SORGHUM Ma_5 AND Ma_6 MATURITY GENES

A Dissertation

by

JEFFREY ALAN BRADY

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

DOCTOR OF PHILOSOPHY

May 2006

Major Subject: Genetics

SORGHUM Ma_5 AND Ma_6 MATURITY GENES

A Dissertation

by

JEFFREY ALAN BRADY

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

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	John E. Mullet
	Forrest L. Mitchell
Committee Members,	Patricia E. Klein
	Alan Pepper
	William L. Rooney
Chair of Genetics Faculty,	James R. Wild

May 2006

Major Subject: Genetics

ABSTRACT

Sorghum *Ma*$_5$ and *Ma*$_6$ Maturity Genes. (May 2006)

Jeffrey Alan Brady, B.A., Tarleton State University;

M.S., Tarleton State University

Co-Chairs of Advisory Committee: Dr. John Mullet
Dr. Forrest Mitchell

The *Ma*$_5$ and *Ma*$_6$ maturity loci in sorghum contain genes interacting epistatically to block flowering until an appropriate daylength is met. Because sorghum is a crop of tropical origin, its critical daylength is close to 12 hours. Sorghums with dominant alleles at these two loci are photoperiod sensitive, extremely late flowering, and ill-suited to cultivation in the temperate U.S. Most sorghum lines grown in the U.S. have been converted to photoperiod insensitive plants that have recessive mutations at the *ma*$_6$ locus. This work describes ongoing efforts to clone the genes responsible for the *Ma*$_5$/*Ma*$_6$ –controlled late flowering response in sorghum. To reach this goal, the two loci were mapped with AFLP and SSR markers that were part of an integrated genetic, physical, and cytogenetic map of the sorghum genome. Genetic markers have been linked to both the *Ma*$_5$ and *Ma*$_6$ loci on chromosomes 2 and 6, respectively. BAC libraries have been screened to identify numerous BACs associated with each locus. Additional work to fine-map each locus and identify potential candidate genes by comparison with the rice genome is ongoing.

ACKNOWLEDGEMENTS

I would like to thank my co-chairs, John Mullet and Forrest Mitchell, for their guidance and financial support in this project. The members of my committee, Patricia Klein, Alan Pepper, and Bill Rooney, each supported this project in different ways, and I express my gratitude for their help. Members of a large sorghum research community at Texas A&M were helpful in various ways during the work on this project, and during my graduate studies in general. This community included: Daryl Morishige, Monica Menz, Ron Salzman, Christina Buchanan, Liz Summer, Kevin Childs, Page Morgan, Scott Finlayson, Bob Klein, Jeong-Soon Kim, David Stelly, Natalie Unruh, Julie McCollum, Jacque Obert, Lucy Haizlip, and fellow graduate students Karen Harris, Sang-Hyun Lim, and Bin Zhou. I also thank the Texas A&M Board of Regents for financial support during my first year of graduate studies. John Calahan and Harold Rathburn started me on this path, and for that I am grateful. Above all, I thank my wife Stacy for her love and support, and my sons Nick and Nate for their enduring patience during these studies.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

AFLP	Amplified Fragment Length Polymorphism
BAC	Bacterial Artificial Chromosome
PI	Photoperiod Insensitive
PS	Photoperiod Sensitive
LD	Long Day
SD	Short Day
PCR	Polymerase Chain Reaction
SSR	Simple Sequence Repeat
FISH	Fluorescence *in situ* Hybridization
STS	Sequence Tagged Site
EST	Expressed Sequence Tag

CHAPTER I

INTRODUCTION: A REVIEW OF FLOWERING PATHWAYS

INTRODUCTION

Sorghum bicolor

Sorghum (*Sorghum bicolor* [L.] Moench) is a small-grain cereal crop native to the semi-arid tropical and subtropical regions of northern Africa[1,2]. In the U.S., sorghum is grown primarily as an animal feed and forage crop, but in portions of Africa, India, and Asia, it is grown as a staple crop for sustaining the human population. In these areas every part of the plant is often used as food, fodder, shelter, or even for the production of beer. Regardless of where it is grown, sorghum is particularly favored as a dryland crop due to its notable drought tolerance[3], and because of this valuable trait, sorghum is most often grown under non-irrigated conditions[4]. In spite of its perception as a low-input subsistence crop, in 2005 sorghum was the 5th most important cereal crop worldwide in terms of metric tonnage produced (http://faostat.fao.org).

In addition to its agronomic importance, sorghum is becoming increasingly important as a potential bridge between rice and other large genome cereals species in comparative cereal genomics. Rice (tribe Oryzae) has the smallest genome of the major cereals (389 Mb), and its sequenced genome[5-7] serves as a model for studying the

---

This thesis follows the style of *Nature Genetics.*

genomes of other cereals. The sorghum genome is intermediate in size (818 Mb)[8,9] between rice and other large genome cereals such as maize (~2,500 Mb)[10] and sugarcane (~3,000 Mb)[10], and sorghum is more closely related to maize and sugarcane (all are in the tribe Andropogoneae) than is rice. Comparative studies of isolated regions of the rice genome with maize and sorghum have shown that microcolinearity (preservation of gene content and gene order) can exist between rice and the maize/sorghum lineage, but that frequent microstructural rearrangements are common, with the differences in genome sizes being largely attributable to an increasing number of retroelements in the larger genome species[11-15].

This work describes efforts to isolate genes controlling the initiation of flowering in sorghum, genes that have a direct and profound impact on productivity of the crop. The two genes at the center of this work, *MATURITY*$_5$ (*Ma*$_5$) and *MATURITY*$_6$ (*Ma*$_6$), interact epistatically to repress flowering until a critical daylength is met. The efforts to isolate these genes utilized map-based, cytogenetic, and comparative approaches.

Sorghum flowering responses

Tropical sorghum varieties and other tropical plants have adapted to their native environment by timing the reproductive process so that it coincides with the end of the local rainy season, which in the tropics is fairly constant from year to year[16]. Plants using this drought-avoidance strategy to adapt to a given locality are able to synchronize flowering at the correct time of year by sensing small changes in photoperiod. Evidence of local adaptation to photoperiod is seen in the strong relationship between the degree

of photoperiod sensitivity and the latitude of origin of different sorghum varieties[17]. Varieties used for cultivation in new latitudes have been selected for photoperiod sensitivities appropriate to the environment in which they are grown. Indeed, when sorghum was first introduced into the U.S., it was still tropically adapted[18,19]. Since sorghum is a short day (SD) species[20], most of the sorghum introduced into the U.S. would not flower until daylength was close to 12 hours, which occurs very late in the growing season in the temperate southern U.S. Thus, sorghum planted in early spring is reported to have flowered up to 8 1/2 months later, producing "giant milo," perhaps up to 5 meters tall, that was very susceptible to lodging in wind and rain[18,19]. In order to produce sorghum adapted to temperate cultivation, the tall, photoperiod sensitive (PS) phenotype was converted to a shorter, earlier-maturing photoperiod insensitive (PI) phenotype. This was accomplished by selecting for mutations in several maturity (*Ma*) genes that influence the time of floral initiation, two of which are the focus of this dissertation, and by selecting for mutations in several genes influencing internode length (*Dw* genes). Before they can be grown in the U.S., valuable tropically adapted sorghum varieties must first go through a conversion program to introduce recessive maturity alleles so that the plants will flower in a timely manner in a temperate latitude[21]. After conversion, the resulting varieties are photoperiod insensitive, earlier flowering, shorter, and thus suitable for combine harvesting when grown in temperate latitudes.

Sorghum maturity genes

Six maturity genes (*Ma1-Ma6*) have been described in sorghum to date. Recessive mutations in *Ma1, Ma2*, and *Ma3* had all been discovered by 1911 in the U.S. milo sorghums, and the resulting earlier-flowering PI plants were selected and increased by growers[19]. The inheritance of these 3 genes and their effect on maturity was first described by Quinby and Karper [22]. The *Ma4* locus was described years later[23], and most recently, *Ma5* and *Ma6* have been genetically characterized[24,25]. There are numerous interactions among the 6 maturity loci, and dominant alleles at all 6 loci delay flowering with one exception: in the presence of recessive *ma1*, dominant *Ma2* causes earlier flowering[26]. Among the first 4 loci, *Ma1* causes the largest delay in flowering time, and causes extreme lateness when coupled with dominant *Ma2*. Both *Ma2* and *Ma4* have been shown to be temperature sensitive[23,27]. In the case of *Ma5* and *Ma6*, both loci must be dominant in order to significantly delay flowering. In almost all crosses, when both *Ma5* and *Ma6* are dominant, flowering is delayed, regardless of the constitution of the first four loci, until daylength is less than 12 hours and 20 minutes[25]. Depending on the planting date, the dominant *Ma5/Ma6* interaction can more than double the delay in flowering caused by dominant alleles at the first 4 *Ma* loci. An allelic series exists at each of the first 4 *Ma* loci[26,28,29], and discovery of one *Ma3* allele turned out to be particularly fortuitous. The recessive Ryer allele, ma_3^R, was much earlier flowering than recessive *ma3*[29]. Phenotypic similarity between ma_3^R/ma_3^R sorghum and plants treated with exogenous gibberellic acid[30,31] and altered levels of gibberellic acid in ma_3^R/ma_3^R plants[32] led to the discovery that a key regulator of gibberellic acid metabolism--

phytochrome--is also abnormal in ma_3^R/ma_3^R plants[33]. The *Ma₃* gene was cloned using a candidate gene and approach and shown to be *PHYTOCHROME B* (*PHYB*)[34]. This is the only sorghum maturity locus for which a gene has been identified.

Sorghum maturity genotypes that vary at the *Ma₁* to *Ma₄* loci have been previously examined for photoperiod responses, both with monthly plantings in Puerto Rico and with growth chamber experiments[27,35]. Both studies found that *Ma₁* was involved in photoperiodic flowering response, while *Ma₂, Ma₃, and Ma₄* were not. As mentioned above, the *Ma₃* gene is *PHYB*, a gene that is obviously connected to photoperiod sensing. The fact that it was not identified as involved in photoperiodic flowering in these studies is surprising. The monthly plantings in Puerto Rico by Miller, *et al.* compared the *Ma₃* allele with the *ma₃* allele[35]. The effect of the *ma₃* mutation is so slight that it was statistically indistinguishable from the effect of the *Ma₃* allele. The growth chamber experiments of Major, *et al.* compared *Ma₃, ma₃,* and the Ryer allele, ma_3^R, which has an extreme effect on flowering time, and allowed cloning of the *Ma₃* gene[34]. The effects of the ma_3^R allele were so extreme in comparison to the *Ma₃* allele that they were interpreted as a shortening of the basic vegetative phase (BVP), the period in which a juvenile plant is insensitive to changes in photoperiod. Whether or not the photoperiod-sensitive maturity loci *Ma₁* and *Ma₃* are functionally connected in any way to *Ma₅* and *Ma₆* remains unclear, but in crosses of EBA-3 (an Argentinean forage/grain sorghum) with sorghum maturity standards, the progeny flowered at about 175 days, regardless of the dominant or recessive constitution of the *Ma₁* and *Ma₃* loci[24]. The crosses made with maturity genotypes show that when both *Ma₅* and *Ma₆* are dominant,

sorghum has an obligate requirement for short photoperiods in order to flower, regardless of which alleles are present at the *Ma1* to *Ma4* loci[24,25].

Sorghum is a reference C4 grass species for genome studies, and the tools to map and clone the genes responsible for important agronomic traits like flowering time have recently been developed[36]. A high density integrated genetic and physical map of sorghum based on AFLP[37] data has been created that incorporates previous SSR and RFLP data[38] so that comparisons can be made with other cereal crops, facilitating gene discovery. The many thousands of BAC clones comprising the current sorghum physical map have been pooled in a six-dimensional cube so that an efficient screening strategy may be employed to link BACs to the genetic map[39]. Additionally, cytogenetic tools for identifying each of the 10 chromosomes and placing a BAC on a specific chromosome are now available[9]. These newly developed tools have been used in this work in an attempt to map-based clone the *Ma5* and *Ma6* sorghum maturity loci.

Pathways regulating flowering

Flowering is the event of central importance in the life cycle of a plant, because it determines whether the genetic complement of a given plant will be passed on or will become a dead end. Given the importance of flowering, it is no surprise that plants possess multiple pathways used to sense both internal and external cues so that the process of flowering can occur at a time when the plant will have the highest chance of producing viable progeny.

Even before the molecular tools to dissect the multiple pathways affecting flowering became available, early plant scientists hypothesized that external or internal cues, or both, might influence when plants flowered. Garner and Allard were the first to show that the duration of exposure to light, referred to as daylength or photoperiod, was the most important factor influencing initiation of flowering, while temperature, light intensity, and wavelength of light also play a role in influencing flowering[40]. This makes sense because photoperiod is the most constant and noise-free environmental cue that could be used to determine time of year. Garner and Allard separated plant species into different groups based on how they responded to photoperiod. Plants that require more hours of daylight than a given critical daylength in order to flower were called long-day (LD) plants, while plants that require less hours of daylight than a given critical daylength were called short-day (SD) plants[40], a third group of plants do not show critical daylength requirements and are referred to as day-neutral plants. In the process of characterizing a number of plants for their photoperiod requirements, Garner and Allard were the first to show that sorghum was a SD plant[20]. They also made the connection between latitude and photoperiod sensitivity in plants. Since daylength varies least near the equator, and varies more at higher latitudes according to the season, they hypothesized that plants of the same species growing at different latitudes may have some physiological differences allowing them to sense changes in daylength and time the flowering process appropriately[40]. Plants of tropical origin are most often SD plants, some of which initiate flowering due to very small changes in daylength (10-20 minutes), while plants of temperate origin are most often LD plants[16]. These differences

in photoperiod response allow tropical or temperate plants to flower under the most favorable conditions, such as the end of the rainy season or in a period of optimal temperature and irradiance, respectively.

Different models of how plants might control the induction of flowering were developed early in the 20th century. The model that is consistent with our current understanding of this regulatory system was developed initially by Bünning, who was the first to suggest that the same mechanisms controlling photoperiodic leaf movement in plants may also sense seasonal changes and thus control flowering responses. His hypothesis was extended and formalized into what is now called the external coincidence model by Pittendrigh and Minnis (reviewed in [41]). In brief, this model holds that external cues are sensed by the plant and that these external cues interact with an internal clock, allowing time measurement and proper control of various plant functions. The external cue in the model is sunlight, and it serves dual functions in the control of flowering. First, it synchronizes or entrains the internal clock, so that the circadian rhythm of internal clock components begins with dawn. Second, depending on length of photoperiod, sunlight either will or will not interact with an internal clock component that is present or absent at various points in the circadian cycle, and either will or will not induce the transition from vegetative to reproductive growth depending on the presence or absence of this interaction[41]. While the photoreceptor phytochrome (described below) had been discovered and was incorporated into Pittendrigh and Minnis' model, discovery of most of the individual components of both the circadian

clock and flowering-time pathways required molecular genetic techniques of later decades.

Plant scientists in the mid 20th century were able to localize the source of the developmental signal to flower. Using plants that required as little as one inductive photoperiod to induce flowering, exposing even a single leaf to an inductive photoperiod could cause the plant to flower even if the rest of the plant were kept in non-inductive conditions. Additionally, grafting a single leaf exposed to photoinductive conditions onto a plant kept in non-inductive conditions could cause the plant to flower. The induction of flowering worked even if the grafted leaf was from a different species, or was grafted between SD and LD plants, suggesting some universality to the floral signal (reviewed in [42]). Since the shoot apical meristem (SAM) was caused to differentiate from vegetative to reproductive growth due to a distant signal generated in leaves, physiologists began searching for a plant hormone, or florigen, that could be synthesized or made active in leaves and that could account for the graft-transmissible properties of the floral-inducing substance. An early candidate for florigen was gibberellic acid (GA_3). GA_3 has been found to cause flowering in a few species, and to hasten flowering in many species, but in other species it has little effect on flowering. Other hypotheses about the nature of the leaf-generated signal included: a balance of florigen/anti-florigen, changes in source-sink relationships, and a multifactorial system of numerous inducers[43]. The fact that no universal promoter of flowering has been found while numerous changes in long-distance signaling are seen in induced plants supports a multifactorial model for the physiological promotion of flowering[43]. Recent demonstration that the

mRNA encoded by the flowering gene *FT* can move from leaves to the shoot apex and induce flowering may provide insight into the molecular basis of florigen[44]. Along with the information about the physiological changes in plants initiating flowering, a great deal is known about the genes involved in flowering-time pathways. This information was generated in large part by studying the model plant *Arabidopsis thaliana*.

The control of flowering in *Arabidopsis thaliana*

A. thaliana is a small plant from the mustard family (Brassicaceae) with a number of features that make it amenable to study. It is widely distributed, with many ecotypes showing adaptation to particular environments. Several rapid-cycling laboratory strains exist that can prolifically produce seed in about 6 weeks while occupying a limited space. It has a relatively small (125 Mb), fully-sequenced genome, with numerous genetic, physical, and cytogenetic resources, including a large number of mutant lines created by various methods. It is easily transformable, and as the center of research for a large number of laboratories, many protocols are readily available for its manipulation.

As regards flowering time, *A. thaliana* is a facultative LD plant, with a requirement for vernalization, or cold treatment, in order to flower most rapidly. Numerous mutants in various flowering-time pathways have been created or identified[45], and these mutants have been used to dissect the signaling pathways that bring about the developmental switch from vegetative to reproductive growth. The genes affecting flowering time in *A. thaliana* are most often placed in one of 4 pathways involved in the switch from vegetative to reproductive growth. These pathways are the photoperiod

pathway, involving responses to changes in daylength; the vernalization pathway, involving responses to prolonged cold temperatures; the autonomous pathway, involving responses that had in the past been described as unlinked to environmental cues, and hence autonomous; and the gibberellin pathway, involving responses to changes in gibberellin levels. The 4 pathways converge to regulate the same set of floral integrator genes described below.

The photoperiod (or long-day) pathway

The photoperiod pathway in *A. thaliana* is a system involved in sensing and responding to photoperiod. The input in the pathway is light energy, which is detected by various photoreceptors. The photoreceptors can interact with circadian clock components to generate output that can serve as a stimulus to flower.

Plant photoreceptors

Plants monitor their light environment by photoreceptors that fall into several different classes: phytochromes, cryptochromes, phototropins, and zeitlupe family photoreceptors. The phytochromes respond predominantly to the red and far-red portions of the electromagnetic spectrum, while the cryptochromes, phototropins, and zeitlupe photoreceptors respond predominantly to the blue/UV portion[46,47].

Phytochromes show homology to prokaryotic, two-component response regulators[48]. Phytochromes are large (120 kDa each monomer), soluble, dimeric proteins characterized by an N-terminal photosensory domain containing covalently

bound tetrapyrrole chromophores, two PAS-related domains (Period circadian protein, Aryl hydrocarbon receptor nuclear translocator protein, and Single-minded protein) involved in protein-protein interaction[49-51], and a histidine kinase-related domain[52]. The biological activity of this molecule is a result of its ability to undergo a light-inducible reversible conformational change between two forms, a biologically inactive form that absorbs red light (Pr), and a biologically active form that absorbs far-red light (Pfr). In darkness, the Pfr form reverts back to the Pr form, so that most phytochrome is in the Pr form at night, while sunlight converts most phytochrome into the active Pfr form during the day[53].

Upon illumination, cytoplasmically localized Pr is converted to Pfr, initiating a signaling cascade that begins with translocation of some of the Pfr pool into the nucleus[54]. Once in the nucleus, phytochrome interacts with a protein identified as an interaction partner in yeast two-hybrid screens, PHYTOCHROME INTERACTING FACTOR3 (PIF3). PIF3 is a basic, helix-loop-helix transcription factor constitutively localized in the nucleus[55] that binds to cis-regulatory promoter elements known as G-boxes that are present in several light regulated genes[56]. Two of the genes activated by the PHYTOCHROME/PIF3 complex are central components of the *A. thaliana* circadian clock, *CCA1* and *LHY*[56], described in association with the circadian clock, below.

The phytochromes are a family of related photoreceptors, and *A. thaliana* has 5; PHYA-PHYE, while sorghum and other monocots have 3; PHYA-PHYC[57]. There are some unique and some overlapping functions for the different phytochromes. In *A. thaliana*, PHYA controls seed germination, cotyledon expansion, and hypocotyl

elongation, while PHYB is involved in hypocotyl elongation, flowering time, leaf morphology, and shade avoidance responses[57]. PHYD is closely related to PHYB, and seems to play a lesser role in the same responses[57]. PHYD and PHYE have been shown to affect flowering more prominently at low temperatures[58], and have other phenotypic effects in different photoperiods. PHYC plays a role in hypocotyl elongation and leaf expansion[59]. Taken together, the phytochromes serve partially overlapping and partially divergent functions. Of particular note with regard to flowering time, the phytochromes serve to entrain the circadian clock, and some phytochrome-null mutants have altered flowering-time phenotypes. In addition to clock entrainment, phytochromes play a role in regulating flowering time through regulation of the *CO* gene (discussed below). The importance of the phytochromes with regard to flowering time is highlighted by the fact that the only maturity gene cloned thus far in sorghum is the Ma_3 gene, shown to be equivalent to *PHYB*[34].

The cryptochromes are another class of plant photoreceptors involved in sensing photoperiod[60]. Three cryptochromes are present in *A. thaliana* (CRY1,2, DASH)[61], while monocots possess two (CRY1,2). Whereas the phytochromes sense light in the red end of the spectrum, the cryptochromes sense light in the blue/UV end of the spectrum. The cryptochromes show homology to bacterial DNA photolyases[62]. Cryptochromes are characterized by an N-terminal domain that binds two chromophores, a flavin and a pterin[62], and a C-terminal domain shown to mediate cryptochrome light responses[63]. While the phytochromes act at the transcriptional level to control plant responses to photoperiod, cryptochromes act post-translationally, targeting proteins for

ubiquitination and degradation via proteasome pathways[64]. The cryptochrome C-terminal domain (CCT) interacts with CONSTITUTIVE PHOTOMORPHOGENESIS1 (COP1)[63,65], a protein with E3 ubiquitin ligase function, and the complex ubiquitinates and targets transcription factors like LONG HYPCOTYL5 (HY5)[66] for degradation. Additionally, cryptochromes act redundantly with other photoreceptors to entrain the circadian clock[67,68], and thus are part of the system that determines when to flower. Mutant *cry* alleles, particularly *cry2* alleles, are late-flowering in inductive photoperiods[45]. While the cryptochromes and phytochromes are important in entraining the circadian clock, they are not fully responsible for entrainment. Quadruple *phyAphyBcry1cry2 A. thaliana* mutants retain some responsiveness to light signals[69], so there are other photoreceptors providing input to the circadian clock.

Besides the phytochromes and cryptochromes, there are plant photoreceptors called phototropins, as well as a family of photoreceptors with homology to phototropins. The phototropins themselves are blue light sensing photoreceptors and have not been implicated in the control of flowering-time or the circadian clock, and so will not be described here. The Zeitlupe family of photoreceptors has some similarity to the phototrophins and has been shown to be involved in clock entrainment and alteration of flowering time. This group of photoreceptors includes ZEITLUPE (ZTL)[70]; FLAVIN-BINDING, KELCH REPEAT, F-BOX1 (FKF1)[71]; and LOV KELCH PROTEIN2 (LKP2)[72,73]. These photoreceptors share motifs, including an N-terminal PAS/LOV domain that may serve to bind a flavin chromophore, an F-box domain that may be involved in a proteosome pathway, and C-terminal kelch repeats that may be

involved in protein-protein interactions. The mode of function for these photoreceptors could be by changing transcription of a clock component, as with the phytochromes, or by targeting clock components for degradation as with the cryptochromes[41]. Regardless of their mode of action, these photoreceptors, along with the phytochromes and cryptochromes, help modulate the central circadian oscillator that serves a number functions, including measuring time of year.

The *A. thaliana* circadian clock

While *A. thaliana* doesn't possess homologs of genes previously described in the circadian clocks of *Neurospora*, mouse, etc., it does possess proteins with key features of circadian clock components[74]. These proteins show the characteristic autoregulatory transcriptional and translational feedback loops associated with the circadian clock components of other species. While the precise mechanism of function for the *A. thaliana* circadian clock is not fully established, at least three genes appear to be involved in the central circadian oscillator mechanism. These genes are *TIMING OF CHLOROPHYLL A/B BINDING PROTEIN 1* (*TOC1*), *LATE ELONGATED HYPOCOTYL* (*LHY*), and *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*). *CCA1* and *LHY* encode MYB-transcription factors that peak in expression shortly after dawn. CCA1 has been shown to require phosphorylation by the CK2 complex[75] in order for the clock to function correctly[76,77]. Like the central clock components in other species, they are part of an autoregulatory loop, each downregulating the expression of both *CCA1* and *LHY*[78,79]. *A. thaliana* with null alleles or with RNAi knockouts of either gene show

altered circadian rhythmicity and early flowering, and double mutant *cca1/lhy* plants show more severe circadian phenotypes. Additionally, the circadian oscillations of these plants disappear after a couple of days in either the light or the dark[80,81], indicating that they are not merely a means of light input to the clock. Both genes contain a G-box in their promoters that serves as a binding site for PIF3[56]. When phytochrome is exposed to light at dawn, it undergoes a rapid change from inactive Pr to the active Pfr form. Pfr is transported to the nucleus where it interacts with PIF3 to increase transcription of *CCA1* and *LHY*, and this is likely what sets the circadian clock at dawn[41]. Another protein modulates the expression of *CCA1* and *LHY*, and the gene encoding it is a target for repression by CCA1 and LHY proteins. The protein is an *A. thaliana* pseudo response regulator (APRR) called TIMING OF CHLOROPHYLL A/B BINDING PROTEIN 1 (TOC1), and its transcript levels peak in the evening, opposite that of *CCA1/LHY*. *TOC1* and a number of other genes contain an evening element in their promoters that serves as a site for binding and repression by CCA1 and LHY. In a model of how the *A. thaliana* circadian oscillator may work, at dawn TOC1 and light signals augment expression of *CCA1/LHY* transcripts. CCA1/LHY proteins then activate the expression of genes needed in daylight (*CAB*) while repressing their own expression as well as that of *TOC1*/other evening genes. As levels of CCA1/LHY proteins decrease towards evening, *TOC1*/evening gene expression resumes and reaches its maximum in preparation for initiating *CCA1/LHY* expression at dawn [82].

