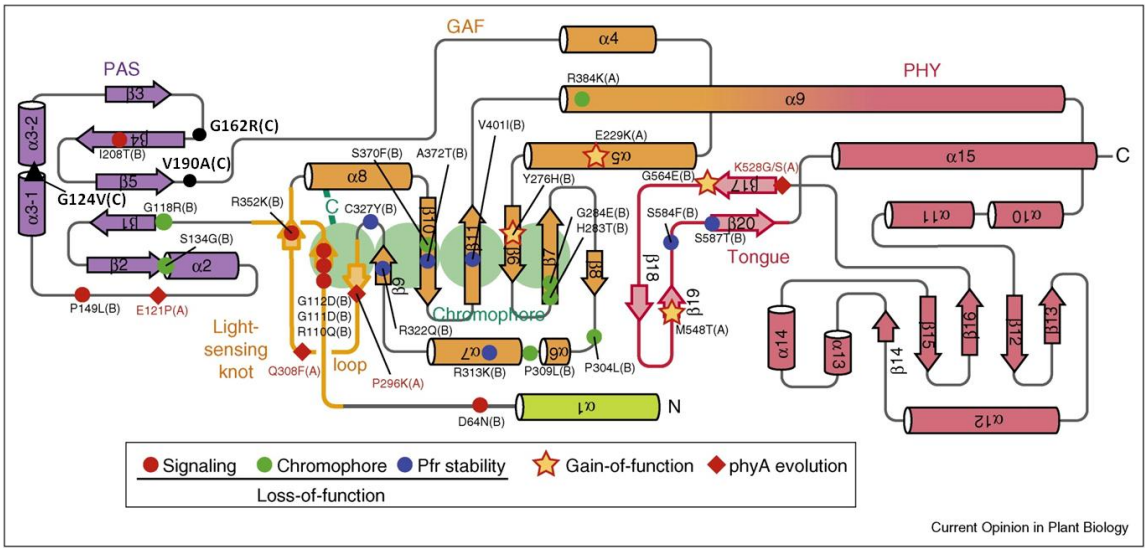


Figure 22. Location of sorghum phytochrome C (PhyC) mutations in the schematic representation of phytochrome N-terminal moiety. The structure representation is based on Arabidopsis PhyA and PhyB. The positions of sorghum PhyC mutations are determined by amino acid alignment. Mutation only identified in R.07007 (*phyC-1* (*ma5*) allele unique) was labeled as triangle; and mutations identified in R.07007 and shared with other genotypes (such as 100M and BTx642) (*PHYC-2* alleles) are labeled as circles. [72]

PhyB modulates expression of *SbPRR37* and *SbGHD7* in long days

Expression of *SbPRR37* and *SbGHD7* in leaves is regulated by light and gating by the circadian clock [15, 113]. The influence of PhyB on *SbPRR37* and *SbGHD7* expression was analyzed using 100M (*PHYB*) and 58M (*phyB-1*) plants grown for 32 days in LD then entrained for 7 days in LD or SD (Fig 23). Following entrainment, leaf samples were collected from plants for one 24 h LD or SD light-dark cycle, then from plants exposed to continuous light and temperature for an additional 48h. In leaves of 100M, *SbPRR37* and *SbGHD7* expression peaked in the morning (arrow) and evening (arrowhead) in LD as previously reported (Fig 23A/C, solid lines). *SbPRR37* and *SbGHD7* RNA abundance continued to oscillate with peaks in the morning and evening when 100M plants were transferred to continuous light and temperature consistent with regulation by the circadian clock (Fig 23, 24-72 h). In leaves of 58M in LD, *SbPRR37* and *SbGHD7* showed an increase in RNA abundance in the morning (arrow) but only a small increase in expression in the evening (arrowhead) compared to 100M (Fig 23A/C, dashed red line). These results indicate that *PHYB* is required for the elevated evening expression of *SbPRR37* and *SbGHD7* in LD observed in 100M.

When 100M and 58M plants were entrained and assayed in SD, the morning peak of *SbPRR37* expression was of similar amplitude in both genotypes and expression of *SbPRR37* was low during the evening (Fig 23B). Similarly, *SbGHD7* expression in SD was highest in the morning and similar in 100M and 58M and lower in the evening, and compared to LD (Fig 23D). These results indicate that in SD, PhyB has a limited effect on *SbPRR37* and *SbGHD7* expression. When 100M plants entrained in SD were

exposed to continuous light, the evening peak of *SbPRR37* and *SbGHD7* expression observed in LD reappeared on the first subjective day, but expression levels increased to some extent in the second subjective day (Fig 23B/D). In 58M, the evening peak of *SbPRR37* and *SbGHD7* reappeared during the first subjective day, however overall expression was attenuated relative to 100M during the second subjective day.

PhyB modulates expression of *CO, Ehd1, SbCN8, SbCN12* and *SbCN15*

In 100M entrained to LD, the sorghum ortholog of *CONSTANS* (*SbCO*) shows peaks of expression at dawn (24h) and in the evening (15h) that are regulated by *SbPRR37*, the circadian clock, and day length [15]. In 58M entrained and sampled in LD, the amplitude of the peak of *SbCO* expression at dawn (24h) was reduced compared to 100M (Fig 24A). The peak of *SbCO* expression at dawn was also reduced and of similar amplitude in plants entrained in SD (Fig 24A, lower graph). These results show that the peak of *SbCO* expression at dawn is dependent on PhyB, most likely because expression of *SbPRR37* in the evening of LD is dependent on PhyB (Figure 23A). In contrast, the evening peak (15h) of *SbCO* expression was similar in both LD and SD in 100M and 58M indicating that PhyB does not significantly modulate *SbCO* expression at this time (15h) of day.

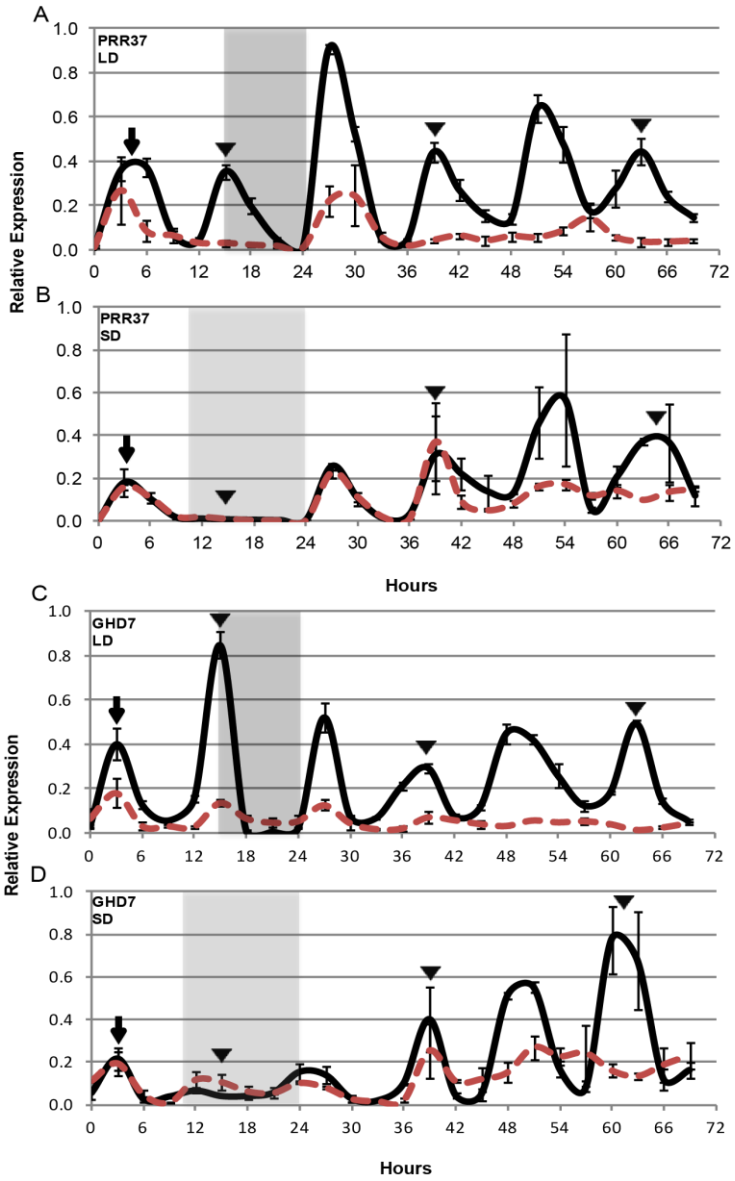


Figure 23. PhyB is the photoreceptor required for light-induced expression of *PRR37*and *GHD7*. 100M (solid black line) and 58M (dashed red line) were treated under LD (14h light/10h dark) or SD (10h light/14h dark) condition and subjected to qRT-PCR analysis for relative expression every 3 hours. First 24 hours are light-dark cycle, following 48 hours continuous light. The gray shade represents dark period. Arrows represent morning expression peak and arrowheads represent evening expression peak. (A) In LD, the second peak (arrowhead) of *PRR37* expression in the evening (~15h) is missing in *phyB* deficient line, 58M. (B) In SD, the second peak (arrowhead) of *PRR37* is absent in both 100M and 58M. (C) A similar expression pattern for *GHD7*. In LD, the second peak (arrowhead) of *GHD7* expression in the evening (~15h) is missing in 58M. (D) In SD, the second peak of *PRR37* is absent in both 100M and 58M.

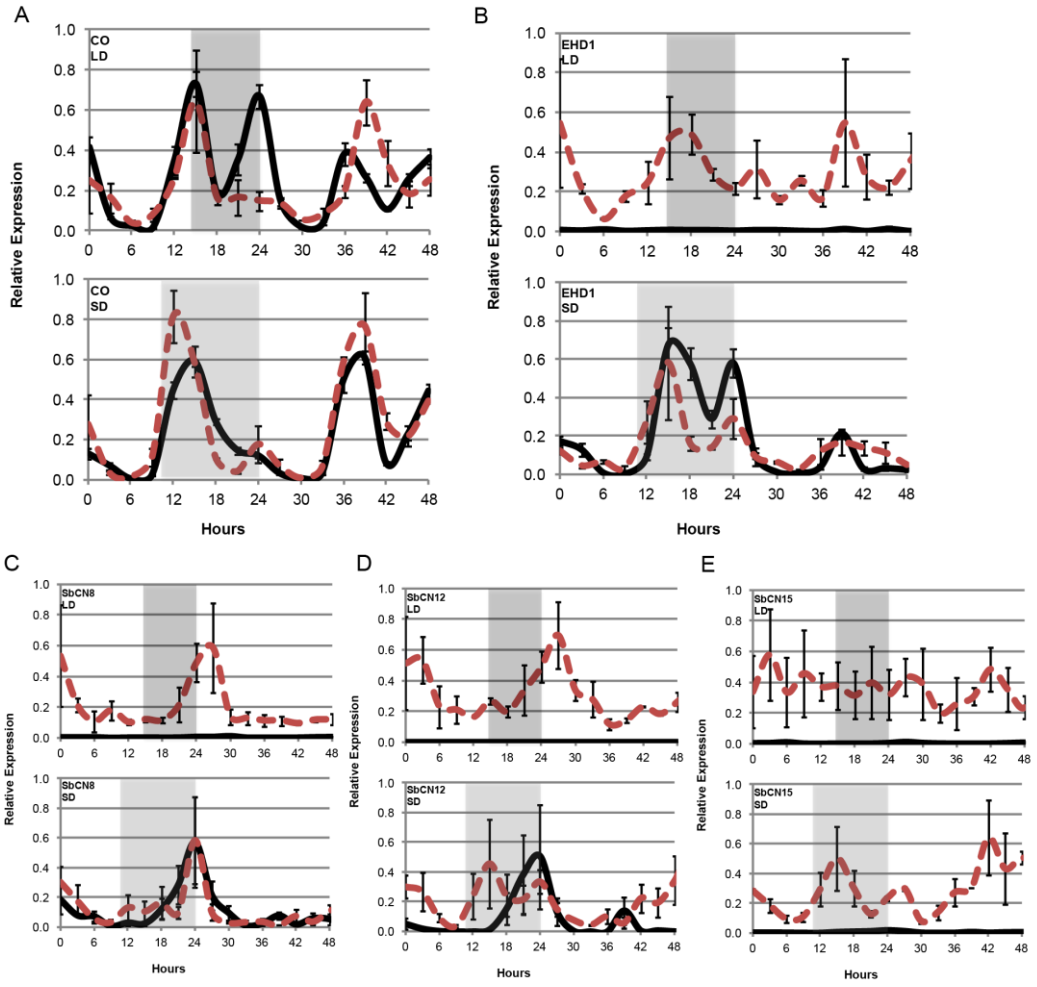


Figure 24. PhyB regulates the expression of downstream genes in flowering time pathway. 100M (solid black line) and 58M (dashed red line) were treated under LD (14h light/10h dark) or SD (10h light/14h dark) condition and subjected to qRT-PCR analysis for relative expression every 3 hours. First 24 hours are light-dark cycle, following 24 hours continuous light. The gray shade represents dark period. (A) *CO*, (B) *EHD1*, (C) *SbCN8*, (D) *SbCN12*, (E) *SbCN15*. Each data point of relative expression corresponds to three technical replicates and three biological replicates. Error bars indicates SEM.

Table 7. *SbCN* genes orthologs in maize and rice.

Sb Gene	Locus ID in Sorghum*		Location	
SbCN8	Sb09g025760		chr_9: 55149243 - 55150558	
SbCN12	Sb03g034580		chr_3: 62753997 - 62755638	
SbCN15	Sb10g003940		chr_10: 3464074 - 3465644	
	Zm Gene	**Locus ID in Maize**	**Os Gene**	**Locus ID in Rice**
SbCN8	*ZCN8*	GRMZM2G179264	*FTL10*	LOC_Os05g44180
SbCN12	*ZCN12*	GRMZM2G103666	*FTL9*	LOC_Os01g54490
SbCN15	*ZCN15*	GRMZM2G051338	*FTL2 (Hd3a)*	LOC_Os06g06320

* Gene Locus IDs in Sorghum, Maize and Rice are all based on Phytozome v8.0 (http://www.phytozome.net/).

EHD1 is an activator of *Hd3a*, one of the florigens in rice [54]. The sorghum ortholog of *Hd3a* is *SbCN15* (Table 7). Expression of *SbEHD1* increases when 100M is transferred from LD to SD in parallel with increased expression of *SbCN8* (ortholog of *ZCN8* [94]) and *SbCN12* (ortholog of *ZCN12*) that have been proposed to encode florigens in sorghum (Table 7). *SbPRR37* and *SbGHD7* repress expression of *SbEHD1* in 100M entrained in LD [15, 113]. Therefore *SbEHD1* expression in 58M and 100M was quantified and compared to determine if PhyB modulates *SbEHD1* expression. In LD, *SbEHD1* RNA abundance peaked in the evening and was up to ~100-fold higher in 58M relative to 100M throughout the time course (Fig 24B, upper; Fig 25A). In SD, expression of *SbEHD1* was high in both genotypes and peaked during the night (Fig 24B, lower; Fig 25A).

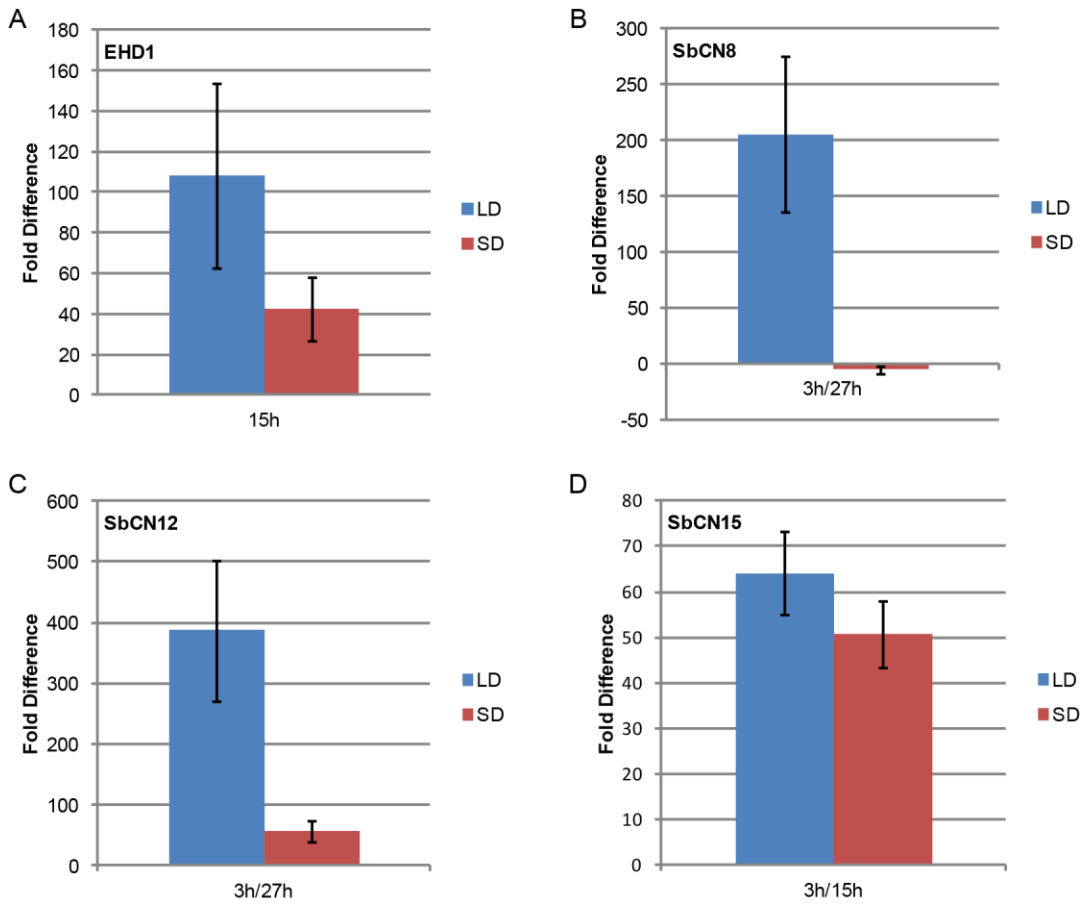


Figure 25. Fold differences of *EHD1*, *SbCN8*, *SbCN12* and *SbCN15* expression peaks of 100M and 58M grown in LD (14h light/10h dark) or SD (10h light/14h dark). Positive fold difference values indicate higher mRNA levels detected in 58M. (A) *EHD1*, (B) *SbCN8*, (C) *SbCN12*, (D) *SbCN15*.

In 58M entrained and analyzed in LD, expression of *SbCN8* (Fig 24C, upper) and *SbCN12* (Fig 24D, upper) peaked early in the morning and the relative abundance of RNA derived from these genes was elevated more than ~100-fold relative to their levels in 100M (Fig 25B/C). In SD, *SbCN8* (Fig 24C, lower) and *SbCN12* (Fig 24D, lower) expression was similar in both genotypes, and elevated ~100-fold relative to levels in 100M in LD (Fig 25B/C). Similarly, *SbCN15* (*Hd3a*) expression was increased up to ~60-fold in 58M compared to 100M in LD and SD (Fig 25D) at all time points assayed indicating that PhyB mediated repression of *SbCN15* expression occurs regardless of photoperiod.

PhyB could be inducing *SbPRR37* and *SbGHD7* expression directly, and/or indirectly by altering output from the circadian clock. To determine if allelic variation in *PHYB* affected clock gene expression, *TOC1* and *LHY/CCA1*, the central oscillators, and *GI*, a mediator of clock output were examined (Fig 26). In LD and SD, *TOC1*, *LHY* and *GI* expression in 58M and 100M peaked at similar times and showed similar amplitude of expression, although expression of GI was approximately 2-fold lower in 58M. *PHYB* and *PHYC* RNA levels were similar in 100M and 58M plants in LD and SD (data not shown).

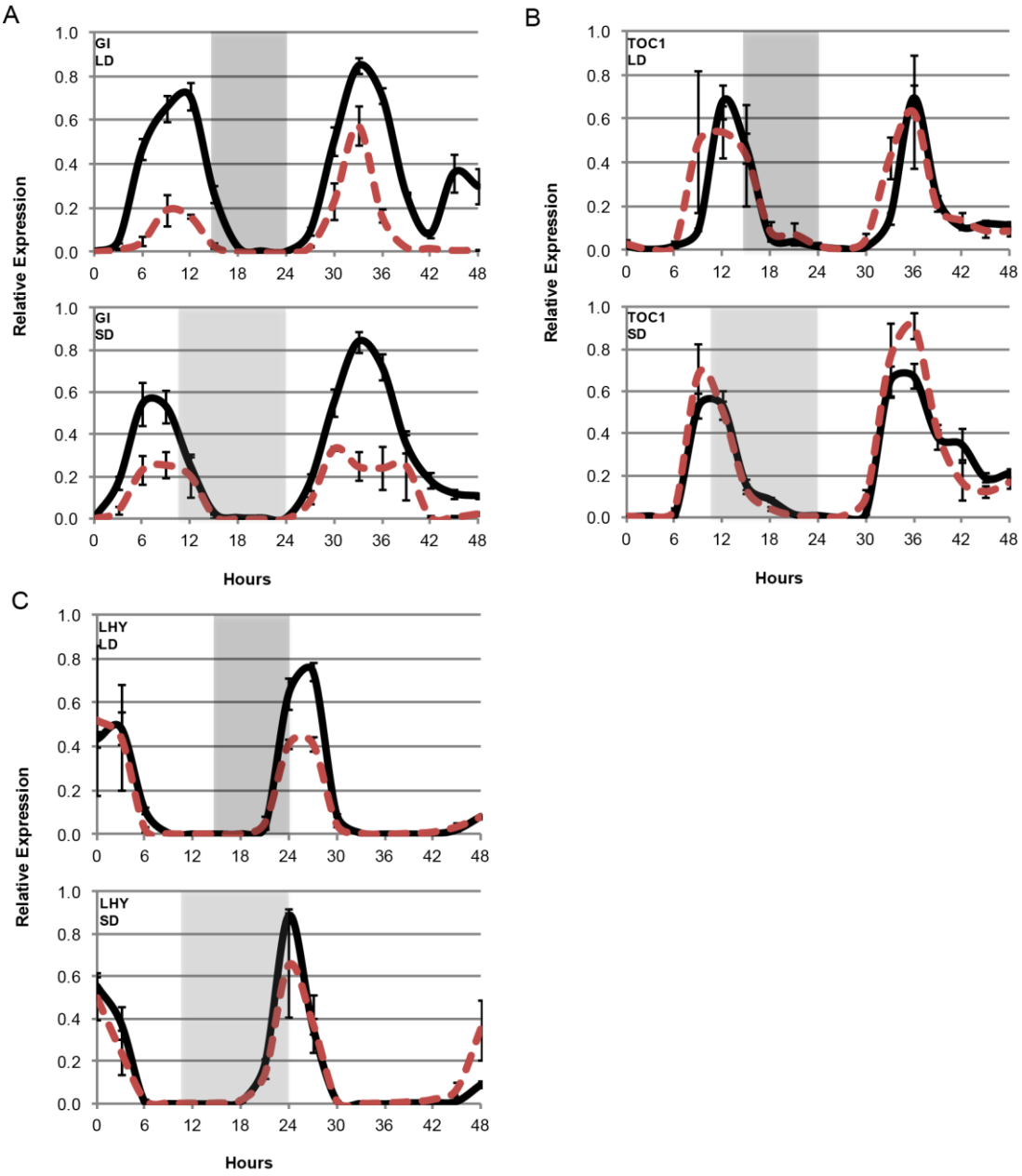


Figure 26. Relative expression level of circadian clock genes and *GI* in 100M (black solid line) and 58M (red dashed line) under either LD (14h light/10h dark) or SD (10h light/14h dark) conditions. The gray shaded area represents the dark period. The first 24h covers one light-dark cycle, followed by 24h of continuous light. (A) *GI*. (B) *TOC1*. (C) *LHY*. Each data point of relative expression corresponds to three technical replicates and three biological replicates. Error bars indicates SEM.

Discussion

Sorghum genotypes used for grain production are typically photoperiod insensitive and flower in 55-75 days when planted in April in locations such as College Station where day lengths increase during the early portion of the growing season. Early flowering in grain sorghum helps avoid adverse weather and insect pressure during the reproductive phase thereby enhancing yield. In contrast, highly photoperiod sensitive energy sorghum genotypes planted in this same location will not initiate flowering for 175 days until mid-September when day lengths decrease to less than 12.2 h [2, 18]. Delayed flowering results in long duration of vegetative growth of energy sorghum increasing biomass yield and nitrogen use efficiency [11]. The importance of optimal flowering time for sorghum productivity led us to investigate the genetic and molecular basis of variation in this trait in sorghum.

Variation of flowering time of sorghum germplasm grown in LD environments is caused principally by differences in photoperiod sensitivity, although shading, GA, temperature, length of the juvenile phase among other factors also affect this trait [12]. A model summarizing information about photoperiod regulation of flowering time in sorghum is shown in Figure 27. In LD, flowering is delayed in photoperiod sensitive sorghum by the additive action of the floral repressors, SbPRR37 (*Ma1*) and SbGhd7 (*Ma6*) [13, 15, 18, 113]. SbPRR37 and SbGHD7 repress expression of the grass specific floral activator, *SbEHD1*. In addition, SbPRR37 inhibits the activity of CO, another activator of flowering in sorghum [116]. The floral activators, SbEhd1 and SbCO, induce expression of *SbCN8* and *SbCN12,* the proposed sources of FT in sorghum.

SbCN15, the ortholog of Hd3a and a source of florigen in rice [46] may also be a source of florigen in sorghum. The circadian clock regulates expression of *SbCO, SbGI, SbPRR37* and *SbGHD7* and light regulates expression of *SbGHD7* and *SbPRR37* [15, 113].

Photoperiod has minimal impact on flowering time in sorghum genotypes such as SM100 that encode null versions of *SbPRR37* and *SbGHD7* [15, 113]. Presence of functional alleles of either gene increases photoperiod sensitivity and a further delay in flowering is observed when both genes are present in a dominant *Ma5* background. Expression of *SbPRR37* and *SbGHD7* is regulated by light and the circadian clock. Both genes show peaks of RNA abundance in the morning and again in the evening in LD and both peaks of RNA are attenuated in darkness. Importantly, the evening peak of expression is attenuated in SD when this phase occurs in darkness, indicating a need for light signaling during the evening to maintain sufficiently high levels expression of *SbPRR37* and *SbGHD7* to inhibit flowering. The dual peak of *SbPRR37* and *SbGHD7* expression observed in sorghum in LD is not observed in all plants. For instance, *PRR7*, the ortholog of *SbPRR37* in Arabidopsis, shows a single peak of clock-regulated expression during the morning [124]. In rice, *SbGHD7* shows a single peak of clock-gated expression in the morning of LD and a single peak during the night in SD [53]. These differences suggest that some features of flowering time regulation in sorghum will be different from rice and Arabidopsis.

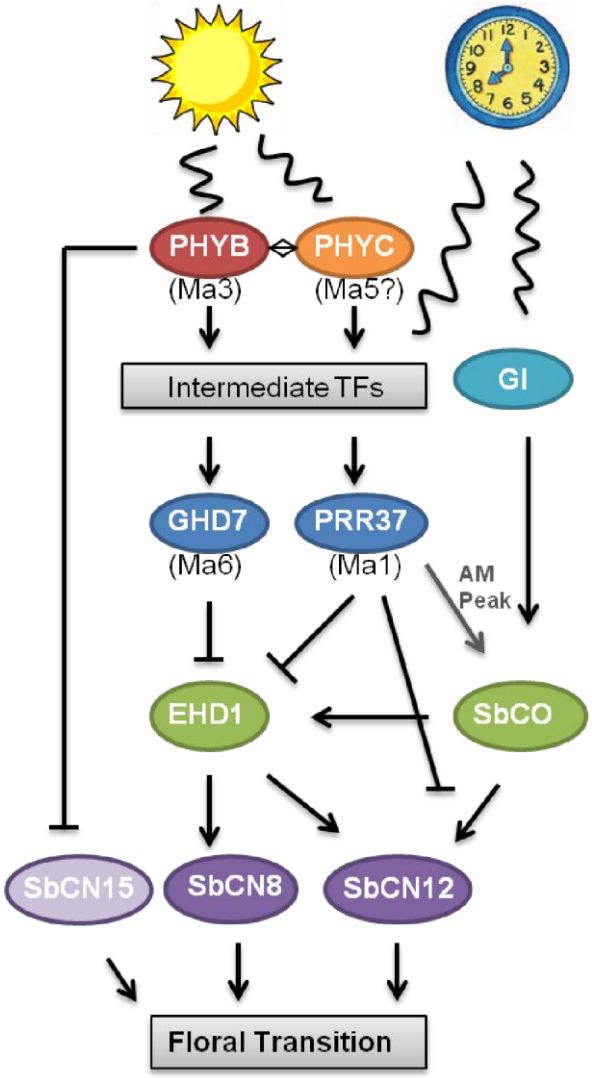


Figure 27. Model of photoperiod flowering time pathway in sorghum (PhyB). Phytochrome B (PhyB) is the photoreceptor signaling light to regulate flowering time in response to photoperiod in sorghum. In LD, PhyB, as a light receptor, may interact with Phytochrome C (PHYC) and up-regulates the expression of *PRR37* and *GHD7*, the central floral repressors, during the evening phase. PRR37 activates dawn peak of *SbCO* mRNA. Together with GHD7, PRR37 represses floral inductors *EHD1*, *SbCN8*, *SbCN12* and *SbCN15* in general, leading to delayed flowering time under long day condition. In SD or 58M, based on external coincidence model, there is no light perception for light-induced expression of *PRR37*and *GHD7*, releasing sorghum for floral transition.

The current study focused on characterizing the light-signaling pathway that regulates *SbPRR37* and *SbGHD7* expression in response to day length. Previous studies showed that sorghum genotypes lacking PhyB (58M, *phyB-1*) flower earlier in LD compared to near isogenic genotypes (100M) expressing *PHYB* demonstrating that light signaling through this photoreceptor is required for photoperiod sensitive variation in flowering time [16]. The current study showed that *PHYB* (*Ma3*) is epistatic to genes encoding the floral repressors SbPRR37 and SbGHD7 and that PhyB is required for photoperiod-regulated expression of these genes. Moreover, 58M, a genotype lacking functional PhyB, showed attenuated expression of *SbPRR37* and *SbGHD7* during the evening of LD compared to 100M (PhyB). In SD, expression of the floral repressors was similar in 58M and 100M. Taken together, these results indicate that in sorghum PhyB is required for light signaling during the evening of LD that results in elevated expression of *SbPRR37* and *SbGHD7* during this time of day.

The molecular basis of PhyB induced expression of *SbPRR37* and *SbGHD7* during the evening of long days is unknown but could involve other photoreceptors and intermediary transcription factors such as PIFs [34]. Detailed studies in rice showed that PhyA, PhyB and PhyC modulate flowering time [120]. PhyC in particular plays a role in natural variation of flowering time in pearl millet [125], Arabidopsis [126], and wheat [127]. In Arabidopsis, a long day plant, PhyB destabilizes CO, an action countered by Cry, PhyA and SPA in LD leading to floral induction [41]. In rice, *phyB* mutants flower early in LD and SD conditions similar to sorghum. Interestingly, rice *phyC* mutants flower early only in LD [120]. In addition, in rice, both PhyB and PhyC are required to

induce *GHD7* expression, where PhyB alone causes some repression of *GHD7* mRNA level [128]. This indicates that PhyB regulates floral induction in both LD and SD while PhyC modifies flowering time selectively in LD. The stability of PhyC is reduced in the absence of PhyB in rice and Arabidopsis [121]. PhyB increases PhyC stability, and chromophore is required for PhyC activity [122]. Therefore, in sorghum the requirement for PhyB in photoperiod sensitive flowering time may be due in part because PhyB increases PhyC stability and through formation of PhyB:PhyC heterodimers.

Genetic analysis of the role of *PHYB* in sorghum was examined using a population dominant for *Ma1* (*SbPRR37*) and segregating for alleles of *PHYB* (*Ma3*), *Ma5*, and *SbGHD7* (*Ma6*). The presence of *Ma1* in all progeny of the population caused delayed flowering in LD unless the expression or activity of *Ma1* (and in some genotypes *Ma1* and *Ma6*) was altered by recessive alleles of *Ma3* or *Ma5*. Flowering was delayed an additional ~30 days in LD in progeny that were dominant for both *Ma1* and *Ma6* in *Ma3_Ma5_* backgrounds consistent with the additive repressing action of *Ma1* and *Ma6*. The analysis showed that plants homozygous for null alleles of *PHYB* (*phyB-1*) in *Ma5_* backgrounds had reduced photoperiod sensitivity and flowered earlier in LD compared to plants encoding PhyB. Similarly, progeny homozygous for recessive alleles of *Ma5*, in *Ma3_* backgrounds, showed reduced photoperiod sensitivity and flowered earlier in LD. The results indicated that both *SbPHYB* and *Ma5* were epistatic to *Ma1* and *Ma6*. Progeny recessive for either gene flowered earlier in LD, but showed a range of flowering times indicating that other genes and/or environmental factors affected flowering time in these backgrounds although with reduced response to

photoperiod. Interestingly, *PHYB* and *Ma5* appear to be co-dependent or acting at a similar point in the regulatory pathway because allelic differences at *Ma5* did not affect flowering time significantly in *phyB-1* backgrounds and vice versa. R.07007 (*Ma3ma5*) and 58M ($ma3^{R}Ma5$) show attenuated expression of *SbPRR37* and *SbGHD7* in the evening of LD [15] (and this study) indicating that both *Ma3* (PhyB) and *Ma5* are required for elevated expression of the sorghum floral repressors during the evening of LD. In searching for an explanation for this co-dependence, we found the *Ma5* locus spans several genes known to affect flowering time including *PHYC* and that the sequence of PhyC in R.07007 (*ma5*) contained amino acid changes that could potentially disrupt the function of this protein. The hypothesis that *Ma5* corresponds to *PHYC* is consistent with studies showing that PhyC modifies flowering in an LD specific manner in rice, similar to *Ma5* [120]. If sorghum PhyC is regulated by PhyB in a similar manner, this would explain why *Ma5* (presumptive *PHYC*) activity is not observed in *phyB-1* backgrounds. Experiments designed to test this hypothesis are currently underway.

In Arabidopsis, *CO* expression peaks once per day in the evening and the amplitude of *CO* expression is regulated by blue light/GI-FKF1-ZTL mediated turnover of CDF1, a repressor of *CO* expression [37]. PRR7 also modifies *CO* expression through repression of *CDF1* expression [28]. In sorghum, *SbCO* expression peaks twice each day, at dawn and again in the evening in LD. The peak of *SbCO* expression at dawn is attenuated in SD [15] and in genetic backgrounds lacking SbPRR37 [116]. It is possible that SbPRR37 modulates *SbCO* expression by repressing sorghum orthologs of

CDF1. The peak of *SbCO* expression at dawn in LD was not observed in the sorghum genotype lacking PhyB (58M). Since PhyB is required for elevated *SbPRR37* expression in the evening of LD, and SbPRR37 has been shown to induce elevated expression of *SbCO* at dawn, it is likely that lack of PhyB induced expression of *SbPRR37* during the evenings of LD in 58M explains this observation.

In rice, *Hd3a*, a member of the PEBP gene family, encodes an FT protein that acts as a florigen [48]. In maize, *ZCN8* and possibly *ZCN12* are sources of florigen [93, 94]. Sorghum encodes orthologs of *Hd3a* (*SbCN15*), *ZCN8* (*SbCN8*) and *ZCN12* (*SbCN12). SbCN8* and *SbCN12* expression is regulated by day length and by alleles of *SbPRR37*, *SbGHD7*, and *SbPHYB* in a manner consistent with these genes being sources of florigen in sorghum. In prior studies, *SbCN15* expression was modulated to only a small extent by variation in photoperiod and in mutants of *SbPRR37* and *SbGHD7* that affect flowering time suggesting that this gene was not an important target of photoperiod regulation [15, 113]. In contrast, in the current study, expression of *SbCN15* was found to be ~60-fold higher in leaves of 58M (*phyB-1*) compared to 100M (*PHYB*) in both LD and SD. If *SbCN15* functions as a source of florigen as in rice, then photoperiod independent repression of *SbCN15* expression by PhyB suggests that this gene may be responsible for early flowering induced by shading [12] or GA pathway. 58M plants exhibit shade avoidance responses including longer leaf blades and sheaths, fewer tillers, narrower leaf blades, less leaf area, and more rapid stem elongation (Fig 28)[12]. In Arabidopsis, light signaling through PhyB represses shade avoidance

responses, and PhyB deficient mutants have elongated stems and an early flowering phenotype associated with “constitutive shade avoidance” [81].

The gibberellin (GA) pathway is one of the four major pathways controlling flowering time in Arabidopsis and phytochromes mediate light-induced changes in GA metabolism [19]. In sorghum, GA levels exhibited diurnal rhythms with peaks at noon or later in 12-hour photoperiod in 100M. PhyB controls GA levels by altering the timing of GA peaks. In 12-hour photoperiods, 58M exhibits a GA1 peak in the morning, which is different from afternoon peak for 100M. In addition, the rhythm of GA level correlates to floral initiation. The presence of a morning peak of GA1 in *phyB* recessive plants or under 10-hour photoperiod conditions is associated with promotion of the floral transition, while the exhibition of afternoon peak through *PHYB* dominant or 18-hour photoperiod conditions is associated with delayed flowering time [12]. The results suggest GAs play a role in flowering time. Another experiment showed that GA synthesis inhibitor treated plants had delayed floral initiation, which confirms GAs’ function in flowering time modulation in sorghum [12].

A model for light and gibberellin control of cell elongation in Arabidopsis was constructed, in which PIF3 and PIF4 are key integration factors [79, 80, 129]. PhyB destabilized PIF3/PIF4 in the light and DELLAs, binding to the DNA-recognition domain of PIFs, inhibit their transcriptional activity. The negative regulation of PIFs is released by GAs through destabilizing DELLA, leading to free PIFs in the nucleus. Since PIFs are also involved in the flowering time pathway, PhyB may regulate floral

initiation through the GA pathway as well: (1) through altering GA levels and rhythms; (2) through PIFs/DELLA pathway.

Information on photoperiod regulated flowering time in sorghum described in this paper will hopefully facilitate analysis of flowering time variation caused by shading and other environmental factors in subsequent studies.

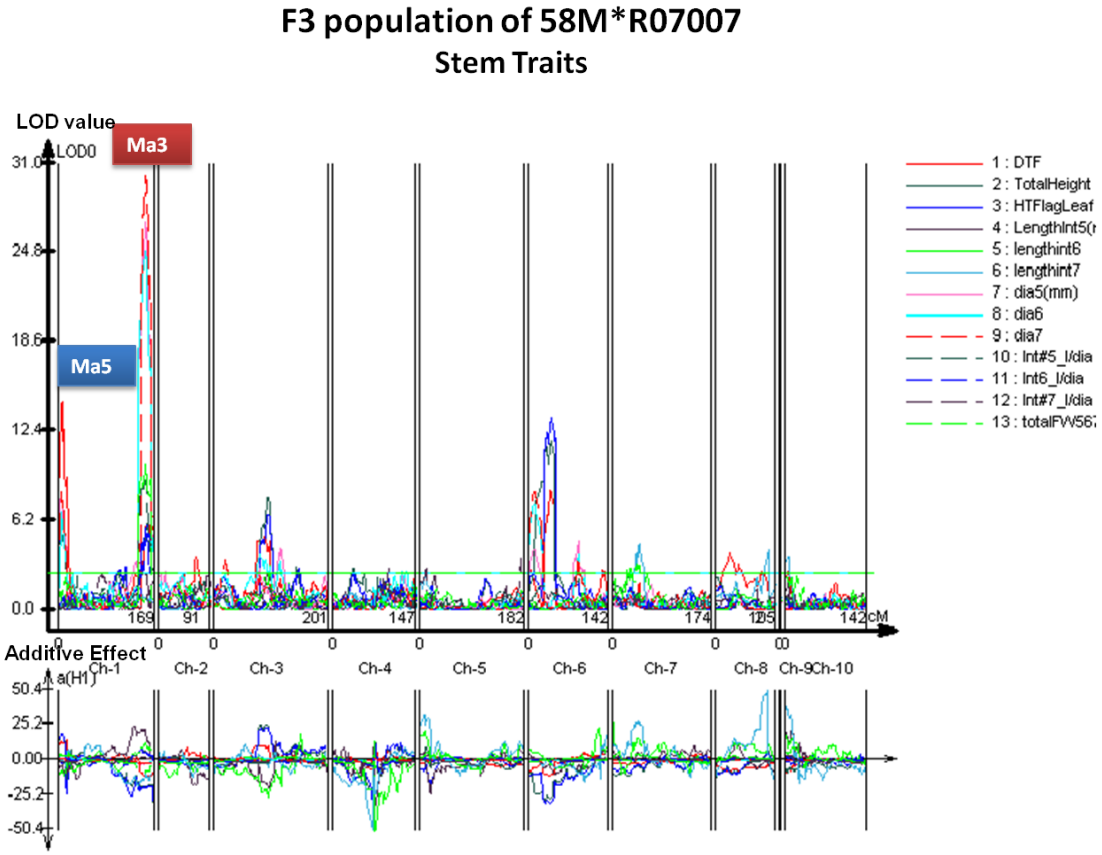


Figure 28. Stem trait QTLs in 58MxR.07007 F3 population under LD condition. Candidate genes in Chromosome 1 (Ch-1) were labeled above each QTL. The right panel indicates stem traits included in QTL analysis: total height (2), height towards flag leaf (3), length of internodes 5 (4), length of internodes 6 (5), length of internodes 7 (6), diameter of internodes 5 (7), diameter of internodes 6 (8), diameter of internodes 7 (9), the ratio of length to diameter for internodes 5 (10), the ratio of length to diameter for internodes 6 (11), the ratio of length to diameter for internodes 7 (12), and total fresh weight of internodes 5, 6, and 7 (13).

Methods

Phenotyping flowering time of sorghum lines

The maturity loci and flowering dates of all sorghum lines used in this study are listed in Table 1. To demonstrate the difference in flowering time between different genotypes and day-length, both 100M and 58M were planted in Metro-Mix 200 (Sunshine MVP; Sun Gro Horticulture) and grown in a greenhouse in LD (14h light/10h dark) and SD (10h light/14h dark) conditions. Days to mid-anthesis were recorded and photos were taken.

Sequencing of *PHYB* alleles

To identify coding alleles in the *PHYB* gene, the full-length genomic sorghum *PHYB* gene from historical sorghum cultivars were amplified as three overlapping segments by PCR (Phusion High-Fidelity DNA polymerase, New England BioLabs, Inc). The amplified PCR products were cleaned and concentrated (QIAquick PCR Purification kit, QIAGEN). PCR products were separated by electrophoresis on 1% agarose gels. Specific PCR products were excised and purified (QIAquick Gel Extraction Kit, QIAGEN). The purified PCR products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and the Applied Biosystems 3130xl Genetic Analyzer. All primers used for sequencing were designed using PrimerQuest^{SM} software (Integrated DNA Technologies, Inc) (Table 8). Sequencher v4.8 (Gene Codes) was used for sequence assembly and alignment with the BTx623 whole genome sequence of *Sorghum bicolor* (version 1.4) downloaded from Phytozome v8.0 (http://www.phytozome.net/).

Table 8. Primer sequences used for *PHYB* alleles amplification and sequencing.

Segment	Primer Sequence
Segment 1 Amplification (N terminal, 2190bp)	F: ATAGCCCACTTCAGCTTTCTCCCA R: TTCTCCCATGGTAGGCTTCTGCTT
Segment 1 Sequencing	ATAGCCCACTTCAGCTTTCTCCCA AGCTTCGACTACTCCCAGTC ACACAGTCGTGGAGCATGTT CCCCCGCCCGCGCGGGACGT CTCGGAGGACGGCGTGGGCG CATGACACGGTCGTAACCCG TTCTCTGACAGCTGATGCGC TTCTCCCATGGTAGGCTTCTGCTT
Segment 2 Amplification (middle, 3693bp)	F: TTTCTCATGCAGGCATTTGGGCTG R: AGGGATTTGAGTTTCCGTGCAAGC
Segment 2 Sequencing	TTTCTCATGCAGGCATTTGGGCTG TTTTGGTTTCGGTCACACAC CATCATCCTTATCCTCAGGG GCTACTCTCACGAGCTTTAA CCGGTGCTTATAACTAGCAG TCGAGAGGTGAAGTTGTTGG GATGATCAGAGGCAGTTCCT TAAGGCTTGCCGCTGATACC AAAAAAGTGCATAGTGGGGG GCACTCCGCAATTCTCATAT CTCCTGTGACTAGCTAACAC GGCCCTTGAGTCGACAAAAA CTTGGACATCTGTTTCTCAC GTACCCAGTGCTGAGAGCA TCAGTTCACAAGGCTATTCC AGGGATTTGAGTTTCCGTGCAAGC
Segment 3 Amplification (C terminal, 4358bp)	F: CAAGCGCTGAAATACAGCAAGCCT R: TTAGCCAGCTTACAGCCCACCATA

Table 8. continued

Segment	Primer Sequence
Segment 3 Sequencing	CAAGCGCTGAAATACAGCAAGCCT TCAAGGCTACTCCAGGCTCAAGTT ATTGGCTGAGAGAGAAGCAGTCCA AAAAAAGTGCATAGTGGGGG TCAGTTCACAAGGCTATTCC GGTTCATTGGTCCAGTGTTT TTACTTCAACCTTTGGCTGG CAGTATCAGAAGGCAACAGT AGAGAGGAAAAGACCTCTGC ACTGACCCCACCTAACTTCT GGCTGACAACAGCATTCATA CTTGCTTGGGACTAAAAGGC CCAATGTGTGAGCTCAACCA GGCTTTAACGTCGACTTTTG TCAGGGACCTAAGTTACCTA TTAGCCAGCTTACAGCCCACCATA

QTL analysis of *PHYB* action

The sorghum cultivar 58M (*Ma1Ma2ma3^{R}Ma4Ma5ma6*) was crossed to R.07007 (*Ma1ma2Ma3Ma4ma5Ma6*). F1 generation plants were self-pollinated to produce F2 populations from which subsequent F3 populations were derived. F2 and F3 populations were planted in the greenhouse and grown under long day conditions (14h light/10h dark). Days to mid-anthesis of the F2 and F3 populations were recorded. For genotyping, genomic DNA of 86 F2 individuals and 132 F3 individuals was extracted from leaf tissue using the FastDNA Spin Kit (MP Biomedicals). Template for sequencing on an Illumina GAIIx sequencer was generated following the standard Digital Genotyping (DG) protocol [119]. Genetic markers of all individuals from both

populations were identified. The genetic map was constructed using the Haldane mapping function in MAPMAKER v3.0 with 285 markers from the F2 population and 653 markers from the F3 population. QTLs were mapped on the genetic map using the Composite Interval Mapping (CIM)) function in WinQTL Cartographer v2.5 [130]. Significant LOD threshold for QTL detection was calculated based on experiment specific permutations with 1000 permutations and α=0.05 [131].

Gene expression assay

Sorghum genotypes 100M and 58M were planted and grown in a greenhouse under long day conditions (14h light/10h dark) for 32 days and then transferred to growth chambers under either LD (14h light/10h dark) or SD (10h light/14h dark) conditions for seven days for entrainment prior to collection of tissue. In the growth chamber, daytime (lights on) temperature was set at 30°C with a light intensity of ~300 $\mu mol \cdot s^{-1} \cdot m^{-2}$ and night (lights off) temperature was set at 23°C. Relative humidity was set at ~50% throughout the course of the experiment. At day 39, leaf segments from the top three leaves from three individual plants of each genotype and photoperiod condition were collected every 3 hours through one 24h light-dark cycle and 48h of continuous light. The leaf tissues at each time point were subjected to total RNA extraction using TRI Reagent (MRC) with the protocol for high polysaccharides samples. RNA was further purified using the RNeasy Mini kit (QIAGEN), including removal of DNA contamination by on-column DNase I digestion before reverse transcription. RNA integrity was examined on 1% MOPS gel. First-strand cDNA synthesis was performed using the SuperScript^{TM} III First-Strand Synthesis System (Invitrogen) with oligo dT and

random hexamer primer mix. After first-strand cDNA synthesis, the reactions were diluted to 10ng/µl of the initial total RNA. Gene-specific qPCR reactions were carried out using Power SYBR Green PCR Master Mix (Applied Biosystems). 18S rRNA was selected as the internal control reference and the reactions were performed using the TaqMan Universal PCR Master Mix Protocol with rRNA Probe (VIC^{TM} Probe) and rRNA Forward/Reverse Primer. All reactions were run on the 7900HT Fast Real-Time PCR System with SDS v2.3 software. The specificity of each gene specific primer set was validated by melting temperature curve analysis. Amplification efficiency of each primer sets was determined by the serial dilution method [132] (Table 9). Relative expression was determined by the comparative cycle threshold (ΔΔCt) method [132] with calibration from most highly expressed samples. The calculated primer efficiencies were used to adjust data for relative quantification by the efficiency correction method [133]. Each relative expression value was derived from an average of three technical replicates and three biological replicates.

Table 9. Primer sequences and amplification efficiency for qRT-PCR *(PHYB)*.