TOC1[83] is a member of a family of genes that are referred to as the *A. thaliana* pseudo response regulator (*APRR*) quintet[84]. The *APRR* genes are expressed in

sequential waves in the order *APRR9, APRR7, APRR5, APRR3*, and *APRR1* (*APRR1* is synonymous with *TOC1*[84]). Since *TOC1* overexpression does not lead to increased expression of *CCA1/LHY*[85], a model of the circadian oscillator including only these 3 genes is incomplete. While single mutants of *aprr9*, *7*, or *5* have little effect on circadian rhythms, double and triple mutant combinations of these genes have dramatic effects, indicating they are part of the *A. thaliana* circadian oscillator[86]. The promotive effects of *TOC1* expression on *CCA1/LHY* expression are then possibly the result of *TOC1* repressing the other members of the APRR quintet with repressive effects on *CCA1/LHY*. Thus a more complete model of the circadian oscillator including the other members of the APRR quintet would involve activation of *CCA1/LHY* genes at dawn. CCA1/LHY would activate genes required in daylight (*CAB*, etc.) as well as the first members of the APRR quintet, and would repress their own expression and that of *TOC1*. A series of waves of APRR expression leading ultimately to the expression of *TOC1* would then repress other members of the APRR quintet and thereby release repression of *CCA1/LHY* in time for expression at dawn[86].

Light input to the circadian oscillator is gated in *A. thaliana*. The EARLY FLOWERING 3 (ELF3) protein oscillates in a circadian manner and is present at highest levels at night, where it serves to block light input to the clock[87-89]. This gating may be necessary in order to prevent moonlight, starlight, or lightning flashes from resetting the clock to dawn. During the day ELF3 levels are low, allowing photoperiod inputs to affect the clock.

Circadian clock output

As described above, central components of the *A. thaliana* circadian clock regulate gene expression either through direct interaction with promoter motifs, or by interaction with multiprotein complexes. In addition to autoregulating their own expression, CCA1 and LHY regulate the expression of *GIGANTEA* (*GI*), a gene encoding a protein with transmembrane motifs. *GI* expression is lowest in the dark and highest in the daylight and peaks at 8-10 hours after dawn in wild-type plants[90], but expression is altered in *cca1* and particularly in *lhy* mutant plants. *gi* mutant plants show lower levels of both *CCA1* and *LHY*, indicating reciprocal regulation between these genes. The gigantea designation for the mutant phenotype is a result of extreme late flowering and continued vegetative growth, hence gigantic plants. The GI protein influences flowering time by modulating the expression of the floral promoter *CONSTANS* (*CO*)[91,92]. GI appears to influence the expression of *CO* through an interaction with SPINDLY (SPY), an O-linked β-N-actetylglucosamine transferase that has been show to interact with GI in two-hybrid screens[93,94]. SPY influences flowering by two separate pathways. It acts to repress flower-promoting signaling by the GA pathway[94,95], and it interacts with the GI protein via the tetratricopeptide domain to downregulate expression of *CO* by an unknown mechanism[93].

CO is a zinc finger transcription factor that accelerates flowering in *A. thaliana*[96]. It is part of a large family of proteins that share N-terminal B-Box domains, probably involved in protein-protein interactions, and C-terminal CCT (Constans, Constans-like, TOC1) domains that probably have multiple functions, including targeting CO to the

nucleus[97]. Other proteins in the photoperiod pathway (TOC1) share the CCT domain. *CO* expression shows circadian periodicity, with expression peaking in the evening; the CO protein functions as a light-dependent activator[98,99] of the floral pathway integrator genes *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*)[92,99,100]. The CO protein is especially important in the understanding of photoperiodic flowering because it has the properties of the internal oscillator described in the external coincidence model. *CO* expression levels begin to drop before dawn and begin rising as the day progresses. In short days, increasing expression of *CO* does not overlap with exposure to light. During inductive long days, increasing expression of *CO* overlaps with exposure to light late in the day. Thus, an internal oscillator, *CO*, overlaps with an external light signal as days grow longer in spring, initiating flowering in LD plants[98]. One problem that remains in placing *CO* as the internal oscillator is that *CO* levels are high at dawn, although they are dropping, and overlap with light in both long and short days. It is possible that an activator of the CO protein cycles and is not present at dawn, thus the coincidence of light and CO at dawn has no affect on flowering. The CO protein has been shown to be degraded at night and activated during the day in a light-dependent manner requiring PHYA, PHYB, CRY1, and CRY2[98,101]. CO shows a second property that ties it to early theories of the substance initiating flowering. The graft-transmissible substance, or florigen, hypothesized to initiate flowering is produced in leaves and then travels to the meristem, initiating the developmental change from vegetative to reproductive growth. Recently, grafting leaves expressing *CO* onto *co* mutant plants in non-inductive photoperiods has shown that *CO* by itself is sufficient to

generate the graft-transmissible signal to flower[102]. *CO* has subsequently been shown to act in a non-cell autonomous manner, in the phloem companion cells but not in meristematic cells, to regulate the synthesis or transport of a flowering signal partially through the activation of the floral pathway integrator *FT*[103]. *CO* represents the final step in the pathway through which photoperiod regulates flowering, acting in specific cells, partially through activation of floral integrators, to control a graft-transmissible substance that promotes flowering.

The gibberellin (or short-day) pathway

The gibberellin pathway in *A. thaliana*, or short-day pathway as it is also called, hastens flowering time in non-inductive short days. Biosynthesis of active gibberellins is initiated in response to a number of factors, including the developmental stage of the plant, the light environment, and crosstalk with other hormonal pathways. The gibberellins themselves are synthesized from geranylgeranyl diphosphate into *ent*-kaurene by cyclization reactions. *Ent*-kaurene can then undergo a number of oxidation, hydroxylation, and ring contraction reactions to form about 126 different GAs seen in plants, fungi, and bacteria (reviewed in [104]). Only a few of the GAs appear to be active in plants, primarily GA_1 and GA_4[105]. While it is clear that the enzymes catalyzing the steps of GA biosynthesis serve as control points for GA-modulated plant responses, the GA signal transduction process remains less clear. Experiments with cereal aleurone systems have indicated that a membrane-localized GA receptor protein is part of the signal transduction cascade (reviewed in [106]). GA-modulated signal transduction in the

flowering process may or may not share components of GA signal transduction in germinating seeds. While the perception and transduction of the GA signal needs to be clarified with regard to flowering, many of the downstream components of GA signaling have been elucidated by studying mutants in *A. thaliana*. Mutations in the *SPINDLY* (*SPY*) gene[107] have been shown to block normal GA signaling[94,95,108,109]. Interestingly, *SPY* interacts with a clock output gene, *GI*, described above, to regulate both flowering (through *CO*) and transpiration in a circadian manner[93,110]. Other GA signal transduction components are proteins belonging to the DELLA family. This gene family includes 5 members in *A. thaliana*: *GA INSENSITIVE* (*GAI*)[111], *REPRESSOR OF ga1-3* (*RGA*)[112], and *RGA LIKE 1/2/3* (*RGA1, RGA2, RGA3*)[113,114]. This gene family is particularly important to agriculture because mutant *gai* genes in wheat produced the 'green revolution' varieties of the 1970's[115]. The DELLA proteins are negative regulators of GA signaling. They possess an N-terminal DELLA domain and a C-terminal GRAS domain that they share with a larger gene family, the GRAS family. GA appears to overcome the negative regulation of DELLA proteins by targeting them for destruction in the 26S proteasome[116-118]. Through an unclear mechanism, DELLA proteins down regulate both *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS* (*SOC1*)[119], and *GAMYB* transcription factor genes with promotive effects on flowering. GAMYB transcription factors bind to GA response elements (GAREs) promoting gene expression. One of the genes induced by GAMYB is the *LEAFY* (*LFY*) gene, one of the floral meristem identity genes in *A. thaliana*[120-122]. The *GAMYB* transcription factor RNAs in *A. thaliana* are also under post-transcriptional regulation. The microRNA miR159

directs cleavage of the *GAMYB* mRNAs, and is itself under positive regulation by the GAMYBs and under negative regulation by the DELLA proteins[123]. Thus, through the action of DELLA proteins on the *SOC1* gene and through the action of GAMYB transcription factors on the *LFY* gene, the GA pathway influences flowering time.

The vernalization pathway

In addition to the need for a promotive long-day photoperiod, *A. thaliana* also has a requirement for vernalization in order to flower. Vernalization is a prolonged cold exposure enabling a plant to flower or accelerating flowering. The vernalization pathway enables certain plants to flower only under the most favorable conditions of spring, following a winter vernalization exposure. Dividing cells in meristematic shoot and root tips are the sites of cold perception in the vernalization response[124]. The mechanism of cold-sensing leads to epigenetic changes in the *FLOWERING LOCUS C* (*FLC*) gene. *FLC* encodes a MADS-box transcriptional regulator that represses the floral integrator genes[125]. *FLC* transcription is stably repressed in response to vernalization. The mechanism of repression involves the proteins VERNALIZATION 1 (VRN1)[126], VERNALIZTION 2 (VRN2)[127], VERNALIZATION INSENSITIVE 3 (VIN3)[128], and perhaps LIKE HETEROCHROMATIN PROTEIN 1 (LHP1)[129]. Vernalization results in modifications of histone H3 at the *FLC* locus, including deacetylation of K9 and K14, followed by dimethylation of K27 and K9[128,130]. These modifications lead to the repression of *FLC* through the formation of heterochromatin at the *FLC* locus. The *VRN2* gene encodes a homolog of the Drosophila Polycomb group

(PcG) protein SUPPRESSOR OF ZESTE 12[131]. This protein is a histone methyltransferase involved in developmental epigenetic switch mechanisms and is opposed by trithorax group (trxG) proteins in *Drosophila*[132]. The VIN3 protein has a PHD-finger motif that is often found in proteins involved in chromatin remodeling[128], while the VRN1 protein has a DNA binding motif[126]. *vrn1* mutants can not maintain stable repression of *FLC*, indicating its role is at least partially involved in maintenance of heterochromatin at the *FLC* locus. Analysis of mutants has revealed that VRN1 likely functions downstream of VRN2, since dimethylation of H3 K27 is lost only in *vrn2* mutants, while H3 K9 dimethylation is lost in both *vrn1* and *vrn2* mutants[130]. *VIN3* is expressed in response to prolonged cold and is localized to root and shoot meristems, the sites of perception for the vernalization response, while *VRN1* and *VRN2* are expressed in a more constitutive manner[126,127], indicating that *VIN3* may be involved in the specific localization of the vernalization response to meristems. Mutant *vin3* plants are the only ones from this group of genes that completely block *FLC* repression during vernalization, and prolonged cold induction of *VIN3* remains in *vrn1* and *vrn2* plants, so the protein acts upstream of VRN1/2. In animal systems, mitotically stable repression by PcG proteins is maintained by binding of the POLYCOMB (PC) protein to methylated H3 K27 residues[132]. While no obvious *PC* ortholog exists in the *A. thaliana* genome[131], there are genes in *A. thaliana* with the critical chromodomain required for binding modified histone residues. The *A. thaliana* gene *LIKE HETEROCHROMATIN PROTEIN 1* (*LHP1*) encodes a protein with a chromodomain and chromo shadow domain and influences flowering time[129]. It may be possible that this gene or one similar

to it plays a role in maintaining *FLC* chromatin in a repressed, heterochromatic state once vernalization has marked it with a specific histone code.

The autonomous pathway

A model of the autonomous flowering-time pathway began with observations that a class of *A. thaliana* mutants flower late in both short and long days, but remain responsive to vernalization[45,133]. Since these plants did not belong to the short day, long day, or vernalization pathways, and since they appeared to be responding to a constitutive internal signal that was not connected to the environment, they were grouped together and referred to as the autonomous pathway[134]. Autonomous pathway mutants were all found to have increased expression of *FLC*, a gene encoding a repressor of floral integrator genes[125]. Since these mutants have elevated *FLC* levels, the genes in the autonomous pathway act to stimulate flowering by repressing *FLC* in various ways. Subsequent to their discovery, it has now been found that at least some of the autonomous pathway genes are not autonomous, but respond to environmental stimuli[135]. The autonomous pathway genes include *FPA*, *FCA*, *FY*, *FVE* (the initials do not stand for phenotypic names)[45,133], *FLOWERING LOCUS D* (*FLD*)[136], *LUMINIDEPENDENS* (*LD*)[137], and *FLOWERING LOCUS K HOMOLOGY DOMAIN* (*FLK*)[138]. The genes *FCA*, *FPA*, *FLK*, and *FY* are involved in post-transcriptional regulation of the *FLC* mRNA transcript while *FLD* and *FVE* exert epigenetic control over the *FLC* locus[139]. LD is a homeodomain containing protein that down regulates *FLC* and up regulates expression of the meristem identity gene *LEAFY* (*LFY*) by an

unexplained mechanism[140]. FCA autoregulates its own expression and down regulates expression of *FLC*[141]. Both of these functions require an interaction with FY, a 3' RNA end processing factor[142]. Both FCA and FPA share RNA recognition motifs[141,143], and FLK also has an RNA binding domain, pointing out the importance of RNA regulation in the repression of *FLC* via the autonomous pathway[138]. The FVE protein has homology to a retinoblastoma-associated protein and is involved in histone deacetylation at the *FLC* locus[144], while FLD is itself a histone deacetylase that acts at the *FLC* locus[145]. No obvious connection exists between the RNA-modifying genes and the histone deacetylation genes in the autonomous pathway other than a common target of repression (FLC). Genes from both portions of the autonomous pathway (*FCA* and *FVE*) have been shown to serve temperature-sensing functions in the control of flowering[135]. Additionally, *FCA* also appears to be involved in photoperiod regulation of flowering[146], which suggests the autonomous pathway is far more interconnected and integrated into environmental inputs than previously suspected.

FLC promotive pathways

In addition to the *FLC*-repressive mechanisms of the autonomous and vernalization pathways, *A. thaliana* also possesses a group of genes with promotive effects on *FLC* expression. Mutations in this group of genes result in lower expression of *FLC*. First among this group of genes is *FRIGIDA* (*FRI*), which encodes a novel protein with coiled-coil domains[147]. In addition to *FRI*, *FRIGIDA-LIKE 1* and *2* (*FRL1* and *2*) are similar to *FRI*, and *FRL1* has been shown to be required for *FRI* mediated upregulation

of *FLC*[148]. Other genes with promotive effects on *FLC* include *VIP3*, which encodes a protein with WD interaction motifs and belongs to a family of proteins[149]; *AERIAL ROSETTE 1* (*ART1*), a gene that interacts synergistically with *FRI* to activate *FLC* expression[150]; *EARLY IN SHORT DAYS 4* (*ESD4*), a gene encoding a protease that processes SMALL UBIQUITIN-RELATED MODIFIER (SUMO), and upregulates *FLC* by unknown mechanisms[151,152]; and *HUA 2* (Hua means flower in Chinese), a gene previously identified as involved in processing the pre-mRNA of the floral organ identity gene AGAMOUS[153], and also involved in upregulation of the floral repressors *FLC*, *FLOWERING LOCUS M* (*FLM*), *MADS AFFECTING FLOWERING 2* (*MAF2*), and *SHORT VEGETATIVE PHASE* (*SVP*)[154].

Another group of proteins that increases *FLC* expression shares homology with the PAF1 chromatin remodeling complex in yeast. As described above in the vernalization section, a PcG-like complex marks and represses *FLC* chromatin. The opposing proteins in Drosophila are proteins belonging to the trithorax group (trxG), a complex whose function is to maintain chromatin in a conformation open to transcription[155]. The *A. thaliana* complex with PAF1 homology serves a function similar to the trxG proteins in *Drosophila*; it maintains the *FLC* locus in an open, transcriptionally active conformation[156]. While modifications of histone H3 at the *FLC* locus, including deacetylation of K9 and K14, followed by dimethylation of K27 and K9 mark *FLC* for repression, trimethylation of H3 at K4 is the histone code for active genes. The yeast PAF1 complex associates with RNA polymerase II during transcription, hence the name (RNA Polymerase II Associated Factor I), and recruits a SET 1

methyltransferase to the transcribed gene. The methyltransferase then generates H3 K4 trimethylation predominantly in the 5' portion of the gene. This methylation pattern serves as a mark of recent gene activity. Several *A. thaliana* homologs of members of the yeast multiprotein PAF1 complex have been found whose mutants have altered flowering time phenotypes and altered histone methylation patterns at the *FLC* locus. *EARLY FLOWERING 7* (*ELF7*) is a homolog of yeast *PAF1*; *EARLY FLOWERING 8* (*ELF8*) is a homolog of yeast *CTR9*; and *VERNALIZATION INDEPENDENCE 4* (*VIP4*) is a homolog of yeast *LEO1*[157]. The SET domain methylase recruited to the *FLC* locus by these PAF1 complex homologs may be EARLY FLOWERING IN SHORT DAYS (EFS), a protein necessary for trimethylation at the FLC locus[158,159]. In addition to trimethylation of H3 K4 residues at sites of transcription, the PAF1 complex in yeast also functions to recruit ISW1p, an ATP-dependent chromatin remodeling enzyme, to sites of active transcription. An *A. thaliana* homolog of ISW1p, *PHOTOPERIOD INDEPENDENT EARLY FLOWERING 1* (*PIE1*), has been shown to be necessary for *FLC* activation[160]. Interestingly, mutation in any of the PAF1 complex genes in *A. thaliana* suppresses the ability of *FRI* and autonomous pathway mutations to increase *FLC* expression, and negate the requirement for vernalization, since *FLC* levels are never high enough to inhibit flowering[157,161]. The *Arabidopsis* PAF1 complex is required for high levels of *FLC* transcription, but it additionally increases transcription of several MADS-box genes related to *FLC* that contain a conserved motif in their 5' UTRs that is a target of H3 K4 trimethyation[157]. *FLOWERING LOCUS M* (*FLM*)[162] and *MADS AFFECTING FLOWERING 2* (*MAF2*)[163,164] are two floral repressor genes

closely related to *FLC* that also show expression mediated by the PAF1 complex[157]. These genes are unaffected by *FRI*, by autonomous pathway genes, or by vernalization, but interact with genes in the photoperiod pathway described above, indicating that the PAF1 complex is active in selectively modifying the expression of several related MADS-box repressors of flowering in response to different inputs[165]. In contrast to *FLM* and *MAF2*, *FLC* expression is affected by vernalization, autonomous pathway genes, and *FRI*. Additionally, *FLC* activation by *FRI* requires a functional PAF1 complex[157]. To date, the inducer(s) of the PAF1 complex in *A. thaliana* have yet to be described.

Other (developmental) repressors of flowering

In addition to the photoperiod, gibberellin, autonomous, and vernalization pathways that affect flowering-time, another set of floral repressors exists that have not been directly linked to any of these pathways. It is possible that some or all of these repressors serve as a developmental block that prevents precocious flowering until the plant has matured to the point that flowering can be successfully accomplished. The genes in this group of repressors typically repress the effects of the floral pathway integrator genes and floral meristem identity genes. One of these genes is *TERMINAL FLOWER 1* (*TFL1*)[166,167]. *TFL1* encodes a phosphotidylethanolamine binding protein that is ~59% identical to the floral pathway integrator FT[168], and similar to the other FT-like proteins TWIN SISTER OF FT (TSF)[169,170], and MOTHER OF FT AND TFL1 (MFT)[170,171]. *TFL1* functions as a repressor of terminal flower formation and as a repressor of flowering in general.

Surprisingly, switching a single amino acid residue between FT and TFL1 causes a reciprocal, albeit partial, change of function. That is, the modified FT switches from floral activator to a weak floral activator and to a repressor of terminal flower formation while the modified TFL1 switches from a floral repressor and repressor of terminal flower formation to a weak floral activator[168]. Following duplication, a single amino acid substitution event in these proteins could create a protein with opposite phenotypic effects. The exact mode of action for FT and TFL1 in influencing floral timing and structure remains unknown. Several more repressors of flowering appear to operate in PcG-like complexes, forming a transcriptionally repressed, heterochromatic environment at various flowering-time loci[172,173]. Members of these repressive complexes include *MULTICOPY SUPPRESSOR OF IRA 1* (*MSI1*)[174], *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*)[175], *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*)[176] , *CURLY LEAF* (*CLF*)[177], *EMBRYONIC FLOWER 2* (*EMF2*)[178], and *SWINGER* (*SWN*)[172,179]. Most of the members of these repressive complexes have additional effects outside of flowering-time repression. Additional repressors of flowering include *EMBRYONIC FLOWER 1* (*EMF1*), a gene that causes *A. thaliana* to completely skip vegetative development when it is non-functional. *EMF1* encodes a potential transcriptional regulator of flowering[180]. Another repressor gene is *EARLY BOLTING IN SHORT DAYS* (*EBS*), a gene that may be involved in a chromatin-remodeling complex that represses *FT*[181,182]. The *FWA* (doesn't stand for a phenotype) gene is interesting in that *fwa* mutants are late-flowering, but show no sequence difference at the *fwa* locus[45,183]. Instead, the late flowering phenotype is caused by loss of methylation in repeats in the

FWA gene, a homeodomain transcription factor, and in its 5' promoter sequence[183]. Several other floral repressor genes that potentially share a similar regulatory mechanism are *TARGET OF EAT 1* and *2* (*TOE1/2*)[184], *SCHLAFMUTZE* (*SMZ*) and *SCHNARCHZAPFEN* (*SNZ*)[185], all of which are *AP2*-like genes that are repressed by microRNA172 probably at the level of translation[184].

An integrated view of flowering pathways

Given the importance of flowering in the life cycle of plants, it is not surprising that multiple controls of the process exist. Layers of floral repressors and activators work simultaneously to hold back or hasten the flowering process. Various environmental and developmental cues come together in the floral pathway integrators SOC1[119,186] and the FT family of proteins[169,171,187], and downstream in proteins controlling meristem identity such as LFY[188,189] and AP1[190]. While the interaction of the genes affecting flowering is complicated (Fig. 1), much of the input for the floral pathway integrators appears to be the antagonistic inputs from two genes, *CO* and *FLC*, a major promoter and repressor of flowering, respectively. Leaving the developmental repressors of flowering aside, flowering can be viewed as largely dependent on whether the promotive effects of *CO* outweigh the repressive effects of *FLC*, or vice versa. *CO* promotion of flowering is clearly associated with photoperiod, while *FLC* repression is modulated by vernalization, temperature, and perhaps slightly by photoperiod. While the promotive and repressive effects of these and many other genes on flowering is more clear now than ever, the picture remains clouded because many genes such as *CO* and *FLC* are

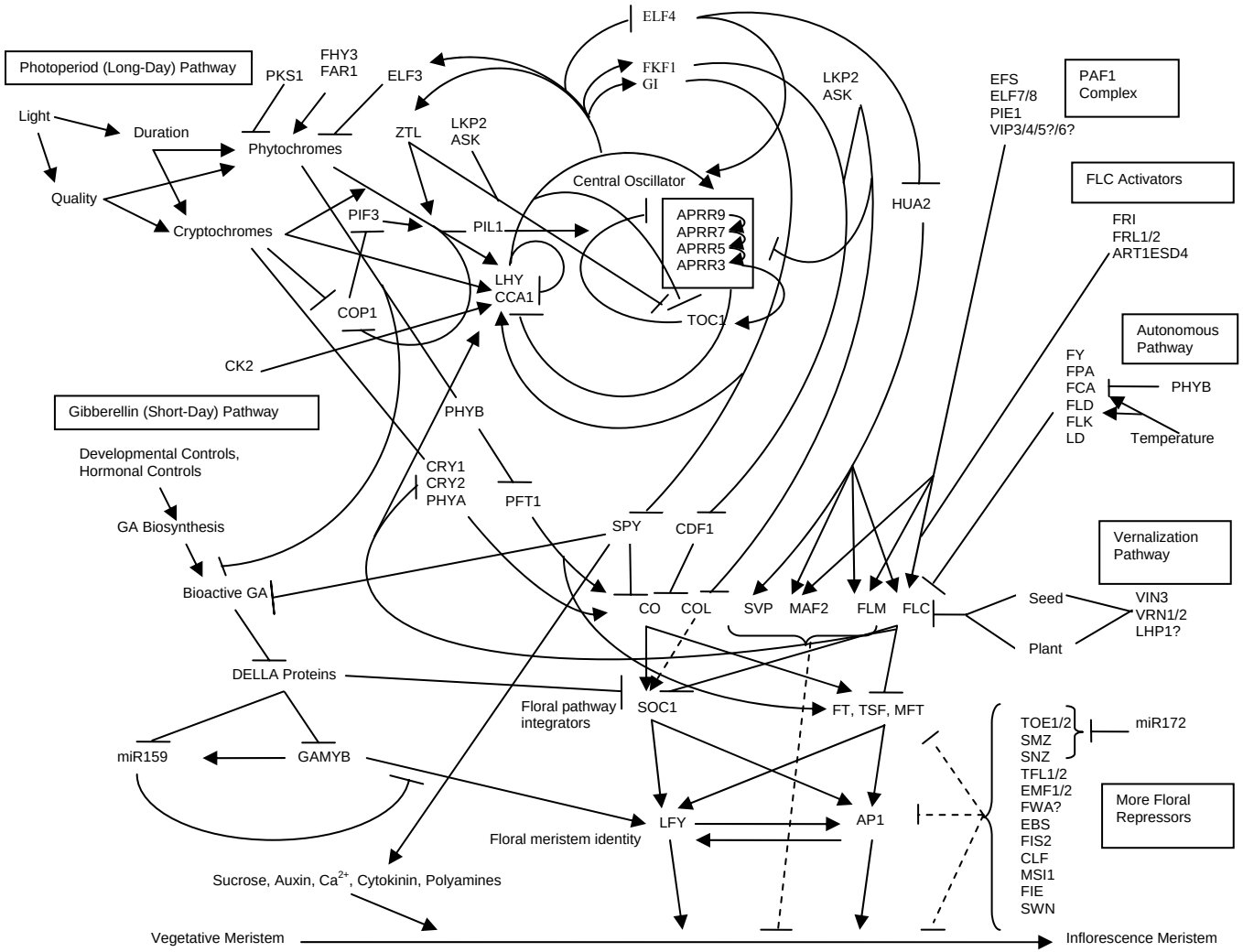

Fig. 1 Genes controlling flowering in the long day plant *Arabidopsis thaliana*. Gene abbreviations are given in the text.

members of large gene families whose members often have partially redundant and partially discrete effects on flowering time and other characters.

Conservation of flowering time genes between dicots and monocots

Knowledge of the flowering time genes and pathways in a eudicot such as *A. thaliana* can be useful in monocot crops only if the mechanisms controlling flowering are conserved between the two groups. Rice is the only monocot with a fully sequenced genome, and thus serves as a model monocot when comparing flowering time genes between monocots and eudicots. Genetic loci controlling flowering in rice, called heading date (*Hd*) or photoperiod sensitivity (*Se*) loci, were originally mapped as QTL[191-195]. As these *Hd* QTL were cloned, a picture of conservation in flowering time pathways began to emerge, since *Hd1* and *Hd3a* genes in rice were homologs of *CO* and *FT*, respectively[196,197]. With the publication of the rice genome sequence in 2002[6,7], many additional homologs of most *A. thaliana* flowering time genes were discovered *in silico*[198]. Since both rice and *A. thaliana* have many of the same flowering time genes, but respond to photoperiod differently— *A. thaliana* is a LD plant while rice is a SD plant—the major problem to resolve is explaining how the two plants can respond to photoperiod differently while using the same basic set of genes. In *A. thaliana*, *CO* serves as an activator of *FT* expression that is dependent on light, *CRY2*, and *PHYA*[98]. Since high levels of *CO* expression are clock-regulated and overlap with light only under LD conditions, *A. thaliana* flowers under LD conditions. In rice, the *CO* homolog *Hd1* has a similar expression pattern but different effects on flowering. *Hd1* accumulates to

high levels late in the day and represses the *FT* homolog *Hd3a* in association with phytochrome and light[199,200]. In the dark, *Hd1* promotes *Hd3a* expression. During short days, expression of *Hd1* doesn't overlap with light, so it acts to promote flowering in short days in rice. The dual activity of *Hd1* as both a repressor in LD and a promoter in SD confers rice with different photoperiod requirements for flowering. Since the same basic set of genes are used to control flowering in *A. thaliana* and rice, a comparative understanding of flowering in *A. thaliana* and rice may aid in identifying the genes responsible for photoperiod sensitive flowering in the Ma_5 and Ma_6 populations in *Sorghum bicolor*.

CHAPTER II

GENETIC MAPPING OF *Ma_5/Ma_6*

INTRODUCTION

Numerous genetic maps of sorghum have been created[201-205]. These RFLP-based maps are highly suited to making genetic comparisons among related plant taxa, but lack of marker density renders these maps ill-suited for use in efforts to positionally isolate and clone individual genes. Therefore, AFLP markers were chosen to map the *Ma_5* and *Ma_6* loci due to the long-range goal of map-based cloning both genes as recently demonstrated[206]. AFLP markers allow efficient genome-wide screening without prior sequence information[37], and an integrated AFLP-based genetic, physical, and cytogenetic map of sorghum was being created simultaneous with the mapping of the *Ma_5* and *Ma_6* loci[38,39,207,208].

MATERIALS AND METHODS

AFLP and SSR analysis

The AFLP markers used to map the *Ma_5* and *Ma_6* loci were produced using a modification of the AFLP procedure for use on LI-COR DNA sequencing instruments (LI-COR Biotechnologies, Lincoln, NE)[37,39]. Sorghum genomic DNA was extracted using a FastDNA kit according to the manufacturer's instructions (Qbiogene, MP Biomedicals, Irvine, CA). The DNA was quantified on a Turner Designs TD-360 fluorometer (Turner Designs, Inc., Sunnyvale, CA) according to the manufacturer's

instructions. AFLP reactions utilizing both *Eco*RI/*Mse*I or *Pst*I/*Mse*I enzyme combinations were performed and the products analyzed on LI-COR dual dye DNA sequencing systems as previously described[37,39]. Screening for sorghum SSR markers was performed on LI-COR gels as described previously[37,38]. Sorghum AFLP and SSR reactions were arrayed on LI-COR gels so that multiple PI, early flowering samples were directly beside multiple PS, late flowering samples. Additionally, the parents of a recombinant inbred mapping population (i.e. BTx623 and IS3620C) used to construct the TAMU-ARS high density sorghum genetic map were included on all gels[38,209-211]. Polymorphic AFLP bands that segregated with either flowering phenotype were scored visually. When an AFLP band that was polymorphic and linked to flowering phenotype in the Ma_5 or Ma_6 populations was also polymorphic in the parents of the recombinant inbred mapping population, the band served as a link to one of the 10 sorghum chromosomes in the genetic reference map[38] (Fig. 2). Other AFLP and SSR markers in the same area of the TAMU-ARS high density genetic map were then examined in the maturity populations described below. In this way, a number of AFLP and SSR bands common to both the recombinant inbred mapping population and the maturity populations were discovered.

Creation of Ma_5 and Ma_6 regional genetic maps using MapMaker

Segregation data for the AFLP and SSR markers were scored manually, entered into a Microsoft Excel spreadsheet (Microsoft, Tacoma, WA), transformed, imported into MapMaker/exp (v3.0) on a Sun Microsystems workstation and used to calculate

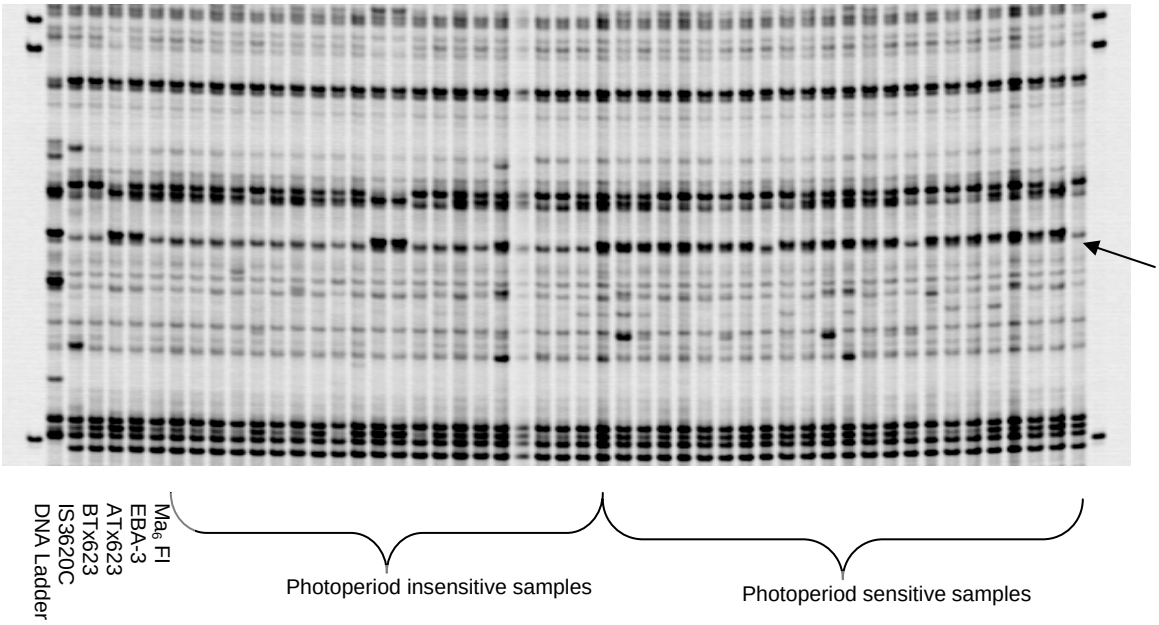

Fig. 2 A representative LI-COR AFLP gel. Parents of the *Ma_6* population and the recombinant inbred population (IS3620C and BTx623) are arrayed on the left side of the gel. The arrow shows *Xtxa3550*, a genetic marker from the recombinant inbred map linked to sorghum chromosome 6 at ~13 cM.

recombination fractions between pairs of linked markers. The Kosambi mapping function[212] was used to calculate centimorgan distances. Initially, the "group" command was used to determine which markers were closest to the maturity locus at a LOD of 6.0 and a maximum centimorgan distance of 10.0. The "lod table" command was used to select 5 markers separated by a minimum of 2.5 cM. These 5 markers were ordered into an initial framework with the "compare" command. The remaining markers were 3-point ordered into the framework map 3 at a time, first at LOD 3.0, then at LOD 2.0 using the "build" command. Three-point local order was assessed using the "ripple"

command in 5 marker intervals with a threshold LOD of 2.0. The regional *Ma5* and *Ma6* maps were produced with Mapmaker Macintosh v2.0.

Creation of the *Ma5* population

The initial populations used for studying the genetic segregation of the *Ma5/Ma6* maturity genes were created in the mid to late 1990's as part of a master's thesis at Texas A&M University under the direction of Dr. William L. Rooney[24]. Briefly, a population of plants for mapping the *Ma5* locus was created by crossing an elite male-sterile sorghum, A3Tx436 (*Ma5Ma5ma6ma6*), with a two-dwarf, forage/grain sorghum from Argentina, EBA-3 (*ma5ma5Ma6Ma6*)[25]. Both parents are PI, while the F_1 is PS, heterozygous at both loci (*Ma5ma5Ma6ma6*), and male-sterile. The F_1 was then backcrossed using EBA-3 as a pollinator to produce a BC_1F_1 mapping population that segregates for the *Ma5* locus producing a 1:1 ratio of PI: PS progeny (*ma5ma5Ma6__*:*Ma5ma5Ma6__*). Subsequently, Dr. Rooney created a second population of plants for mapping the *Ma5* locus that more closely resembles the genetics of the *Ma6* population (see below). In this second population, ATx623 (*Ma5Ma5ma6ma6*) was crossed with EBA-3 (*ma5ma5Ma6Ma6*). The male-fertile F_1 (*Ma5ma5Ma6ma6*) was then backcrossed as a pollinator onto a male-sterile version of EBA-3, A3EBA-3 (*ma5ma5Ma6Ma6*), to produce a BC_1F_1 population segregating at the *Ma5* locus in a 1:1 ratio (*ma5ma5Ma6__*: *Ma5ma5Ma6__*). Unless otherwise noted, the *Ma5* data in this dissertation was produced from the A3Tx436*EBA-3 cross, which is the cross referred to as the *Ma5* population.

Collecting and phenotyping *Ma_5* samples, summer of 2001

Approximately 300 plants that were planted on 1 April 2001 were phenotyped and collected from the *Ma_5* population in College Station, Texas on 15 August 2001. Since there was variation in flowering time in this population, the plants were placed into 4 different phenotypic groups: 1) photoperiod insensitive early (PI-E), including plants that flowered early and had dried seed in the head at the collection date; 2) photoperiod insensitive late (PI-L), including plants that had flowered close to the collection date, having soft, green seed in the head; 3) differentiated (D), including plants that were just booting and plants that revealed a floral meristem on dissection of the shoot apical meristem; and 4) photoperiod sensitive (PS), including plants that had not produced a floral meristem at the collection date.

Screening the 2001 *Ma_5* population with AFLP and SSR markers

A small subset of *Ma_5* plants (10 PI-E and 10 PS) were screened with 82 AFLP primer combinations and 7 SSR markers as described above in order to place the locus on one of the 10 sorghum linkage groups and to quickly and efficiently identify molecular markers from the reference map [38] that could be used in the *Ma_5* population. SSR markers were used to screen plants in the maturity population in PCR reactions of 10 µl total volume containing the following components: 1 µl of 10X PCR buffer, 1 µl of 25 mM $MgCl_2$, 0.8 µl of a 2.5 mM mixture of each dNTP, 1 µl of IRD-labeled SSR forward primer at 1 pmol/µl, 1 µl of unlabeled SSR reverse primer at 1 pmol/µl, 0.04 µl of Taq

polymerase, 3.16 µl of sterile water, and 2.0 µl of genomic template DNA at 2.5 ng/µl (reagents from Promega, Madison, WI). The SSR reactions were carried out with the following cycling program: 94° C for 10 minutes, followed by 33 cycles of 94° C for 1 minute, annealing temperature of primers for 1 minute, 72° C for 1 minute, with a final extension step of 72° C for 10 minutes and a 4° C hold. SSR primer sequences have been previously described[209,210]. Following the pilot screening of 10 PI and 10 PS plants, a larger set of 202 *Ma5* BC_1F_1 plants (104 PI-E, 98PS) were then screened with markers identified as informative in the small population.

Collecting and phenotyping *Ma5* samples, summer of 2003

In order to fine map the *Ma5* locus, a BC_1F_1 population segregating for the *Ma5* locus was planted in two locations on 4 April 2003, both in College Station, Texas, and flowering plants were collected at weekly or bi-weekly intervals in the summer of 2003. All of these plants were from the original *Ma5* population (A3Tx436*EBA-3)*EBA-3. Together, 2915 plants that varied in flowering time were collected from the two locations from May through November of 2003. Although an attempt was made to collect every plant in the 2003 growing season so that genotypic and phenotypic ratios could be determined, a large number of the late-flowering plants from this growing season were not collected or phenotyped due to death from insects, diseases, lodging, etc. Many of the plants from the final collection date had yet to flower, but were collected because the plants left in the field were dying. The death of these very late flowering plants was likely due to the cold November temperatures. The total estimated

number of *Ma5* plants growing in the 2003 growing season was 4200. All of the early flowering plants were collected in 2003.

Screening the 2003 *Ma5* plants with AFLP and SSR markers

Ninety-six plants from the *Ma5* BC_1F_1 population were screened with markers previously found to flank the *Ma5* locus as described above. This screening was intended to confirm the position of the *Ma5* locus as discovered in mapping with populations from previous years, and to discover plants with potential crossovers between two markers flanking the *Ma5* locus.

Creation of the *Ma6* population

A population of plants segregating for the *Ma6* locus was created by crossing an elite male-sterile sorghum, ATx623 (*Ma5Ma5ma6ma6*) with EBA-3 (*ma5ma5Ma6Ma6*)[24,25]. Both parents are PI, and the F_1 is PS, heterozygous at both maturity loci (*Ma5ma5Ma6ma6*), and is male-fertile in this cross. ATx623 was backcrossed using the F_1 as a pollinator to produce a BC_1F_1 mapping population that segregates for the *Ma6* locus producing a 1: 1 ratio of PI: PS progeny (*Ma5__ma6 ma6*: *Ma5__Ma6 ma6*).

Collecting and phenotyping *Ma6* samples, summer of 2000

There is less variation in flowering time in the *Ma6* population than in the *Ma5* population. Therefore, the BC_1F_1 plants collected in mid-August were grouped into only PI and PS phenotypes. Initially, a group of 83 plants that was grown in College Station,

Texas, was collected and phenotyped by a lab technician in the summer of 2000 for the purpose of mapping the *Ma*$_6$ locus.

Collecting and phenotyping *Ma*$_6$ samples, summer of 2002

In the summer of 2002, *Ma*$_6$ BC_1F_1 plants planted about 1 April, 2002 in College Station, Texas were phenotyped and collected on 25 July, 2002. This group of 506 plants was early-flowering and phenotyped as PI. On 21 August, 2002 a group of 365 late-flowering, PS plants was collected, bringing the total number of *Ma*$_6$ plants collected in 2002 to 871. Every plant within the 2002 *Ma*$_6$ BC_1F_1 population was collected.

Collecting and phenotyping *Ma*$_6$ samples, summer of 2003

In order to fine map the *Ma*$_6$ locus, a large BC_1F_1 population segregating for the *Ma*$_6$ locus was planted in two locations in College Station, Texas on 4 April 2003, then phenotyped and collected at weekly or bi-weekly intervals in the summer of 2003. Together, 2000 plants that varied in flowering time were collected from the two *Ma*$_6$ locations from May through November of 2003. Although an attempt was made to collect every plant in the 2003 growing season so that genotypic and phenotypic ratios could be determined, many of the late-flowering plants from these populations were not collected or phenotyped due to death from insects, diseases, lodging, etc. The very latest collection date included many plants that had yet to flower. Since many plants were dying at the late collection date, probably due to the cold November temperatures, a

decision was made to collect all remaining plants at that time. An estimation had been previously made that there were 2300 *Ma*$_6$ BC_1F_1 plants grown in the 2003 season, so approximately 300 PS, late flowering plants were not collected in the 2003 season.

Genetic screening of the 2003 *Ma*$_6$ samples

Due to large sample numbers, a rapid and inexpensive method of DNA extraction was developed for screening the 2003 *Ma*$_6$ PS plants. Five ¼ inch leaf punches were placed in 500 µl of Tris-buffered saline solution (50 mM Tris, 150 mM NaCl, pH 7.5) in 1.2 ml 96 well tubes with strip caps (Fisher Scientific International, Hampton, NH., bulk tubes 07-200-317, racked tubes 07-200-319, caps 07-200-323). The samples were disrupted in a GenoGrinder (BT&C/OPS Diagnostics**,** Bridgewater, NJ) at 1,750 strokes/minute for 2 minutes with steel dowel pins (1/8" X 1/2" Small Parts Inc., Miami FL, DWX-02-08). The samples were heated to 65° C for 1 hour, and then centrifuged at 2,000 X g for 5 minutes to pellet debris. The supernatant was transferred to fresh tubes and diluted 1:10 in sterile water. Two microliters of the dilution were used as template in PCR reactions with markers *Xtxp434* and *12255.Contig1* as described above. Any samples that showed a potential crossover by screening with this method were then used to make AFLP template and were screened with AFLP markers linked to the *Ma*$_6$ locus as described above.

RESULTS AND DISCUSSION

Genetic mapping of *Ma$_5$* in 2001

Twenty *Ma$_5$* plants (10 PI-E and 10 PS) were screened with 82 AFLP primer combinations and 7 SSR markers in order to place the *Ma$_5$* locus on one of the 10 sorghum linkage groups and to quickly and efficiently identify molecular markers from the reference map[38] that could be used in the *Ma$_5$* maturity population. Several AFLP markers linked to flowering phenotype were located on sorghum chromosome 2 on TAMU-ARS high density sorghum genetic map [38].

Following the pilot screening of 10 PI-E and 10 PS plants, a larger set of 202 *Ma$_5$* BC_1F_1plants (104 PI-E, 98 PS) were then screened with markers identified as informative in the small population. The markers were manually scored to produce a genetic map of the *Ma$_5$* locus with the program MapMaker as described above (Fig 3). The markers closest to the *Ma$_5$* locus were *Xtxa3424* at 0.5 cM above the locus and *Xtxp100* at 1.0 cM below the locus. When comparing the regional map of the *Ma$_5$* locus with the TAMU-ARS high density genetic map[38], the *Ma$_5$* regional map shows less recombination around the locus (Fig 4). Lower levels of recombination around the *Ma$_5$* locus could be due to several factors, such as linkage disequilibrium caused by multiple flowering-time genes residing at the *Ma$_5$* locus, population structure differences between the two populations, or it could simply be an artifact resulting from subsampling only the earliest and latest flowering plants from the population in the 2001 growing season. Most of the PI early flowering recessive *ma$_5$* plants had an EBA-3/EBA-3 genotype between *Xtxa3424* and *Xtxp100* on chromosome 2, while most of the PS late flowering

plants had an A3Tx436/EBA-3 genotype between *Xtxa3424* and *Xtxp100* as expected if A3Tx436 contains the dominant allele[25]

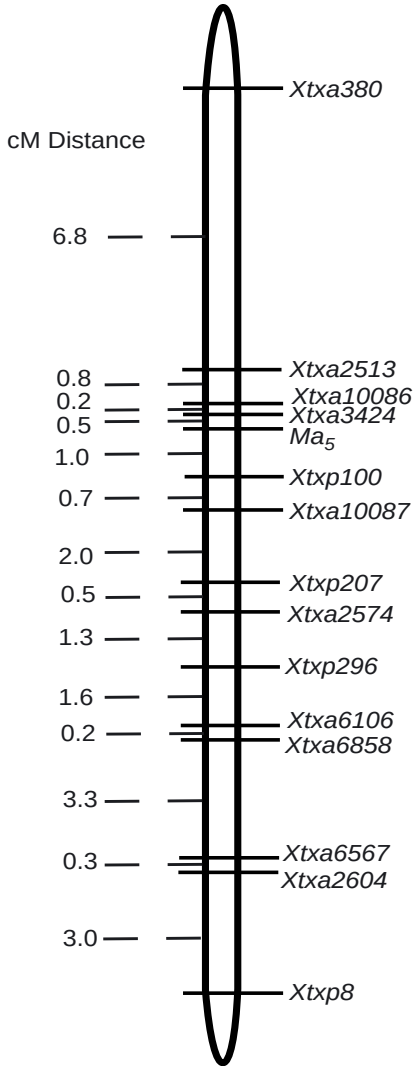

Fig. 3 A genetic map of the *Ma*$_5$ locus. This map was created from 202 BC_1F_1 samples (104 PI and 98 PS) collected in 2001. Many of the markers shown are linked to the TAMU-ARS high density genetic map used as a reference.