Gene	Locus ID in Sorghum*	Forward Primer	Reverse Primer
PRR37	Sb06g014570	AACAGGACGGAACTGGAGAGAGAT	CCAAAGCAATCTTGCTAGAGGCGA
GHD7	Sb06g000570	TCAGGACAACGATGACCACCAAGA	ATCAACCTCAAAGGTGAGCCCGTT
CO	Sb10g010050	TAGTCCCAGACAACATGGCAACGA	AGGTCAAGTGGAGTGGCATCTGAA
EHD1	Sb01g019980	CGTCAGGGAAGCAATGTCCTTCAT	CTTCAGTTGGAAAGCACACATCGC
SbCN8	Sb09g025760	AACTGTCAAAGGGAAGGTGGATCG	GACTAAGCTCTCAACCCTTCAAGTC
SbCN12	Sb03g034580	TGCATGCATGAATATCGTCGTCT	CCCGGGTAGTACATATAAGGTGGT
SbCN15	Sb10g003940	GCTAGCTTATCCCGCATATTACCC	CCACCCAAACTGCATCCACTCTTGAA
GI	Sb03g003650	ATGCACCCGCTTCCTAGTCATCTT	TTCAGGGCTGTCATGGTTCCTCAT
TOC1	Sb04g026190	GAGTGCAGATGATTACTGCTCACTTTG	TGCTGCCTTGTTGCCAGTAGAAGA
LHY	Sb07g003870	GGCCTGCCTCTACCATGAAGTTTA	GCACTGCATTGCAAGGTTTGAAGTCC
	Amplification Efficiency in 100M		**Amplification Efficiency in 58M**
PRR37	1.16		1.16
GHD7	1.19		1.18
CO	1.25		1.18
EHD1	1.19		1.20
SbCN8	1.12		1.19
SbCN12	1.22		1.16
SbCN15	1.11		1.13
GI	1.15		1.22
TOC1	1.18		1.11
LHY	1.15		1.14

***** Gene Locus IDs in Sorghum, Maize and Rice are all based on Phytozome v8.0 (http://www.phytozome.net/).

CHAPTER III

CONSTANS IS A PHOTOPERIOD REGULATED ACTIVATOR OF FLOWERING TIME IN SORGHUM [*]

Background

Optimal regulation of the timing of floral transition is critically important for reproductive success and crop yield. The C4 grass *Sorghum bicolor* is widely adapted and grown as an annual crop from 0 to >40 degrees N/S latitude. Sorghum crops have been selected for a range of flowering times depending on growing location and used as a source of grain, sugar, forage, or biomass [2]. Grain sorghum is generally selected for early flowering (60-80 days) to enhance grain yield stability by avoiding drought, adverse temperatures, and insect pressure during the reproductive phase. In contrast, energy sorghum hybrids are designed with high photoperiod sensitivity in order to delay flowering and extend the duration of vegetative growth, resulting in more than 2-fold increases in biomass production [2, 11]. The stage of plant development, signals from photoperiod, temperature, gibberellins and other factors are integrated to regulate flowering time in sorghum [12].

The genetic architectures of photoperiod-responsive flowering-time regulatory pathways have been characterized [21, 22, 39, 41, 43, 46, 49, 50, 96, 134]. For Arabidopsis, the floral transition is promoted in long-days (LD) by coincidence of light

[*] Part of the data reported in this chapter is reprinted with permission from "CONSTANS is a photoperiod regulated activator of flowering in sorghum" by Yang S, Weers B, Morishige D, Mullet J, 2014. BMC Plant Biology,14:148, Copyright [2014] by Yang S, Weers B, Morishige D, Mullet J.

signaling and circadian clock output, thus allowing the plant to sense and respond to seasonal changes in photoperiod. Clock output to the flowering pathway is mediated in part by GI. GI is regulated by the central clock oscillator comprised of TOC1, CCA1 and LHY. In long days, GI activates *CO* expression in conjunction with FKF1 by inducing degradation of CDF1 repressors of *CONSTANS* transcription. CO accumulates in LD due to stabilization mediated by cryptochromes (Cry1/2), phytochrome A (PhyA) and SPA1 that counteract degradation of CO mediated by phytochrome B (PhyB): COP1 [41, 43]. Increased CO protein levels in long days leads to the activation of *FT* expression and production of florigen that moves from leaves to shoot apical meristems (SAM) where it binds to FD and induces floral transition.

The GI-CO-FT regulatory pathway identified in Arabidopsis, a long day (LD) plant, is also present in rice, a short day (SD) plant [46]. When rice is exposed to inductive SD, *Hd1*, the ortholog of *CO*, activates expression of the *FT*-like gene *Hd3a*, one of two sources of florigen in rice. In non-inductive LD, Hd1 functions as a repressor of *Hd3a* and flowering [47]. Thus, photoperiod sensitivity in rice depends in part on differences in the activity of CO (Hd1) in long days and short days. Two modulators of flowering time unique to grasses were identified in rice: *EHD1* [54] and *GHD7* [62, 117]. EHD1 activates the expression of *Hd3a* and *RFT1*, a source of florigen in long days. GHD7 represses flowering by down-regulating expression of *EHD1* and *Hd3a* in LD in rice [53] and *SbEHD1* and *SbCN8* in sorghum [113].

The effect of photoperiod on flowering time varies extensively among and within grass species. Barley and wheat are LD plants, while rice and sorghum are SD plants.

Maize is classified as a day-neutral plant that flowers after a set number of degree days; however tropical maize is a photoperiod sensitive short day plant [96]. Sorghum is a short day plant, although grain sorghum is usually photoperiod insensitive (day neutral), whereas forage and energy sorghum genotypes are photoperiod sensitive [2]. More than 40 QTL for flowering time have been identified in sorghum [118]. The *Ma1-Ma4* loci were discovered while breeding for early flowering photoperiod insensitive grain sorghum in the U.S. (1920-1960) [13]. *Ma1* corresponds to *PSEUDORESPONSE REGULATOR PROTEIN 37* (*SbPRR37*), a repressor of flowering in LD [15]. *Ma3* encodes *PHYTOCHROME B* (PHYB), a red-light photoreceptor that plays an important role in photoperiod sensing and repression of flowering [16, 114, 115]. *Ma6* encodes *Sb*Ghd7, a repressor of *SbEHD1* expression and flowering in long days [113]. *Ma2, Ma4*, and *Ma5* are flowering time loci that enhance photoperiod sensitivity in sorghum [13, 18].

CONSTANS (CO) was initially identified as a transcriptional activator of *FT* and flowering in Arabidopsis [82]. CO belongs to a family of transcription factors unique to plants that contain one or two N-terminal zinc finger B-box domains and a C-terminal CCT domain. Two conserved cysteine and histidine amino acids in the Zn finger domain are essential for CO activity [43]. Arabidopsis mutants with amino acid substitutions at these positions have late flowering phenotypes. Extensive gene duplication events have occurred in this gene family, resulting in ~17 CO family members in Arabidopsis, ~16 in rice and ~9 in barley [89, 90]. The ortholog of CONSTANS in rice, Hd1, plays a key role in photoperiod regulation of flowering, by activating flowering in SD and repressing

flowering in LD [47]. Alleles of *Hd1* account for ~44% of the variation in flowering time observed in cultivated rice [61]. *Hd1* transcript and protein levels are similar in LD and SD, consistent with the finding that Hd1 activity is modulated post-transcriptionally by PhyB [97] and PRR37 [98, 99].

Results

Identification of flowering time QTL

Flowering time QTL were mapped in a RIL population derived from a cross of BTx642 and Tx7000, genotypes used in U.S. grain sorghum breeding programs as sources of drought tolerance [4]. A RIL population (n=90) derived from these genotypes was previously used to map QTL for flowering time and the stay-green drought tolerance trait using a genetic map based on RFLP markers [5]. The genomes of BTx642 and Tx7000 were recently sequenced and analyzed for variation in DNA polymorphisms that distinguish these genotypes [135]. Digital Genotyping was used to create a high-resolution genetic map aligned to the genome sequence based on this RIL population [119, 135]. Digital Genotyping identified 1,462 SNP markers segregating in the RIL population and data on recombination frequency was used to create a 1139cM genetic map spanning the 10 sorghum chromosomes [135]. Flowering time QTL were mapped in this population by phenotyping the RIL population for days to half pollen shed in greenhouses in 14h long days (LD), 10h short days (SD), and under field conditions where day length increases following plant emergence in mid-April from 12.6h to 14.3h in July. Tx7000 flowered in 73 days and BTx642 flowered approximately 4 days later under field conditions in College Station, Texas. When grown in a

greenhouse at constant 14h day lengths (LD) during the summer, Tx7000 flowered in 84 days and BTx642 flowered ~19 days later (Fig 29A). When Tx7000 and BTx642 were grown in a greenhouse under 10h day lengths (SD) during the winter, Tx7000 flowered in 54 days whereas BTx642 flowered ~11days later.

The BTx642/Tx7000 RIL population was grown and assayed for days to flowering under field conditions in 2008-2010, in LD greenhouses in 2009 and 2010, and in a SD greenhouse during the winter of 2011. WinQTL Cartographer was used to identify flowering time QTL using flowering time data collected from each location/year and the genetic map generated by Evans et al [135]. Three QTL for flowering time were observed in every environment and two additional QTL were identified in only one environment (Table 10).

Three flowering time QTL were identified when RILs were screened in LD greenhouse conditions (Fig.29B) (Fig.30). The QTL on SBI-01 (19.2-22.0Mbp) explained 12.3% of the phenotypic variance for flowering time in this environment. *SbEHD1*, an activator of flowering in grasses located on SBI-01 (Sb01g019980, 21921315-21925396) was found in a one LOD interval spanning this QTL. *SbEHD1* was previously identified as a floral activator in sorghum based on sequence similarity to rice *EHD1* and observed changes of *SbEHD1* expression in LD compared to SD, consistent with this function [15]. There were no amino acid differences between the SbEhd1 protein sequences from BTx623 and Tx7000. However, comparison of SbEhd1 from BTx642 and Tx7000 revealed two amino acid substitutions, Asp144Asn and Thr157Ile (Table 11). The differences in Ehd1 protein sequences occur in a GARP

domain that is highly conserved among *OsEHD1*, *SbEHD1* and *ARABIDOPSIS RESPONSE REGULATOR 1/2* (*ARR1/2*). The *SbEHD1* allele in BTx642 (tentatively designated *Sbehd1-2*) delays flowering in LD and SD relative to Tx7000 (*SbEHD1-1*), consistent with the hypothesis that the amino acid changes in *Sbehd1-2* reduce the activity of this floral activator and explain why a flowering time QTL was detected in this region of SBI-01.

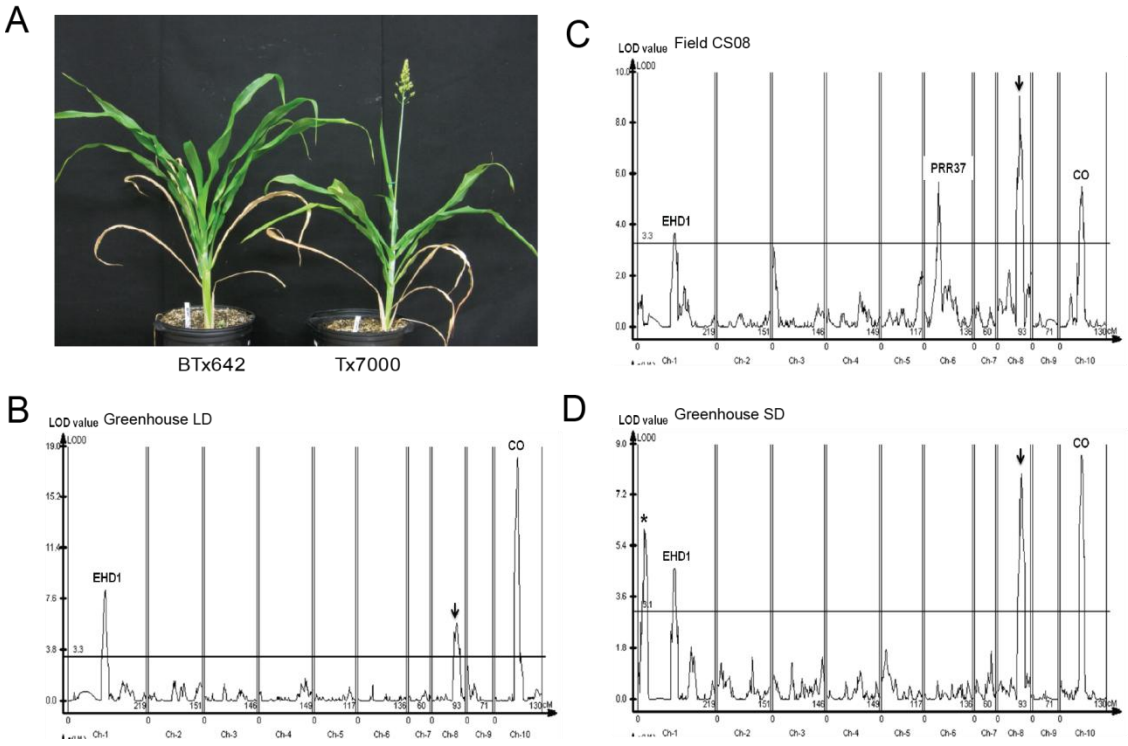


Figure 29. Genetic basis of flowering time variation in the BTx642/Tx7000 RIL population. A. Flowering time phenotypes of BTx642 and Tx7000 in LD. (Days to flowering for BTx642 and Tx7000 are 103 and 84.) Flowering time QTL identified when RIL population were grown in a LD greenhouse (B), under field conditions in 2008 (C) and in a SD greenhouse (D). Permutation tests were carried out to identify 95% confidence thresholds and significant threshold of LOD score is presented as a horizontal line. Candidate genes associated with main affect QTL are noted above peaks.

Table 10. Parameters of flowering time QTLs in BTx642/Tx7000 RILs population.

Greenhouse LD (14h)							
QTL	Candidate Gene	Chr #	Position (cM) [a]	LOD Score	Peak Coordinate [b]	Additive Effect [c]	R^{2} [d]
1	*EHD1*	Chr_01	102.7	8.31	22012456-22012527	-6.25	0.12
2	ND[e]	Chr_08	67.9	5.82	50255989-50256060	-5.02	0.08
3	*CO*	Chr_10	61.7	18.43	13696999-13697070	-12.69	0.40
Field LD condition CS08							
QTL	Candidate Gene	Chr #	Position (cM) [a]	LOD Score	Peak Coordinate	Additive Effect [c]	R^{2} [d]
1	*EHD1*	Chr_01	102.7	3.74	22012456-22012527	-1.09	0.09
2	*PRR37*	Chr_06	42.0	5.71	40201054-40201125	1.53	0.15
3	ND	Chr_08	60.2	9.09	49290307-49290378	-1.80	0.26
4	*CO*	Chr_10	59.7	4.11	10080053-10080126	-1.50	0.16
Greenhouse SD (10h)							
QTL	Candidate Gene	Chr #	Position (cM)	LOD Score	Peak Coordinate	Additive Effect	R^{2}
1	ND	Chr_01	16.3	6.00	7208344-7208415	2.18	0.09
2	*EHD1*	Chr_01	102.7	4.92	22012456-22012527	-1.80	0.07
3	ND	Chr_08	65.1	7.96	49797259-49797330	-2.46	0.14
4	*CO*	Chr_10	59.7	8.70	10080053-10080126	-3.30	0.17

a. Position of likelihood peak (highest LOD score).

b. Peak Coordinate: physical coordinate of the likelihood peak.

c. Additive Effect: A positive value means the delay of flowering time due to Tx7000 allele. A negative value means the delay of flowering time due to BTx642 allele.

d. R^{2} (coefficient of determination): percentage of phenotypic variance explained by the QTL.

e. ND: Candidate gene is not determined.

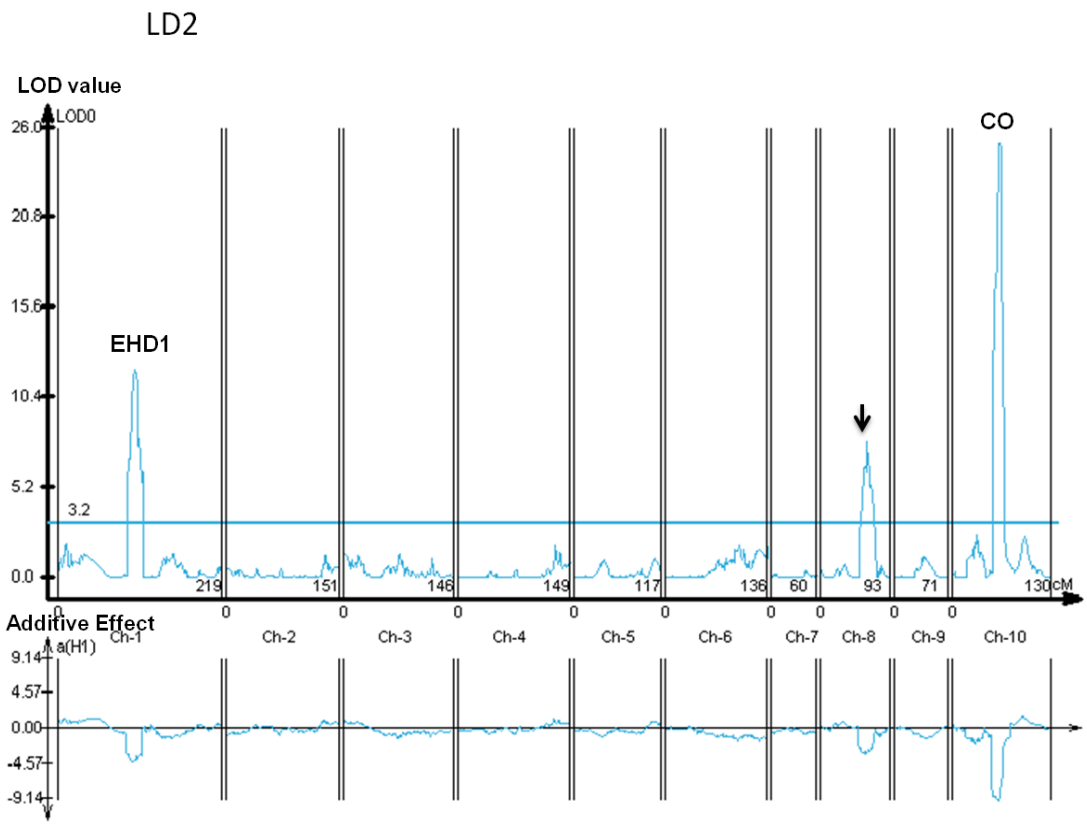


Figure 30. Flowering time QTLs in BTx642/Tx7000 RILs population for days to anthesis under LD greenhouse condition (2nd replicate). Significant threshold of LOD score presents as a horizontal line. Candidate gene was labeled above each QTL. Positive value of additive effect represents allele from Tx7000 contributes to delayed flowering time.

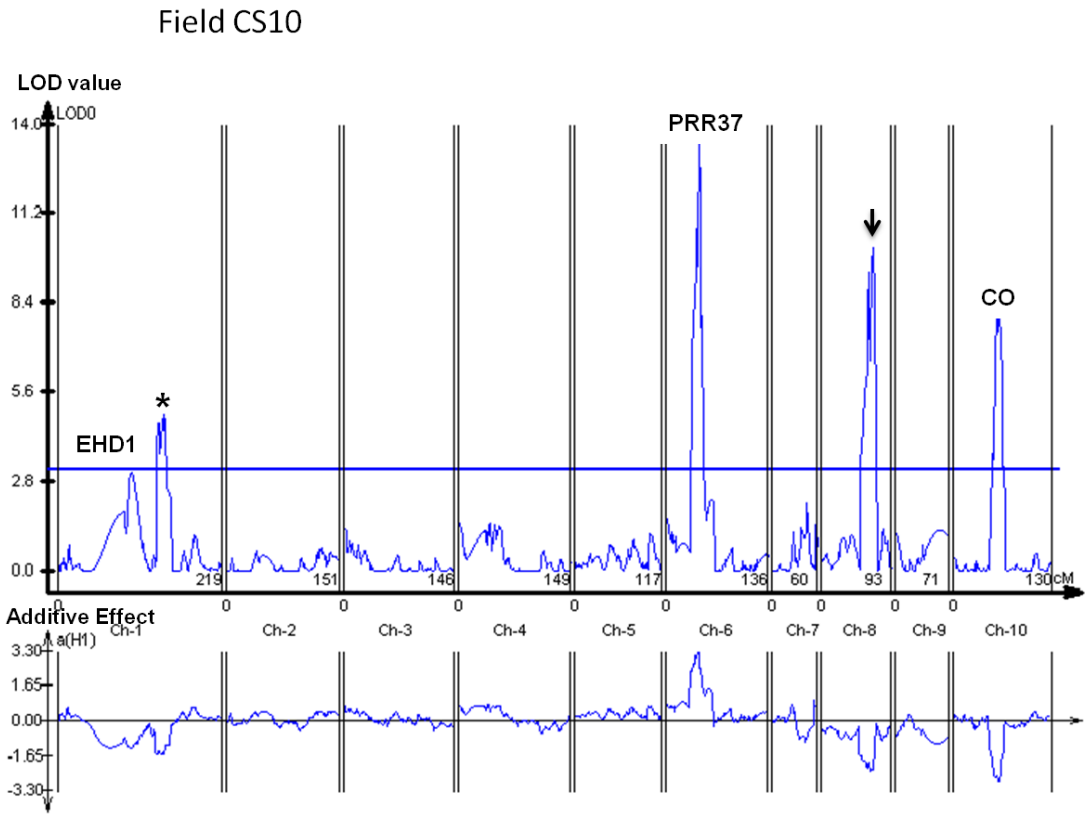


Figure 31. Flowering time QTLs in BTx642/Tx7000 RILs population for days to anthesis under LD field condition (College Station, 2010). Significant threshold of LOD score presents as a horizontal line. Candidate gene was labeled above each QTL. Positive value of additive effect represents allele from Tx7000 contributes to delayed flowering time.

Table 11. Characterization of *SbEHD1* alleles.

SNP #	**1**	**2**	**3**	**4**	**5**
Location (SBI-01)	21921341	21922009	21922977	21923172	21923861
Nucleotide Variation	G>A	C>G	G>A	C>T	T>C
Protein Modification	N/A	Intron	Intron	Intron	Intron
EHD1 Domain					
SbEhd1-1 **(Tx7000, BTx623)**	-	-	-	-	-
Sbehd1-2 **(BTx642)**	+	+	+	+	+
	6	**7**	**8**	**9**	
Location (SBI-01)	21923883	21924404	21924444	21925032	
Nucleotide Variation	C>T	G>A	C>T	G>A	
Protein Modification	Intron	Asp144Asn	Thr157Ile	Intron	
EHD1 Domain		GARP	GARP		
SbEhd1-1 **(Tx7000, BTx623)**	-	-	-	-	
Sbehd1-2 **(BTx642)**	+	+	+	+	

A flowering time QTL located on SBI-10 (10.1-13.7Mbp) was observed in all environments and spanned a region that encodes a homolog of *CONSTANS* and *Hd1* (Sb10g010050, 12275128-12276617), an important regulator of flowering in Arabidopsis and rice, respectively (Fig 29B-D) (Fig 31). The QTL spanning the sorghum homolog of *CONSTANS* explained ~40% of the variance in flowering time in LD greenhouses, and 16-17% when plants were grown in the field or SD greenhouses.

A flowering time QTL located on SBI-08 (48.1-50.3Mbp) was observed in LD, SD and under field conditions. This QTL explained 8-14% of the phenotypic variance in LD and SD and 18-22% of the variance in field environments. Additional analysis will be required to identify the gene corresponding to this flowering time QTL. A QTL located at the end of SBI-01(~7.2Mbp) was observed only when the BTx642/Tx7000 RIL population was grown in the SD greenhouse (Fig 29D). A QTL on SBI-06 (~40.2Mbp) explaining ~15% variance of flowering time was identified when RILs were grown in the field (Fig 29C). *Ma1/PRR37*, a repressor of flowering in LD, was located in the flowering time QTL on SBI-06. Sequence analysis showed that BTx642 contains a truncated version of PRR37 (*Sbprr37-1*) and that Tx7000 encodes a full-length version of PRR37 that contains a Lys162Asn change in the pseudo-response regulator domain resulting in the allele designated *Sbprr37-2* [15]. Genotypes containing *Sbprr37-2* flowered later than genotypes encoding *Sbprr37-1* (null) under field conditions, indicating that *Sbprr37-2* is an active but weak allele of *SbPRR37*. This conclusion is consistent with analysis of a flowering time QTL aligned to PRR37 identified in a RIL population derived from crossing the genotypes Rio and BTx623 [136]. Sequence analysis of *SbPRR37* alleles showed that Rio encodes *Sbprr37-2* and BTx623 contains *Sbprr37-3*, a null allele [15]. The *Ma1* allele from Rio (*Sbprr37-2*) delayed flowering relative to BTx623 in field conditions in College Station in a manner similar to the delayed flowering attributed to the same allele in Tx7000 compared to BTx642, which encodes a different null allele of *SbPRR37* (*Sbprr37-1*).

Identification of sorghum *CONSTANS*

The hypothesis that the flowering time QTL on SBI-10 was caused by alleles of *CONSTANS/Hd1* was investigated further. The amino acid sequence of rice Hd1 was used to identify homologs in sorghum, maize, barley and Arabidopsis using data from Phytozome v9.1 [http://www.phytozome.net/]. Sb10g010050 (score = 71.9), GRMZM2G405368_T01 (score = 80.7), AF490468 (score = 63.2) and AT5G15850 (score = 40.5) had the highest similarity to Hd1 in each species (Fig 32). GRMZM2G405368_T01 and AF490468 were previously identified as the maize *CONSTANS*-like gene, *conz1* [107] and barley *CONSTANS*-like gene, *HvCO1*[99], respectively, while AT5G15850 encodes *CO* in Arabidopsis [82]. The phylogenetic tree was constructed to investigate the evolutionary relationship (Fig 33).Multiple sequence alignment of the CO homologs showed that Sb10g010050 has all of the characteristic protein domains found in CONSTANS-like gene families (Fig 34), including an N-terminal B-box1 (residues 35-76), B-box2 (residues 77-120) domains and a C-terminal CCT domain (residues 339-381).