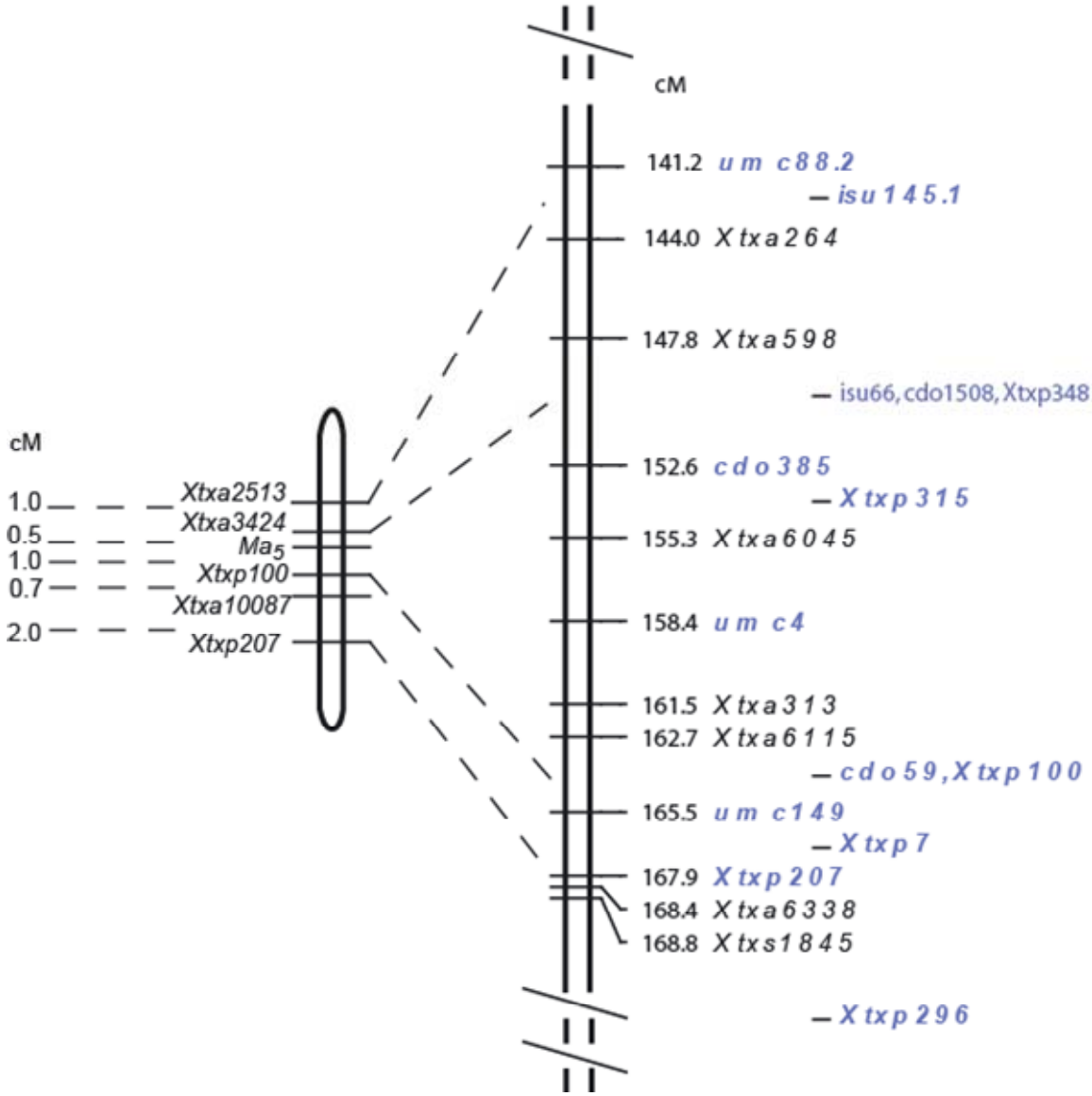

Fig. 4 *Ma*$_5$ regional map compared to the TAMU-ARS high density genetic map. A regional map of the *Ma*$_5$ locus from 2001 (left) shows lower recombination when compared to the reference map (right).

Distribution of flowering in the 2003 *Ma5* BC_1F_1 population

The distribution of flowering in the 2003 *Ma5* population indicates that multiple genes controlling flowering time are likely segregating in *Ma5* BC_1F_1 plants. The total estimated number of plants growing in the *Ma5* population in 2003 was 4200. All of the early flowering plants were collected in 2003, while many of the late flowering plants were not collected due to premature death. Plotting the number of plants flowering by days after planting (Fig. 5) reveals a bimodal distribution of early and late flowering plants. Unlike the *Ma6* population, in which flowering ceased for several weeks, individual plants in the *Ma5* population transitioned to floral growth throughout the growing season, with a lull in flowering in the first week of August, 2003, about 125 days after planting (DAP). If 7 August, 2003 (125 DAP) is taken as the date dividing early and late flowering, 1173 out of 4200 *Ma5* plants, or about 28% of the *Ma5* population was early flowering, and 3027 plants were PS and late flowering. These numbers are far different from the 1:1 ratio expected if only one dominant repressor of flowering is segregating in this *Ma5* BC_1F_1 population [χ^2=818.4.6, and $P(\chi^2_{1df}>10.83)<0.001$]. Even if the estimated 1300 or so uncollected late flowering plants are not considered in the chi-square analysis, the segregation ratio of this population remains significantly different from a population segregating in a 1:1 ratio. Additionally, the continuous variation in flowering seen in the 2003 *Ma5* populations indicates that more than 1 flowering-time gene is segregating in the population (Fig. 5). It has not escaped our notice that the flowering pattern in the 2003 *Ma5* population of 72% photoperiod sensitive late flowering: 28% photoperiod insensitive early flowering is very close to a

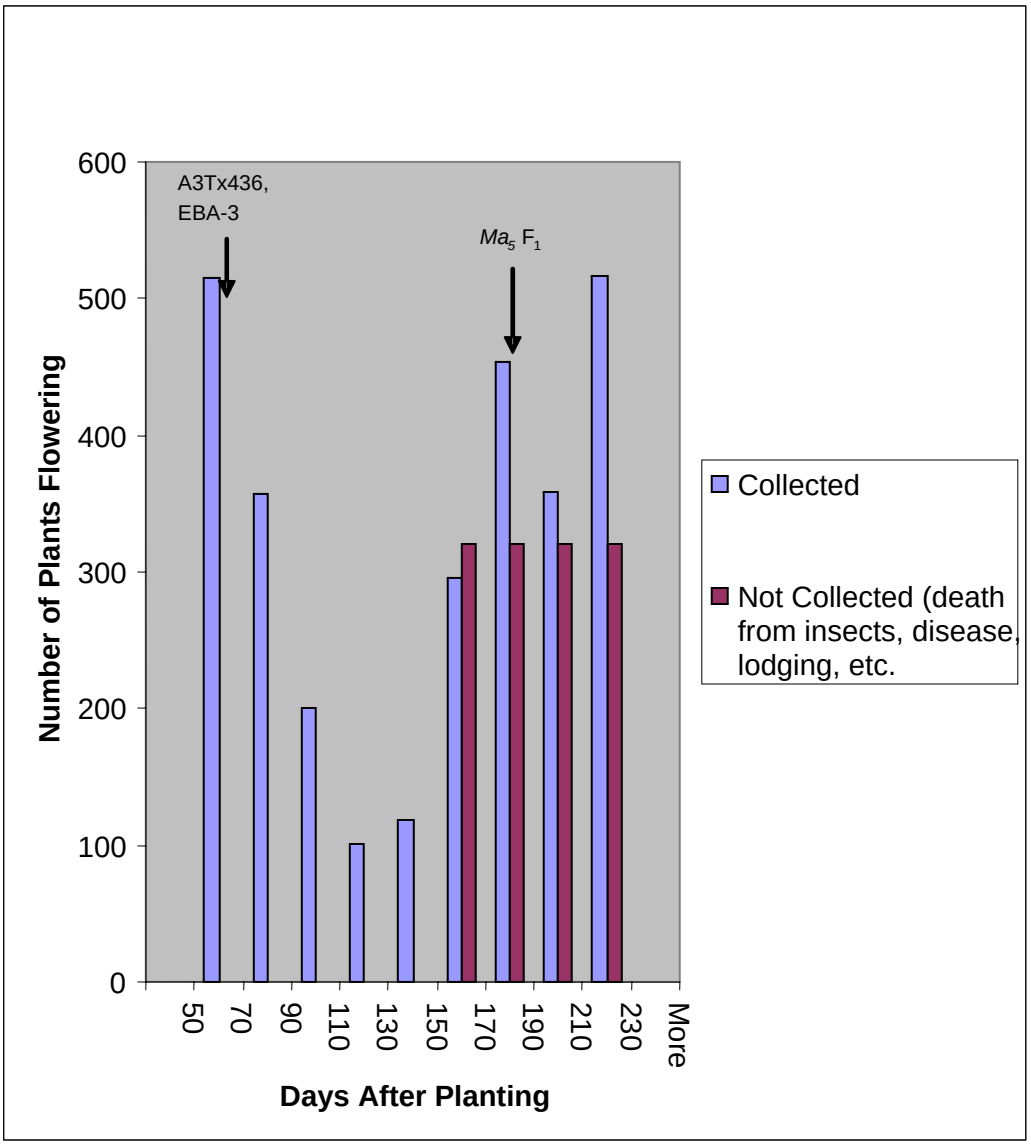

Fig. 5 Distribution of flowering in the 2003 *Ma5* BC_1F_1 population. The material was planted on 4 April 2003 in College Station, Texas. The flowering time of the inbred parents and F_1 are indicated. The number of plants not collected is based on an estimation of population size at the beginning of the growing season. Many plants collected at the latest date had not flowered, but were collected because most of the plants remaining in the population were dying, probably because of cold weather.

A3Tx436 X EBA-3

Ma_5-1Ma_5-1Ma_5-2Ma_5-2ma_6ma_6 $ma_{5\text{-}1}ma_{5\text{-}1}ma_{5\text{-}2}ma_{5\text{-}2}$***$Ma_6Ma_6$***

F_1

Ma_5-1$ma_{5\text{-}1}$***Ma_5-2***$ma_{5\text{-}2}$***Ma_6***ma_6

EBA-3 gametes
only $ma_{5\text{-}1}ma_{5\text{-}2}$***$Ma_6$***

BC_1F_1 progeny genotypes, ratios, and expected phenotypes

F_1 gametes		
Ma_5-1Ma_5-2Ma_6	1) 1/8 ***Ma_5-1***$ma_{5\text{-}1}$***Ma_5-2***$ma_{5\text{-}2}$***Ma_6Ma_6***	PS
Ma_5-1Ma_5-2ma_6	2) 1/8 ***Ma_5-1***$ma_{5\text{-}1}$***Ma_5-2***$ma_{5\text{-}2}$***Ma_6***ma_6	PS
Ma_5-1$ma_{5\text{-}1}$***Ma_6***	3) 1/8 ***Ma_5-1***$ma_{5\text{-}1}ma_{5\text{-}2}ma_{5\text{-}2}$***$Ma_6Ma_6$***	PS
Ma_5-1$ma_{5\text{-}1}ma_6$	4) 1/8 ***Ma_5-1***$ma_{5\text{-}1}ma_{5\text{-}2}ma_{5\text{-}2}$***$Ma_6$***$ma_6$	PS
$ma_{5\text{-}1}$***Ma_5-2Ma_6***	5) 1/8 $ma_{5\text{-}1}ma_{5\text{-}11}$***$Ma_5$-2***$ma_{5\text{-}2}$***$Ma_6Ma_6$***	PS
$ma_{5\text{-}1}$***Ma_5-2***ma_6	6) 1/8 $ma_{5\text{-}1}ma_{5\text{-}11}$***$Ma_5$-2***$ma_{5\text{-}2}$***$Ma_6$***$ma_6$	PS
$ma_{5\text{-}1}ma_{5\text{-}2}$***$Ma_6$***	7) 1/8 $ma_{5\text{-}1}ma_{5\text{-}1}ma_{5\text{-}2}ma_{5\text{-}2}$***$Ma_6Ma_6$***	PI
$ma_{5\text{-}1}ma_{5\text{-}2}ma_6$	8) 1/8 $ma_{5\text{-}1}ma_{5\text{-}1}ma_{5\text{-}2}ma_{5\text{-}2}$***$Ma_6$***$ma_6$	PI

Fig. 6 Two gene *Ma_5* model. A simple explanation of the 3:1 photoperiod sensitive: photoperiod insensitve ratio seen in the 2003 *Ma_5* population. Either one of two *Ma_5* genes with overlapping function can interact with a dominant *Ma_6* allele to repress flowering. Interaction between the alleles of the 3 genes can account for the quantitative flowering response. Dominant alleles are in boldface.

3:1 ratio. A simple explanation for this segregation ratio is that there are two *Ma_5* genes, at least partially redundant in function, segregating in the *Ma_5* population. One gene is on chromosome 2 in the vicinity of *Xtxa3424* and the other is at another unknown and unlinked location in the genome. In this model (Fig. 6), possession of either *Ma_5* gene would delay flowering. It is unlikely that a large number of genes controlling flowering time are segregating in the *Ma_5* population due to the large number of plants exhibiting

the most extreme early-flowering phenotype in the 2003 population. In that year, there were 500 out of 4200 plants that flowered together at the earliest collection date (Fig. 5). This is close to $1/8^{th}$ of the population as a whole. While there was continuous variation in flowering time in that population, gene dosage between two *Ma_5* genes and one *Ma_6* gene could account for the variation seen.

Genetic mapping of *Ma_5* in 2003

Ninety-six *Ma_5* BC_1F_1 plants collected in the summer of 2003 were screened with the two closest markers flanking the locus as determined from previous mapping efforts in the summer of 2001. The two closest flanking markers at that time were *Xtxa3424* at 0.5 cM above the *Ma_5* locus, and *Xtxp100* at 1.0 cM below the *Ma_5* locus. Just as in the 2001 population, most of the PI early flowering recessive *ma_5* plants had an EBA-3/EBA-3 genotype between *Xtxa3424* and *Xtxp100*, while most of the PS late flowering plants had an A3Tx436/EBA-3 genotype between *Xtxa3424* and *Xtxp100*. These 96 samples were not the earliest flowering and latest flowering as in the 2001 screening; rather, they equally represented samples flowering throughout the growing season. Unlike the plants from 2001, the 2003 plants showed roughly equivalent recombination levels at the *Ma_5* locus with those of the recombinant inbred population used as a reference map[38] (data not shown). The differences in recombination rates between the 2001 population and the recombinant inbred population at the *Ma_5* locus (Fig 4) may therefore be due to subsampling the population in 2001.

Since 14 crossover plants were discovered among 96 plants genotyped at the *Xtxa3424* and *Xtxp100* loci from 2003 (data not shown), no further efforts to screen the remaining 2,819 Ma_5 plants were made. Effort was instead aimed at developing additional molecular markers in the *Xtxa3424* to *Xtxp100* genetic interval so that the Ma_5 locus could be narrowed.

Genetic mapping of Ma_6 in 2000

Forty-eight different AFLP primer combinations were used to screen 83 plants (42 PI and 41 PS) for markers linked to the Ma_6 locus. A number of AFLP markers were found to be linked to the flowering time phenotype and those that were also polymorphic in the TAMU-ARS genetic map were all located near the top of sorghum chromosome 6[38]. While markers near the top of chromosome 6 were linked to flowering-time phenotype, the scores did not closely approach complete linkage. Eleven of the plants in the original data set were then thrown out as duplicated samples collected from tillers off of the same plant, based on identical crossover patterns at multiple loci (data not shown). Since sample collection and phenotyping was performed by another student within the laboratory and it was therefore not possible to verify the phenotypic data, 8 other plants (6 PI, 2 PS) were removed from the data set as being incorrectly phenotyped. The 6 PI plants had a dominant EBA-3 allele while the 2 PS plants had recessive ATx623 alleles at the Ma_6 locus. The two PS plants are almost certainly phenotyping errors, since hundreds of Ma_6 PS plants in subsequent growing seasons all had dominant EBA-3 alleles at the Ma_6 locus. Some or all of the 6 PI plants likely belonged to a 'modified'

class of early flowering PI plants with a dominant EBA-3 allele at the *Ma$_6$* locus discussed at length in chapter III below. When these plants were removed from the data set, the markers *Xtxa7*, *Xtxa3550*, and *Xtxa4001* all show complete linkage to the late flowering phenotype.

A regional map of the *Ma$_6$* locus was created that contained 7 markers linked to the TAMU-ARS reference map[38] and 8 novel AFLP markers (Fig. 7). Three of the markers were completely linked to the *Ma$_6$* gene, so additional crossover plants were needed to narrow the genetic interval containing the *Ma$_6$* gene.

Distribution of flowering in the *Ma$_6$* population grown in 2002

Every plant in the *Ma$_6$* population grown in 2002 was collected so that the phenotypic ratio of PI:PS plants could be determined. In total, there were 506 PI early flowering plants and 365 PS late flowering plants, for a total of 871 plants in the 2002 *Ma$_6$* population. The 506:365 ratio is significantly different from the expected 1:1 ratio that would be seen if one dominant repressor of flowering is segregating in this *Ma$_6$* population [χ^2=263, and P(χ^2_{1df}>10.83)< 0.001].

Genetic mapping of *Ma$_6$* in 2002

All 871 plants in this *Ma$_6$* population were screened with the markers found to flank the *Ma$_6$* locus in the 2000 population. This analysis confirmed that a gene responsible for the late flowering PS phenotype exists in close proximity to the markers *etaccaa184* and *Xtxa2124* at the top of chromosome 6. Sixty-two plants found to have crossovers

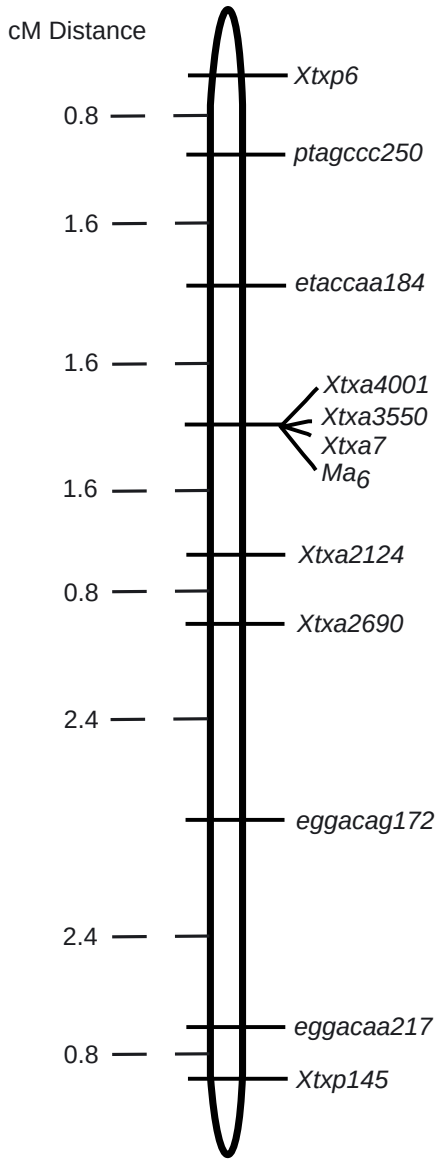

Fig. 7 A genetic map of the *Ma*$_6$ locus. This map was created from *Ma*$_6$ BC_1F_1 samples collected in the summer of 2000. Markers with *Txa* or *Txp* designations are linked to the TAMU-ARS high density genetic map.

between these two flanking markers (50 PI and 12 PS plants) were then screened with primers for the markers *Xtxa4001, Xtxa7*, and *Xtxa3550*, all of which had been found in the 2000 screening to be completely linked to the Ma_6 gene. Out of the 12 PS plants, all had an ATx623/EBA-3 genotype at the markers *Xtxa4001, Xtxa7*, and *Xtxa3550*. Therefore, in the PS class, screening additional plants had failed to break the linkage disequilibrium between these 3 markers and the Ma_6 locus. In the PI class, 393 out of 506 plants had an ATx623/ATx623 genotype at the markers *Xtxa4001, Xtxa7*, and *Xtxa3550*. However, 113 PI plants had an ATx623/EBA-3 genotype at this locus, indicating that there was more than 1 gene controlling flowering time segregating in the Ma_6 BC_1F_1 population. The genes causing plants to flower early in spite of dominant repressive genes present at both the Ma_5 and Ma_6 loci were referred to as modifiers, and are discussed at length in chapter III below. Among the Ma_6 PI class in 2002, there were plants with crossovers between the markers *Xtxa4001, Xtxa7*, and *Xtxa3550*. However, these crossover plants were uninformative in narrowing the Ma_6 locus because a determination could not be made as to whether individual plants were in the PI class due to a recessive ma_6 allele or due to a modifier gene elsewhere in the genome.

Distribution of flowering in the Ma_6 population grown in 2003

The distribution of flowering in the Ma_6 population grown in 2003 indicates that more than one gene is segregating in Ma_6 BC_1F_1 plants. Together, 2000 plants that varied in flowering time were collected from the Ma_6 population from May through November of 2003. Although an attempt was made to collect every plant in this population so that

genotypic and phenotypic ratios could be determined, many of the late flowering plants from this population were not collected or phenotyped due to death from insects, diseases, lodging, etc. The very latest collection date included many plants that had yet to flower. Since many plants were dying at the late collection date, probably due to the cold November temperatures, a decision was made to collect all remaining plants at that time. An estimation had been previously made that there were 2300 plants in the *Ma$_6$* population planted in 2003. Using this total number, an examination of the distribution of flowering in the *Ma$_6$* population reveals a clear bimodal distribution with a period of several weeks without flowering separating the PI early flowering and the PS late flowering classes (Fig. 8). If 150 DAP is the cutoff between early and late flowering (the last week of August), a total of 1475 plants were PI and early, and approximately 825 plants were PS and late flowering. These numbers are significantly different from the 1:1 segregation ratio expected if only 1 dominant repressor of flowering is segregating in this population [χ^2=91.9, and P(χ^2_{1df}>10.83)< 0.001]. In this population, about 64% of the plants were PI and early flowering.

Genetic mapping of *Ma$_6$* in 2003

A total of 96 plants from the *Ma$_6$* population were screened with the markers *etaccaa184* and *Xtxa2124* that had previously been found to flank the *Ma$_6$* locus. The 96 samples chosen for mapping were composed of 80 early flowering PI and 16 late flowering PS plants. Again, the PI samples were predominantly ATx623/ATx623 and the PS samples were all ATx623/EBA-3 at the *Ma$_6$* locus, between the markers *etaccaa184* and

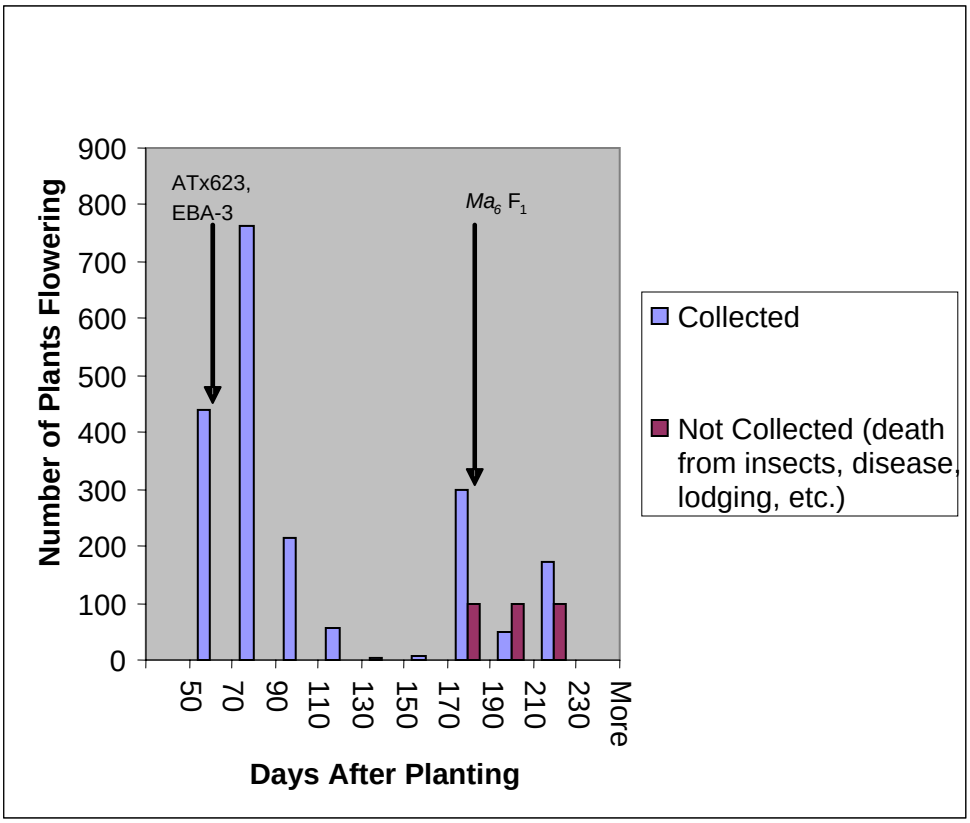

Fig. 8 Distribution of flowering in the 2003 *Ma*$_6$ BC_1F_1 population. The population was planted on 4 April 2003 in College Station, Texas. The flowering time of the inbred parents and F_1 are indicated. The number of plants not collected is based on an estimation of population size at the beginning of the growing season. Many plants collected at the latest date had not flowered, but were collected because most of the plants remaining in the population were dying, probably because of cold weather.

Xtxa2124, as expected. As described above, the modified flowering time phenotype was observed in the *Ma*$_6$ population grown in 2003. Since the samples were collected on a weekly or biweekly schedule, a comparison of flowering time can be made between plants with different genotypes at the *Ma*$_6$ locus. The ATx623/ATx623 PI plants started flowering at 65 days after planting (DAP), were the only class flowering for the first 20

days of flowering in this population, and the latest flowering of these samples was at 114 DAP among the 80 PI samples initially tested. The ATx623/EBA-3 modified class first flowered at 85 DAP. The modified plants comprised the major class of flowering samples from 91 DAP onward and fully represented the flowering plants from 120 DAP to the cessation of early flowering at 150 DAP. Late flowering plants all had an ATx623/EBA-3 genotype at the *Ma_6* locus among this group of 96 plants.

The late flowering PS plants from the *Ma_6* population grown in 2003 were screened with AFLP markers to look for plants with crossovers between the markers *etaccaa184* and *Xtxa2124*. Approximately 300 PS plants died and were not collected or screened. A total of 525 PS plants were collected and screened by a quick DNA extraction/PCR method described above using markers just below the *Ma_6* locus. These two SSR markers, *Xtxp434* and *12225.Contig1*, were developed from EST and sorghum methyl-filtered genomic sequences[213], respectively.