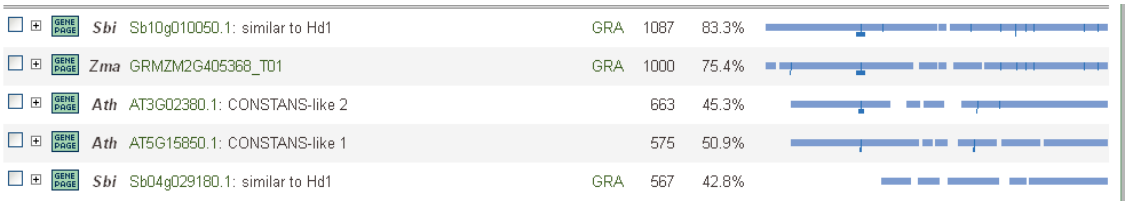
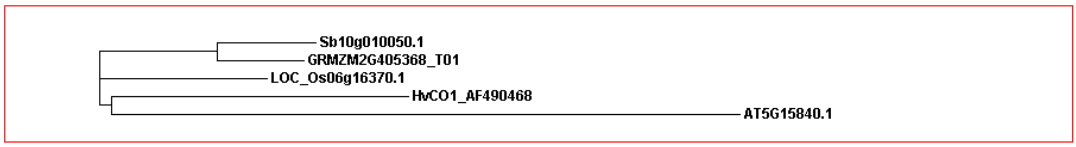


Figure 32. The amino acid sequence similarity of rice Hd1 homologs in sorghum, maize and Arabidopsis from Phytozome v9.1.

A



B

SeqA	Name	Length	SeqB	Name	Length	Score
1	AT5G15840.1	373	2	Sb10g010050.1	410	37.53
1	AT5G15840.1	373	3	LOC_Os06g16370.1	395	40.48
1	AT5G15840.1	373	4	GRMZM2G405368_T01	428	38.34
1	AT5G15840.1	373	5	HvCO1_AF490468	383	35.92
2	Sb10g010050.1	410	3	LOC_Os06g16370.1	395	71.9
2	Sb10g010050.1	410	4	GRMZM2G405368_T01	428	80.73
2	Sb10g010050.1	410	5	HvCO1_AF490468	383	60.57
3	LOC_Os06g16370.1	395	4	GRMZM2G405368_T01	428	67.85
3	LOC_Os06g16370.1	395	5	HvCO1_AF490468	383	63.19
4	GRMZM2G405368_T01	428	5	HvCO1_AF490468	383	61.36

Figure 33. Phylogenetic analysis of CONSTANS homologs. (A) Phylogenetic tree of CONSTANS homologs from Sorghum (Sb10g010050, SbCO), Maize (GRMZM2G405368_T01, conz1), Rice (Os06g16370, OsHd1), Barley (AF490468, HvCO1) and Arabidopsis (AT5G15850, AtCO). (B) Pairwise comparison of amino acid sequences of CONSTANS homologs used in phylogenetic tree construction.

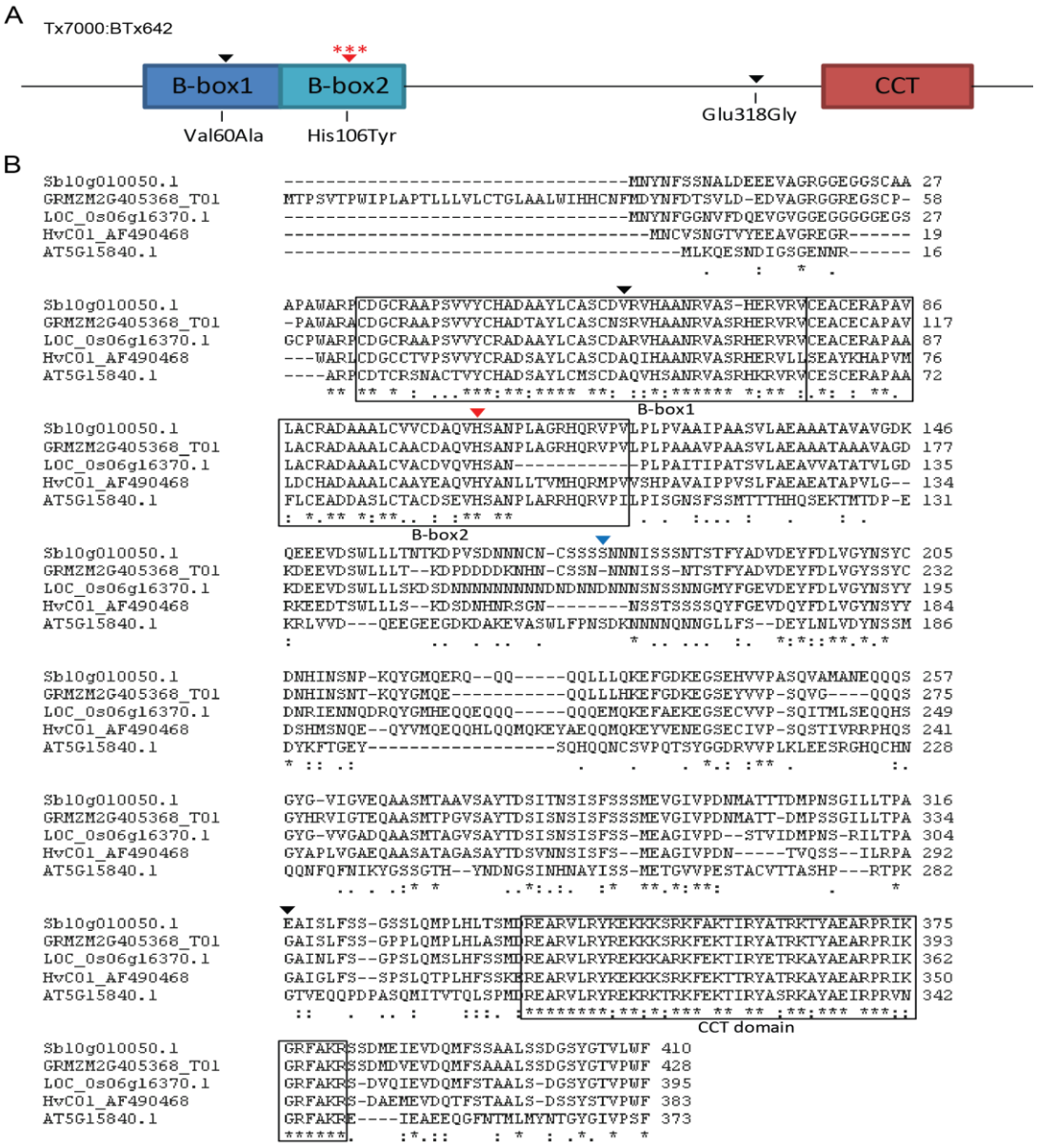


Figure 34. Multiple alignment analysis of CONSTANS homologs. A. Protein structure of SbCO showing domains characteristic of CONSTANS-like gene families: B-box1, B-box2 and CCT domain are boxed. Red asterisks above the His106Tyr mutation indicate that this functional mutation was also identified in rice and Arabidopsis. B. Multiple sequence alignments of CO homologs from sorghum (Sb10g010050, SbCO), maize (GRMZM2G405368_T01, conz1), rice (Os06g16370, OsHd1), barley (AF490468, HvCO1) and Arabidopsis (AT5G15850, AtCO). The sorghum sequence used for alignment was derived from BTx623 (SbCO-1). Protein residues conserved among all 5 species are underscored by asterisks. Amino acid residues underscored by a colon indicate residues of strongly conserved properties, while residues underscored by a period indicate residues with more weakly similar properties. One amino acid substitution distinguishes BTx623 (SbCO-1) and Tx7000 (SbCO-2) (marked with blue arrow). Unique amino acid substitutions that distinguish BTx623 and BTx642 (Sbco-3) are marked with black arrows (tolerant) and a red arrow (intolerant).

The sorghum homolog of *CONSTANS* (Sb10g010050) is located on SBI-10 and rice *Hd1* (Os06g16370) is located on the homeologous rice chromosome 6, suggesting that these genes may be orthologs. The sequences encoding these genes and adjacent sequences in each chromosome were aligned to determine if *SbCO* and *OsHd1* were in a region of gene colinearity. The sorghum sequences flanking Sb10g010050 were downloaded from Phytozome and aligned with sequences from rice chromosome 6 flanking *Hd1* using GEvo (http://genomevolution.org/CoGe/GEvo.pl). Three genes and *Hd1* were aligned and in the same relative order in a 100kbp region in the two chromosomes, consistent with the identification of Sb10g010050 as an ortholog of rice *Hd1* (Fig 35). The similar results were also obtained when comparing sorghum to maize (Fig 36A) and rice to maize (Fig 36B). Therefore, based on sequence similarity and colinearity, Sb10g010050 was designated as an ortholog of rice *Hd1* and a probable ortholog of Arabidopsis *CO* and termed "*SbCO*".

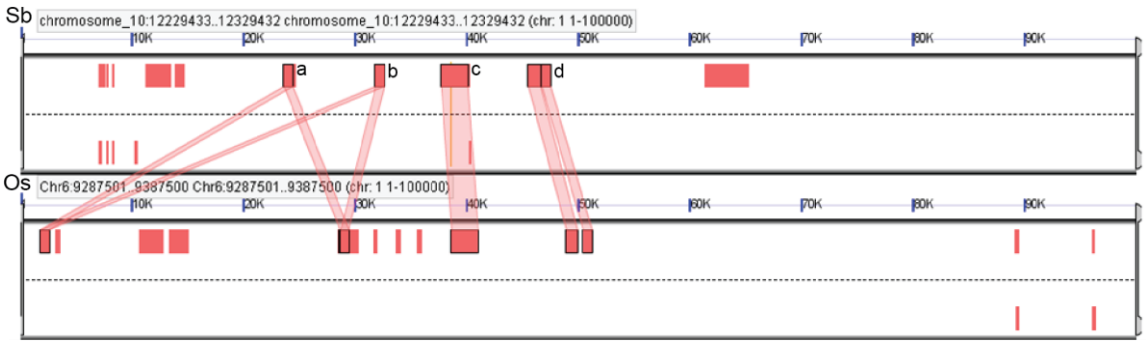


Figure 35. Colinearity of rice *Hd1* and sorghum *CONSTANS*. Reference genome sequences including sorghum *SbCO* (upper panel) and rice *OsHd1* (lower panel) were analyzed for sequences that align (red boxes) Colinear genes within the aligned region are connected by red lines. a-d represent four colinear genes in rice and sorghum (Sb10g010020- Sb10g010050) including *SbCO* (Sb10g010050, d).

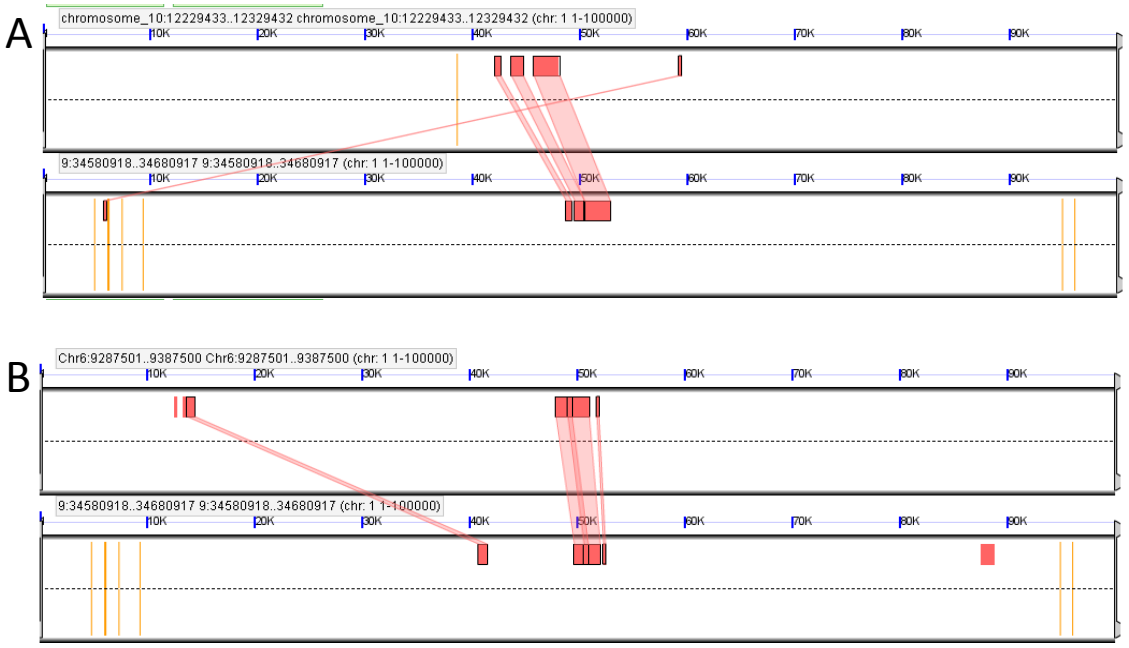


Figure 36. Synteny analysis of *CONSTANS* homolog genes. 100kbp reference genome sequences containing *CO* homolog genes of each species are shown. Aligned genome sequences are labeled as red box. Corresponding aligned sequences by position (collinearity) are connected by red lines. (A) Collinearity of region spanning *SbCO* (upper) and region spanning maize *conz1* (lower). *SbCO* locates ~46K and maize *conz1* locates ~50K. (B) Collinearity of region spanning rice *OsHd1* (upper) and region spanning maize *conz1* (lower). *OsHd1* locates ~50K.

The hypothesis that the flowering time QTL on SBI-10 was caused by different alleles of *SbCO* in BTx642 and Tx7000 was investigated further by comparing the *SbCO* sequences from these genotypes. The comparison revealed one difference in intron sequence and four differences in the coding region, three of which cause changes in amino acid sequence (Table 12). The amino acid change Val60Ala, occurs in B-box1 (Fig 34, black arrow), a conservative change in amino acid sequence that is expected to be tolerated based on SIFT analysis [123]. The amino acid change Glu318Gly occurs

outside the B-boxes and CCT-domain (Fig 34, black arrow) and was also predicted to be tolerated based on SIFT analysis. However, the His106Tyr change in BTx642 located in B-box2 (Fig 34, red arrow) is predicted to disrupt CO function. In the wild type version of *CONSTANS,* His106 is required for zinc coordination and protein activity [43]. The BTx642 allele of CONSTANS was designated *Sbco-3* because the Arabidopsis allele *co-3* has the same His106Try substitution that disrupts function [82]. The wild type alleles of CO in BTx623 and Tx7000 had identical CO protein sequences except for a Ser177Asn substitution in Tx7000 (Fig 34B, blue arrow), a modification that does not affect the B-boxes or the CCT domain, and is predicted by SIFT to have minimal impact on CO function. Based on this analysis, the *CONSTANS* alleles in BTx623 and Tx7000 were designated as *SbCO-1* and *SbCO-2*, respectively, and the allele in BTx642 as *Sbco-3*. BTx642 (*Sbco-3)* flowers later than Tx7000 (*SbCO-2*) in both long and short days, suggesting that inactivation of CO causes a general delay in flowering in sorghum, irrespective of day length.

Table 12. Characterization of *SbCO* alleles from BTx623, Tx7000 and BTx642.

SNP #	1	2	3	4	5	6
Location (SBI-10)	12275306	12275331	12275443	12275657	12276109	12276334
Nucleotide Variation	T>C	T>G	C>T	G>A	C>T	A>G
Protein Modification	Val60Ala	N/A	His106Tyr	Ser177Asn	Intron	Glu318Gly
CONSTANS Domain	B-box1		B-box2			
SIFT Score	tolerant	N/A	intolerant	tolerant	N/A	tolerant
***SbCO-1* (BTx623)**	-	-	-	-	-	-
***SbCO-2* (Tx7000)**	-	-	-	+	-	-
***Sbco-3* (BTx642)**	+	+	+	+	+	+

SbCO alleles modulate expression of genes in the flowering time pathway

The influence of *SbCO* alleles on the expression of other genes in the flowering-time regulatory pathway was analyzed to further understand how *SbCO* affects flowering time. RILs were identified that differ in alleles of *SbCO* but not at the other main loci that affect flowering time. RIL105 and RIL112 are homozygous for BTx642 DNA for the flowering time QTL on SBI-01 (spanning *Sbehd1-2*), SBI-06 (spanning *Sbprr37-1*), and SBI-08. BTx642 encodes a null allele of *Ma1* (*Sbprr37-1*; LG-06), a gene that contributes to photoperiod sensitivity. Tx7000 contains a weak allele of *Ma1* (*Sbprr37-2*) that encodes a full-length protein that inhibits flowering based on QTL analysis [15,

136]. Therefore, RIL105 and RIL112 were selected for expression studies because both contain DNA from BTx642 on SBI-06 from 0-42Mbp, ensuring that these genotypes are null for *Ma1* (*Sbprr37-1*).

When grown in a LD greenhouse, RIL105 (*SbCO-2*) flowered in ~75 days, whereas RIL112 (*Sbco-3*) flowered in 113 days consistent with the hypothesis that SbCO functions as an activator of flowering (Fig 37A). *SbCO* expression in RIL105 (*SbCO-2*) was analyzed using qRT-PCR during a 24h LD cycle followed by 24h of continuous light and temperature (LL). *SbCO* expression decreased at dawn and remained at low levels during most of the day and then increased to a peak in the evening, approximately 15h after dawn, followed by a decline and second smaller peak at dawn (Fig 37B). The peaks of *SbCO* expression in the evening and near dawn were previously observed in sorghum [15] and for *conz1* in maize [107]. The increase in *SbCO* expression in the evening also occurred in continuous light (LL), consistent with prior studies showing that light and the circadian clock modulate this peak of *CO* expression. The pattern of *SbCO* expression in RIL112 (*Sbco-3*) was similar to RIL105 (*SbCO-2*) although with slightly higher (<2-fold) levels of expression (Fig 38).

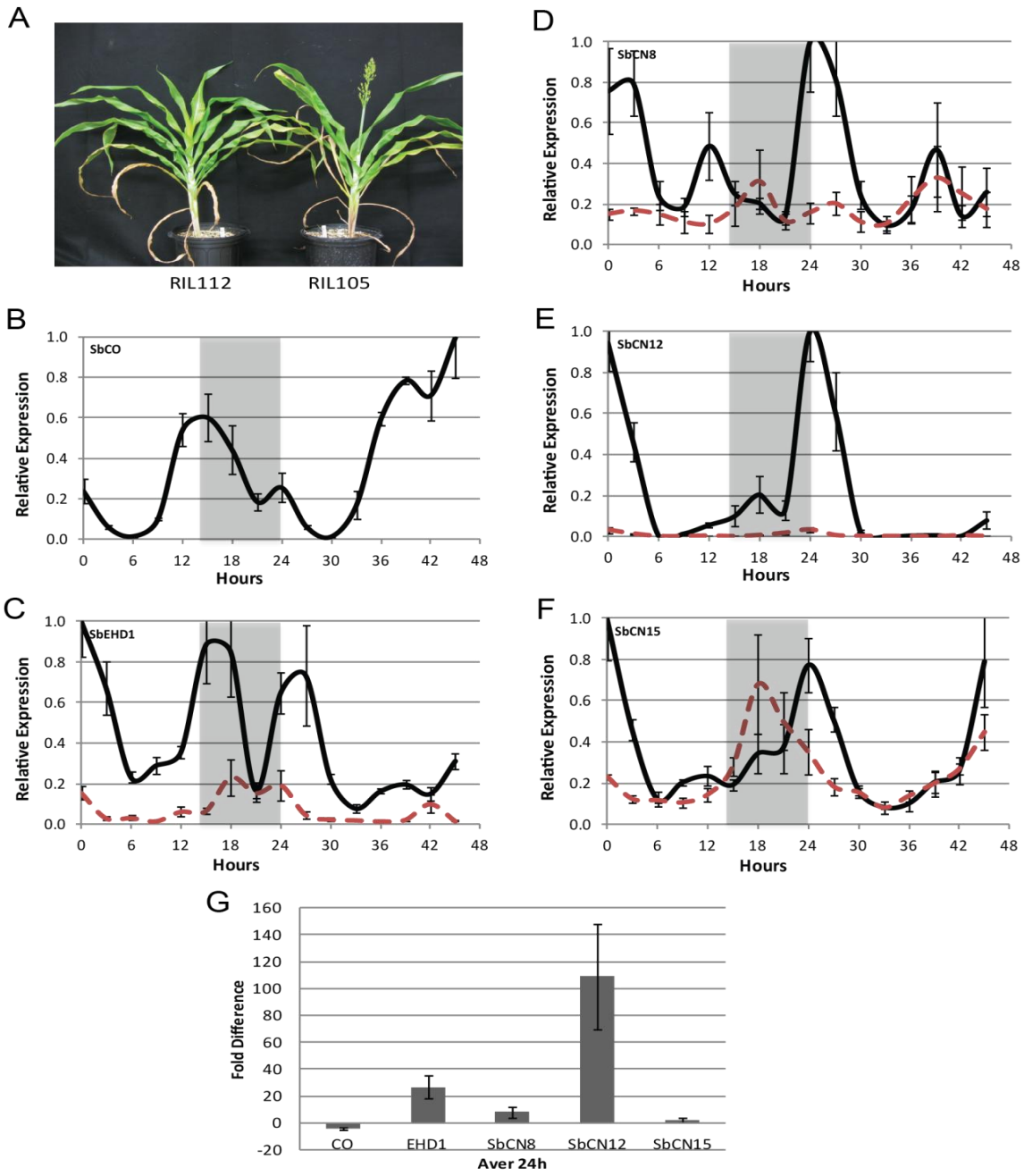


Figure 37. SbCO promotes flowering by inducing *SbEHD1* and *FT*-like genes in LD (14h light/10h dark). A. Flowering time phenotype of RIL112 and RIL105. (Days to flowering for RIL112 and RIL105 are 113 and 75.) B-F. Relative expression levels of flowering time genes in RIL105 (black solid line) and RIL112 (red dashed line). Gray shading denotes the dark/night portion of each 24h cycle. The first 24h covers one light-dark cycle, followed by 24h of continuous light. B. *SbCO*. C. *SbEHD1*. D. *SbCN8*. E. *SbCN12*. F. *SbCN15*. G. Average fold differences between the mRNA levels of each gene in RIL105 and RIL112 is plotted. Positive values represent higher expression detected in RIL105. Each expression data point corresponds to three technical replicates within three biological replicates.

RIL105 (*SbCO-2*) and RIL112 (*Sbco-3*) were used to analyze how alleles of *CONSTANS* affect expression of other genes in the sorghum flowering time regulatory pathway. Expression of the clock genes *TOC1, LHY* and *GI* were similar in RIL105 and RIL112, indicating that these genes are not affected by *SbCO* alleles as expected for genes upstream of *SbCO* (Fig. 39). In Arabidopsis CO activates flowering by inducing expression of *FT* and in rice Hd1 activates *Hd3a/RFT1*, genes encoding PEBP (phosphatidylethanolamine-binding) domain protein 'florigens' that move from the leaf to the shoot apical meristem where they interact with FD and induce floral transition. In rice, two members of the PEBP gene family, *Hd3a* and *RFT1* were identified as encoding florigens [46]. In maize, *ZCN8*, a different member of the PEBP gene family, was identified as a source of florigen [93, 94]. In sorghum, *SbCN8* and *SbCN12* are potential sources of florigen because expression of both genes is regulated by photoperiod, modulated by *Ma1* alleles, and induction of expression occurs coincident with floral initiation [15]. The sorghum orthologs of maize *ZCN8* (*SbCN8*), *ZCN12* (*SbCN12*) and rice *Hd3a* (*SbCN15*) were identified and qRT-PCR primers specific to each gene were designed to enable analysis of gene expression (Table 7). No ortholog of *RFT1* is present in the sorghum genome.

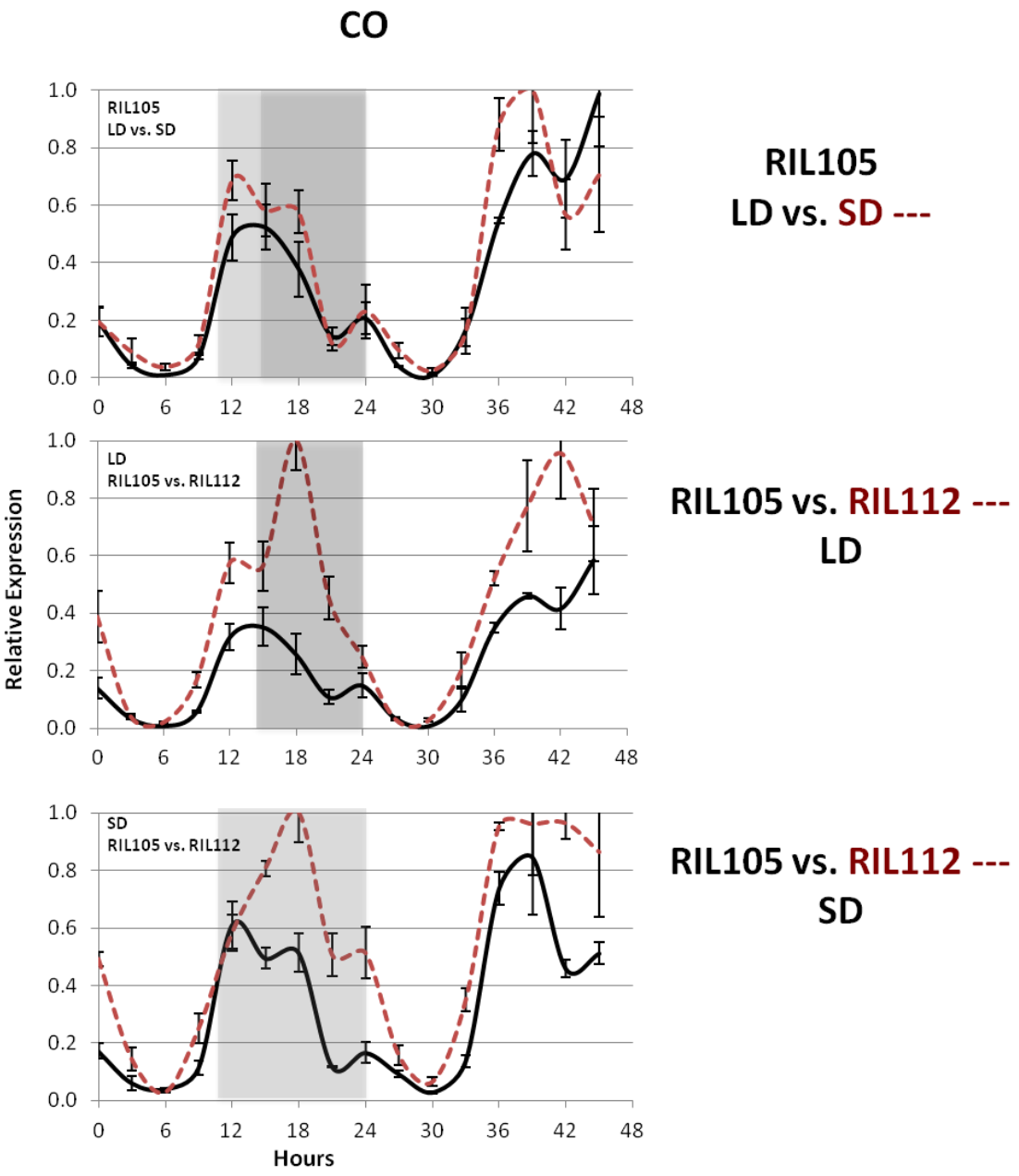


Figure 38. Relative expression level of *SbCO* in RIL105 and RIL 112 under LD and SD conditions. The gray shade represents dark period. First 24h is light-dark cycle, following 24h continuous light. Each expression data corresponds to three technical replicates within three biological replicates.