A total of 22 PS plants showed potential crossovers by the quick screening method with primers for *Xtxp434* and *12225.Contig1*. DNA from these potential crossover plants was then extracted using a FASTDNA kit, made into AFLP template as described above and screened with the AFLP markers *etaccaa184* and *Xtxa2124*, two markers closely flanking the *Ma_6* locus, and the AFLP markers *Xtxa4001*, *Xtxa7*, and *Xtxa3550*, all of which had been previously shown to be completely linked to the *Ma_6* locus. None of these 22 plants displayed a crossover between the markers *Xtxa4001*, *Xtxa7*, and *Xtxa3550*.

Since this quick screening method failed to detect any plants with crossovers between the markers *Xtxa4001*, *Xtxa7*, and *Xtxa3550* in the 525 PS plants from the Ma_6 populations grown in 2003, all 525 PS plants were used to make AFLP template as described above and were screened with the AFLP marker *etaccaa184,* which is located above the Ma_6 locus, to detect any crossover plants missed by initially screening with markers only below the Ma_6 locus. Again, no plants were detected with crossovers in the *Xtxa4001*, *Xtxa7*, and *Xtxa3550* genetic interval (data not shown).

Once again, screening a large number of PS plants had failed to break the linkage disequilibrium between the group of markers most closely associated with the Ma_6 locus. On the TAMU-ARS high density genetic map used as the reference map, the markers *Xtxa3550* and *Xtxa4001* both fall into a marker bin from 11.2 to 14.2 cM, while *Xtxa7* is a framework marker placed at 14.2 cM[38].

CHAPTER III

MODIFICATION OF THE *Ma$_5$/Ma$_6$* INTERACTION

INTRODUCTION

Phenotyping in the initial characterization of the *Ma$_5$/Ma$_6$* interaction indicated that delayed flowering was possibly due to complimentary dominant epistatic interaction between two genes, and these two genes were designated *Ma$_5$* and *Ma$_6$*[24,25]. Several lines of evidence are now indicating that the genetic cause of late flowering in the *Ma$_5$* and *Ma$_6$* populations may be due to more than two genes.

Phenotypic ratios in *Ma$_5$* and *Ma$_6$* populations

The original characterizations of the *Ma$_5$/Ma$_6$* interaction involved crossing EBA-3 with more than 10 different inbred lines and following segregation of photoperiod sensitivity/insensitivity in F_2 and BC_1F_1 progeny, and $F_{2:3}$ families[24,25]. Of the crosses made, all but two groups of progeny fit segregation ratios expected of a two gene interaction by χ^2 test[24,25], indicating that the genetic interaction producing photoperiod sensitivity in those crosses could be due to two gene complimentary dominant epistasis. Crosses with sorghum maturity genotypes revealed that the maturity differences in EBA-3 and most U.S. germplasm are not likely to be due to allelic differences at any of the 4 previously characterized sorghum maturity loci, *Ma$_1$-Ma$_4$*, hence the designation of 2 new maturity loci, *Ma$_5$* and *Ma$_6$*[24,25]. Most U.S. germplasm was hypothesized to be

$Ma_5Ma_5ma_6ma_6$, while EBA-3 was hypothesized to be *$ma_5ma_5Ma_6Ma_6$*. The interaction of these two genes could account for the segregation ratios reported in that work[24,25].

Phenotypic ratios from plants collected in the *Ma_6* population, that is, the ATx623*(ATx623*EBA-3) cross, consistently showed significant deviation from expected ratios in the populations collected in order to map the *Ma_6* locus in the present work. Out of 871 plants collected in the summer of 2002, 506 were photoperiod insensitive (PI), while 365 were photoperiod sensitive (PS) and late flowering. Every sorghum plant in the *Ma_6* population was sampled in that summer. The 506:365 ratio differs significantly from the 1:1 ratio expected in this BC_1F_1 population if one dominant floral repressor is segregating in this population [χ^2=22.8, and $P(\chi^2_{1df}>10.83)< 0.001$].

Among the PS plants from the *Ma_6* population grown in 2002, all 365 possess the genotype ATx623/EBA-3 at the *Ma_6* locus on chromosome 6, between the markers *Xtxa4001*, *Xtxa3550*, and *Xtxa7* confirming that a gene responsible for a portion of the photoperiod sensitive response comes from the EBA-3 parent and resides in that genetic interval. The story is not as clear in the PI plants from the same year. Out of 506 PI plants, 393 had the expected genotype at the *Ma_6* locus given the hypothesis that one floral repressor was segregating (ATx623/ATx623 at the markers *Xtxa4001*, *Xtxa3550*, and *Xtxa7*). One hundred and thirteen PI plants had an ATx623/EBA-3 genotype at this locus, and since this is a backcross population, the early flowering was not due to a recessive *ma_5* allele, since all plants had a dominant *Ma_5* gene on chromosome 2 from the recurrent parent ATx623. Therefore, there are more than 2 genes controlling photoperiod sensitivity in the *Ma_6* population. The additional genes have been

designated 'modifiers' of the *Ma5/Ma6* interaction. The modifier genes change the phenotype of dominant *Ma5/Ma6* plants from PS to PI. The evidence that there are at least two modifiers segregating in this cross will be presented below.

As mentioned previously, 506 out of 871 plants in the *Ma6* population grown in 2002 were photoperiod insensitive and early flowering, and these numbers deviated significantly from the expected 1:1 ratio under a two gene model. In that year, 58% of the plants in the population were PI and early flowering. In 2003, *Ma6* BC_1F_1 segregating plants were grown in two separate locations and phenotyped at weekly to biweekly intervals. While an attempt was made to collect every plant as it flowered, late flowering plants were underrepresented due to death from insect infestation, disease, lodging, etc. An estimate was made of the total number of plants in each of the two locations, however, and this number can be used along with the number of early flowering plants, all of which were collected, to compare flowering patterns from year to year in the Ma_6 population. In the combined *Ma6* locations from 2003, an estimated 2300 plants were grown, 1475 of which were PI and early flowering, while 825 plants were PS and late flowering. These numbers were significantly different from a 1:1 ratio by χ^2 analysis [χ^2=183.7, and P(χ^2_{1df}>10.83)< 0.001]. The percentage of early flowering plants in the 2003 *Ma6* population (64%) is close to the percentage seen in 2002 (58%), but slightly higher, and thus even farther from the 1:1 ratio of PI:PS plants expected if two genes control the photoperiod sensitive response in the *Ma5/Ma6* interaction, only one of which is segregating in this particular backcross. The PI: PS ratio has been found in this work to repeatedly differ from the expected 1:1 segregation

ratio. The numbers of PI and PS plants collected in this study are in disagreement with the ratios found in the original cross of ATx623*(ATx623*EBA-3), in which there were 106 PI and 130 PS, which is 106/236, or about 45% early flowering PI plants, a number that was not significantly different from a 1:1 PI:PS ratio in a χ^2 test[24]. The disagreement in segregation between the original segregation studies and this work may be due to differences in phenotyping. In this work, phenotyping was carried out at weekly to bi-weekly intervals, whereas in the original study phenotyping was carried out at a single point in time late in the growing season (mid-September)[24].

The modified *Ma6* plants from the 2002 population represent almost exactly 1/8th of the total population (113 PI/871 total). Since F_1 plants from the *Ma6* population grown in the years 2000, 2002, and 2003 were always PS, the modifier or modifiers must be inactive when heterozygous, or must require exposure to an imprinted allele if genetic imprinting is involved. These facts suggest at least two simple models to account for the modified phenotype. In the first model, a single modifier influences the *Ma5/Ma6* gene interaction, displacing 1/8th of the population from the PS class into the PI class (Fig. 9). This one gene model depends on gene dosage at the *Ma5* locus, and is a testable model given the markers developed in mapping the *Ma5* and *Ma6* loci. The second model involves two genes, and proof would require developing markers linked to both modifiers (Fig. 10). There are, of course, other more complicated models that could account for the modified phenotype, such as an allelic interaction involving epigenetic imprinting. The simple models suggested here assume no genetic linkage between the modifiers and the *Ma5* or *Ma6* maturity genes.

ATx623 X EBA-3
Ma*$_5$*Ma*$_5$ma*$_6$*ma*$_6$*A1A1*** ***ma*$_5$*ma*$_5$*Ma*$_6$*Ma*$_6$*A2A2***

F_1
Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A2

F_1 gametes	ATx623 gametes	BC_1F_1 progeny genotypes, ratios, and expected phenotypes	
Ma*$_5$*Ma*$_6$*A1	only ***Ma*$_5$*ma*$_6$*A1***	1) 1/8 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A1***	PS
Ma*$_5$*Ma*$_6$*A2		2) 1/8 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A2***	PS
Ma*$_5$*ma*$_6$*A1		3) 1/8 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A1***	PI
Ma*$_5$*ma*$_6$*A2		4) 1/8 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A2***	PI
ma$_5$***Ma*$_6$*A1***		5) 1/8 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A1***	PS (PI)
ma$_5$***Ma*$_6$*A2***		6) 1/8 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A2***	PS
ma$_5$*ma*$_6$*A1*		7) 1/8 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma*$_6$*A1A1***	PI
ma$_5$*ma*$_6$*A2*		8) 1/8 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma*$_6$*A1A2***	PI

Fig. 9 Single gene modifier model. This model is a gene dosage model. At least one dominant *Ma*$_5$ and one dominant *Ma*$_6$ allele are required for the PS response. *A1* is a modifying allele from a modifier gene. *A2* is a non-modifying allele of the same gene. #5 above is the modified class. A single copy of the *Ma*$_5$ floral repressor is rendered inactive by the modifier when homozygous, shifting plants expected to be PS into the PI class, whereas two functional Ma_5 alleles can't be overridden by the modifier and class #1 remains PS. This would give the 1/8th modified ratio seen in the 2002 *Ma*$_6$ population, and approximately 62.5% early flowering seen in the 2002 and 2003 *Ma*$_6$ BC_1F_1 plants.

ATx623 X EBA-3
Ma*$_5$*Ma*$_5$ma*$_6$*ma*$_6$*A1A1B1B1*** ***ma*$_5$*ma*$_5$*Ma*$_6$*Ma*$_6$*A2A2B2B2***

F_1
Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A2B1B2

ATx623 gametes
only ***Ma*$_5$*ma*$_6$*A1B1***

F_1 gametes

Ma*$_5$*Ma*$_6$*A1B1
Ma*$_5$*Ma*$_6$*A1B2
Ma*$_5$*Ma*$_6$*A2B1
Ma*$_5$*Ma*$_6$*A2B2
Ma*$_5$*ma*$_6$*A1B1
Ma*$_5$*ma*$_6$*A1B2
Ma*$_5$*ma*$_6$*A2B1
Ma*$_5$*ma*$_6$*A2B2
ma$_5$***Ma*$_6$*A1B1***
ma$_5$***Ma*$_6$*A1B2***
ma$_5$***Ma*$_6$*A2B1***
ma$_5$***Ma*$_6$*A2B2***
ma$_5$*ma*$_6$*A1B1*
ma$_5$*ma*$_6$*A1B2*
ma$_5$*ma*$_6$*A2B1*
ma$_5$*ma*$_6$*A2B2*

BC_1F_1 progeny genotypes, ratios, and expected phenotypes

1) 1/16 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A1B1B1***	PS (PI)
2) 1/16 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A1B1B2***	PS (PI)
3) 1/16 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A2B1B1***	PS
4) 1/16 ***Ma*$_5$*Ma*$_5$*Ma*$_6$*ma*$_6$*A1A2B1B2***	PS
5) 1/16 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A1B1B1***	PI
6) 1/16 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A1B1B2***	PI
7) 1/16 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A2B1B1***	PI
8) 1/16 ***Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$*A1A2B1B2***	PI
9) 1/16 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A1B1B1***	PS (PI)
10) 1/16 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A1B1B2***	PS (PI)
11) 1/16 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A2B1B1***	PS
12) 1/16 ***Ma*$_5$*ma*$_5$*Ma*$_6$*ma*$_6$*A1A2B1B2***	PS
13) 1/16 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma*$_6$*A1A1B1B1***	PI
14) 1/16 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma*$_6$*A1A1B1B2***	PI
15) 1/16 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma2A1A2B1B1***	PI
16) 1/16 ***Ma*$_5$*ma*$_5$*ma*$_6$*ma*$_6$*A1A2B1B2***	PI

Fig. 10 Two gene modifier model. At least one dominant *Ma*$_5$ and one dominant *Ma*$_6$ allele are required for the PS response. One modifier has to come from ATx623 (an AFLP marker, gen340, has already been linked to it). The second modifier could come from either parent. So the active modifier combination would be either *A1A1B1B1* or *A1A1B1B2*. In the first case, #'s 1 and 9 above would be added to produce 1/8th total modified plants. In the second case, #'s 2 and 10 above would be added to produce 1/8th total modified plants. Either case would give the 1/8th modified ratio seen in the 2002 *Ma*$_6$ population, and approximately 62.5% early flowering seen in the 2002 and 2003 *Ma*$_6$ mapping populations.

Mapping modifier genes

The fact that segregation ratios are distorted from a 1:1 ratio in this work indicates that more than 1 gene is segregating that influences flowering time, particularly in the *Ma₆* BC_1F_1 progeny. A non-segregating gene that maps to the *Ma₅* locus on chromosome 2 near *Xtxa3424*, and a segregating gene that maps to the *Ma₆* locus on chromosome 6 between *Xtxa4001* and *Xtxa3550* can account for the flowering time phenotype of most plants that arise from the *Ma₆* backcross. If additional genes exist that significantly influence flowering time in the *Ma₆* BC_1F_1 population**,** these genes should be linked to AFLP markers mapping to other locations within the genome. If these AFLP markers are not already linked to the TAMU-ARS sorghum reference map[37], the AFLP bands can be excised from the gel and sequenced in an attempt to link the marker to a genetic and/or physical locus.

MATERIALS AND METHODS

AFLP mapping of the *Ma₆* modifier

Random AFLP *Eco*RI+3/*Mse*I-C+2 primer combinations (described above) were examined in 27 unmodified *Ma₅_Ma₆ma₆* PS plants versus 27 modified *Ma₅_Ma₆ma₆* PI plants from the *Ma₆* population in order to genetically map the modifying factors as described above. A total of 54 *Eco*RI+3/*Mse*I-C+2 primer combinations were examined in this subset of 54 plants from the 2002 *Ma₆* population.

Cloning of genetic and epigenetic AFLP markers

Genetic and epigenetic AFLP markers that were identified as described above and were associated with the modified or unmodified phenotype were rerun on a second LI-COR gel and isolated either with a LI-COR Odyssey scanner according to the manufacturer's instructions, or by a manual method described below. Since the LI-COR sequencing instrument scans the gel as it runs in real-time, the marker appears on the computer screen only a few seconds after passing through the middle of the laser scanner window in the instrument. When the marker first appeared, the instrument was shut down, and the middle of the scanning window was marked on the glass plate with a sharpie to give the vertical position of the marker. The horizontal position of the marker on the gel was identified at the beginning of the gel run by placing two white paper strips on the long glass plate, then starting and stopping the instrument and moving the paper strips until the lane containing the marker had been exactly bracketed by the strips in the scanning window. The paper strips are visualized as solid black objects on the gel image. Once the gel was stopped and the marker position was identified in the scanning window, the gel assembly was removed from the instrument, the marker position was identified on both glass plates with a sharpie, the two glass plates were opened, and a small piece of acrylamide gel was excised at the position of the marker. The DNA marker was liberated from the excised gel by crushing it with a micropestle in a 1.5 ml microtube containing 50 μl of TE buffer, pH 8.0. Two microliters of this solution were then used for marker reamplification using unlabeled AFLP primers containing the same selective bases used to initially amplify the marker. PCR reaction conditions were the same as

those used in the original AFLP reaction, except the reaction volume was increased to 50 μl. The AFLP fragment was then separated from other potential PCR products on a 1.5% agarose gel, cut from the gel and cleaned with a Qiaquick gel extraction kit (Qiagen, Inc. Valencia, CA) according to the manufacturer's instructions, cloned into the pCR4-TOPO cloning vector (Invitrogen, Inc., Carlsbad CA) according to the manufacturers instructions, and sequenced with T3 or T7 primers. Sequencing reactions were performed as described previously and analyzed using Applied Biosystems instruments[214] (Applied Biosystems, Foster City, CA). Each AFLP fragment produced a single band on the agarose gel following reamplification.

Epigenetic marker sequence

Sequences of the epigenetic markers were used in several ways. The sequences were first used in BLAST analyses against GenBank databases to find similar sequences at an *E* value cutoff of < 1 e-10.. When genic sequences were obtained these were examined in an attempt to provide information on putative modifier function. Secondly, the marker sequences were compared to the rice genome sequence in an effort to identify modifier candidate genes, and link the modifier back to the colinear location in the sorghum genome. Thirdly, the sequence information was used to determine whether the epigenetic band was the result of hyper or hypomethylation based on the presence or absence of cryptic internal *Eco*RI restriction sites. Finally, the epigenetic marker sequences were used to create new STS (sequence tagged site) markers for PCR-based

screening the sorghum BAC libraries to potentially identify both the physical and genetic locations of the modifier genes as previously described[214].

Epigenetic bands in maturity standards

The bands linked to the modified or unmodified phenotype that segregated in an epigenetic fashion in the *Ma*$_6$ BC_1F_1 plants were examined in a number of sorghum maturity standards. These maturity standards included 60M, 80M, 90M, 100M, and Hegari, with the maturity genotypes listed in Table 1. Leaf tissue from vegetatively growing and flowering plants was collected from each of these genotypes and used to produce AFLP template as described above. Primer combinations producing epigenetic bands in the *Ma*$_6$ population were used to screen the maturity standards.

Table 1 · Genotypes of sorghum maturity standards (adapted from [26])

Maturity standard	Genotype	Days to flower
100M	*Ma*$_1$*Ma*$_2$*Ma*$_3$*Ma*$_4$	90
90M	*Ma*$_1$*Ma*$_2$*ma*$_3$*Ma*$_4$	82
80M	*Ma*$_1$*ma*$_2$*Ma*$_3$*Ma*$_4$	68
60M	*Ma*$_1$*ma*$_2$*ma*$_3$*Ma*$_4$	64
Hegari	*Ma*$_1$*Ma*$_2$*Ma*$_3$*ma*$_4$	70

RESULTS AND DISCUSSION

Mapping modifier genes

The fact that segregation ratios are distorted from a 1:1 ratio in this work indicates that there is more than 1 gene segregating that influences flowering time particularly in the *Ma₆* BC_1F_1 progeny. A non-segregating gene that maps to the *Ma₅* locus on chromosome 2 near *Xtxa3424*, and a segregating gene that maps to the *Ma₆* locus on chromosome 6 between *Xtxa4001* and *Xtxa3550* can account for the flowering time phenotype of almost 90% of the plants that arise from the *Ma₆* backcross. There are 113/871 plants in the 2002 *Ma₆* population that flower early in spite of dominant alleles at the *Ma₅* and *Ma₆* loci. Random AFLP primer combinations were examined in 27 unmodified *Ma₅__Ma₆ ma₆* PS plants versus 27 modified *Ma₅__Ma₆ ma₆* PI plants in order to genetically map the potential modifier genes. A total of 54 AFLP primer combinations were examined in this subset of plants from the 2002 *Ma₆* population. The 54 primer combinations yielded 66 markers linked to the recombinant inbred map[38], and about 90 additional markers not linked to the recombinant inbred map, for a total of 156 markers. Although numerous AFLP markers appeared to be linked to the modified phenotype or the unmodified phenotype, only one marker linked to phenotype was inherited in a predictable genetic fashion. That is, only one marker appeared in one of the parents, also in the F_1, and predominantly in the *Ma₅Ma₆* PS plants (data not shown). The remaining markers linked to the modified class or the unmodified class showed an epigenetic pattern of inheritance. That is, the marker would be absent in both parents, absent in the F_1, and would be present in the BC_1F_1 generation, and would segregate

with either the modified PI class or the unmodified PS class (Fig. 11). In one case an epigenetic marker displayed a light/dark difference between flowering and vegetative

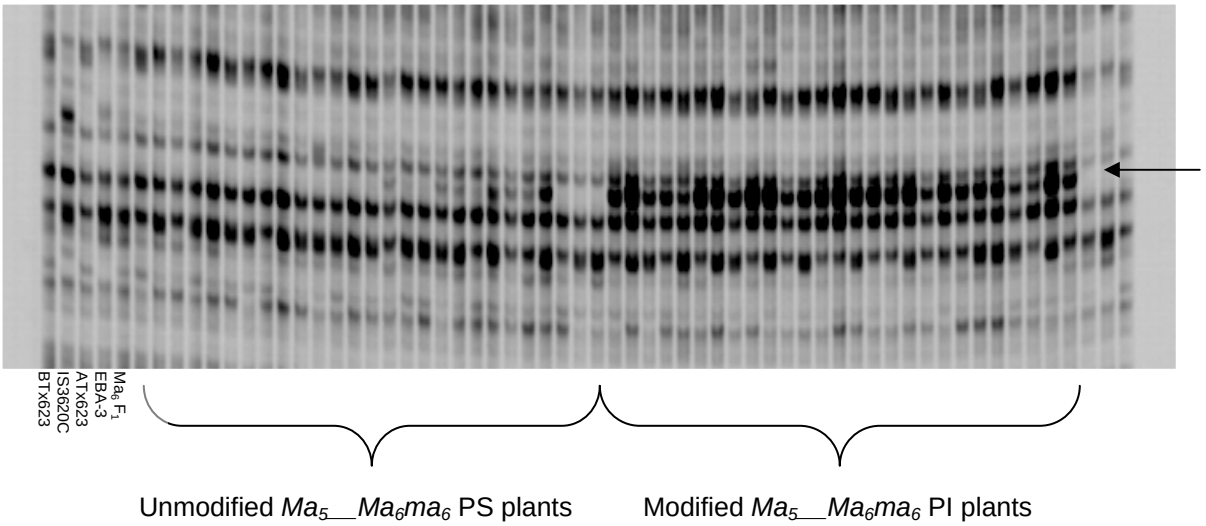

Fig. 11 Gel image of epigenetic AFLP band. The *epi155* marker shows up in all modified plants, but only 8/27 unmodified plants. It is faintly present in the IS3620C line, but not the parents (ATx623 and EBA-3) nor the *Ma*$_6$ F_1 plants that gave rise to the segregating modified and unmodified BC_1F_1 progeny. The unmodified plants with the band all had a very tall phenotype as well.

samples and was present in the parents and F_1 plants, and the light/dark difference was also seen in the *Ma*$_6$ BC_1F_1 modified and unmodified dominant *Ma*$_6$ plants. Since the AFLP markers were created by *Eco*RI/*Mse*I digests of genomic DNA, the assumption was made that these epigenetic markers were the result of the impaired or blocked ability of *Eco*RI to cut sites that had a methylated cytosine in the recognition sequence GAATTC[215-217]. In addition to effects on flowering time, the epigenetic bands were also associated with a very tall phenotype, so at least some of these

methylation/demethylation events are associated with pleiotropic phenotypic effects. One or more of the epigenetic AFLP markers could be modifiers of the *Ma5/Ma6* interaction, or they could simply be additional epigenetic targets sites that are differentially methylated in the transition to flowering in the *Ma5/Ma6* population. A process of demethylation has for some time been associated with induction of flowering, but this process is associated with vernalization[218]. It seems unlikely that a crop of tropical origin such as sorghum would retain a cold-stimulated flowering response. However, it is well known that plants carry out maintenance and *de novo* methylation[219], and that these changes in cytosine methylation are intimately connected to chromatin dynamics involved in epigenetic developmental switch mechanisms[220].

Epigenetic bands in maturity standards

A randomly chosen subset of the AFLP primer sequences that had produced the epigenetically inherited markers associated with the modified phenotype were used to screen vegetative and floral induced samples from the maturity standards 60M, 80M, 90M, 100M, and Hegari, as well as the parents and F_1 plants from both the *Ma5* and *Ma6* populations. The epigenetic markers *epi155*, *epi225*, and *epi395* were present primarily in the modified class of *Ma6* plants. These markers did not show up in vegetative or flowering samples from the maturity standards (data not shown). The epigenetic phenomena they represent may therefore be specific to flowering in the *Ma6* population. The *epi40* marker was also present primarily in the modified class, but appeared less predictably than the other epigenetic bands when examined in the maturity standards. It

was not present in the parents of the *Ma6* population, but was present in the A3Tx436 parent of the *Ma5* population in flowering samples only, and was inconsistently present in F_1's of the *Ma5* and *Ma6* populations (data not shown). This marker is associated with the modified phenotype, but appears to be additionally involved in epigenetic phenomena that are not connected to flowering. The *epi285* marker was different than the other epigenetic markers in that it was present in both the modified and unmodified *Ma6* plants, and in both parents and F_1 plants, but showed differences in intensity between different plants. It was light in the modified class and dark in the unmodified class. The modified class was therefore assumed to be more highly methylated at this marker locus, since the gel band was less intense in modified samples and the cloned sequence revealed no internal *Eco*RI sites (data not shown). In all maturity standards, *Ma5*, and *Ma6* genotypes tested, this band was lighter in flowering samples and darker in vegetative samples. This marker locus is possibly a site that all sorghum varieties methylate at the transition from vegetative to floral growth. The fact that this band is segregating as light/dark among modified and unmodified plants, respectively, may be due to the fact that the modified plants had flowered and most of the unmodified plants had not at the time of collection for plants in the 2002 *Ma6* population.