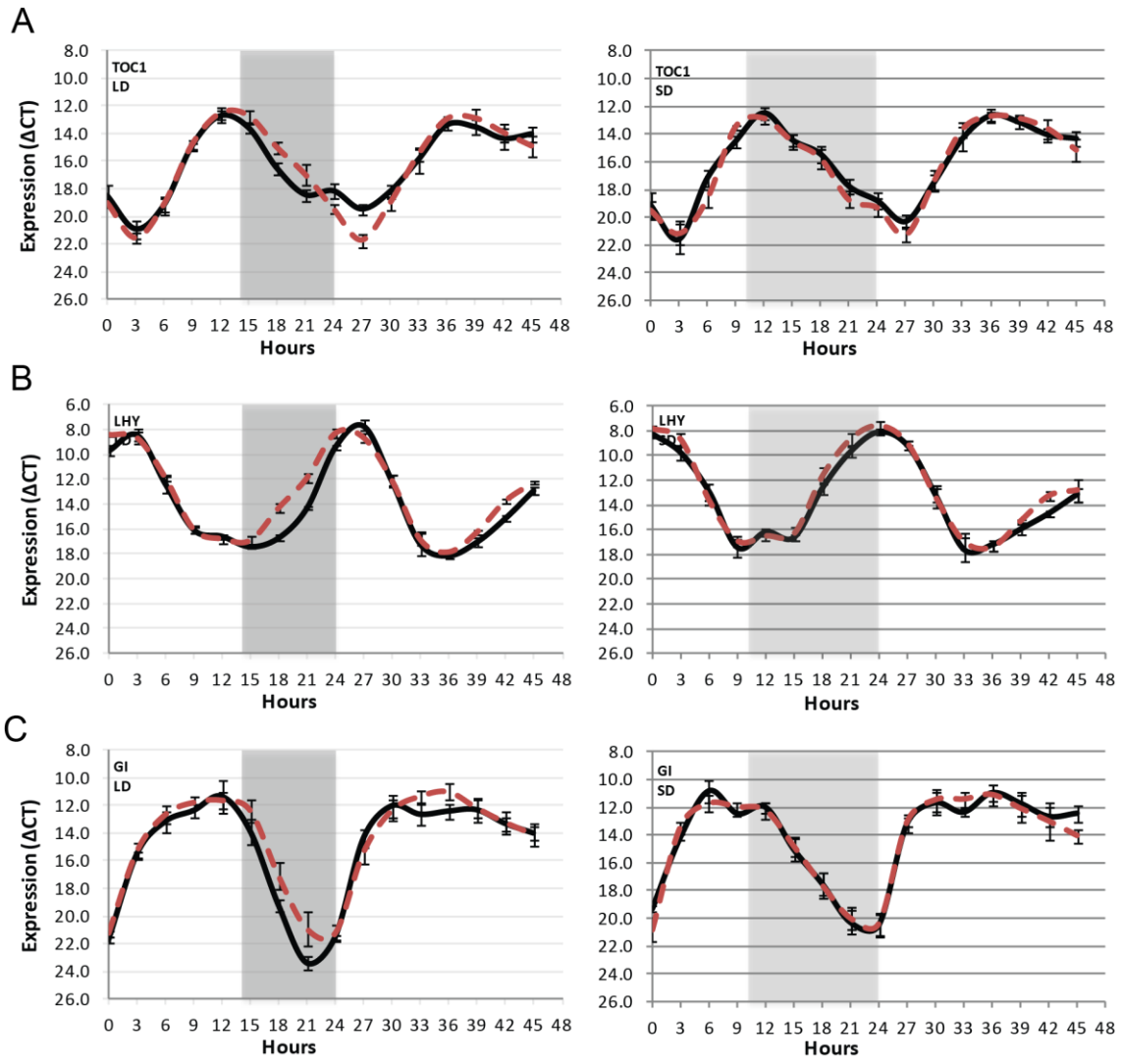


Figure 39. Expression level (ΔCt) of circadian clock genes and *GI* in RIL105 (black solid line) and RIL112 (red dashed line) under either LD (14h light/10h dark) or SD (10h light/14h dark) conditions. The gray shaded area represents the dark period. The first 24h covers one light-dark cycle, followed by 24h of continuous light. A. *TOC1*. B. *LHY*. C. *GI*. Each expression data point corresponds to three technical replicates within three biological replicates.

In leaves of RIL105 (*SbCO-2*) grown in LD, *SbEHD1* expression was high at dawn and then declined during the day before increasing in the evening approximately

15h after dawn (Fig 37C, black solid line), with a pattern similar to *SbEHD1* expression in 100M (*Ma1*) in short days [15]. *SbEHD1* expression was ~20-fold higher in RIL105 (*SbCO-2*) compared to RIL112 (*Sbco-3*) over the 24h LD cycle indicating that CO activates expression of *SbEHD1* (Fig 37C, RIL112 = red dashed line). In leaves of RIL105 (*SbCO-2*) grown in LD, *SbCN8* and *SbCN12* mRNA levels were highest at dawn, then decreased during the day with a second smaller peak of expression approximately 12-18h after dawn (Figs 37D and E, black solid line). In RIL112 (*Sbco-3*), the same pattern of expression was observed; however, *SbCN8* and *SbCN12* mRNA levels were much lower (Fig 37D and E, red dashed line). Expression of *SbCN8* was ~10-fold higher in RIL105 (*SbCO-2*) compared to RIL 112 (*Sbco-3*) (Fig 37C) and expression of *SbCN12* was ~100-fold higher in RIL105 (*SbCO-2*) compared to RIL112 (*Sbco-3*) (Fig 37D) over a 24h LD cycle in the *SbCO-2* background (Fig 37G). In contrast, the mRNA level of *SbCN15* (*Hd3a*) was similar in the two genotypes, although the gene's peak of expression was at dawn in RIL105 (*SbCO-2*) and at 18h in RIL112 (*Sbco-3*) (Fig **37F**). Together, these results are consistent with the hypothesis that SbCO promotes flowering by inducing expression of *SbEHD1*, *SbCN8,* and *SbCN12*, with *SbCN12* showing the largest CO-mediated increase in expression in LD and SD (see below).

Regulation of SbCO floral promoting activity in SD and LD

Comparison of flowering time and flowering pathway gene expression in RIL105 (*ma1, ma6, CO*) and RIL112 (*ma1, ma6, co-3*) showed that SbCO activates *SbCN8/12* expression and flowering in LD. The next question addressed was whether photoperiod

alters SbCO activity in sorghum. To this end, the relative ability of SbCO to activate expression of *SbCN12* and *SbCN8* in SD and LD was compared using RIL105 (*ma1, ma6, CO*) and RIL112 (*ma1, ma6, co-3*) (Fig 40). In RIL 105, *SbCN12* and *SbCN8* had significantly higher expression in SD compared to LD, especially during the night when both genes showed their highest expression (Fig 40 A and D; SD = red dashed line, LD=solid line). The difference between *SbCN12* mRNA levels in SD and LD ranged up to 100-fold depending on time of day, with the largest differences occurring during the night, peaking at 18h (Fig 40B). A similar pattern was observed for *SbCN8* where expression in SD was 20-40 fold higher during the night in SD, peaking between 18-21h (Fig 40E). When a similar comparison of *SbCN8/12* expression in SD/LD was done using RIL112 (*Sbco-3*), <10-fold difference in expression in SD vs. LD was observed (Fig 40 C and F). These results are consistent with CO having greater activity and causing higher expression of *SbCN8/12* in SD compared to LD in genetic backgrounds that contain null alleles of *Ma1* and *Ma6*. Complete figures showing the expression of *SbCN8, SbCN12, SbCN15* and *EHD1* in RIL105 and RIL 112 under LD and SD conditions present in Fig41, Fig42, Fig43, Fig44, and Fig45.

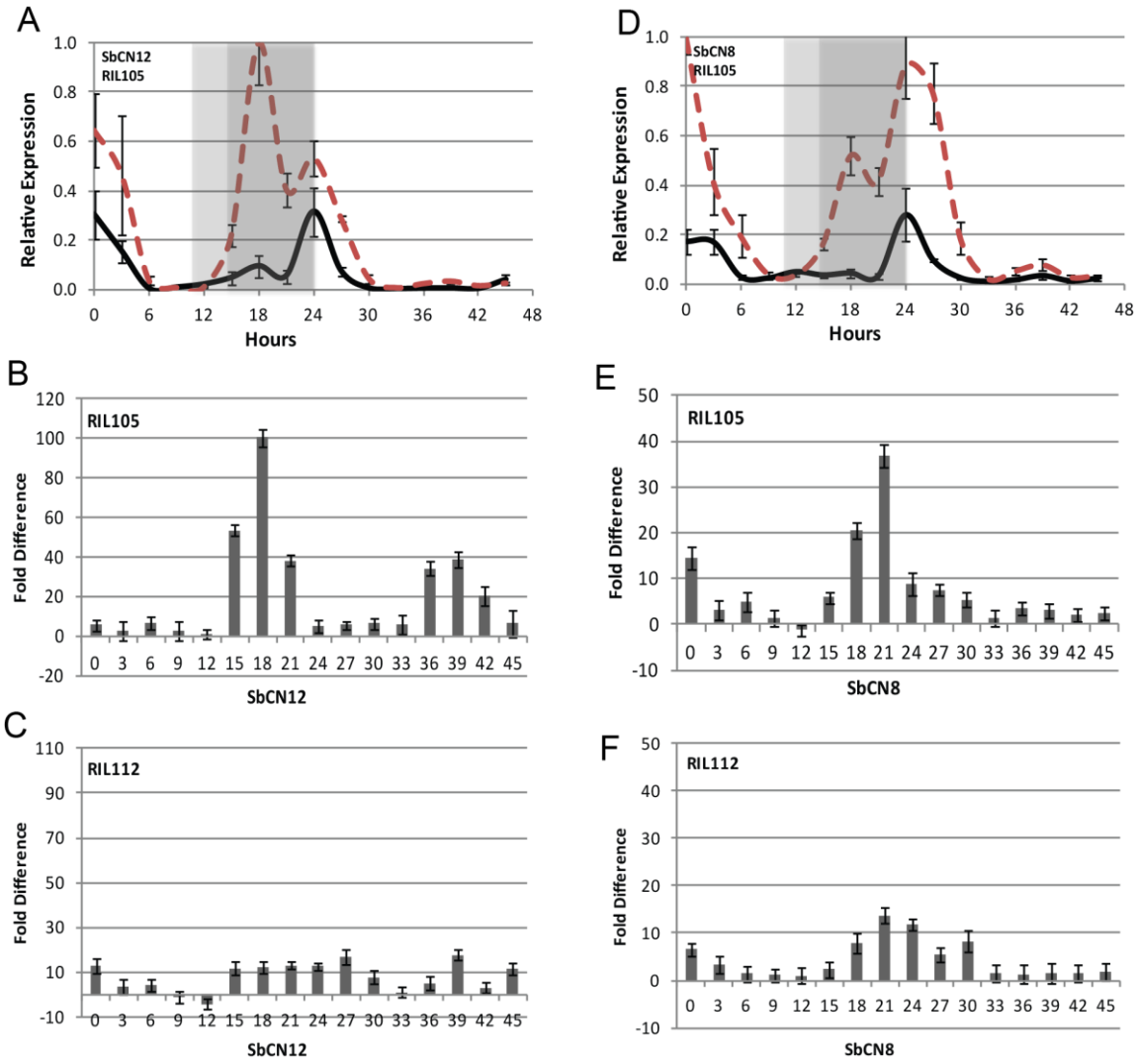


Figure 40. Relative expression levels and fold differences of *SbCN8* and *SbCN12* mRNA in plants grown in LD (14h light/10h dark) or SD (10h light/14h dark). Black solid lines represent relative expression in LD and red dashed lines represent relative expression in SD. Positive fold difference values indicates higher mRNA levels detected under SD condition. A-C. *SbCN12*. D-F. *SbCN8*.

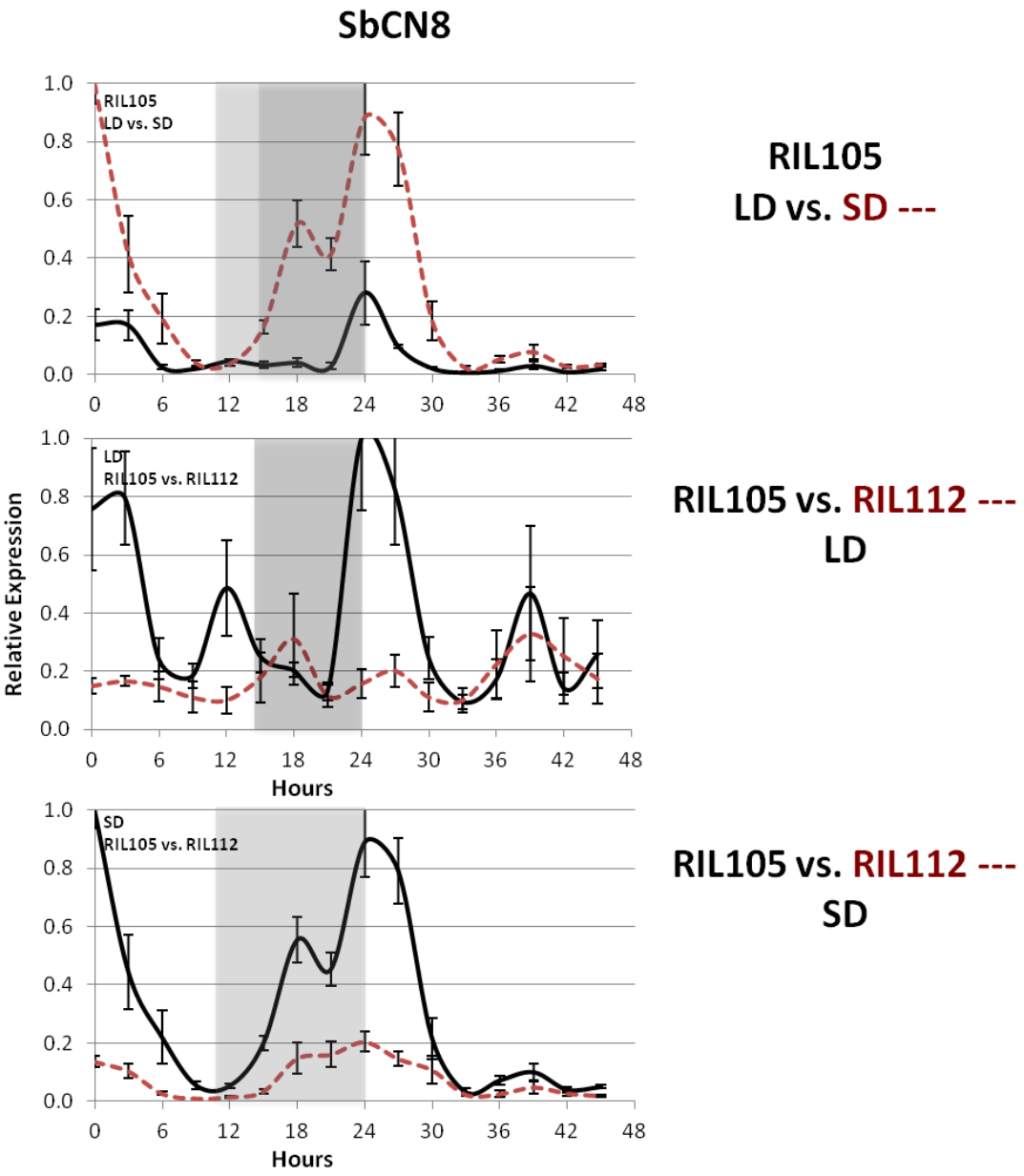


Figure 41. Relative expression level of *SbCN8* in RIL105 and RIL 112 under LD and SD conditions. The gray shade represents dark period. First 24h is light-dark cycle, following 24h continuous light. Each expression data corresponds to three technical replicates within three biological replicates.

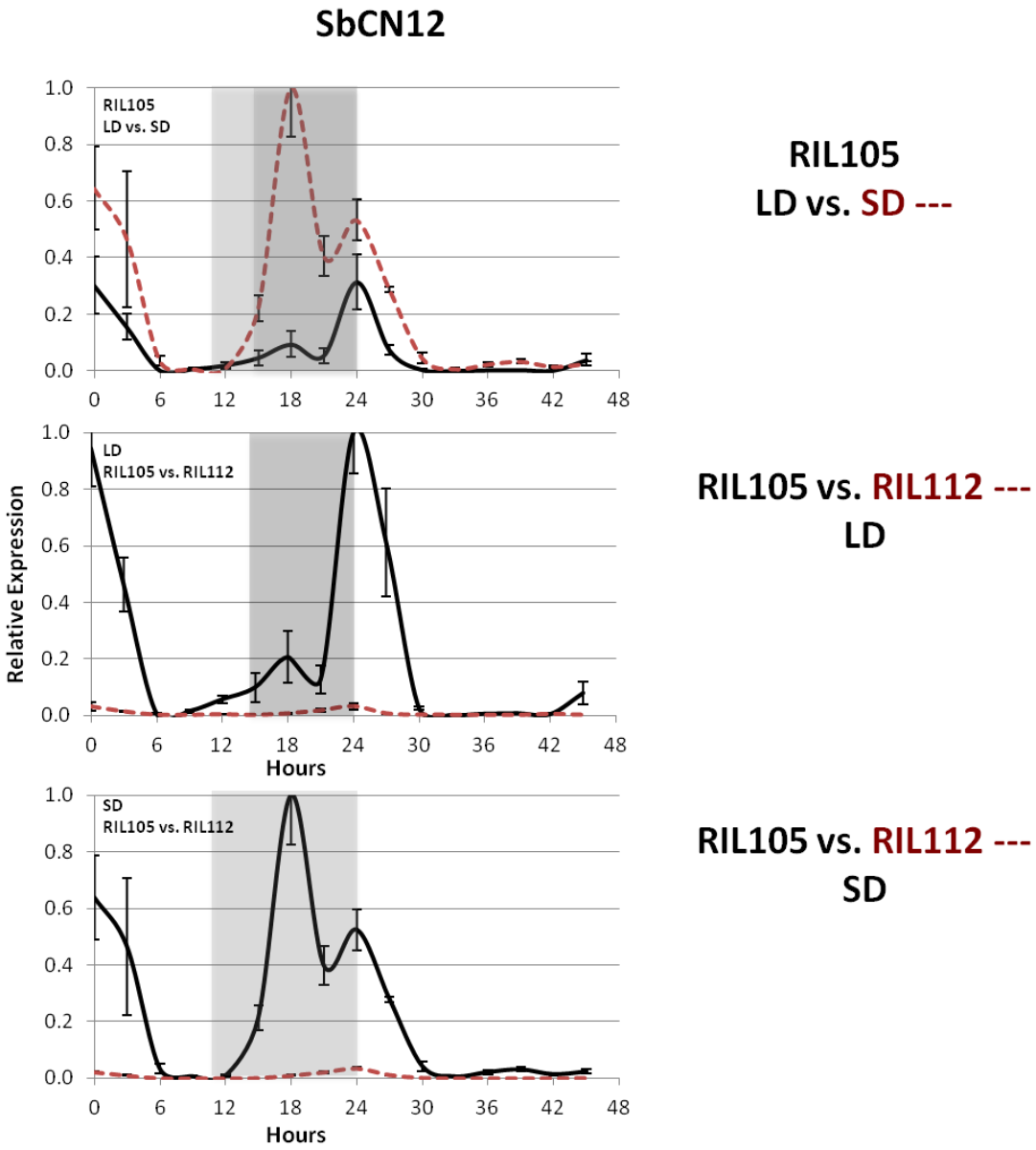


Figure 42. Relative expression level of *SbCN12* in RIL105 and RIL 112 under LD and SD conditions. The gray shade represents dark period. First 24h is light-dark cycle, following 24h continuous light. Each expression data corresponds to three technical replicates within three biological replicates.

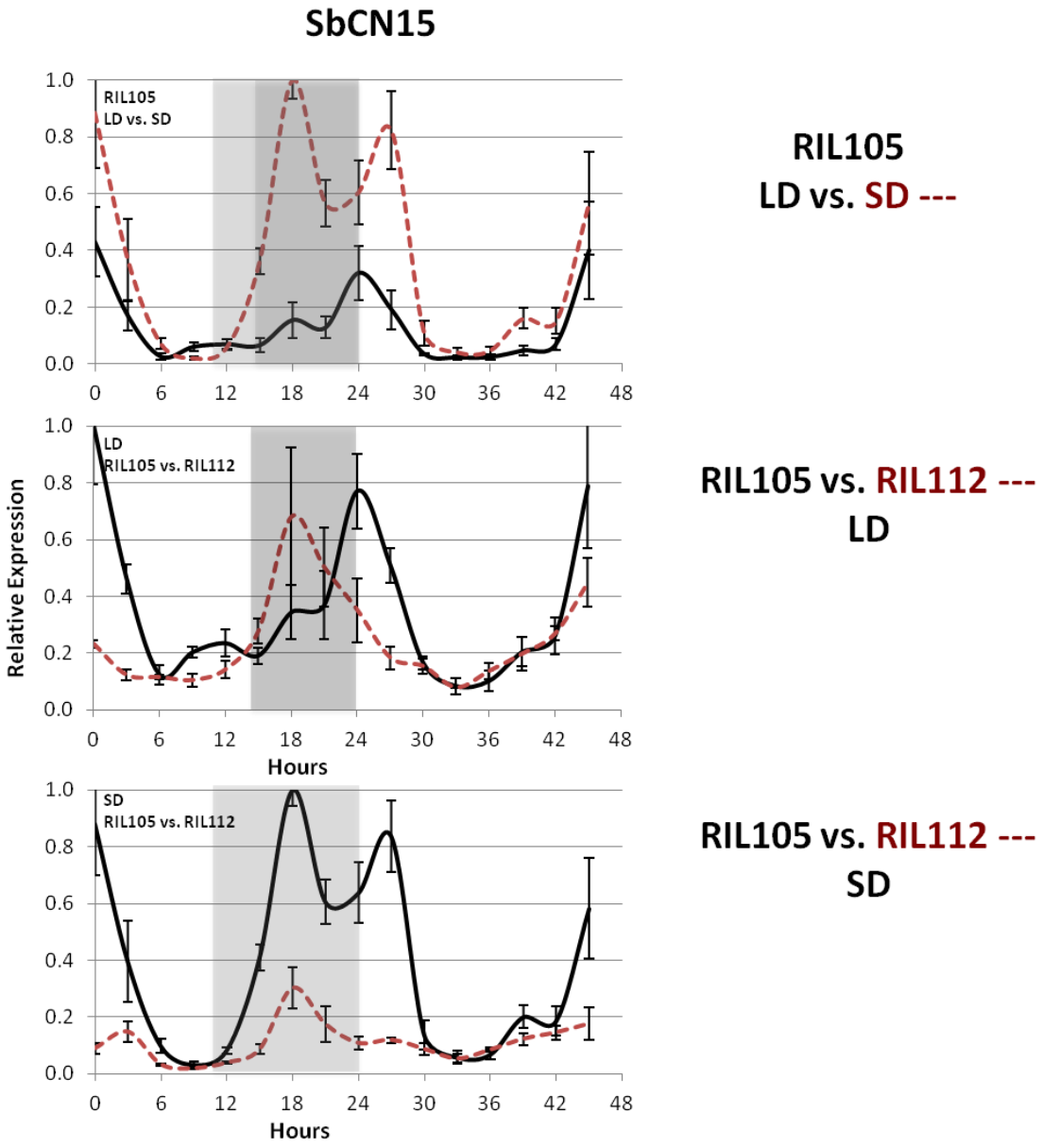


Figure 43. Relative expression level of *SbCN15* in RIL105 and RIL 112 under LD and SD conditions. The gray shade represents dark period. First 24h is light-dark cycle, following 24h continuous light. Each expression data corresponds to three technical replicates within three biological replicates.