Epigenetic marker sequence

Four epigenetic AFLP bands were excised from acrylamide gels and cloned. An AFLP band may be composed of a single PCR amplicon, or it may be composed of multiple PCR amplicons that happen to be of equivalent lengths, and thus comigrate in an acrylamide gel. Two of the cloned AFLP bands represented single PCR amplicons, one

of the bands was composed of two amplicons, and one of the bands was composed of three amplicons, so that 7 sequence tags were generated by cloning 4 AFLP bands. Four of the 7 cloned epigenetic bands contained internal *Eco*RI sites, indicating that these epigenetic bands were the result of impaired restriction at these internal recognition sites, while 3 of the 7 cloned bands had no internal *Eco*RI sites, and thus were likely produced as a result of increased restriction at a flanking *Eco*RI site. The epigenetic marker sequences were used to create STS markers for screening the sorghum BAC pools to potentially identify both the physical and genetic locations of the modifier genes. In total, 4 epigenetic AFLP markers and 1 genetic AFLP marker were cut from LI-COR acrylamide gels, cloned, and sequenced.

The smallest epigenetically inherited band that was cloned and sequenced was the *epi155* band that was ~155 bp in size, and produced with the AFLP primers *Eco*RI-CTG/*Mse*I-CGA. This sequence had one internal *Eco*RI site. BLASTN analysis of this sequence produced no nucleotide alignments over 19 bp in length. Additionally, TBLASTX analysis against all plant species produced no significant alignments. There are only short stretches of homology between this sorghum *epi155* sequence and any sorghum ESTs. Primers produced from this sequence for PCR analysis of the BTx623 and IS3620C sorghum BAC pools were Epi155F-CCCGACTTTCGTTCACGTAG, and Epi155R-AGTCGGCTTTCTTGGGAACT. These primers produced a fragment in almost every BAC pool, indicating this sequence is most likely repetitive in nature.

The second epigenetic marker cloned, *epi225*, was ~225 bp size and produced with the primer combination *Eco*RI+GAA/*Mse*I+CTT. This sequence had one internal

*Eco*RI site. BLAST analysis of this sequence produced no significant hits. Primers designed to reamplify this sequence, Epi225F-TTGAATGAATTCCTAAGAACTCGTAAT, and Epi225R-CGTGTCTGGTTGTAGTTCTTTGAG, also amplified almost every BAC pool, indicating that this sequence was also likely repetitive.

The third epigenetic marker cloned was the *epi285* marker, which was ~285 bp in size and was produced with the primers *Eco*RI+ACC/*Mse*I+CAG. When cloned, this marker was found to be composed of two different sequences, which were referred to as *epi285a* and *epi285b*. Neither of these sequences had internal *Eco*RI sites. TBLASTX analysis against rice aligned the *epi285a* marker sequence with a rice BAC on chromosome 3 at 87.4 cM (*E* value 1e-26). This rice BAC, AC133003, contains a carpel factory-like gene (AAT76308) in the area of alignment with the *epi285a* marker. There is also alignment at an *E* value of 3e-26 with a rice chromosome 9 BAC at 21.4 cM, AP005782. The area of alignment on chromosome 9 contains another carpel factory-like gene. The sorghum cDNA BG948578 also shares 100% homology with the *epi285a* sequence for 155 bp. Screening sorghum BAC pools with primers designed for *epi285a* marker sequence, Epi285aF-AAGGACCATCCATTGTCTGC, and Epi285aR-TGGAGGTCAGTGATGCCATA, revealed five IS3620C BACs potentially containing this sequence: 50N20, 50N22, 51J15, 70F20, and 70F22. The BACs 51J15 and 70F22 have been individually confirmed to have a PCR band of the correct size when amplified with these primers (data not shown). These primers produced no positives in the BTx623 BAC pools. The *epi285b* sequence, although different from the *epi285a*

sequence, has homology (*E* value 1e-40) with RNA helicases on rice chromosome 3 at 137.9 cM, (AC092558), and on rice chromosome 2 at 0 cM (AP004851) when analyzed by TBLASTX against rice. The *epi285b* sequence had no significant homology to any sorghum cDNAs, and primers designed to amplify this sequence, Epi285bF-AATATGCCAAACGCTTCGAC, and Epi285bR-GTGCTGGCAAAACAAATGTC, identified only one sorghum IS3620C BAC, 53C14, and no BTx623 BACs when used to screen the IS3620C and BTx623 BAC libraries (data not shown).

The fourth and final epigenetic marker cloned was *epi395*, a 395 bp marker produced with the primer combination *Eco*RI+CTG/*Mse*I+CTC. This band was found to be composed of 3 different sequences, *epi395a*, *epi395b*, and *epi395c*. The *epi395a* sequence had two internal *Eco*RI sites and the *epi395b* sequence had one, while the *epi395c* sequence had no internal *Eco*RI sites. The *epi395a* sequence had homology only with sorghum leviathan retroelements (*E* value 3e-24) by TBLASTX analysis of all plants and was not used to screen sorghum BAC libraries. The *epi395b* sequence also had homology with the same sorghum retroelements at an *E* value of 4e-12 and was not used to screen sorghum BAC libraries. TBLASTX of the *epi395c* sequence against all organisms returned a number of homologous sequences, notably a rice chromosome 2 BAC at 50 cM, AP005398 (9e-33), an uncharacterized region of a sorghum BAC, AY542311 (1e-50), and a maize transcriptional activator, AY078063 (2e-36). The *epi395* sequence has not yet been used to screen the BTx623 or IS3620C BAC libraries.

Genetic marker sequence

A single genetically inherited AFLP band was found to be linked to the modified phenotype. This band, called *gen340*, produced with the primer combination *Eco*RI+TGA/*Mse*I+CAT, was ~340 bp in size and was present in the EBA-3 parent, in the F_1, in 25/27 $Ma_5\_\_Ma_6ma_6$ unmodified PS plants, and in 1/27 $Ma_5\_\_Ma_6ma_6$ PI modified plants. Because of the close linkage of the EBA-3 allele at this locus with the unmodified phenotype, the modifying allele at this locus comes from the ATx623 parent. This marker is not present in the TAMU-ARS mapping population, so the marker had to be cloned and sequenced in order to attempt to locate it on the genetic map, just as the epigenetic bands described above. The marker inserted into the plasmid pCR4-TOPO as a concatemer of 2 sequences of equivalent size. The first half of the concatemer, *gen340a*, has homology to sorghum retroelements and sorghum cDNA BE596570 (*E* value 1e-83). Primers designed to amplify this sequence, Gen340aF-GCTCATACTTCGCCTTCCAG, and Gen340aR-AAGCATATTCACCGCAAGGT failed to identify unique BACs when screening the BAC libraries (data not shown). The second half of the concatemer, *gen340b*, has little homology with any GenBank sequences. The *gen340a* sequence had no internal *Eco*RI sites, while the *gen340b* sequence had one. Primers based on *gen340b*, Gen340bF-CAAACCAGCGAGCCATATTT, and Gen340bR-AGGAATTGCGTGACTTCCAC identified one BTx623 BAC from the sorghum BAC library, sbb6323 (66g11), and 3 IS3620C BACs from the library, 69D9, 69D11, and 69C12. The BTx623 BAC 66g11 was digested with *Eco*RI and *Xho*I, subcloned and 96 clones were sample sequenced as

previously described[214]. The sequences from this BAC produced multiple alignments with the rice genome in BAC AL662935 on rice chromosome 4 at 3.1 cM, and alignments to other areas of the rice genome as well. The rice chromosome 4 BAC is colinear with the top of sorghum chromosome 6 just about 10 cM above the *Ma$_6$* locus on the sorghum recombinant inbred map[37], and about 47 cM above the area colinear with the *Ma$_6$* locus on the rice genetic map[221,222]. Scanning the annotation of the region in rice reveals no obvious candidates for modifiers of the *Ma$_5$/Ma$_6$* interaction. If the genetic modifier from the ATx623 parent is located at the top of chromosome 6, the *Ma$_5$__Ma$_6$ma$_6$* modified plants should always have ATx623 alleles in that area of the genome if there is only one modifier, or should have ATx623 alleles in that area more often than expected by chance if more than one modifier exists. In order to examine these possibilities, the segregation of markers at the *Ma$_6$* locus (*etaccaa184* at the *Ma$_6$* locus and *Xtxp6* above the *Ma$_6$* locus) was examined in 27 *Ma$_5$__Ma$_6$ma$_6$* PI modified plants. *Xtxp6* is located at 0 cM on the very top of sorghum chromosome 6 and *etaccaa184* is located at about 10 cM on sorghum chromosome 6, just above the *Ma$_6$* locus. The genotype at these two marker loci was also determined for 27 *Ma$_5$__Ma$_6$ma$_6$* unmodified PS plants. Among the 27 unmodified plants, all were heterozygous ATx623/EBA-3 for that 10 cM block of the genome. Among the 27 modified plants 18/27 were heterozygous, and 9/27 were homozygous ATx623/ATx623 at the *Xtxp6* locus at 0 cM, and 27/27 were heterozygous ATx623/EBA-3 at the *etaccaa184* locus about 10 cM below, and very close to *Ma$_6$*. There are at least four important pieces of evidence regarding the modifier genes in this test: 1) none of the plants in the

unmodified class had a crossover, while there should have been about 3 crossover plants in that 10 cM interval if no genes above the *Ma*$_6$ locus on chromosome 6 affect flowering time; 2) the 9 crossovers out of 27 plants in the modified class in this 10 cM interval is about 3 times higher than would be expected, suggesting that an allele that modifies flowering time may be located in this genetic interval; 3) the number of crossover plants in the modified class is less than 100%, suggesting that if the genetic modifier from ATx623 is in fact located at the top of chromosome 6, it cannot fully account for the modified phenotype, and additional modifier genes must be involved, and 4) the modifier must be homozygous ATx623/ATx623 in order to displace dominant *Ma*$_5$*/Ma*$_6$ PS plants into the modified PI class, because every plant in the *Ma*$_6$ population was a result of a backcross to ATx623 and thus had at least one ATx623 allele for every gene, but not every plant showed the modified phenotype. Additionally, F_1 plants never show the modified phenotype and are all heterozygous at this locus. While it is possible that the number of crossovers between *etaccaa184* and *Xtxp6* could have been affected by interference from crossovers below these markers, the possibility is doubtful. The genotypes of the 27 modified and 27 unmodified plants were actually determined at marker *Xtxa2124* as well, located at 16.8 cM on the reference map and below the *etaccaa184* marker. All 27 modified plants were heterozygous at *Xtxa2124*. Three of the 27 unmodified plants had crossovers between *Xtxa2124* and *etaccaa184*. Two of these crossovers were between *Xtxa2124* and *Xtxa7*, so that everything above the *Ma*$_6$ locus was heterozygous in these two plants. One unmodified plant had a crossover between *Xtxa4001* and *etaccaa184*, and also had a second crossover between

etaccaa184 and *Xtxp6* near the modifier, making it a double crossover plant in this genetic interval. Like all of the other unmodified $Ma_5\_\_Ma_6ma_6$ plants, this plant had an ATx623/EBA-3 genotype at *Xtxa4001*, *Xtxa7*, and *Xtxa3550* at the Ma_6 locus, and it had the same genotype at *Xtxp6* near the modifier locus.

The IS3620C BAC 69D9 was also sequence scanned, and several of the sequences aligned with a rice chromosome 9 BAC at 88.2 cM, AP006548. This BAC contains genes with DNA binding domains, but no obvious candidates for a modifier of the Ma_5/Ma_6 interaction exist in this area of the rice genome (data not shown).

Testing the one-gene modifier model

In order to test the one gene modifier model described above and in Fig. 9, segregation of the *gen340* marker that was found to be linked to a modifier was examined in a plate of 27 $Ma_5\_\_Ma_6ma_6$ PS plants and in 27 $Ma_5\_\_Ma_6ma_6$ PI modified plants, along with the segregation of two codominant markers flanking the Ma_5 locus, *Xtxa2513* and *Xtxp100*, and genetic markers flanking the Ma_6 locus (*etaccaa184* and *Xtxa2124*). Overall, the Ma_6 population should segregate in a 1:1 ratio of Ma_5Ma_5 and Ma_5ma_5 plants (see population descriptions above). The one gene modifier model depends on gene dosage at the Ma_5 locus. In the model, homozygous dominant Ma_5 plants ($Ma_5Ma_5Ma_6ma_6$) would be PS, while heterozygous plants ($Ma_5ma_5Ma_6ma_6$) would be the modified class and PI. An examination of the genotype of modified plants at the Ma_5 locus reveals they show segregation in the expected 1:1 ratio [χ^2=0.25, and $P(\chi^2_{1df}>3.84)< 0.05$]. In other words, the modified phenotype is not due to an interaction

between gene dosage at the *Ma*$_5$ locus and a modifier gene from ATx623 in the area of the *gen340* marker.

In summary, the modifiers in the *Ma*$_6$ population represent genes that influence flowering time to a lesser extent than the genes at the *Ma*$_5$ and *Ma*$_6$ loci. But these genes are capable of displacing a portion of the plants with dominant *Ma*$_5$ and *Ma*$_6$ genes into the photoperiod insensitive, early-flowering class. One of these modifier genes may be linked to the *Ma*$_6$ locus, approximately 10 cM above the *Ma*$_6$ gene, and located at the very end of the p arm of sorghum chromosome 6. A second modifier gene may exist in another locus elsewhere in the sorghum genome.

CHAPTER IV

PROGRESS TOWARDS MAP-BASED IDENTIFICATION OF *Ma_5/Ma_6*

INTRODUCTION

While genetic mapping of the *Ma_5* and *Ma_6* loci with AFLP and SSR markers was significant, it only placed these genes in specific locations within the sorghum genome. Identification of the *Ma_5* and *Ma_6* genes would require cytogenetic, physical, and comparative genetic approaches. The genetic markers flanking these genes narrowed the genetic interval in which they reside to about 3 centimorgans for both the *Ma_5* and *Ma_6* loci. Depending on the amount of recombination around these loci, 3 cM could represent either a very large or a very small physical distance. In order to estimate the physical distance between the nearest markers flanking each locus, fluorescent *in situ* hybridization (FISH) was performed on each locus by Jeong-Soon Kim, a student in the laboratory of Dr. David Stelly. This method involves fluorescently labeling DNA sequences linked to flanking markers (in this case BACs) and hybridizing the fluorescently labeled sequence to sorghum chromosome spreads. The distance between the fluorescent tags on the chromosome provides an estimate of physical distance spanned by the locus. Cloning the *Ma_5* and *Ma_6* genes would also be facilitated by building a contiguous physical sequence of DNA comprised of overlapping BACs (a BAC contig) that spanned both loci in order to search for candidate genes in the genetic

interval. In order to build BAC contigs spanning these two loci, a comparative genetic approach was employed[214]. Sorghum sequences that aligned with rice chromosomal regions colinear to the *Ma5* and *Ma6* loci were used to screen sorghum BAC libraries. Sorghum BACs containing homologous sequences were fingerprinted by restriction enzyme digestion and placed in a growing contig until the loci were spanned (P. Klein, personal communication). Additionally, when a rice gene colinear to *Ma5* or *Ma6* was involved in a flowering response in other plants, the sorghum homolog was examined for expression differences by qRT-PCR in the parents and F_1 of the *Ma5* and *Ma6* populations.

MATERIALS AND METHODS

FISH analysis

FISH analysis of the *Ma5* and *Ma6* loci was performed by Dr. Jeong-Soon Kim in the Laboratory of David Stelly according to previously described methods[207,223].

Screening sorghum BAC pools

Six dimensional sorghum BAC pools constructed from the genotypes BTx623 and IS3620C were screened for sorghum sequences aligned with the rice genome or for flowering homolog sequences as previously described[39,214].

Quantitative RT-PCR

RNA was isolated from flowering and non-flowering sorghum by grinding individual meristems under liquid nitrogen with a mortar and pestle. RNA was extracted using a Trizol-based RNA extraction method (Molecular Research Center, Cincinnati). RNAs were converted to cDNA template for qRT-PCR using random hexamer primers and Multiscribe reverse transcriptase (Applied Biosystems, Foster City, CA). qRT-PCR was performed in duplicate 10 μL reactions using Sybr Green mastermix (Applied Biosystems) for the sample reactions and TAQMAN Universal PCR mastermix (Applied Biosystems) with VIC probe labeling for ribosomal control reactions. No-template control reactions using untranscribed RNA controls confirmed that no interfering products derived from genomic DNA were present. Primers for amplifying genes of interest were designed using Primer Express (Applied Biosystems).
Amplification specificity was determined by dissociation curve analysis. Mean induction folds were calculated as $2^{(\Delta\Delta CT)}$, and SD range of replicate reactions was calculated by: upper error bar = $2^{(\Delta\Delta CT + s)}$, lower error bar = $2^{(\Delta\Delta CT - s)}$, where: $\Delta\Delta CT = (\Delta CT_{\text{control cDNA}}) - (\Delta CT_{\text{treatment cDNA}})$, $\Delta CT = (\text{mean CT cDNA}_{\text{test primers}}) - (\text{mean CT cDNA}_{\text{ribosomal primers}})$, $S = \sqrt{[(\text{sd of CT}_{\text{test primers}}^2) + (\text{sd of CT}_{\text{ribosomal primers}}^2)]}$.

RESULTS AND DISCUSSION

FISH analysis of the *Ma₅* locus was successful, while FISH analysis of the *Ma6* locus was not. Repeated attempts to create fluorescent probes on the p arm of chromosome 6

near the *Ma$_6$* locus have produced probes that bind in many areas of the genome in a manner characteristic of repetitive DNA (J.-S. Kim personal communication), so FISH analysis of the *Ma$_6$* locus provided no clue as to the physical distance between closely linked flanking markers. FISH analysis of the *Ma$_5$* locus did provide an estimate of the physical distance between closely linked flanking markers. The AFLP markers *Xtxa3424* and *Xtxp100* are linked to the BACs sbb11773 and sbb4217, respectively (P. Klein, personal communication). These BACs hybridized to sorghum pachytene chromosome spreads indicating a physical distance of roughly 5 Mb between these markers on chromosome 2 (Fig. 12). Sorghum BACs containing the AFLP marker *Xtxa3424* have been linked to rice chromosome 7 BAC AP004299 at 60.8 cM. BACs containing the SSR marker *Xtxp100* have been linked to rice chromosome 7 BAC AP004674 at 73.2 cM (P. Klein, personal communication). The genetic distance in rice between these two BACs agrees closely with the genetic distance in the TAMU-ARS sorghum map between markers *Xtxa3424* and *Xtxp100,* which is ~10-15 cM[38]. Additionally, the physical distance between these two BACs in rice is about 3.5 Mb, which is similar to the estimated physical distance of about 5 Mb in sorghum. Some difference in size is expected due to the size difference between the genomes of rice (389 Mb)[5] and sorghum (818 Mb)[9,10]. A BAC contig spanning 5 Mb would consist of well over 30 BACs with an average insert size of 150 kb, so efforts were shifted into marker development around the locus in order to narrow the interval in which the *Ma$_5$* gene resides.

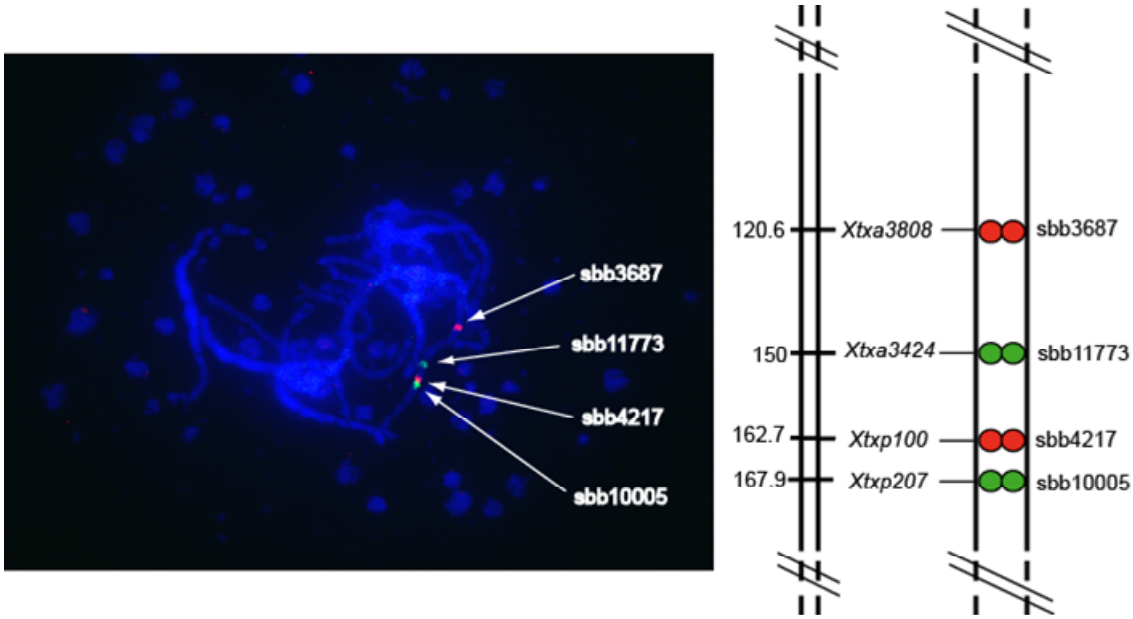

Fig. 12 FISH analysis of markers linked to the *Ma*$_5$ locus. *Sorghum bicolor* BACs (sbb #s) linked to genetic markers around the *Ma*$_5$ locus were used to probe sorghum pachytene chromosome spreads. The *Ma*$_5$ locus is near the markers *Xtxa3424* and *Xtxp100*. Figure kindly provided by Dr. Jeong-Soon Kim.

Screening BAC pools for flowering time candidate genes

Initially, primers for several flowering-time genes identified in other plant species were used to screen the BTx623 and IS3620C sorghum BAC pools comprising an integrated genetic and physical map of sorghum[38,39] in an attempt to identify BACs that contained flowering time genes in the *Ma*$_5$ and *Ma*$_6$ chromosomal regions. While this effort was successful in identifying BACs containing flowering-time gene homologs, none of the

BACs containing these genes were linked to a chromosomal region near either maturity locus (data not shown).

Several primer sets for genes known to influence flowering time in *A. thaliana* or rice were used to screen BTx623 and IS3620C BAC libraries. Although homologs of several flowering time genes were discovered, none of these genes were extremely close to areas colinear with the Ma_5 or Ma_6 loci. *Hd1* is a homolog of the *A. thaliana* gene *CO*[197], which is part of a family of *CO*-like genes[224]. One of these genes was found on sorghum BAC sbb22641 on chromosome 4 near *Xtxp327*. Other flowering time gene homologs identified in this effort include a *FRI*-like gene or genes on several BACs that were unlinked to the genetic map; a *TFL1*-like gene also unlinked to the genetic map; 2 genes that flank *CRY1* in rice were both found on sorghum chromosome 6 about 35 cM away from the Ma_6 locus; and a *LD*-like gene was found on sorghum chromosome 3 at approximately 10 cM (data not shown).

Examination/Development of cDNA/EST/STS markers

Simultaneous with the initial use of flowering-time gene sequences to screen the BAC pools for Ma_5 and Ma_6 candidate genes, a sorghum cDNA sequencing project and a rice whole genome sequencing project were in progress. Because of the colinearity among cereal genomes, the sorghum cDNA information (or any other sorghum sequence information) can be comparatively aligned with the rice genome. By screening the sorghum BAC pools with PCR primers complementary to these short cDNA sequences, or Expressed Sequence Tags (ESTs), BACs containing the ESTs can be identified and

potentially assigned a chromosomal location based on colinearity with rice and/or linkage to the integrated genetic and physical map of sorghum. In a separate ongoing project in the sorghum genomics laboratory, the 10 sorghum chromosomes were aligned to the rice genome by conducting sequence scans of genetically mapped sorghum BACs and aligning the gene sequences obtained to the rice pseudomolecule (P. Klein, personal communication). This sequence information along with EST screening of the 6D BAC pools was also being used for sorghum physical map construction. In this method, low or single copy sorghum ESTs that align *in silico* to regions of the rice genome where a gap occurs in the sorghum physical map are amplified in the BAC pools to identify colinear sorghum BACs. These BACs are then fingerprinted using a modified version of high information content fingerprinting (HICF) to aid in gap filling[214]. In the current work, STSs in the Ma_5 and Ma_6 regions were used to identify BACs located within these two loci to aid in physical map construction across these regions[39,214]. Additionally, the sequences aligned with these two loci that contained SSRs or SNPs were used to identify new polymorphisms within the maturity populations and the recombinant inbred mapping population[38,204]. Although the sequences aligned with the Ma_5 colinear region on rice chromosome 7 were useful in identifying sorghum colinear BACs and extending existing sorghum BAC contigs, those aligned with the Ma_6 colinear region on rice chromosome 4 were seldom of utility in comparative mapping. As mentioned above, attempts to use FISH probes in the area around the Ma_6 locus have shown that this region is repetitive and heterochromatic in nature, and has been shown to be at the border of a heterochromatic region[225]. Sorghum EST sequences aligned with rice

chromosome 4 in the area colinear to the *Ma*$_6$ locus often amplified all BAC pools, indicating that the ESTs were part of repetitive sequence. Additionally, BACs identified by low copy ESTs aligned with the *Ma*$_6$ colinear region on rice chromosome 4 seldom align with the *Ma*$_6$ BAC contig produced by fingerprinting methods (P. Klein, personal communication). The inability of the two methods to align the *Ma*$_6$ region of sorghum chromosome 6 and rice chromosome 4 raises the possibility that the *Ma*$_6$ locus is less colinear to the rice genome than is the *Ma*$_5$ locus.