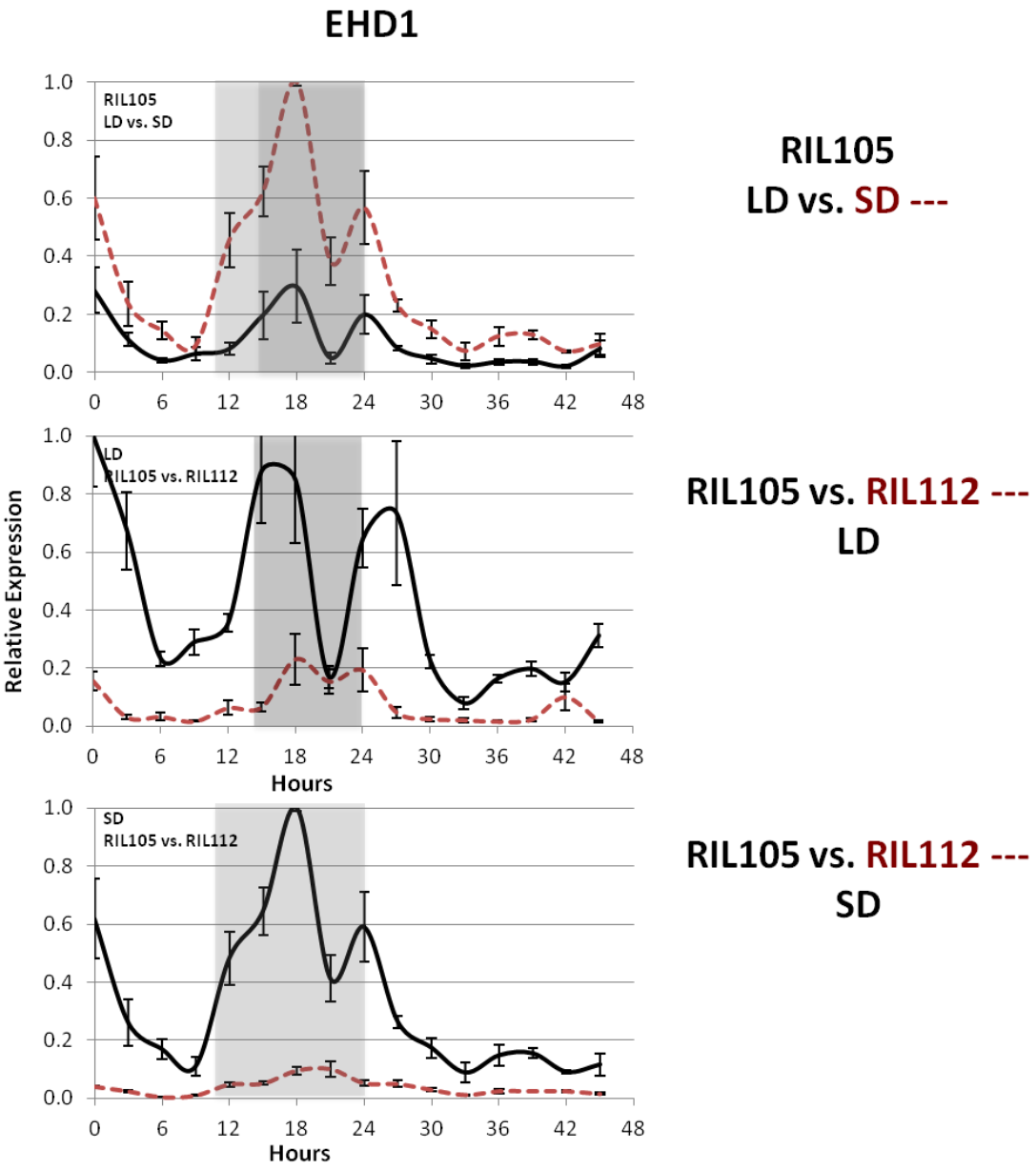


Figure 44. Relative expression level of *SbEHD1* in RIL105 and RIL 112 under LD and SD conditions. The gray shade represents dark period. First 24h is light-dark cycle, following 24h continuous light. Each expression data corresponds to three technical replicates within three biological replicates.

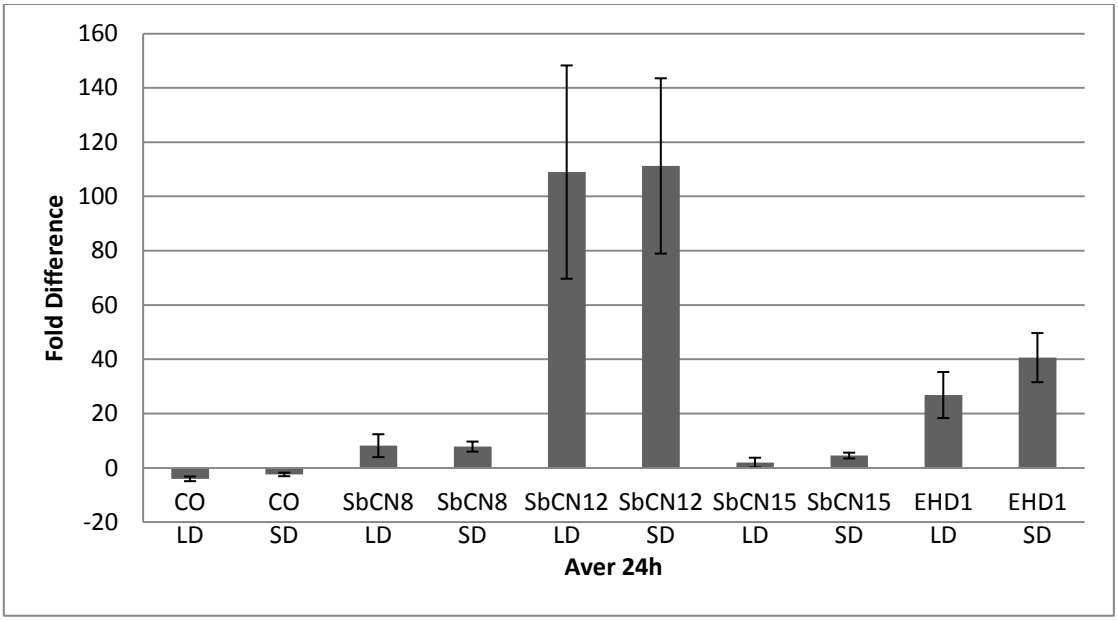


Figure 45. Average fold difference of expression level for first 24h (light-dark cycle) comparing RIL105 and RIL112 under LD and SD conditions. Positive value represents higher expression detected in RIL105.

Post-transcriptional inhibition of SbCO activity by SbPRR37 (*Ma1*)

In sorghum, *Ma1* (*SbPRR37*) increases photoperiod sensitivity by repressing expression of *SbEHD1* and *SbCN8/12*, resulting in delayed flowering in LD but with minimal effect in SD [15]. The ability of SbPRR37 to inhibit expression of *SbCN8* and *SbCN12* could be due to direct inhibition of SbEHD1 or SbCO, activators of *SbCN8* and *SbCN12* expression, and/or by direct inhibition of SbCN8 and SbCN12. A flowering time QTL coincident with *Ma1* was identified in the BTx642/Tx7000 RIL population grown under field conditions in 2008, 2009 and 2010 as well as in Lubbock, Texas (data not shown). This QTL was not observed in SD conditions, as expected, because

SbPRR37 has minimal impact on flowering under these conditions. As noted above, BTx642 encodes a null allele *Sbprr37-1*, however, the *Ma1* allele in Tx7000, *Sbprr37-2*, encodes a full-length protein with one amino acid substitution Lys62Asn with sufficient activity to delay flowering time under field conditions.

If SbPRR37 delays flowering by inhibiting SbCO, and SbCO increases expression of *SbEhd1* and *SbCN8/12*, then epistatic interaction between SbPRR37 and SbCO may be detected in the RIL population. *SbPRR37* and *SbCO* allelic interactions were examined by first sorting the RILs into lines that contain *Sbprr37-1* (null) or *Sbprr37-2*, and then analyzing the influence of *SbCO* and *SbEHD1* alleles on flowering time in each background. (Fig 47A,B; Fig 48 A,B) In the portion of the population containing the null version of *Sbprr37-1*, the QTL corresponding to *SbCO*/*Sbco-3* (LOD=13) explained 48% of the phenotypic variance for flowering time in the field (Fig46A) and 60% and 50% in LD greenhouse and SD greenhouse, respectively (Fig 49 A,B). In contrast, in the portion of the RIL population containing the active allele of *Ma1* (*Sbprr37-2*), no QTL corresponding to *SbCO* was observed. In this portion of the RIL population (*Sbprr37-2)*, the QTL corresponding to *SbEhd1-1*/*Sbehd1-2* explained ~20% of the phenotypic variance (Fig 49C). This result indicates that Sbprr37-2 inhibits SbCO-mediated induction of flowering. If this hypothesis is correct, then the ability of *Sbprr37-2* to inhibit flowering could be dependent on an active allele of *SbCO*. To test this hypothesis, the RIL population was sorted into lines that contained *SbCO-2* and lines that contained *Sbco-3*, and flowering time QTL were identified in each background (Fig 50A,B; Fig 51A,B) (Fig 46B). This analysis showed that *SbPRR37* alleles affected

flowering time in the *SbCO-2* background but not in the genetic background containing *Sbco-3* alleles, indicating that the ability of SbPRR37 to inhibit flowering is dependent on SbCO.

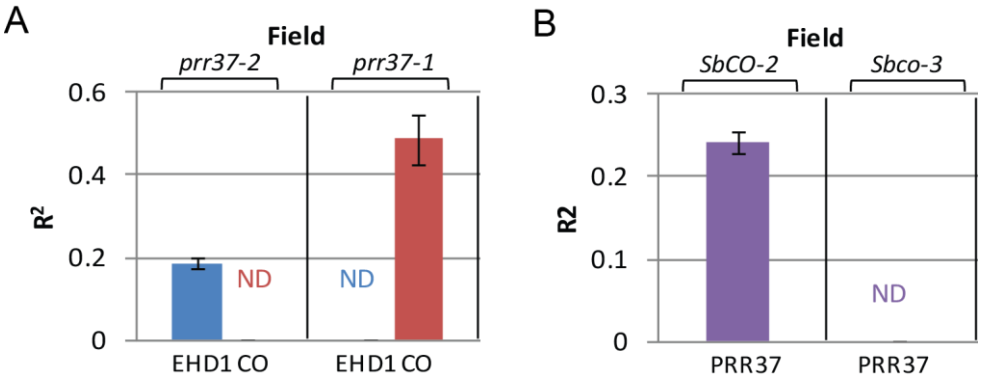


Figure 46. Epistasis analysis of *SbPRR37* and *SbCO* QTL in BTx642/Tx7000 RIL population under field conditions. A. Proportion of phenotypic variance (R^2) explained by the QTL corresponding to *SbCO-2*/*Sbco-3* in the portion of the population homozygous for *Sbprr37-1* (right) or *Sbprr37-2* (left). B. Proportion of the phenotypic variance explained by the QTL corresponding to *Sbprr37-1/Sbprr37-2* in the portion of the population homozygous for *SbCO-2* (left) or *Sbco*-3 (right). Each R^2 value represents the average obtained under field conditions in three years.

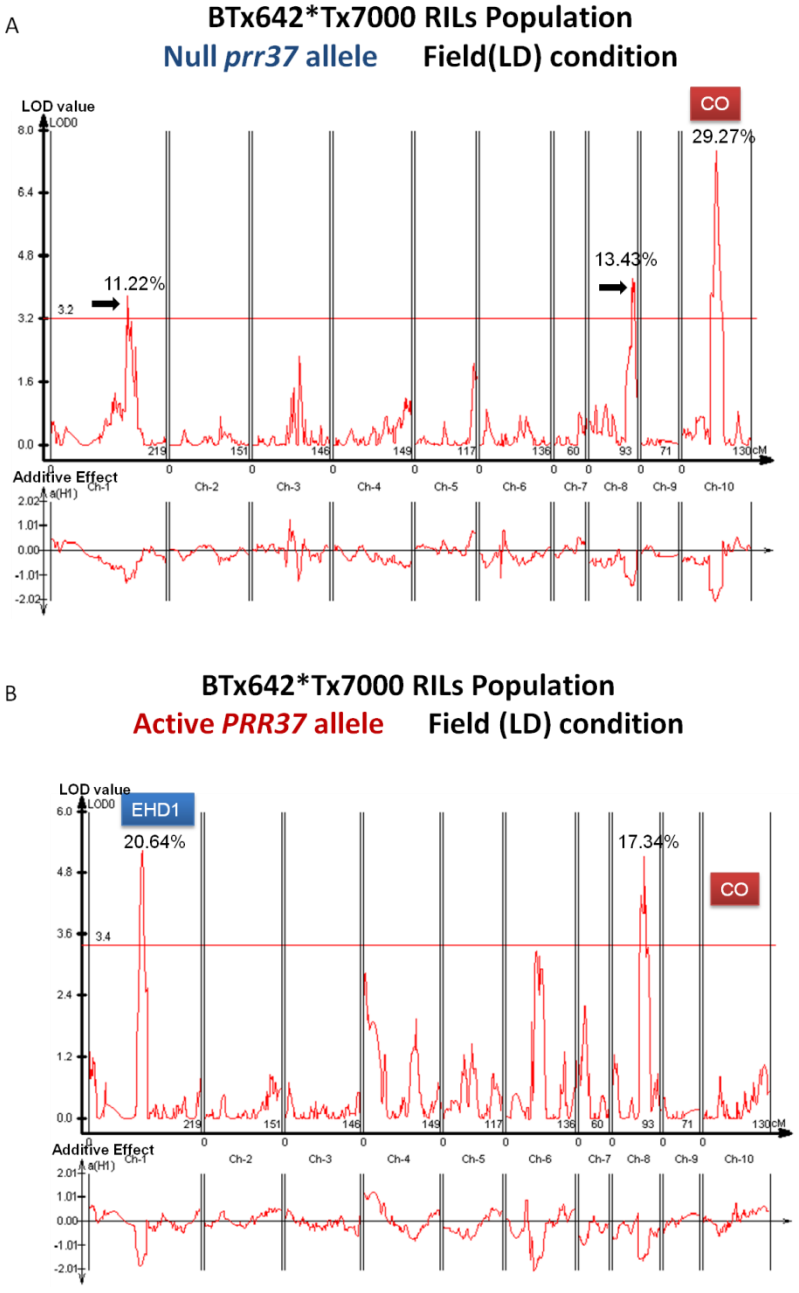


Figure 47. Epistasis analysis of flowering time QTLs in BTx642/Tx7000 RILs population under field LD condition. Candidate gene with R^2 value (proportion of phenotypic variance explained by the QTL) was labeled above each QTL. Arrow represents the direction of peak shift from original position. (A) QTLs detected in homozygous null *prr37* subpopulation. (B) QTLs detected in homozygous active *PRR37* subpopulation.

Figure 48. Epistasis analysis of flowering time QTLs in BTx642/Tx7000 RILs population under greenhouse SD condition. Candidate gene with R^2 value was labeled above each QTL. Arrow represents the direction of peak shift from original position. (A) QTLs detected in homozygous null *prr37* subpopulation. (B) QTLs detected in homozygous active *PRR37* subpopulation.

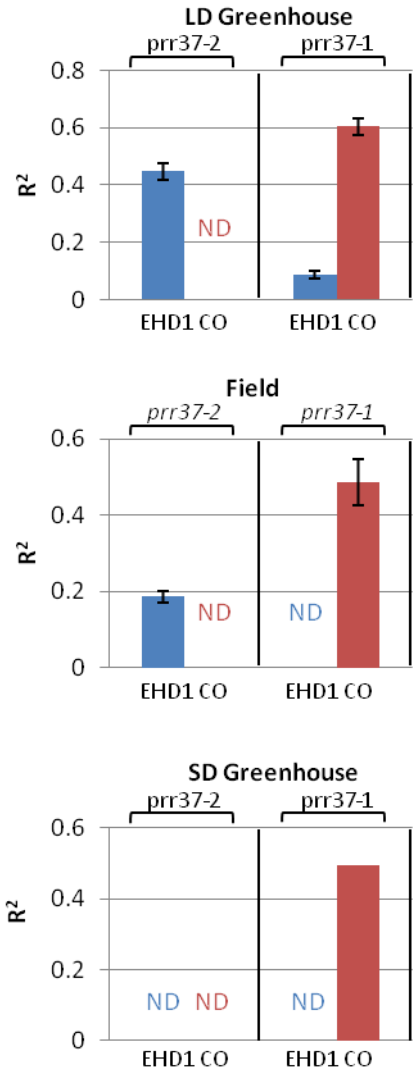


Figure 49. Epistasis analysis of *SbPRR37*, *SbCO* and *SbEHD1* QTL in BTx642/Tx7000 RIL population under (A) greenhouse LD, (B) greenhouse SD, (C) field conditions. Proportion of phenotypic variance (R^2) explained by the QTL corresponding to *SbEHD1* and *SbCO* in the portion of the population homozygous for *Sbprr37-1* (null, right) or *Sbprr37-2* (active, left).

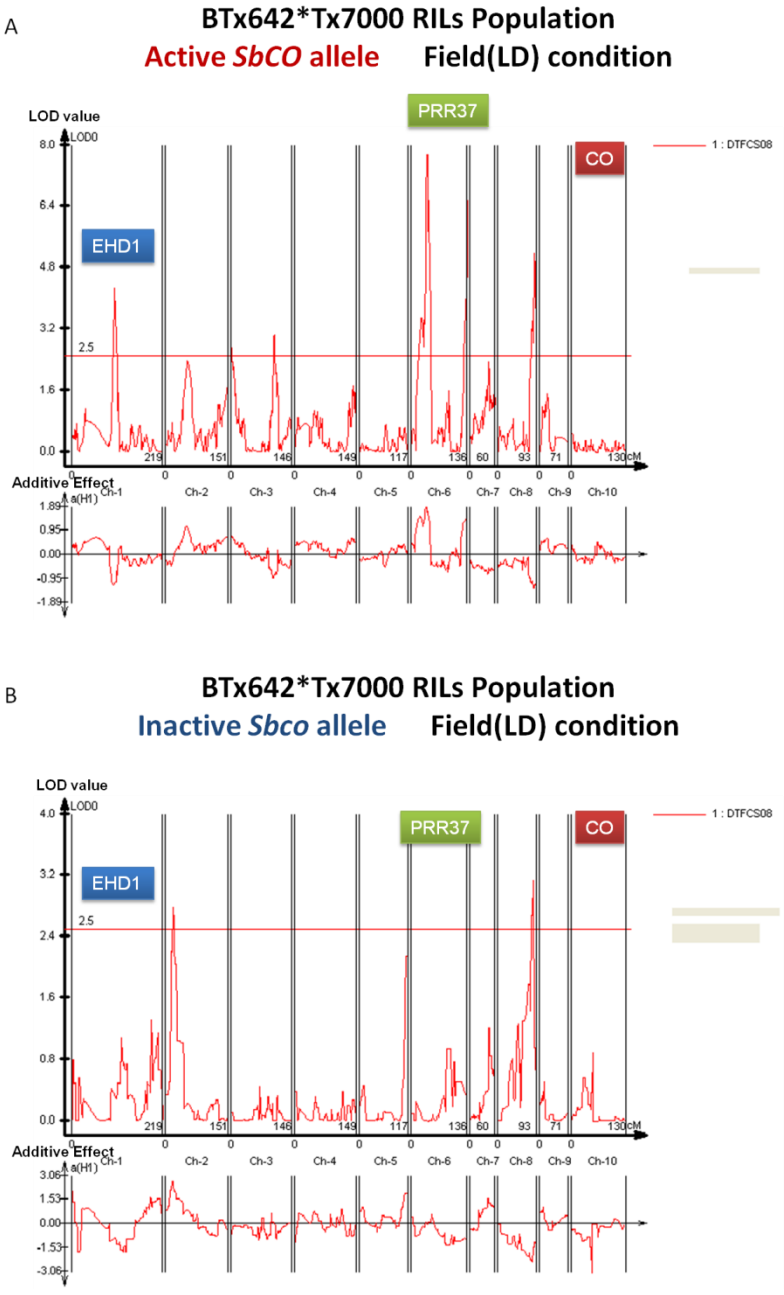


Figure 50. Epistasis analysis of flowering time QTLs in BTx642/Tx7000 RILs population under field LD condition. Candidate gene with R^2 value was labeled above each QTL. Arrow represents the direction of peak shift from original position. (A) QTLs detected in homozygous active *SbCO* subpopulation. (B) QTLs detected in homozygous inactive *Sbco* subpopulation.

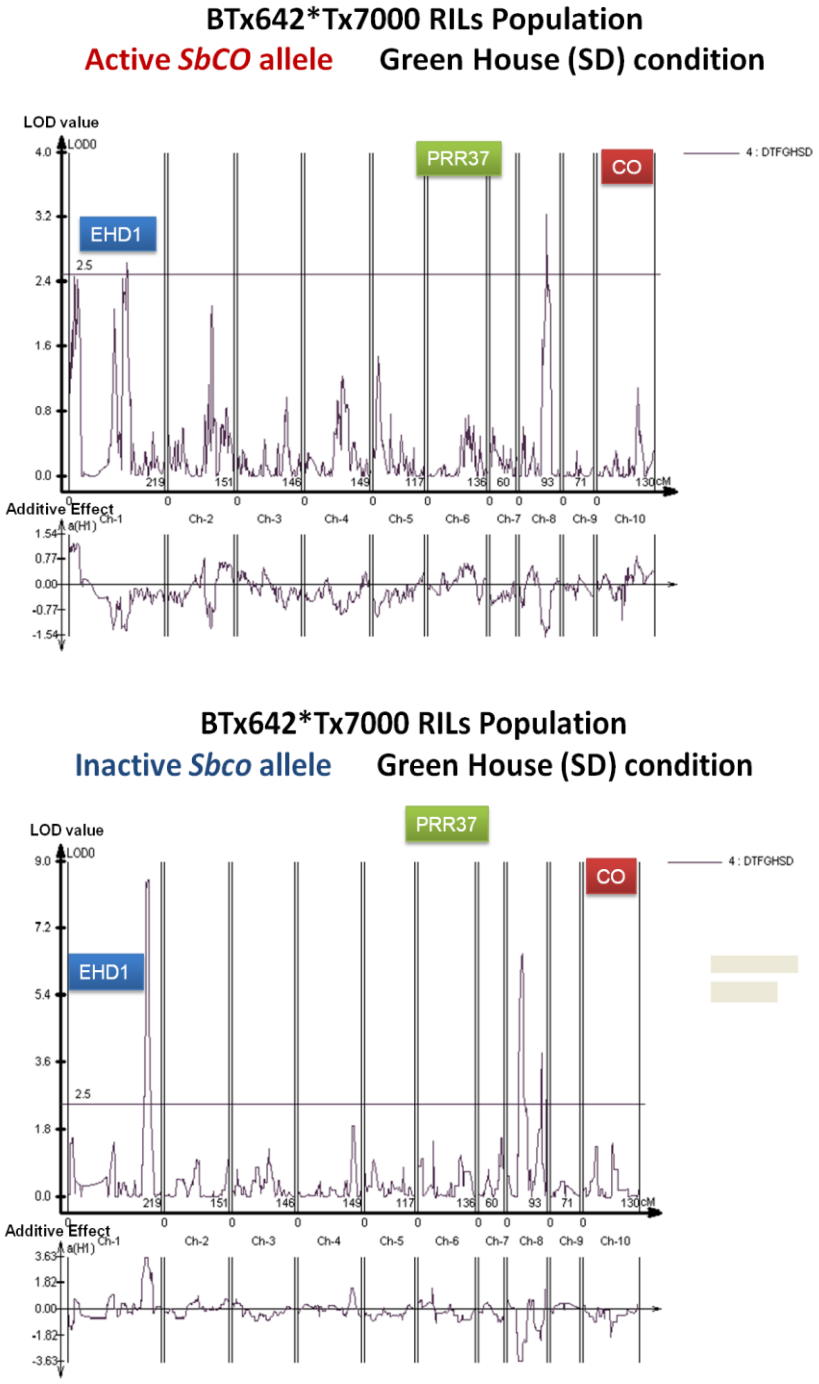


Figure 51. Epistasis analysis of flowering time QTLs in BTx642/Tx7000 RILs population under greenhouse SD condition. Candidate gene with R^2 was labeled above each QTL. Arrow represents the direction of peak shift from original position. (A) QTLs detected in homozygous active *SbCO* subpopulation. (B) QTLs detected in homozygous inactive *Sbco* subpopulation.

Discussion

Sorghum accessions exhibit a wide range of flowering times when plants are grown in long days (i.e., 48d to >175d under field conditions in College Station, Texas) [113]. A large extent of this variation is caused by differences in photoperiod sensitivity mediated by floral repressors encoded by *Ma1* and *Ma6* that inhibit flowering in long days [15, 18]. Much less is known about floral activators in sorghum. The grass specific floral activator *SbEHD1* was previously identified based on the gene's sequence similarity to rice *EHD1* and activation of *SbEHD1* expression coincident with floral initiation [113]. In this study we identify and characterize a second activator of sorghum flowering *SbCO*, a homolog of the floral activator *CONSTANS* in Arabidopsis and an ortholog of *Hd1* in rice.

Flowering time QTL spanning *SbCO*, *SbEHD1*, *SbPRR37* and two other regions of the sorghum genome were identified using a RIL population derived from a cross of BTx642 and Tx7000. The identification of flowering pathway gene alleles located in QTL was aided by prior acquisition of the BTx642 and Tx7000 genome sequences and a high-resolution genetic map generated Digital Genotyping and aligned to the physical map[135]. Phenotyping RILs for flowering time in LD, SD and under field conditions aided QTL detection and the identification and characterization of alleles of known flowering time genes located in each QTL. The present analysis showed that Tx7000 DNA spanning the QTL on SBI-06 (*SbPRR37*) delayed flowering under field conditions. Similarly, BTx642 alleles of the QTL on SBI-01 (*SbEHD1*), SBI-08 and SBI-10 (*SbCO*) resulted in delayed flowering. Coding alleles of *SbEHD1* in BTx642 and Tx7000 were

identified with activity consistent with the known function of *SbEHD1* as an activator of flowering. Analysis of the flowering time QTL on LG-06 led to the identification of coding alleles of *SbPRR37* consistent with QTL action. Tx7000 encodes *Sbprr37-1*, a null allele, and BTx642 encodes *Sbprr37-2*, a weak allele that acts as a repressor of flowering under field conditions. A candidate gene in the flowering time QTL on SBI-08 was not identified. However, coding alleles of CONSTANS were identified through analysis of the flowering time QTL on SBI-10. Results showed that SbCO functions as an activator of flowering in LD and SD in sorghum genotypes with null versions of *Sbprr37-1*. The *Sbco-3* allele in BTx642 was remarkable because it contained a His106Tyr amino acid substitution that also inactivates *CO* function in Arabidopsis [82]. Sorghum and Arabidopsis genotypes containing the inactive His106Tyr *co-3* allele flower late in long days, as well as late in short days in sorghum, indicating that CONSTANS functions as an activator of flowering in both species. SbCO shares a conserved CCT (CO, CO-like, TOC1) domain with TOC1, PRR37, Ghd7, and HEME ACTIVATOR PROTEINS (HAP or NF-Y proteins). Yeast two-hybrid screens showed that CO can interact with HAP3 and HAP5 subunits through its CCT-domain, forming CCAAT-binding CBF-complexes that bind to FT promoters and activate transcription [84, 86]. In sorghum, SbCO was found to activate transcription of *SbEHD1*, *SbCN8* and *SbCN12*, consistent with its role as an activator of flowering, presumably through formation of CBF-complexes, but possibly through direct binding to DNA [87].

The ability of *SbCO* alleles to induce flowering pathway gene expression and flowering was examined in RIL genetic backgrounds that contained null alleles of *Ma1*

(*Sbprr37-1*) and *Ma6* (*Sbghd7)* to eliminate the influence of these LD floral repressors. In this null genetic background, SbCO promoted early flowering in LD and SD and increased the expression of *SbEHD1* (~25-fold), *SbCN8* (~10-fold), *SbCN12* (~100-fold) and *SbCN15* (~5-fold) relative to their expression in lines carrying the inactive *Sbco-3* allele. This information is summarized in a flowering time regulatory model shown in Fig 52. The model includes three members of the PEBP-gene family that could be sources of florigen in sorghum, *SbCN8*, *SbCN12* and *SbCN15*. *SbCN8* is an ortholog of maize *ZCN8*, with a pattern of gene expression consistent with the demonstrated role of *ZCN8* as a source of florigen in maize [94]. *SbCN12* expression is repressed by PRR37, induced in leaves in SD, and induced by *SbCO* (this study), indicating that this gene is also a likely source of florigen in sorghum. In rice, *Hd3a* and *RFT1* have been identified as sources of florigen; therefore expression of *SbCN15,* the ortholog of *Hd3a*, was analyzed. *SbCN15* showed relatively small changes in gene expression in response to photoperiod and mutations in SbPRR37 and SbCO. The sorghum genome does not apparently encode an ortholog of *RFT1*, a source of florigen in rice in LD.

The results indicate that there has been a significant change in the complement of *FT*-like genes that function as the main sources of florigen in sorghum (*SbCN8*, *SbCN12*) and rice (*Hd3a =SbCN15*; *RFT1*, no sorghum ortholog), therefore regulation of flowering time could also differ, even though both grass species are short day plants. SbCO activates expression of *SbCN8* and *SbCN12*, although *SbCN12* was induced to a significantly greater extent. SbCO also increased expression of *SbEHD1*, an activator of *Hd3a* expression in rice. *SbEHD1* expression is repressed by SbPRR37 and SbGhd7 and

induced when photoperiod sensitive sorghum grown in LD is transferred to SD [15, 113]. Increases in *SbEHD1* expression occur in parallel with increases in *SbCN8* and *SbCN12* expression, suggesting that SbEhd1 can induce the expression of these genes as shown in Fig 52. However, the extent and specificity of this proposed activity of SbEhd1 will require further analysis in backgrounds where SbCO has minimal influence on the expression of these genes (*Sbco-3* backgrounds). The results in this paper show that SbCO increases expression of *SbEHD1* and it is proposed that SbEhd1 can activate expression of *SbCN8* and *SbCN12* and flowering. In contrast, rice Hd1 has not been reported to increase expression of *EHD1* [54, 61]. Of interest is the finding that SbCO activates expression of *SbCN12* to a much greater extent than *SbEHD1*. Therefore we conclude that SbCO increases *SbCN12* transcription and that this may be the most important way that SbCO activates flowering.

The finding that SbCO can activate flowering in LD and SD in sorghum genotypes that are null for *Ma1* and *Ma6* raised the question as to how the activity of this gene is regulated by photoperiod in this short day plant. *SbCO* expression is low during the day, then increases in the evening to a peak at ~15h after dawn, followed by a decrease and a second peak of expression at dawn (Fig. 37B). In Arabidopsis, a similar increase in *CO* expression in the evening is due to the interaction of GI and blue light-activated FKF1, resulting in degradation of CDF-factors that inhibit *CO* expression [41]. This mechanism may also explain the evening peak of *SbCO* expression in sorghum.

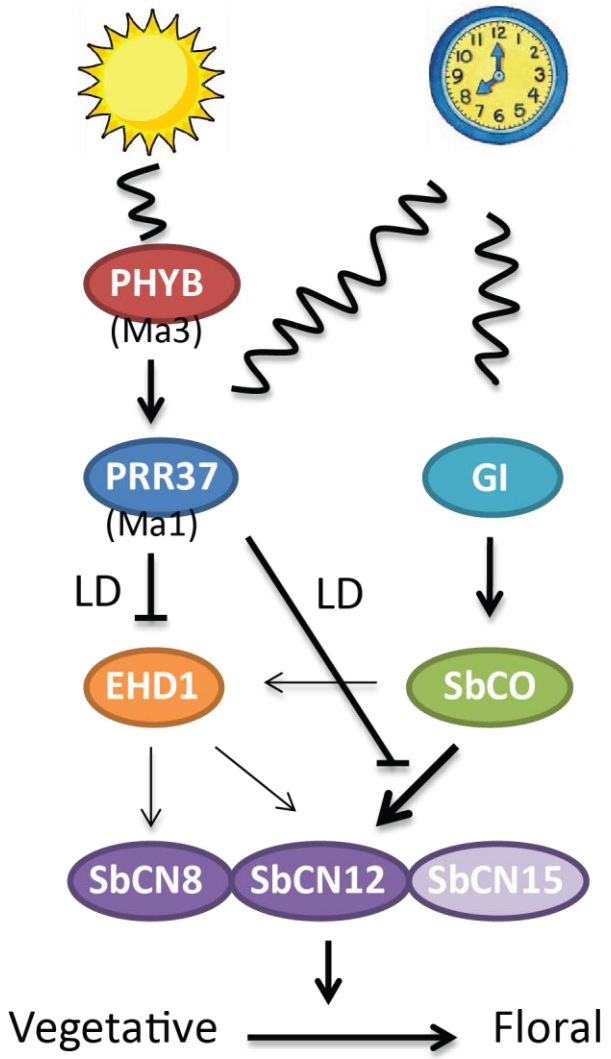


Figure 52. Model of photoperiod flowering-time regulation in sorghum (SbCO). Phytochrome B mediates light signaling, providing information about day length required for photoperiod detection. Light and the circadian clock regulate expression of PRR37 resulting in higher expression in LD vs. SD. PRR37, a floral repressor, inhibits the activity of SbCO, a floral activator resulting in delayed flowering in LD. In inductive SD, SbCO activates expression of *SbEHD1*, *SbCN8* and *SbCN12* genes thereby promoting flowering. Thin lines represent possible mechanism of flowering time regulation.

The second less prominent peak of *SbCO* expression at dawn is modulated by alleles of *SbPRR37* and enhanced in LD [15]. The function of the peak of *SbCO* expression at dawn is not currently understood, although production of SbCO at this time could help activate *SbEHD1* expression in the morning.

Functional alleles of *SbCO* increased the expression of *SbCN8* and *SbCN12* to a greater extent in SD relative to LD (Fig 40). SbCO expression levels peak maximally during the night between 18-21h, helping to explain why *SbCN8/12* expression increases at night as well. Since expression of *SbCO* was not altered significantly by photoperiod, increased activity of SbCO in SD is most likely due to an increase in protein level or activity. In Arabidopsis, a long day plant, CO levels are higher in LD due to COP1-SPA1-Cry2 stabilization of the protein [43]. This stabilization module may be missing or attenuated in sorghum. Reduced PhyB/C-mediated degradation of SbCO in SD, relative to LD, could result in greater SbCO-mediated activation of SbCN8/12 in SD. In sorghum genotypes containing active alleles of *SbPRR37* the evening peak of *SbCO* expression is not altered, the peak of expression at dawn increases, but the activity of SbCO is strongly attenuated. Expression of *SbPRR37* is high in the evening in plants grown in LD, but low in SD. Therefore, higher levels of *SbPRR37* expression in LD, and SbPRR37 repression of SbCO activity under these conditions, is predicted to prevent SbCO from activating flowering in LD.

SbPRR37 is a CCT-domain protein, like CO that has been shown to interact with HAP3 [85]. Therefore, SbPRR37 may be a competitive inhibitor of SbCO binding to the HAP complex. SbPRR37 may also directly bind to DNA in a fashion similar to TOC1

and other PRR-proteins [88]. TOC1 binding to its cognate motif in the promoter of LHY/CCA1 is mediated by its CCT-domain, resulting in PRR-domain mediated repression of transcription. If PRR37 binds to the *SbCN12* promoter in a similar manner, it could directly repress transcription, block SbCO binding to the HAP complex, and/or interact with CO or other proteins in order to repress *SbCN12* transcription. Recent results on the PRR37 ortholog in rice (Hd2) concluded that PRR37 directly represses *Hd3a* transcription [98]. Further genetic and biochemical analysis will be required to distinguish among these possibilities.

In barley, a long day plant, HvCO1 activates flowering in LD, and activation is dependent on Ppd-H1, an ortholog of SbPRR37. Overexpression of *HvCO1* induced flowering in both LD and SD, but photoperiod sensitivity mediated by *Ppd-H1* was still observed in this background [99]. Ppd-H1 does not directly affect expression of *HvCO1*, but potentiates the ability of HvCO1 to activate *HvFT1* expression in LD. It is interesting to note that Ppd-H1 increases HvCO1 activity in LD, whereas SbPRR37 inhibits the floral promoting activity of SbCO in LD. The expression and activity of SbPRR37 and Ppd-H1 increase in LD and both affect CO's ability to modulate FT-gene expression, but in an opposite manner, consistent with barley being a long day plant and sorghum a short day plant. The difference in activity of PRR37 could be due to differences in direct binding of PRR37 to the promoters of *SbCN12* and *HvFT1* (homolog of *Hd3a*, *SbCN15*), or more indirectly by interaction of PRR37/Ppd-H1 with activators, repressors, HAP subunits, HvCO1 and SbCO. In rice, Koo et al. suggest that phosphorylation of PRR37 by Hd6 may cause PRR37, in conjunction with Hd1, to

become a repressor of *Hd3a* expression in LD [98]. The possibility that PRR37 can form a co-repression complex with CO is consistent with results in sorghum and rice. However in the absence of PRR37, CO functions as an activator of *FT* expression in sorghum and rice. While the biochemical basis of variation in PRR37 activity remains to be elucidated, taken together, the results suggest that interaction between PRR37 and CO on the promoters of specific florigen-related PEBP-genes establish the fundamental differences in photoperiod-sensitive flowering between LD and SD grasses.

The flowering time model in Fig 6 shows that SbEhd1 and SbCO can independently induce flowering by activating *SbCN8* and/or *SbCN12*. EHD1 and Hd1 have been shown to independently activate Hd3a (FT) and flowering in rice [54]. In sorghum we show that there is cross-talk between these pathways because SbPRR37 activates *SbCO* expression at dawn in LD, while SbCO induces *SbEHD1* expression in SD. Rice EHD1, a B-type response regulator, is controlled by several upstream modulators including the repressors GHD7, GRAIN NUMBER, PLANT HEIGHT AND HEADING DATE 8 (GHD8), OsLEC1 and FUSCA-LIKE1 (OsLFL1), OsMADS56 and the activators GI, EARLY HEADING DATE 2 (EHD2) and OsMADS50 [50, 59]. The pathway regulating *SbEHD1* will be subject of a subsequent study in order to better understand its role in flowering time regulation in sorghum. The existence of two parallel pathways that can activate flowering in sorghum provides for a wide range of responses to diverse environmental factors, contributing to sorghum's wide geographical adaptation. Sorghum crop breeders are utilizing different alleles of key genes in these

parallel pathways to generate early flowering grain sorghum hybrids and late flowering energy sorghum hybrids.

Methods

Plant materials

The BTx642/Tx7000 RIL population (n=90) and parental lines were grown under field conditions in a replicated randomized block design at Texas A&M Research Farm near College Station Texas in 2008, 2009 and 2010 with planting between April 1-14. Days to mid-anthesis (pollen shed) were determined as a measure of flowering time. In the field, day-lengths increased from ~12.6h in April to 14.3 h in July, with an average daily maximum temperature of 31.7°C and an average daily minimum temperature of 20.0°C. Ten plants of each RIL and the parental lines were grown in a greenhouse in 10h day lengths (SD, 2011) or 14h day lengths (LD, 2009 and 2010) and phenotyped for flowering time in a similar manner as the populations grown in the field. RIL105 and RIL112 correspond to 4_6 and 12_14 in original BTx642/Tx7000 RIL population [5].

Genotyping by sequencing and QTL analysis

Genotyping by sequencing was carried out using Digital Genotyping (DG) [119] on the 90 RILs derived from BTx642 and Tx7000 [135]. A genetic linkage map was constructed using data generated from 1462 polymorphic DG markers using Mapmaker/EXP ver. 3.0b where recombination frequency was calculated using the Kosambi mapping function. QTLs were detected using Composite Interval Mapping (CIM) in WinQTL Cartographer v2.5 [130]. Significant LOD thresholds for QTL detection were determined based on experiment-specific permutations with 1000 repeats

at α=0.05 [131]. In QTL-based epistasis analysis, the 90 RILs were categorized into sub-populations based on alleles of *SbPRR37* or alleles of *SbCO* respectively. Sub-populations homozygous for each allele of *SbPRR37* and each allele of *SbCO* were then subjected to QTL analysis.

Phylogenetic and colinearity analysis

The amino acid sequence of rice Hd1(Os06g16370) was used to search Phytozome v9.1 (http://www.phytozome.net/) for homologs of CONSTANS in rice, maize, barley and sorghum. Multiple sequence alignment, alignment scores and phylogentic analysis were performed using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) using protein sequences of Sb10g010050 (Sorghum CO), GRMZM2G405368_T01 (Maize conz1), Os06g16370 (Rice Hd1), AF490468 (Barley HvCO1) and AT5G15850 (Arabidopsis CO). Rice and sorghum genome sequences (Phytozome v9.1, 100kbp) spanning homologs of CO were used for synteny/colinearity analysis. Colinearity was determined by GEvo (http://genomevolution.org/CoGe/GEvo.pl) , a high-resolution sequence analysis tool of genomic regions from CoGe (Accelerating Comparative Genomics) tool kit (http://genomevolution.org/CoGe/). A similar phylogenetic/colinearity analysis was performed for EHD1.

Allele characterization

SNPs in candidate genes were identified by comparing DNA sequences derived from BTx623, BTx642 and Tx7000. The BTx623 Sbi1 assembly and Sbi1.4 gene annotation were used as the reference genome sequence (Phytozome). BTx642 and

Tx7000 genome sequence assemblies used for analysis were obtained previously [135]. SNPs were called using the SNP Detection function of CLC Genomics Workbench 4.9. Minimum coverage for a variant call was set at 5, and maximum was set at 150. Allele types were designated based on SNPs. The SIFT algorithm (sorting intolerant from tolerant) [123] was utilized to predict whether an amino acid substitution affects protein function based on the degree of conservation of amino acid residues in sequence alignments derived from closely related gene sequences. RIL105 and RIL112 were selected from the BTx642/Tx7000 RIL population using DG markers flanking each QTL peak and spanning each candidate gene. RIL112 contains BTx642 haplotypes spanning all of the flowering time QTL, including the QTL on SBI-10 (*Sbco-3*). RIL105 contains BTx642 haplotypes spanning QTL on SBI-01, SBI-06, SBI-08 and the Tx7000 haplotype spanning the QTL on SBI-10 (*SbCO-2*).

LD, SD and circadian experiments

For circadian rhythm experiments, RIL105 and RIL112 were grown in the greenhouse in LD (14h light) for 32 days. For entrainment, the plants were transferred to growth chambers set for LD (14h light/10h dark) or SD (10h light/14h dark) treatment for one week prior to collection of tissue for expression analysis. In the growth chamber, daytime temperature was 30°C at a light intensity of ~300 $\mu mol \cdot s^{-1} \cdot m^{-2}$ and night (dark) temperature was 23°C with ~60% relative humidity. At day 39, the fully expanded portion of the top three leaves from three different plants were sampled from each genotype and photoperiod every 3 hours through a 24h light-dark cycle followed by 24h of continuous light (continuous 30°C). RNA was extracted from leaf tissues using TRI

Reagent (MRC) using the protocol for tissues with high polysaccharide content. RNA was cleaned up using RNeasy Mini Kits (QIAGEN), including DNA removal by on-column DNase I digestion. RNA integrity was examined on 1% MOPS gels. First-strand cDNA synthesis was performed using the SuperScript® III First-Strand Synthesis System (Invitrogen) with oligo dT and random hexamer primer mix. After first-strand cDNA synthesis, the reactions were diluted to a final concentration of 10ng/µl of the initial total RNA. Gene-specific qRT-PCR was performed using Power SYBR Green PCR Master Mix (Applied Biosystems). 18S rRNA was selected as internal control and detected using the TaqMan Universal PCR Master Mix (Applied Biosystems), rRNA Probe (VIC® Probe) and rRNA Forward/Reverse Primer. All reactions were run on a 7900HT PCR System with SDS v2.3 software (Applied Biosystems). The specificity of each qRT-PCR primer set was validated using melting temperature curve analysis. Amplification efficiency of each primer set was determined by the series dilution method [132], which can be calculated by the slope of the curve made from each Ct value and the dilution factor (Table 13). Relative expression was determined using the comparative cycle threshold (ΔΔCt) method with calibration using samples with the highest levels of RNA. The primer efficiency was employed to adjust data for relative quantification following the efficiency correction method [133]. Each expression data point was derived from analysis of three technical replicates within three biological replicates.

Table 13. Primer sequences and amplification efficiency of qRT-PCR (*SbCO*).

Gene	Locus ID in Sorghum*	Forward Primer	Reverse Primer
SbCO	Sb10g010050	CGGTACGGTTCTATGGTTCTG	AACTACTTGTACACTCGACCTTATC
SbEHD1	Sb01g019980	CGTCAGGGAAGCAATGTCCTTCAT	CTTCAGTTGGAAAGCACACATCGC
SbCN8	Sb09g025760	AACTGTCAAAGGGAAGGTGGATCG	GACTAAGCTCTCAACCCTTCAAGTC
SbCN12	Sb03g034580	TGCATGCATGAATATCGTCGTCT	CCCGGGTAGTACATATAAGGTGGT
SbCN15	Sb10g003940	GCTAGCTTATCCCGCATATTACCC	CCACCCAAACTGCATCCACTCTTGAA
SbGI	Sb03g003650	ATGCACCCGCTTCCTAGTCATCTT	TTCAGGGCTGTCATGGTTCCTCAT
SbTOC1	Sb04g026190	GAGTGCAGATGATTACTGCTCACTTTG	TGCTGCCTTGTTGCCAGTAGAAGA
SbLHY	Sb07g003870	GGCCTGCCTCTACCATGAAGTTTA	GCACTGCATTGCAAGGTTTGAAGTCC
	Amplification Efficiency in 100M		Amplification Efficiency in 58M
SbCO	0.82		0.79
SbEHD1	0.88		0.81
SbCN8	0.85		0.84
SbCN12	1.03		0.91
SbCN15	0.83		0.80
SbGI	0.85		0.84
SbTOC1	0.78		0.79
SbLHY	0.81		0.82

* Gene Locus IDs in Sorghum are from Phytozome v9.1 (http://www.phytozome.net/).

CHAPTER IV

CONCLUSIONS

Photoperiod sensitivity is a highly regulated process that has a large impact on plant adaption and reproductive success. Knowledge of photoperiod regulated flowering time in the C4 grass *Sorghum bicolor* is beneficial for grain production, biomass production, and also provides information on the evolutionary development of the photoperiod response. Like rice but unlike barley and wheat, sorghum is a SD plant. Sorghum shows extensive phenotypic diversity for photoperiod sensitivity due in part to the species wide latitudinal adaptation and selection for different end uses. Grain sorghum has been selected for early flowering to benefit grain production, whereas energy sorghum is designed for delayed flowering to provide a longer period of vegetative growth to improve biomass production. This phenotypic diversity has led researchers to identify six maturity loci (*Ma1-Ma6*) that control flowering time in sorghum. Four major loci (*Ma1-Ma4*) were discovered by Quinby and his colleagues in the 1960's. *Ma1* corresponds to *PSEUDORESPONSE REGULATOR PROTEIN (PRR37)* and *Ma3* encodes PHYTOCHROME B (PHYB). *PRR37* is homolog of *PHOTOPERIOD1 (PPD1)* in barley and functions as a central floral repressor in LD in sorghum. Two additional maturity loci (*Ma5* and *Ma6*) were identified by Rooney and Aydin in 1999, in which *Ma6* corresponds to *GHD7*, another floral repressor sensitive to photoperiod.

SbPRR37, corresponding to *Ma1,* has the largest influence on flowering time. This gene was proposed to act as the central floral repressor for *EHD1*, *SbCN15*, *SbCN8* and *SbCN12* delaying floral initiation in long days in sorghum. In addition to SbPRR37, GHD7 also increases photoperiod sensitivity by inhibiting *EHD1* and *SbCN8* in an additive manner. The external coincidence model suggests a physiological response, such as flowering time, is triggered when light perception coincides with the time when output from the circadian clock exceeds a required threshold. The expression pattern of *SbPRR37* and *GHD7* is consistent with the external coincidence model. This study concluded that the photoreceptor mediating light signaling that regulates SbPRR37 and GHD7, the central floral repressors, in response to photoperiod is phytochrome B. In LD, PhyB up-regulates *SbPRR37* and *GHD7* expression during the evening phase. PRR37 activates CO, a repressor of flowering in rice, at dawn and together with GHD7 represses floral inductors *EHD1*, *SbCN8*, *SbCN12* and *SbCN15*, leading to delayed flowering time in long days. In SD or in the 58M ($ma3^{R}ma3^{R}$) genetic background, there is a reduction in light induced expression of *SbPRR37*and *GHD7*, releasing sorghum from floral inhibition.

SbCO was found to be an activator of flowering in both LD and SD and contributes to photoperiod sensitivity through interaction with the LD repressor SbPRR37. This research identified four new major flowering time QTL in a BTx642 x TX7000 Recombinant Inbred Line (RIL) population and showed that the QTL on chromosome 10 corresponds to *Sorghum bicolor CONSTANS* (*SbCO*). Genetic analyses and gene expression studies demonstrate that SbCO is an activator of floral transition in

long and short days in genotypes lacking active *SbPRR37* and *GHD7* alleles. SbCO activates the floral transition by inducing *FT*-like gene expression independently of EHD1 and confers photoperiod sensitivity through interaction with SbPRR37. SbPRR37 blocks CO-mediated floral activation in long days. In non-inductive LD, PhyB up-regulates expression of *SbPRR37* especially during the evening, which inhibits the activity of two activators SbCO and EHD1. In contrast, without repression from SbPRR37 under inductive SD condition, SbCO promotes expression of *FT*-like genes, including *SbCN8*, *SbCN12* and *SbCN15*. Together with EHD1, CO activates flowering time in sorghum.

In summary, the current study provides a better understanding of how phytochrome B and SbCO regulates flowering time in response to variation photoperiod in sorghum. By manipulating these genes, transition from vegetative growth to flowering can be controlled by sorghum breeders, allowing production of photoperiod-insensitive grain sorghum or photoperiod-sensitive energy sorghum hybrids with higher yield of biomass.

REFERENCES

1. Dahlberg J, Berenji J, Sikora V, Latkovic D: **Assessing sorghum [*Sorghum bicolor* (L) Moench] germplasm for new traits: food, fuels & unique uses.** *Maydica* 2011, **56:**85-92.