The STS markers aligned with the rice genome that were used either to screen the sorghum BAC pools or the *Ma*$_5$ and *Ma*$_6$ loci for polymorphisms are listed in Tables A1 and A2, respectively (appendix). BACs that produced a positive signal when screened with these STS markers and other markers around the *Ma*$_5$ and *Ma*$_6$ loci were then fingerprinted by Dr. Klein's laboratory to aid in contig construction. Some of the STS markers used in this effort and their alignment with the rice genome in the areas colinear to the *Ma*$_5$ locus are shown in Fig. 13. Additionally, several STS markers were used to aid contig construction around a putative modifier locus on sorghum chromosome 6 (data not shown).

The fingerprinting efforts performed by Dr. P. Klein's laboratory using BACs linked to the two maturity loci by STS, SSR, and AFLP markers resulted in BAC contigs that contain both the *Ma*$_5$ locus on chromosome 2 and the *Ma*$_6$ locus on chromosome 6 (Fig. 14). Newly developed molecular markers indicated that the *Ma*$_5$ gene is very close to the AFLP marker *Xtxa3424* (data not shown).

qRT-PCR of candidate genes

Once the *Ma*$_5$ and *Ma*$_6$ maturity loci had been generally aligned with the rice genome, an ongoing effort began to scan the published rice sequence in these areas for potential *Ma*$_5$ and *Ma*$_6$ candidate genes. Gene expression differences in potential candidate genes between the vegetative meristems and floral meristems of the parents and F_1's from both the *Ma*$_5$ and *Ma*$_6$ populations were quantified using qRT-PCR. To date, none of the candidate genes screened have shown a differential pattern of expression that could explain the late-flowering response in the *Ma*$_5$ and *Ma*$_6$ backcross populations. It is possible that the late flowering response in one or both populations is not due to differences in gene expression, but instead is due to interaction differences, post-translational modification, etc., that would not be detected by qRT-PCR.

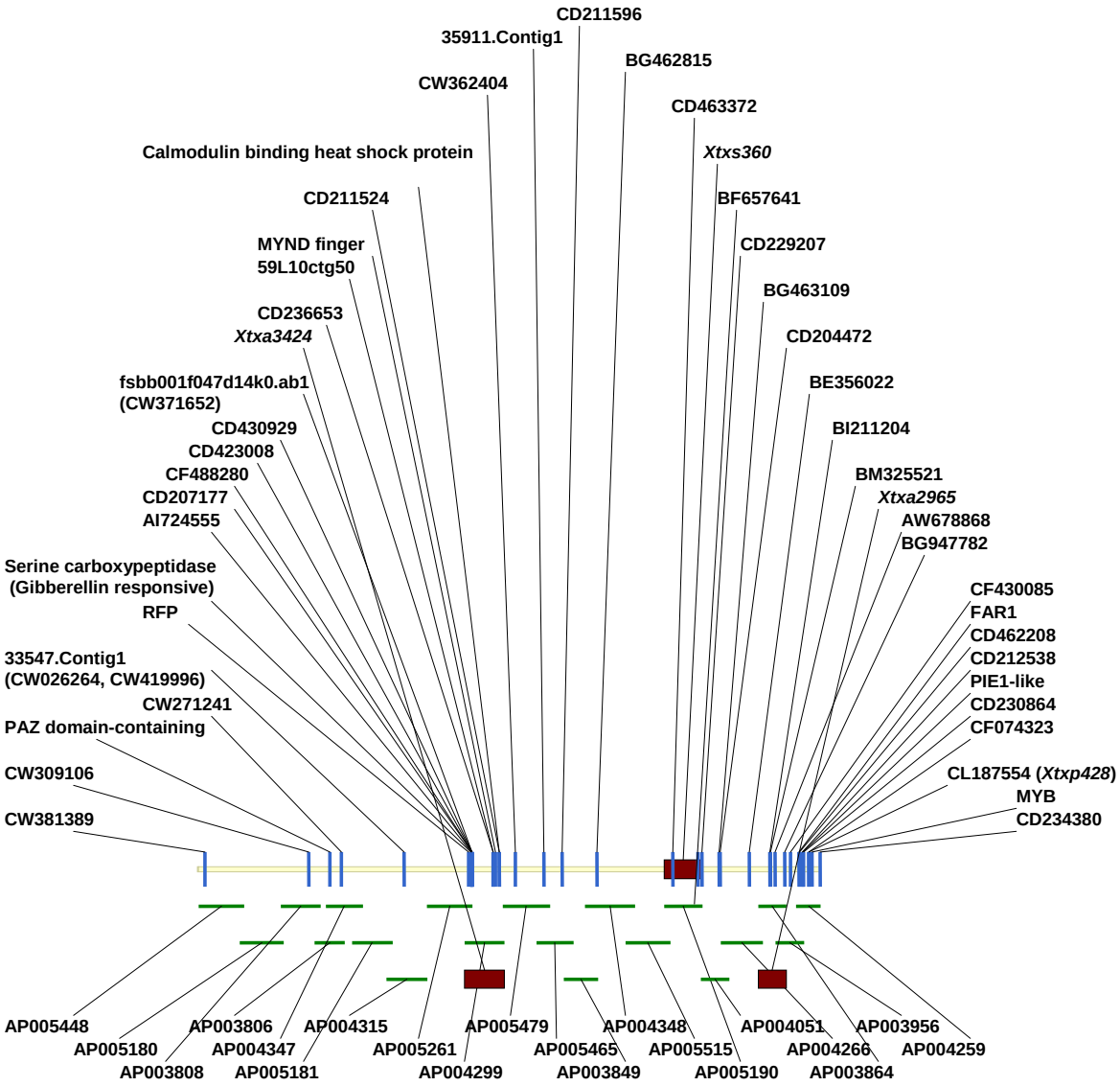

Fig. 13 *Ma*$_5$ colinear region in rice (2,277,629 bp). Rice BACs are labeled below the diagram. Sorghum sequences aligned with rice, candidate genes, and sorghum markers aligned with this portion of the rice genome are labeled above the diagram. The *Ma*$_5$ gene is near sorghum marker *Xtxa3424*.

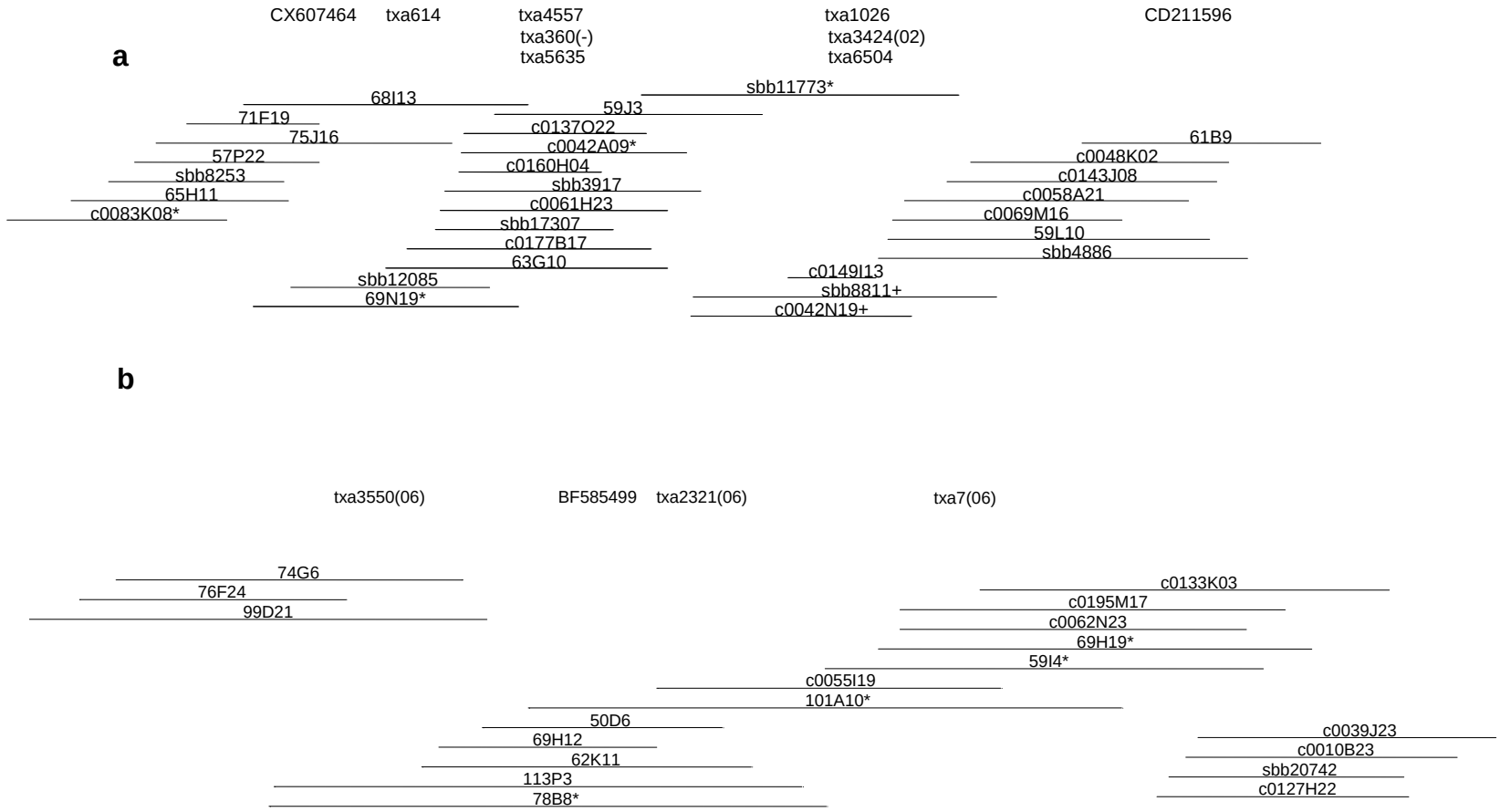

Fig. 14 BAC contigs at the *Ma*$_5$ and *Ma*$_6$ maturity loci. **a**. BACs near the *Ma*$_5$ locus. **b**. BACs near the *Ma*$_6$ locus.

CHAPTER V

CONCLUSION

The photoperiod dependent delay in flowering seen in the progeny of elite U.S. sorghum cultivars crossed to the Argentinean cultivar EBA-3 has been hypothesized to be due to the epistatic interaction of two genes, *Ma*$_5$ and *Ma*$_6$. When both genes are dominant, flowering is delayed until daylength is less than 12 hours and 20 minutes[25]. The U.S. sorghum cultivars have the genotype *Ma*$_5$*Ma*$_5$*ma*$_6$*ma*$_6$, while EBA-3 has the genotype *ma*$_5$*ma*$_5$*Ma*$_6$*Ma*$_6$. Previous studies have detected photoperiod sensitivity QTLs at the *Ma*$_5$ locus but not the *Ma*$_6$ locus[226,227]. This may reflect a unique genetic constitution at the *Ma*$_6$ locus in the EBA-3 parent. This work has linked molecular markers to major loci controlling flowering time for both the *Ma*$_5$ and *Ma*$_6$ populations. Based on the variation in flowering time seen in a large population, *Ma*$_5$ may be two genes with overlapping function with regard to flowering time, one located on sorghum chromosome 2 near the marker *Xtxa3424* at approximately 145 cM, while the chromosomal location of a potential second *Ma*$_5$ gene remains uncertain. The major locus controlling the *Ma*$_6$ late-flowering response has been located on sorghum chromosome 6 between the markers *etaccaa184* and *Xtxa2124* at approximately 12 cM. A second locus with a minor effect on the *Ma*$_6$ response may be located at 0 cM on chromosome 6. In the *Ma*$_5$ population, the flowering behavior of the most extreme early flowering and most extreme late flowering plants can be explained by the genetic constitution of the major *Ma*$_5$ and *Ma*$_6$ loci. In the *Ma*$_6$ population, the flowering

behavior of almost 90% of the population can be explained by the genetic constitution of the *Ma*$_5$ and *Ma*$_6$ loci on sorghum chromosomes 2 and 6, respectively. This work has also confirmed an epistatic interaction between the *Ma*$_5$ and *Ma*$_6$ loci. All photoperiod sensitive, extreme late flowering plants had dominant alleles at both the major *Ma*$_5$ and *Ma*$_6$ loci. However, plants flowering at intermediate times show that other genes affecting flowering are also segregating in these populations.

The populations in which *Ma*$_5$ is segregating show continuous flowering during a period of over 200 days of growth (Fig. 5). The flowering of the population as a whole does show an early and late bimodal distribution, but some plants were initiating floral meristems at all times from about 65 days after planting (DAP) until the last of the plants had to be collected due to increasingly cold weather well over 200 DAP in November, in College Station, Texas. Had only one gene controlling flowering been segregating in the *Ma*$_5$ population, a less continuous distribution of flowering would have occurred.

In the *Ma*$_6$ population, genes outside of the *Ma*$_5$/*Ma*$_6$ interaction were segregating and clearly affected flowering phenotype. While all late flowering plants had the genotype *Ma*$_5$__*Ma*$_6$*ma*$_6$, and most of the early flowering, photoperiod insensitive plants had the genotype *Ma*$_5$__*ma*$_6$*ma*$_6$, a portion of the early flowering plants were dominant at both of these loci, *Ma*$_5$__*Ma*$_6$*ma*$_6$. These early flowering dominant *Ma*$_5$/*Ma*$_6$ plants were designated a modified phenotype. The genes responsible for displacing dominant *Ma*$_5$/*Ma*$_6$ plants into the photoperiod insensitive, early flowering class were referred to as modifiers. One of these modifier genes may be located at the very top of chromosome 6, about 10 cM above the *Ma*$_6$ locus. This one modifier does not fully account for the

modified phenotype, so there is at least one more modifier gene segregating in the *Ma*$_6$ BC_1F_1 population, and its location is unknown. Interaction between gene dosage at the *Ma*$_5$ locus and presence of the modifier gene at the top of chromosome 6 does not account for the modified PI phenotype.

In association with the modified phenotype, a series of epigenetic modifications were present in the *Ma*$_6$ population. These epigenetic modifications were detected as differentially methylated *Eco*RI sites that either were or were not cut when digesting genomic DNA samples for the production of AFLP template DNA. Since a novel epigenetic AFLP band associated with the modified phenotype could be created either due to increased or decreased methylation, this method does not indicate whether flowering is associated with increased or decreased methylation at the epigenetically modified loci. Cloning and sequencing epigenetic bands associated with the modified or unmodified phenotype showed that some arose due to increased methylation, while others probably arose due to decreased methylation of *Eco*RI recognition sites. The epigenetic bands were, however, most often present in the modified *Ma*$_5$__*Ma*$_6$*ma*$_6$ PI class of plants. Cloning and sequencing of several of these bands also showed that a CAF-like protein may be involved in the epigenetic modification phenomena seen in the *Ma*$_6$ population, and that other potentially unrelated loci are involved as well.

While several maturity genes are segregating in the *Ma*$_5$ and *Ma*$_6$ populations, most of the effects on flowering are caused by two major loci. These two loci have both been genetically mapped. The physical distance between closely linked markers has been estimated by FISH for the *Ma*$_5$ locus, and numerous BACs linked to both loci have

been identified. Comparison of both of these loci to the completely sequenced colinear regions in rice has provided a number of candidate genes, and the expression of these candidate genes has been examined by qRT-PCR for differences, without success to date. The major loci controlling the *Ma*$_5$ and *Ma*$_6$ late flowering response have been narrowed and efforts to map-base clone the genes responsible are ongoing. It is possible that the genetic interaction between the *Ma*$_5$ and *Ma*$_6$ genes to repress flowering also involves chromatin remodeling at several distinct loci

REFERENCES

1. Doggett, H. *Sorghum* (Wiley, New York, 1988).
2. Kimber, C. Origins of domesticated sorghum and its early diffusion to India and China. in *Sorghum: Origin, History, Technology, and Production.* (eds. Smith, C.W. & Frederiksen, R.A.) 3-98 (Wiley, New York, 2000).
3. Kebede, H., Subudhi, P.K., Rosenow, D.T. & Nguyen, H.T. Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* L. Moench). *Theor. Appl. Genet.* **2**, 266-276 (2001).
4. Sanchez, A.C., Subudhi, P.K., Rosenow, D.T. & Nguyen, H.T. Mapping QTLs associated with drought resistance in sorghum (*Sorghum bicolor* L. Moench). *Plant Mol. Biol.* **48**, 713-726 (2002).
5. Project, I.R.G.S. The map-based sequence of the rice genome. *Nature* **436**, 793-800 (2005).
6. Goff, S.A., Ricke, D., Lan, T.-H., Presting, G., Wang, R. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. *ssp. japonica*). *Science* **296**, 92-100 (2002).
7. Yu, J., Hu, S., Wang, J., Wong, G.K.-S., Li, S. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. *ssp. indica*). *Science* **296**, 79-92 (2002).
8. Price, H.J., Dillon, S.L., Hodnett, G., Rooney, W.L., Ross, L. *et al.* Genome evolution in the genus *Sorghum* (Poaceae). *Ann. Bot.* **95**, 219-227 (2005).

9. Kim, J.-S., Klein, P.E., Klein, R.R., Price, H.J., Mullet, J.E. *et al.* Chromosome identification and nomenclature of *Sorghum bicolor*. *Genetics* **169**, 1169-1173 (2005).

10. Arumuganathan, K. & Earle, E. Nuclear DNA content of some important plant species. *Plant Mol. Biol. Report.* **9**, 208-218 (1991).

11. Ilic, K., SanMiguel, P.J. & Bennetzen, J.L. A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes. *Proc. Natl. Acad. Sci. USA* **100**, 12265-12270 (2003).

12. Ma, J., SanMiguel, P., Lai, J., Messing, J. & Bennetzen, J.L. DNA rearrangement in orthologous *Orp* regions of the maize, rice and sorghum genomes. *Genetics* **170**, 1209-1220 (2005).

13. Morishige, D.T., Childs, K.L., Moore, L.D. & Mullet, J.E. Targeted analysis of orthologous *Phytochrome A* regions of the sorghum, maize, and rice genomes using comparative gene-island sequencing. *Plant Physiol.* **130**, 1614-1625 (2002).

14. Swigonova, Z., Bennetzen, J.L. & Messing, J. Structure and evolution of the *r/b* chromosomal regions in rice, maize and sorghum. *Genetics* **169**, 891-906 (2005).

15. Tikhonov, A.P., SanMiguel, P.J., Nakajima, Y., Gorenstein, N.M., Bennetzen, J.L. *et al.* Colinearity and its exceptions in orthologous *adh* regions of maize and sorghum. *Proc. Natl. Acad. Sci. USA* **96**, 7409-7414 (1999).

16. Summerfield, R.J., Ellis, R.H., Craufurd, P.Q., Aiming, Q., Roberts, E.H. *et al.* Environmental and genetic regulation of flowering of tropical annual crops. *Euphytica* **96**, 83-91 (1997).

17. Craufurd, P.Q., Mahalakshmi, V., Bidinger, F.R., Mukuru, S.Z., Chantereau, J. *et al.* Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. *Theor. Appl. Genet.* **99**, 900-911 (1999).

18. Karper, R.E. & Quinby, J.R. Additional information concerning the introduction of milo into the United States. *J. Am. Soc. Agron.* **39**, 937-938 (1947).

19. Karper, R.E. & Quinby, J.R. The history and evolution of milo in the United States. *J. Am. Soc. Agron.* **38**, 441-453 (1946).

20. Garner, W.W. & Allard, H.A. Further studies on photoperiodism, the response of plants to relative length of day and night. *J. Agric. Res.* **23**, 871-920 (1923).

21. Stephens, J.C., Miller, F.R. & Rosenow, D.T. Conversion of alien sorghums to early combine genotypes. *Crop Sci.* **7**, 396 (1967).

22. Quinby, J.R. & Karper, R.E. The inheritance of three genes that influence time of floral initiation and maturity date in milo. *J. Am. Soc. Agron.* **37**, 916-936 (1945).

23. Quinby, J.R. Fourth maturity gene locus in sorghum. *Crop Sci.* **6**, 516-518 (1966).

24. Aydin, S. The genetic control of a photoperiod sensitive response in sorghum. Texas A&M University (1998).

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

26. Quinby, J.R. The maturity genes of sorghum. *Adv. Agron.* **19**, 267-305 (1967).

27. Major, D.J., Rood, S.B. & Miller, F.R. Temperature and photoperiod effects mediated by the sorghum maturity genes. *Crop Sci.* **30**, 305-310 (1990).

28. Quinby, J.R. The genetic control of flowering and growth in sorghum. *Adv. Agron.* **25**, 125-162 (1973).

29. Quinby, J.R. & Karper, R.E. Inheritance of duration of growth in the milo group of sorghum. *Crop Sci.* **1**, 8-10 (1961).

30. Pao, C.I. & Morgan, P.W. Genetic regulation of development in *Sorghum bicolor*. II. Effect of the ma_3^R allele mimicked by GA_3. *Plant Physiol.* **82**, 581-584 (1986).

31. Williams, E.A. & Morgan, P.W. Floral initiation in sorghum hastened by gibberellic acid and far-red light. *Planta* **145**, 269-272 (1979).

32. Beall, F.D., Morgan, P.W., Mander, L.N., Miller, F.R. & Babb, K.H. Genetic regulation of development in *Sorghum bicolor*. V. The ma_3^R allele results in gibberellin enrichment. *Plant Physiol.* **95**, 116-125 (1991).

33. Childs, K.L., Pratt, L.H. & Morgan, P.W. Genetic regulation of development in *Sorghum bicolor*. VI. The ma_3^R allele results in abnormal phytochrome physiology. *Plant Physiol.* **97**, 714-719 (1991).

34. Childs, K.L., Miller, F.R., Cordonnier-Pratt, M.M., Pratt, L.H., Morgan, P.W. *et al.* The sorghum photoperiod sensitivity gene, Ma_3, encodes a phytochrome B. *Plant Physiol.* **113**, 611-619 (1997).

35. Miller, F.R., Barnes, D.K. & Cruzado, H.J. Effect of tropical photoperiods on the growth of sorghum when grown in 12 monthly plantings. *Crop Sci.* **8**, 499-502 (1968).

36. Mullet, J.E., Klein, R.R. & Klein, P.E. *Sorghum bicolor* - an important species for comparative grass genomics and a source of beneficial genes for agriculture. *Curr. Opin. Plant Biol.* **5**, 118-121 (2002).

37. Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T. *et al.* AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**, 4407-4414 (1995).

38. Menz, M.A., Klein, R.R., Mullet, J.E., Obert, J.A., Unruh, N.C. *et al.* A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol. Biol.* **48**, 483-499 (2002).

39. Klein, P.E., Klein, R.R., Cartinhour, S.W., Ulanch, P.E., Dong, J. *et al.* A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. *Genome Res.* **10**, 789-807 (2000).

40. Garner, W.W. & Allard, H.A. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *J. Agric. Res.* **18**, 553-606 (1920).

41. Yanovsky, M.J. & Kay, S.A. Living by the calendar: how plants know when to flower. *Nat. Rev. Mol. Cell Biol.* **4**, 265-276 (2003).

42. Corbesier, L. & Coupland, G. Photoperiodic flowering of *Arabidopsis*: integrating genetic and physiological approaches to characterization of the floral stimulus. *Plant Cell Environ.* **28**, 54-66 (2005).

43. Bernier, G., Havelange, A., Houssa, C., Petitjean, A. & Lejeune, P. Physiological signals that induce flowering. *Plant Cell* **5**, 1147-1155 (1993).

44. Huang, T., Bohlenius, H., Eriksson, S., Parcy, F. & Nilsson, O. The mRNA of the *Arabidopsis* gene *FT* moves from leaf to shoot apex and induces flowering. *Science* **309**, 1694-1696 (2005).

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

46. Devlin, P.F. Signs of the time: environmental input to the circadian clock. *J. Exp. Bot.* **53**, 1535-1550 (2002).

47. Guo, H., Yang, H., Mockler, T.C. & Lin, C. Regulation of flowering time by *Arabidopsis* photoreceptors. *Science* **279**, 1360-1363 (1998).

48. Fankhauser, C. The phytochromes, a family of red/far-red absorbing photoreceptors. *J. Biol. Chem.* **276**, 11453-11456 (2001).

49. Hoffman, E.C., Reyes, H., Chu, F.-F., Sander, F., Conley, L.H. *et al.* Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science* **252**, 954-958 (1991).

50. Huang, Z.J., Edery, I. & Rosbash, M. PAS is a dimerization domain common to *Drosophila* Period and several transcription factors. *Nature* **364**, 259-262 (1993).