2. Rooney WL, Blumenthal J, Bean B, Mullet JE: **Designing sorghum as a dedicated bioenergy feedstock.** *Biofuels, Bioproducts and Biorefining* 2007, **1:**147-157.

3. Doggett H: *Sorghum* (2nd Ed). New York: John Wiley; 1988.

4. Rosenow DT, Clark LE: **Drought and lodging resistance for a quality sorghum crop.** In *Proceedings of the 50th Annual Corn and Sorghum Industry Research Conference*; 6-7 Dec.; Chicago, IL. American Seed Trade Association; 1995: 82-97.

5. Xu W, Subudhi PK, Crasta OR, Rosenow DT, Mullet JE, Nguyen HT: **Molecular mapping of QTLs conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench).** *Genome* 2000, **43:**461-469.

6. Harris K, Subudhi PK, Borrell A, Jordan D, Rosenow D, Nguyen H, Klein P, Klein R, Mullet J: **Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence.** *Journal of Experimental Botany* 2007, **58:**327-338.

7. Smith CW, Frederiksen RA: **History of cultivar development in the United States: from "memoirs of A.B. Maunder-sorghum breeder".** In *Sorghum:*

Origin, History, Technology, and Production. Edited by Smith CW, Frederiksen RA. New York: John Wiley & Sons; 2000: 191-223

8. Doebley J, Durbin M, Golenberg EM, Clegg MT, Ma DP: **Evolutionary analysis of the large subunit of carboxylase (rbcL) nucleotide sequence among the grasses (*Gramineae*).** *Evolution* 1990, **44:**1097-1108.

9. Kimber CT: **Origins of domesticated sorghum and its early diffusion to India and China.** In *Sorghum: Origins, History, Technology, and Production.* Edited by Smith CW, Frederiksen RA. New York: John Wiley & Sons Inc.; 2000: 3-98

10. Stock R: *Africa South of the Sahara: A Geographical Interpretation* (Third Edition). New York: The Guilford Press; 2012.

11. Olson SN, Ritter K, Rooney W, Kemanian A, McCarl BA, Zhang Y, Hall S, Packer D, Mullet J: **High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops.** *Biofuels, Bioproducts and Biorefining* 2012, **6:**640-655.

12. Morgan PW, Finlayson SA: **Physiology and genetics of maturity and height.** In *sorghum: origin, history, technology, and production.* Edited by Smith CW, Frederiksen RA. New York: Wiley Series in Crop Science; 2000: 240-242

13. Quinby JR: *Sorghum Improvement and the Genetics of Growth.* College Station: Texas A&M University Press; 1974.

14. Dahlberg J: **Collection, conversion, and utilization of sorghum.** In *Sorghum: Origins, History, Technology, and Production.* Edited by Smith CW, Frederiksen RA. New York: John Wiley & Sons, Inc.; 2000: 309-328

15. Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, Dugas DV, Klein PE, Mullet JE: **Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108:**16469-16474.

16. Childs KL, Miller FR, Cordonnier-Pratt MM, Pratt LH, Morgan PW, Mullet JE: **The sorghum photoperiod sensitivity gene, *Ma3*, encodes a phytochrome B.** *Plant Physiology* 1997, **113:**611-619.

17. Quail PH, Briggs WR, Chory J, Hangarter RP, Harberd NP, Kendrick RE, Koornneef M, Parks B, Sharrock RA, Schafer E, et al: **Spotlight on phytochrome nomenclature.** *The Plant Cell* 1994, **6:**468-471.

18. Rooney W, Aydin S: **Genetic control of a photoperiod-sensitive response in *Sorghum bicolor* (L.) Moench.** *Crop Science* 1999, **39:**397-400.

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

20. Welch SM, Dong Z, Roe JL: **Modelling gene networks controlling transition to flowering in Arabidopsis.** In *Proceedings of the 4th International Crop Science Congress*; 26 Sep – 1 Oct Brisbane, Australia. CDROM; 2004: 1-20.

21. Greenup A, Peacock WJ, Dennis ES, Trevaskis B: **The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals.** *Annals of Botany* 2009, **103:**1165-1172.

22. Andres F, Coupland G: **The genetic basis of flowering responses to seasonal cues.** *Nature Reviews Genetics* 2012, **13:**627-639.

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

24. Mauseth JD: *Botany: An Introduction to Plant Biology* (5 edition). Sudbury, MA: Jones & Bartlett Learning; 2012.

25. Hamner KC, Bonner J: **Photoperiodism in relation to hormones as factors in floral initiation and development.** *Botanical Gazette* 1938, **100:**388-431.

26. Hamner KC: **Interrelation of light and darkness in photoperiodic induction.** *Botanical Gazette* 1940, **101:**658-687.

27. Song YH, Ito S, Imaizumi T: **Similarities in the circadian clock and photoperiodism in plants.** *Current Opinion in Plant Biology* 2010, **13:**594-603.

28. Imaizumi T: **Arabidopsis circadian clock and photoperiodism: time to think about location.** *Current Opinion in Plant Biology* 2010, **13:**83-89.

29. Nozue K, Covington MF, Duek PD, Lorrain S, Fankhauser C, Harmer SL, Maloof JN: **Rhythmic growth explained by coincidence between internal and external cues.** *Nature* 2007, **448:**358-361.

30. Jiao Y, Lau OS, Deng XW: **Light-regulated transcriptional networks in higher plants.** *Nature Reviews Genetics* 2007, **8:**217-230.

31. Kami C, Lorrain S, Hornitschek P, Fankhauser C: **Light-regulated plant growth and development.** In *Current Topics in Developmental Biology.* Volume 91: Elsevier Inc.; 2010: 29-66.

32. Palagyi A, Terecskei K, Adam E, Kevei E, Kircher S, Merai Z, Schafer E, Nagy F, Kozma-Bognar L: **Functional analysis of amino-terminal domains of the photoreceptor phytochrome B.** *Plant Physiology* 2010, **153:**1834-1845.

33. Castillon A, Shen H, Huq E: **Phytochrome interacting factors: central players in phytochrome-mediated light signaling networks.** *Trends in Plant Science* 2007, **12:**514-521.

34. Leivar P, Quail PH: **PIFs: pivotal components in a cellular signaling hub.** *Trends in Plant Science* 2011, **16:**19-28.

35. Pruneda-Paz JL, Kay SA: **An expanding universe of circadian networks in higher plants.** *Trends in Plant Science* 2010, **15:**259-265.

36. Niwa Y, Ito S, Nakamichi N, Mizoguchi T, Niinuma K, Yamashino T, Mizuno T: **Genetic linkages of the circadian clock-associated genes, *TOC1, CCA1* and *LHY*, in the photoperiodic control of flowering time in *Arabidopsis thaliana*.** *Plant & Cell Physiology* 2007, **48:**925-937.

37. Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA: **FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis.** *Science* 2005, **309:**293-297.

38. Fornara F, Panigrahi KC, Gissot L, Sauerbrunn N, Ruhl M, Jarillo JA, Coupland G: **Arabidopsis DOF transcription factors act redundantly to reduce *CONSTANS* expression and are essential for a photoperiodic flowering response.** *Developmental Cell* 2009, **17:**75-86.

39. Imaizumi T, Kay SA: **Photoperiodic control of flowering: not only by coincidence.** *Trends in Plant Science* 2006, **11:**550-558.

40. Lian HL, He SB, Zhang YC, Zhu DM, Zhang JY, Jia KP, Sun SX, Li L, Yang HQ: **Blue-light-dependent interaction of cryptochrome 1 with SPA1 defines a dynamic signaling mechanism.** *Genes & Development* 2011, **25:**1023-1028.

41. Turck F, Fornara F, Coupland G: **Regulation and identity of florigen: *FLOWERING LOCUS T* moves center stage.** *Annual Review of Plant Biology* 2008, **59:**573-594.

42. Ishikawa M, Kiba T, Chua NH: **The Arabidopsis *SPA1* gene is required for circadian clock function and photoperiodic flowering.** *The Plant Journal* 2006, **46:**736-746.

43. Valverde F: ***CONSTANS* and the evolutionary origin of photoperiodic timing of flowering.** *Journal of Experimental Botany* 2011, **62:**2453-2463.

44. Liu B, Zuo Z, Liu H, Liu X, Lin C: **Arabidopsis cryptochrome 1 interacts with SPA1 to suppress COP1 activity in response to blue light.** *Genes & Development* 2011, **25:**1029-1034.

45. Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G: **FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis.** *Science* 2007, **316:**1030-1033.

46. Tsuji H, Taoka K, Shimamoto K: **Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation.** *Current Opinion in Plant Biology* 2011, **14:**45-52.

47. Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T: **Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis fowering time gene *CONSTANS*.** *The Plant Cell* 2000, **12:**2473-2483.

48. Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K: **Hd3a protein is a mobile flowering signal in rice.** *Science* 2007, **316:**1033-1036.

49. Tsuji H, Taoka K, Shimamoto K: **Florigen in rice: complex gene network for florigen transcription, florigen activation complex, and multiple functions.** *Current Opinion in Plant Biology* 2013, **16:**228-235.

50. Itoh H, Izawa T: **The coincidence of critical day length recognition for florigen gene expression and floral transition under long-day conditions in rice.** *Molecular Plant* 2013, **6:**635-649.

51. Komiya R, Ikegami A, Tamaki S, Yokoi S, Shimamoto K: **Hd3a and RFT1 are essential for flowering in rice.** *Development* 2008, **135:**767-774.

52. Komiya R, Yokoi S, Shimamoto K: **A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice.** *Development* 2009, **136:**3443-3450.

53. Itoh H, Nonoue Y, Yano M, Izawa T: **A pair of floral regulators sets critical day length for Hd3a florigen expression in rice.** *Nature Genetics* 2010, **42:**635-638.

54. Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A: **Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1.** *Genes & Development* 2004, **18:**926-936.

55. Yan WH, Wang P, Chen HX, Zhou HJ, Li QP, Wang CR, Ding ZH, Zhang YS, Yu SB, Xing YZ, Zhang QF: **A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice.** *Molecular Plant* 2011, **4:**319-330.

56. Lee YS, Jeong DH, Lee DY, Yi J, Ryu CH, Kim SL, Jeong HJ, Choi SC, Jin P, Yang J, et al: **OsCOL4 is a constitutive flowering repressor upstream of Ehd1 and downstream of OsphyB.** *The Plant Journal* 2010, **63:**18-30.

57. Peng LT, Shi ZY, Li L, Shen GZ, Zhang JL: **Overexpression of transcription factor *OsLFL1* delays flowering time in *Oryza sativa*.** *Journal of Plant Physiology* 2008, **165:**876-885.

58. Ryu CH, Lee S, Cho LH, Kim SL, Lee YS, Choi SC, Jeong HJ, Yi J, Park SJ, Han CD, An G: **OsMADS50 and OsMADS56 function antagonistically in regulating long day (LD)-dependent flowering in rice.** *Plant, Cell & Environment* 2009, **32:**1412-1427.

59. Bentley AR, Jensen EF, Mackay IJ, Hönicka H, Fladung M, Hori K, Yano M, Mullet JE, Armstead IP, Hayes C, et al: **Flowering time.** In *Genomics and Breeding for Climate-resilient Crops.* Volume 2. Edited by Kole C. Verlag Berlin Heidelberg: Springer; 2013: 1-66.

60. Brambilla V, Fornara F: **Molecular control of flowering in response to day length in rice.** *Journal of Integrative Plant Biology* 2013, **55:**410-418.

61. Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K: **Variations in Hd1 proteins, Hd3a promoters, and Ehd1 expression levels contribute to diversity of flowering time in cultivated rice.** *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106:**4555-4560.

62. Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q: **Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice.** *Nature Genetics* 2008, **40:**761-767.

63. Cesari F: **Plant cell biology: Shedding light on plant growth.** *Nature Reviews Molecular Cell Biology* 2008, **9:**187-187.

64. Moglich A, Yang X, Ayers RA, Moffat K: **Structure and function of plant photoreceptors.** *Annual Review of Plant Biology* 2010, **61:**21-47.

65. Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K, Essen LO, van der Horst GT, Batschauer A, Ahmad M: **The cryptochromes: blue light photoreceptors in plants and animals.** *Annual Review of Plant Biology* 2011, **62:**335-364.

66. Christie JM: **Phototropin blue-light receptors.** *Annual Review of Plant Biology* 2007, **58:**21-45.

67. Rizzini L, Favory JJ, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schafer E, Nagy F, Jenkins GI, Ulm R: **Perception of UV-B by the Arabidopsis UVR8 protein.** *Science* 2011, **332:**103-106.

68. Lamparter T: **Evolution of cyanobacterial and plant phytochromes.** *FEBS Letters* 2004, **573:**1-5.

69. Rockwell NC, Su YS, Lagarias JC: **Phytochrome structure and signaling mechanisms.** *Annual Review of Plant Biology* 2006, **57:**837-858.

70. Mathews S, Sharrock R: **Phytochrome gene diversity.** *Plant, Cell & Environment* 1997, **20:**666-671.

71. Chen M, Chory J, Fankhauser C: **Light signal transduction in higher plants.** *Annual Review of Genetics* 2004, **38:**87-117.

72. Nagatani A: **Phytochrome: structural basis for its functions.** *Current Opinion in Plant Biology* 2010, **13:**565-570.

73. Oka Y, Matsushita T, Mochizuki N, Suzuki T, Tokutomi S, Nagatani A: **Functional analysis of a 450-amino acid N-terminal fragment of phytochrome B in Arabidopsis.** *The Plant Cell* 2004, **16:**2104-2116.

74. Khanna R, Huq E, Kikis EA, Al-Sady B, Lanzatella C, Quail PH: **A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors.** *The Plant Cell* 2004, **16:**3033-3044.

75. Leivar P, Monte E, Al-Sady B, Carle C, Storer A, Alonso JM, Ecker JR, Quail PH: **The Arabidopsis phytochrome-interacting factor PIF7, together with PIF3 and PIF4, regulates responses to prolonged red light by modulating phyB levels.** *The Plant Cell* 2008, **20:**337-352.

76. Kunihiro A, Yamashino T, Nakamichi N, Niwa Y, Nakanishi H, Mizuno T: **Phytochrome-interacting factor 4 and 5 (PIF4 and PIF5) activate the homeobox ATHB2 and auxin-inducible IAA29 genes in the coincidence mechanism underlying photoperiodic control of plant growth of *Arabidopsis thaliana*.** *Plant & Cell Physiology* 2011, **52:**1315-1329.

77. Lorrain S, Allen T, Duek PD, Whitelam GC, Fankhauser C: **Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors.** *The Plant Journal* 2008, **53:**312-323.

78. Niwa Y, Yamashino T, Mizuno T: **The circadian clock regulates the photoperiodic response of hypocotyl elongation through a coincidence mechanism in *Arabidopsis thaliana*.** *Plant & Cell Physiology* 2009, **50:**838-854.

79. de Lucas M, Daviere JM, Rodriguez-Falcon M, Pontin M, Iglesias-Pedraz JM, Lorrain S, Fankhauser C, Blazquez MA, Titarenko E, Prat S: **A molecular framework for light and gibberellin control of cell elongation.** *Nature* 2008, **451:**480-484.

80. Lau OS, Deng XW: **Plant hormone signaling lightens up: integrators of light and hormones.** *Current Opinion in Plant Biology* 2010, **13:**571-577.

81. Franklin KA, Quail PH: **Phytochrome functions in Arabidopsis development.** *Journal of Experimental Botany* 2010, **61:**11-24.

82. Robson F, Costa MM, Hepworth SR, Vizir I, Pineiro M, Reeves PH, Putterill J, Coupland G: **Functional importance of conserved domains in the flowering-time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants.** *The Plant Journal* 2001, **28:**619-631.

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

84. Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G: **CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis.** *The Plant Cell* 2006, **18:**2971-2984.

85. Li C, Distelfeld A, Comis A, Dubcovsky J: **Wheat flowering repressor VRN2 and promoter CO2 compete for interactions with NUCLEAR FACTOR-Y complexes.** *The Plant Journal* 2011, **67:**763-773.

86. Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E: **The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA.** *The Plant Journal* 2006, **46:**462-476.

87. Tiwari SB, Shen Y, Chang HC, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, et al: **The flowering time regulator CONSTANS is**

recruited to the *FLOWERING LOCUS T* promoter via a unique cis-element. *The New Phytologist* 2010, **187:**57-66.

88. Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA: **Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor.** *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109:**3167-3172.

89. Griffiths S, Dunford RP, Coupland G, Laurie DA: **The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis.** *Plant Physiology* 2003, **131:**1855-1867.

90. Ballerini ES, Kramer EM: **In the light of evolution: A reevaluation of conservation in the CO-FT regulon and its role in photoperiodic regulation of flowering time.** *Frontiers in Plant Science* 2011, **2:**81.

91. Chardon F, Damerval C: **Phylogenomic analysis of the PEBP gene family in cereals.** *Journal of Molecular Evolution* 2005, **61:**579-590.

92. Faure S, Higgins J, Turner A, Laurie DA: **The *FLOWERING LOCUS T*-like gene family in barley (*Hordeum vulgare*).** *Genetics* 2007, **176:**599-609.

93. Danilevskaya ON, Meng X, Hou Z, Ananiev EV, Simmons CR: **A genomic and expression compendium of the expanded PEBP gene family from maize.** *Plant Physiology* 2008, **146:**250-264.

94. Meng X, Muszynski MG, Danilevskaya ON: **The FT-like *ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize.** *The Plant Cell* 2011, **23:**942-960.

95. Distelfeld A, Li C, Dubcovsky J: **Regulation of flowering in temperate cereals.** *Current Opinion in Plant Biology* 2009, **12:**178-184.

96. Colasanti J, Coneva V: **Mechanisms of floral induction in grasses: something borrowed, something new.** *Plant Physiology* 2009, **149:**56-62.

97. Ishikawa R, Aoki M, Kurotani K, Yokoi S, Shinomura T, Takano M, Shimamoto K: **Phytochrome B regulates Heading date 1 (Hd1)-mediated expression of rice florigen Hd3a and critical day length in rice.** *Molecular Genetics and Genomics* 2011, **285:**461-470.

98. Koo BH, Yoo SC, Park JW, Kwon CT, Lee BD, An G, Zhang Z, Li J, Li Z, Paek NC: **Natural Variation in OsPRR37 Regulates Heading Date and Contributes to Rice Cultivation at a Wide Range of Latitudes.** *Molecular Plant* 2013, **6:**1877-1888.

99. Campoli C, Drosse B, Searle I, Coupland G, von Korff M: **Functional characterisation of HvCO1, the barley (*Hordeum vulgare*) flowering time ortholog of CONSTANS.** *The Plant Journal* 2012, **69:**868-880.

100. Turner A, Beales J, Faure S, Dunford RP, Laurie DA: **The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley.** *Science* 2005, **310:**1031-1034.

101. Nakamichi N, Kita M, Ito S, Sato E, Yamashino T, Mizuno T: **The Arabidopsis pseudo-response regulators, PRR5 and PRR7, coordinately play essential roles for circadian clock function.** *Plant & Cell Physiology* 2005, **46:**609-619.

102. Beales J, Turner A, Griffiths S, Snape JW, Laurie DA: **A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*Triticum aestivum* L.).** *Theoretical and Applied Genetics* 2007, **115:**721-733.

103. Trevaskis B, Hemming MN, Peacock WJ, Dennis ES: **HvVRN2 responds to daylength, whereas HvVRN1 is regulated by vernalization and developmental status.** *Plant Physiology* 2006, **140:**1397-1405.

104. Trevaskis B, Hemming MN, Dennis ES, Peacock WJ: **The molecular basis of vernalization-induced flowering in cereals.** *Trends in Plant Science* 2007, **12:**352-357.

105. Wang CL, Cheng FF, Sun ZH, Tang JH, Wu LC, Ku LX, Chen YH: **Genetic analysis of photoperiod sensitivity in a tropical by temperate maize recombinant inbred population using molecular markers.** *Theoretical and Applied Genetics* 2008, **117:**1129-1139.

106. Ducrocq S, Giauffret C, Madur D, Combes V, Dumas F, Jouanne S, Coubriche D, Jamin P, Moreau L, Charcosset A: **Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10.** *Genetics* 2009, **183:**1555-1563.

107. Miller TA, Muslin EH, Dorweiler JE: **A maize CONSTANS-like gene, *conz1*, exhibits distinct diurnal expression patterns in varied photoperiods.** *Planta* 2008, **227:**1377-1388.

108. Higgins JA, Bailey PC, Laurie DA: **Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses.** *Plos One* 2010, **5:**e10065.

109. Huijser P, Schmid M: **The control of developmental phase transitions in plants.** *Development* 2011, **138:**4117-4129.

110. Poethig RS: **Small RNAs and developmental timing in plants.** *Current Opinion in Genetics & Development* 2009, **19:**374-378.

111. Zhu QH, Helliwell CA: **Regulation of flowering time and floral patterning by miR172.** *Journal of Experimental Botany* 2011, **62:**487-495.

112. Alabadi D, Blazquez MA: **Molecular interactions between light and hormone signaling to control plant growth.** *Plant Molecular Biology* 2009, **69:**409-417.

113. Murphy RL, Morishige DT, Brady JA, Rooney WL, Yang S, Klein PE, Mullet JE: **Ghd7 (Ma6) represses flowering in long days: a dey trait in energy sorghum hybrids.** *The Plant Genome* 2014, **In press**.

114. Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K: **Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice.** *Genes & Development* 2002, **16:**2006-2020.

115. Hanumappa M, Pratt LH, Cordonnier-Pratt MM, Deitzer GF: **A photoperiod-insensitive barley line contains a light-labile phytochrome B.** *Plant Physiology* 1999, **119:**1033-1040.

116. Yang S, Weers B, Morishige D, Mullet J: **CONSTANS is a photoperiod regulated activator of flowering in sorghum.** *BMC Plant Biology* 2014, **14:**148.

117. Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J: **The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization.** *Science* 2004, **303:**1640-1644.

118. Mace ES, Hunt CH, Jordan DR: **Supermodels: sorghum and maize provide mutual insight into the genetics of flowering time.** *Theoretical and Applied Genetics* 2013, **126:**1377-1395.

119. Morishige DT, Klein PE, Hilley JL, Sahraeian SM, Sharma A, Mullet JE: **Digital genotyping of sorghum - a diverse plant species with a large repeat-rich genome.** *BMC Genomics* 2013, **14:**448.

120. Takano M, Inagaki N, Xie X, Yuzurihara N, Hihara F, Ishizuka T, Yano M, Nishimura M, Miyao A, Hirochika H, Shinomura T: **Distinct and cooperative functions of phytochromes A, B, and C in the control of deetiolation and flowering in rice.** *The Plant Cell* 2005, **17:**3311-3325.

121. Monte E, Alonso JM, Ecker JR, Zhang Y, Li X, Young J, Austin-Phillips S, Quail PH: **Isolation and characterization of *phyC* mutants in Arabidopsis reveals complex crosstalk between phytochrome signaling pathways.** *The Plant Cell* 2003, **15:**1962-1980.

122. Clack T, Shokry A, Moffet M, Liu P, Faul M, Sharrock RA: **Obligate heterodimerization of Arabidopsis phytochromes C and E and interaction with the PIF3 basic helix-loop-helix transcription factor.** *The Plant Cell* 2009, **21:**786-799.

123. Kumar P, Henikoff S, Ng PC: **Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.** *Nature Protocal* 2009, **4:**1073-1081.

124. Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua NH, Sakakibara H: **PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock.** *The Plant Cell* 2010, **22:**594-605.

125. Vigouroux Y, Mariac C, De Mita S, Pham JL, Gerard B, Kapran I, Sagnard F, Deu M, Chantereau J, Ali A, et al: **Selection for earlier flowering crop associated with climatic variations in the Sahel.** *Plos One* 2011, **6:**e19563.

126. Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J, Weigel D: **The phytochrome C photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*.** *Nature Genetics* 2006, **38:**711-715.

127. Distelfeld A, Dubcovsky J: **Characterization of the maintained vegetative phase deletions from diploid wheat and their effect on VRN2 and FT transcript levels.** *Molecular Genetics and Genomics* 2010, **283:**223-232.

128. Osugi A, Itoh H, Ikeda-Kawakatsu K, Takano M, Izawa T: **Molecular dissection of the roles of phytochrome in photoperiodic flowering in rice.** *Plant Physiology* 2011, **157:**1128-1137.

129. Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, Chen L, Yu L, Iglesias-Pedraz JM, Kircher S, et al: **Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins.** *Nature* 2008, **451:**475-479.

130. Wang S, Basten CJ, Zeng Z-B: **Windows QTL Cartographer 2.5.** In *Windows QTL Cartographer 2.5.* City; 2012.

131. Churchill GA, Doerge RW: **Empirical threshold values for quantitative trait mapping.** *Genetics* 1994, **138:**963-971.

132. Bookout AL, Mangelsdorf DJ: **Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways.** *Nuclear Receptor Signaling* 2003, **1:**1-7.

133. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Research* 2001, **29:**2002-2007.

134. Jackson SD: **Plant responses to photoperiod.** *The New Phytologist* 2009, **181:**517-531.

135. Evans J, McCormick RF, Morishige D, Olson SN, Weers B, Hilley J, Klein P, Rooney W, Mullet J: **Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes.** *Plos One* 2013, **8:**e79192.

136. Murray SC, Sharma A, Rooney WL, Klein PE, Mullet JE, Mitchell SE, Kresovich S: **Genetic improvement of sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates.** *Crop Science* 2008, **48:**2165-2179.