51. Lindebro, M.C., Poellinger, L. & Whitelaw, M.L. Protein-protein interaction via PAS domains: Role of the PAS domain in positive and negative regulation of the bHLH/PAS dioxin receptor-Arnt transcription factor complex. *EMBO J.* **14**, 3528-3539 (1995).

52. Quail, P.H. Phytochrome photosensory signalling networks. *Nat. Rev. Mol. Cell Biol.* **3**, 85-93 (2002).

53. Schafer, E. & Bowle, C. Phytochrome-mediated photoperception and signal transduction in higher plants. *EMBO Reports* **3**, 1042-1048 (2002).

54. Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K. *et al.* Light quality–dependent nuclear import of the plant photoreceptors phytochrome A and B. *Plant Cell* **11**, 1445-1456 (1999).

55. Ni, M., Tepperman, J.M. & Quail, P.H. PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell* **95**, 657-667 (1998).

56. Martinez-Garcia, F., J., Huq, E. & Quail, P.H. Direct targeting of light signals to a promoter element-bound transcription factor. *Science* **288**, 859-863 (2000).

57. Mathews, S. & Sharrock, R.A. Phytochrome gene diversity. *Plant Cell Environ.* **20**, 666-671 (1997).

58. Halliday, K.J. & Whitelam, G.C. Changes in photoperiod or temperature alter the functional relationships between phytochromes and reveal roles for phyD and phyE. *Plant Physiol.* **131**, 1913-1920 (2003).

59. Qin, M., Kuhn, R., Moran, S. & Quail, P.H. Overexpressed phytochrome C has similar photosensory specificity to phytochrome B but a distinctive capacity to enhance primary leaf expansion. *Plant J.* **12**, 1163-1172 (1997).

60. Lin, C. & Shalitin, D. Cryptochrome structure and signal transduction. *Annu. Rev. Plant Biol.* **54**, 469-496 (2003).

61. Brudler, R., Hitomi, K., Daiyasu, H., Toh, H., Kucho, K.-i. *et al.* Identification of a new cryptochrome class: structure, function, and evolution. *Mol. Cell* **11**, 59-67 (2003).

62. Cashmore, A.R., Jarillo, J.A., Wu, Y.-J. & Liu, D. Cryptochromes: blue light receptors for plants and animals. *Science* **284**, 760-765 (1999).

63. Yang, H.-Q., Wu, Y.-J., Tang, R.-H., Liu, D., Liu, Y. *et al.* The C termini of *Arabidopsis* cryptochromes mediate a constitutive light response. *Cell* **103**, 815-827 (2000).

64. Cashmore, A.R. Cryptochromes: enabling plants and animals to determine circadian time. *Cell* **114**, 537-543 (2003).

65. Wang, H., Ma, L.-G., Li, J.-M., Zhao, H.-Y. & Deng, X.W. Direct interaction of *Arabidopsis* cryptochromes with COP1 in light control development. *Science* **294**, 154-158 (2001).

66. Oyama, T., Shimura, Y. & Okada, K. The *Arabidopsis HY5* gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes Dev.* **11**, 2983-2995 (1997).

67. Devlin, P.F. & Kay, S.A. Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. *Plant Cell* **12**, 2499-2510 (2000).

68. Somers, D.E., Devlin, P.F. & Kay, S.A. Phytochromes and cryptochromes in the entrainment of the *Arabidopsis* circadian clock. *Science* **282**, 1488-1490 (1998).

69. Yanovsky, M.J., Mazzella, M.A. & Casal, J.J. A quadruple photoreceptor mutant still keeps track of time. *Curr. Biol.* **10**, 1013-1015 (2000).

70. Somers, D.E., Schultz, T.F., Milnamow, M. & Kay, S.A. ZEITLUPE encodes a novel clock-associated PAS protein from *Arabidopsis*. *Cell* **101**, 319-329 (2000).

71. Nelson, D.C., Lasswell, J., Rogg, L.E., Cohen, M.A. & Bartel, B. FKF1, a clock-controlled gene that regulates the transition to flowering in *Arabidopsis*. *Cell* **101**, 331-340 (2000).

72. Kiyosue, T. & Wada, M. LKP1 (LOV kelch protein 1): a factor involved in the regulation of flowering time in Arabidopsis. *Plant J.* **23**, 807-815 (2000).

73. Schultz, T.F., Kiyosue, T., Yanovsky, M., Wada, M. & Kay, S.A. A role for LKP2 in the circadian clock of Arabidopsis. *Plant Cell* **13**, 2659-2670 (2001).

74. Harmer, S.L., Panda, S. & Kay, S.A. Molecular bases of circadian rhythms. *Annu. Rev. Cell Dev. Biol.* **17**, 215-253 (2001).

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

76. Daniel, X., Sugano, S. & Tobin, E.M. CK2 phosphorylation of CCA1 is necessary for its circadian oscillator function in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **101**, 3292-3297 (2004).

77. Sugano, S., Andronis, C., Ong, M.S., Green, R.M. & Tobin, E.M. The protein kinase CK2 is involved in regulation of circadian rhythms in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **96**, 12362-12366 (1999).

78. Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J. *et al.* The *late elongated hypocotyl* mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**, 1219-1229 (1998).

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

80. Alabadi, D., Yanovsky, M.J., Mas, P., Harmer, S.L. & Kay, S.A. Critical role for CCA1 and LHY in maintaining circadian rhythmicity in *Arabidopsis*. *Curr. Biol.* **12**, 757-761 (2002).

81. Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M. *et al. LHY* and *CCA1* are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. *Dev. Cell* **2**, 629-641 (2002).

82. Alabadi, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Mas, P. *et al.* Reciprocal regulation between *TOC1* and *LHY*/*CCA1* within the *Arabidopsis* circadian clock. *Science* **293**, 880-883 (2001).

83. Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E. *et al.* Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* **289**, 768-771 (2000).

84. Matsushika, A., Makino, S., Kojima, M. & Mizuno, T. Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in *Arabidopsis thaliana*: insight into the plant circadian clock. *Plant Cell Physiol.* **41**, 1002-1012 (2000).

85. Makino, S., Matsushika, A., Kojima, M., Yamashino, T. & Mizuno, T. The APRR1/TOC1 quintet implicated in circadian rhythms of *Arabidopsis thaliana*: I. characterization with APRR1-overexpressing plants. *Plant Cell Physiol.* **43**, 58-69 (2002).

86. Mizuno, T. & Nakamichi, N. *Pseudo*-response regulators (PRRs) or *true* oscillator components (TOCs). *Plant Cell Physiol.* **46**, 677-685 (2005).

87. Hicks, K.A., Albertson, T.M. & Wagner, D.R. *EARLY FLOWERING3* encodes a novel protein that regulates circadian clock function and flowering in Arabidopsis. *Plant Cell* **13**, 1281-1292 (2001).

88. Liu, X.L., Covington, M.F., Fankhauser, C., Chory, J. & Wagner, D.R. *ELF3* encodes a circadian clock-regulated nuclear protein that functions in an Arabidopsis *PHYB* signal transduction pathway. *Plant Cell* **13**, 1293-1304 (2001).

89. McWatters, H.G., Bastow, R.M., Hall, A. & Millar, A.J. The *ELF3 zeitnehmer* regulates light signalling to the circadian clock. *Nature* **408**, 716-720 (2000).

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

91. Park, D.H., Somers, D.E., Kim, Y.S., Choy, Y.H., Lim, H.K. *et al.* Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis GIGANTEA* gene. *Science* **285**, 1579-1582 (1999).

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

93. Tseng, T.-S., Salome, P.A., McClung, C.R. & Olszewski, N.E. SPINDLY and GIGANTEA interact and act in *Arabidopsis thaliana* pathways involved in light responses, flowering, and rhythms in cotyledon movements. *Plant Cell* **16**, 1550-1563 (2004).

94. Swain, S.M., Tseng, T.-S., Thornton, T.M., Gopalraj, M. & Olszewski, N.E. SPINDLY is a nuclear-localized repressor of gibberellin signal transduction expressed throughout the plant. *Plant Physiol.* **129**, 605-615 (2002).

95. Swain, S.M., Tseng, T.-s. & Olszewski, N.E. Altered expression of *SPINDLY* affects gibberellin response and plant development. *Plant Physiol.* **126**, 1174-1185 (2001).

96. Putterill, J., Robson, F., Lee, K., Simon, R. & Coupland, G. The *CONSTANS* gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* **80**, 847-857 (1995).

97. Robson, F., Costa, M.M.R., Hepworth, S.R., Vizir, I., Pineiro, M. *et al.* Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. *Plant J.* **28**, 619-631 (2001).

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

99. Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z. *et al.* Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. *Science* **288**, 1613-1616 (2000).

100. Onouchi, H., Igeno, M.I., Perilleux, C., Graves, K. & Coupland, G. Mutagenesis of plants overexpressing *CONSTANS* demonstrates novel interactions among Arabidopsis flowering-time genes. *Plant Cell* **12**, 885-900 (2000).

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

102. Ayre, B.G. & Turgeon, R. Graft transmission of a floral stimulant derived from *CONSTANS*. *Plant Physiol.* **135**, 2271-2278 (2004).

103. An, H., Roussot, C., Suarez-Lopez, P., Corbesier, L., Vincent, C. *et al.* CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of *Arabidopsis*. *Development* **131**, 3615-3626 (2004).

104. Olszewski, N., Sun, T.-p. & Gubler, F. Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell* **14**, S61-80 (2002).

105. Hedden, P. & Phillips, A.L. Gibberellin metabolism: new insights revealed by the genes. *Trends Plant Sci.* **5**, 523-530 (2000).

106. Lovegrove, A. & Hooley, R. Gibberellin and abscisic acid signalling in aleurone. *Trends Plant Sci.* **5**, 102-110 (2000).

107. Jacobsen, S.E., Binkowski, K.A. & Olszewski, N.E. SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **93**, 9292-9296 (1996).

108. Tseng, T.-S., Swain, S.M. & Olszewski, N.E. Ectopic expression of the tetratricopeptide repeat domain of SPINDLY causes defects in gibberellin response. *Plant Physiol.* **126**, 1250-1258 (2001).

109. Jacobsen, S.E., Olszewski, N.E. & Meyerowitz, E.M. *SPINDLY*'s role in the gibberellin response pathway. *Symp. Soc. Exp. Biol.* **51**, 73-78 (1998).

110. Sothern, R.B., Tseng, T.S., Orcutt, S.L., Olszewski, N.E. & Koukkari, W.L. *GIGANTEA* and *SPINDLY* genes linked to the clock pathway that controls circadian characteristics of transpiration in *Arabidopsis*. *Chronobiol. Int.* **19**, 1005-1022 (2002).

111. Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J. *et al.* The *Arabidopsis GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev.* **11**, 3194-3205 (1997).

112. Silverstone, A.L., Ciampaglio, C.N. & Sun, T.-p. The Arabidopsis *RGA* gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell* **10**, 155-170 (1998).

113. Dill, A. & Sun, T.-p. Synergistic derepression of gibberellin signaling by removing RGA and GAI function in *Arabidopsis thaliana*. *Genetics* **159**, 777-785 (2001).

114. Sanchez-Fernandez, R., Ardiles-Diaz, W., Van Montagu, M., Inze, D. & May, M.J. Cloning of a novel A*rabidopsis thaliana RGA-like* gene, a putative member of the VHIID-domain transcription factor family. *J. Exp. Bot.* **49**, 1609-1610 (1998).

115. Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M. *et al.* 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* **400**, 256-261 (1999).

116. Fu, X., Richards, D.E., Ait-ali, T., Hynes, L.W., Ougham, H. *et al.* Gibberellin-mediated proteasome-dependent degradation of the barley DELLA protein SLN1 repressor. *Plant Cell* **14**, 3191-3200 (2002).

117. McGinnis, K.M., Thomas, S.G., Soule, J.D., Strader, L.C., Zale, J.M. *et al.* The Arabidopsis *SLEEPY1* gene encodes a putative F-Box subunit of an SCF E3 ubiquitin ligase. *Plant Cell* **15**, 1120-1130 (2003).

118. Sasaki, A., Itoh, H., Gomi, K., Ueguchi-Tanaka, M., Ishiyama, K. *et al.* Accumulation of phosphorylated repressor for gibberellin signaling in an F-box mutant. *Science* **299**, 1896-1898 (2003).

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

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

121. Blazquez, M.A. & Weigel, D. Integration of floral inductive signals in *Arabidopsis*. *Nature* **404**, 889-892 (2000).

122. Gocal, G.F.W., Sheldon, C.C., Gubler, F., Moritz, T., Bagnall, D.J. *et al.* *GAMYB-like* genes, flowering, and gibberellin signaling in Arabidopsis. *Plant Physiol.* **127**, 1682-1693 (2001).

123. Achard, P., Herr, A., Baulcombe, D.C. & Harberd, N.P. Modulation of floral development by a gibberellin-regulated microRNA. *Development* **131**, 3357-3365 (2004).

124. Wellensiek, S.J. Dividing cells as the prerequisite for vernalization. *Plant Physiol.* **39**, 832-835 (1964).

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

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

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

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

129. Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G. *et al.* Mutations in *LIKE HETEROCHROMATIN PROTEIN 1* affect flowering time and plant architecture in *Arabidopsis*. *Development* **128**, 4847-4858 (2001).

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

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

132. Ketel, C.S., Andersen, E.F., Vargas, M.L., Suh, J., Strome, S. *et al.* Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol. Cell. Biol.* **25**, 6857-6868 (2005).

133. Martinez-Zapater, J.M. & Somerville, C.R. Effect of light quality and vernalization on late-flowering mutants of *Arabidopsis thaliana*. *Plant Physiol.* **92**, 770-776 (1990).

134. Koornneef, M., Alonso-Blanco, C., Peeters, A.J.M. & Soppe, W. Genetic control of flowering time in Arabidopsis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 345-370 (1998).

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

136. Sanda, S.L. & Amasino, R.M. Ecotype-specific expression of a flowering mutant phenotype in *Arabidopsis thaliana*. *Plant Physiol.* **111**, 641-644 (1996).

137. Lee, I., Aukerman, M.J., Gore, S.L., Lohman, K.N., Michaels, S.D. *et al.* Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in Arabidopsis. *Plant Cell* **6**, 75-83 (1994).

138. Lim, M.-H., Kim, J., Kim, Y.-S., Chung, K.-S., Seo, Y.-H. *et al.* A new Arabidopsis gene, *FLK*, encodes an RNA binding protein with K homology motifs and regulates flowering time via *FLOWERING LOCUS C*. *Plant Cell* **16**, 731-740 (2004).

139. Simpson, G.G., Quesada, V., Henderson, I.R., Dijkwel, P.P., Macknight, R. *et al.* RNA processing and *Arabidopsis* flowering time control. *Biochem. Soc. Trans.* **32**, 565-566 (2004).

140. Aukerman, M.J., Lee, I., Weigel, D. & Amasino, R.M. The Arabidopsis flowering-time gene *LUMINIDEPENDENS* is expressed primarily in regions of cell proliferation and encodes a nuclear protein that regulates *LEAFY* expression. *Plant J.* **18**, 195-203 (1999).

141. Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R. *et al. FCA*, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell* **89**, 737-745 (1997).

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

143. Schomburg, F.M., Patton, D.A., Meinke, D.W. & Amasino, R.M. *FPA*, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. *Plant Cell* **13**, 1427-1436 (2001).

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

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

146. El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A.J.M., Wagemaker, C., Weller, J.L. *et al.* The role of cryptochrome 2 in flowering in Arabidopsis. *Plant Physiol.* **133**, 1504-1516 (2003).

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

148. Michaels, S.D., Bezerra, I.C. & Amasino, R.M. *FRIGIDA*-related genes are required for the winter-annual habit in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **101**, 3281-3285 (2004).

149. Zhang, H., Ransom, C., Ludwig, P. & Nocker, S. Genetic analysis of early flowering mutants in Arabidopsis defines a class of pleiotropic developmental regulator required for expression of the flowering-time switch Flowering Locus C. *Genetics* **164**, 347-358 (2003).

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

151. Murtas, G., Reeves, P.H., Fu, Y.-F., Bancroft, I., Dean, C. *et al.* A nuclear protease required for flowering-time regulation in Arabidopsis reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. *Plant Cell* **15**, 2308-2319 (2003).

152. Reeves, P.H., Murtas, G., Dash, S. & Coupland, G. *early in short days 4*, a mutation in *Arabidopsis* that causes early flowering and reduces the mRNA abundance of the floral repressor *FLC*. *Development* **129**, 5349-5361 (2002).

153. Cheng, Y., Kato, N., Wang, W., Li, J. & Chen, X. Two RNA binding proteins, HEN4 and HUA1, act in the processing of *AGAMOUS* pre-mRNA in *Arabidopsis thaliana*. *Dev. Cell* **4**, 53-66 (2003).

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

155. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the polycomb and trithorax group proteins. *Annu. Rev. Genet.* **38**, 413-443 (2004).

156. Oh, S., Zhang, H., Ludwig, P. & van Nocker, S. A mechanism related to the yeast transcriptional regulator Paf1c is required for expression of the Arabidopsis *FLC/MAF* MADS box gene family. *Plant Cell* **16**, 2940-2953 (2004).

157. He, Y., Doyle, M.R. & Amasino, R.M. PAF1-complex-mediated histone methylation of *FLOWERING LOCUS C* chromatin is required for the vernalization-responsive, winter-annual habit in *Arabidopsis*. *Genes Dev.* **18**, 2774-2784 (2004).

158. He, Y. & Amasino, R.M. Role of chromatin modification in flowering-time control. *Trends Plant Sci.* **10**, 30-35 (2005).

159. Soppe, W.J., Bentsink, L. & Koornneef, M. The early-flowering mutant *efs* is involved in the autonomous promotion pathway of *Arabidopsis thaliana*. *Development* **126**, 4763-4770 (1999).

160. Noh, Y.-S. & Amasino, R.M. *PIE1*, an ISWI family gene, is required for *FLC* activation and floral repression in Arabidopsis. *Plant Cell* **15**, 1671-1682 (2003).

161. Zhang, H. & Van Nocker, S. The *VERNALIZATION INDEPENDENCE 4* gene encodes a novel regulator of *FLOWERING LOCUS C*. *Plant J.* **31**, 663-673 (2002).

162. Scortecci, K.C., Michaels, S.D. & Amasino, R.M. Identification of a MADS-box gene, *FLOWERING LOCUS M*, that represses flowering. *Plant J.* **26**, 229-236 (2001).

163. Alvarez-Buylla, E.R., Pelaz, S., Liljegren, S.J., Gold, S.E., Burgeff, C. *et al.* An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc. Natl. Acad. Sci. USA* **97**, 5328-5333 (2000).

164. Ratcliffe, O.J., Kumimoto, R.W., Wong, B.J. & Riechmann, J.L. Analysis of the Arabidopsis *MADS AFFECTING FLOWERING* gene family: *MAF2* prevents vernalization by short periods of cold. *Plant Cell.* **15**, 1159-1169 (2003).

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

166. Ohshima, S., Murata, M., Sakamoto, W., Ogura, Y. & Motoyoshi, F. Cloning and molecular analysis of the *Arabidopsis* gene *Terminal Flower 1*. *Mol. Gen. Genet.* **254**, 186-194 (1997).

167. Shannon, S. & Meeks-Wagner, D.R. A mutation in the Arabidopsis *TFL1* gene affects inflorescence meristem development. *Plant Cell* **3**, 877-892 (1991).

168. Hanzawa, Y., Money, T. & Bradley, D. A single amino acid converts a repressor to an activator of flowering. *Proc. Natl. Acad. Sci. USA* **102**, 7748-7753 (2005).

169. Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K. *et al.* Activation tagging of the floral inducer *FT*. *Science* **286**, 1962-1965 (1999).

170. Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M. & Araki, T. A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**, 1960-1962 (1999).

171. Yoo, S.Y., Kardailsky, I., Lee, J.S., Weigel, D. & Ahn, J.H. Acceleration of flowering by overexpression of *MFT* (*MOTHER OF FT AND TFL1*). *Mol. Cells* **17**, 95-101 (2004).

172. Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.-H. *et al.* Interaction of polycomb-group proteins controlling flowering in *Arabidopsis*. *Development* **131**, 5263-5276 (2004).

173. Hennig, L., Bouveret, R. & Gruissem, W. MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. *Trends Cell Biol.* **15**, 295-302 (2005).

174. Ach, R.A., Taranto, P. & Gruissem, W. A conserved family of WD-40 proteins binds to the retinoblastoma protein in both plants and animals. *Plant Cell* **9**, 1595-1606 (1997).

175. Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D. *et al.* Mutations in *FIE*, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407-416 (1999).

176. Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J. *et al.* Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **96**, 296-301 (1999).

177. Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M. *et al.* A polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* **386**, 44-51 (1997).

178. Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J. *et al.* EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell* **13**, 2471-2481 (2001).

179. Luo, M., Bilodeau, P., Dennis, E.S., Peacock, W.J. & Chaudhury, A. Expression and parent-of-origin effects for *FIS2*, *MEA*, and *FIE* in the endosperm and embryo of developing *Arabidopsis* seeds. *Proc. Natl. Acad. Sci. USA* **97**, 10637-10642 (2000).

180. Aubert, D., Chen, L., Moon, Y.-H., Martin, D., Castle, L.A. *et al.* EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis. *Plant Cell* **13**, 1865-1875 (2001).

181. Gomez-Mena, C., Pineiro, M., Franco-Zorrilla, J.M., Salinas, J., Coupland, G. *et al. early bolting in short days*: an Arabidopsis mutation that causes early flowering and partially suppresses the floral phenotype of *leafy*. *Plant Cell* **13**, 1011-1024 (2001).

182. Pineiro, M., Gomez-Mena, C., Schaffer, R., Martinez-Zapater, J.M. & Coupland, G. EARLY BOLTING IN SHORT DAYS is related to chromatin remodeling factors and regulates flowering in Arabidopsis by repressing *FT*. *Plant Cell* **15**, 1552-1562 (2003).

183. Soppe, W.J.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T. *et al.* The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6**, 791-802 (2000).

184. Aukerman, M.J. & Sakai, H. Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-Like target genes. *Plant Cell* **15**, 2730-2741 (2003).

185. Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R. *et al.* Dissection of floral induction pathways using global expression analysis. *Development* **130**, 6001-6012 (2003).

186. Lee, H., Suh, S.-S., Park, E., Cho, E., Ahn, J.H. *et al.* The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev.* **14**, 2366-2376 (2000).

187. Michaels, S.D., Himelblau, E., Kim, S.Y., Schomburg, F.M. & Amasino, R.M. Integration of flowering signals in winter-annual Arabidopsis. *Plant Physiol.* **137**, 149-156 (2005).

188. Schultz, E.A. & Haughn, G.W. *LEAFY*, a homeotic gene that regulates inflorescence development in Arabidopsis. *Plant Cell* **3**, 771-781 (1991).

189. Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F. & Meyerowitz, E.M. *LEAFY* controls floral meristem identity in Arabidopsis. *Cell* **69**, 843-859 (1992).

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

191. Yano, M., Harushima, Y., Nagamura, Y., Kurata, N., Minobe, Y. *et al.* Identification of quantitative trait loci controlling heading date in rice using a high-density linkage map. *Theor. Appl. Genet.* **95**, 1025-1032 (1997).

192. Lin, S.Y., Sasaki, T. & Yano, M. Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa* L., using backcross inbred lines. *Theor. Appl. Genet.* **96**, 997-1003 (1998).

193. Yamamoto, T., Kuboki, Y., Lin, S.Y., Sasaki, T. & Yano, M. Fine mapping of quantitative trait loci *Hd-1*, *Hd-2* and *Hd-3*, controlling heading date of rice, as single Mendelian factors. *Theor. Appl. Genet.* **97**, 37-44 (1998).

194. Lin, H.X., Yamamoto, T., Sasaki, T. & Yano, M. Characterization and detection of epistatic interactions of 3 QTLs, *Hd1*, *Hd2*, and *Hd3*, controlling heading date in rice using nearly isogenic lines. *Theor. Appl. Genet.* **101**, 1021-1028 (2000).

195. Yamamoto, T., Lin, H., Sasaki, T. & Yano, M. Identification of heading date quantitative trait locus *Hd6* and characterization of its epistatic interactions with *Hd2* in rice using advanced backcross progeny. *Genetics* **154**, 885-891 (2000).

196. Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T. *et al. Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* **43**, 1096-1105 (2002).

197. Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L. *et al. Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to

the Arabidopsis flowering time gene *CONSTANS*. *Plant Cell* **12**, 2473-2484 (2000).

198. Izawa, T., Takahashi, Y. & Yano, M. Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. *Curr. Opin. Plant Biol.* **6**, 113-120 (2003).

199. Hayama, R., Yokoi, S., Tamaki, S., Yano, M. & Shimamoto, K. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**, 719-722 (2003).

200. Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M. *et al.* Phytochrome mediates the external light signal to repress *FT* orthologs in photoperiodic flowering of rice. *Genes Dev.* **16**, 2006-2020 (2002).

201. Crasta, O.R., Xu, W.W., Rosenow, D.T., Mullet, J. & Nguyen, H.T. Mapping of post-flowering drought resistance traits in grain sorghum: association between QTLs influencing premature senescence and maturity. *Mol. Gen. Genet.* **262**, 579-588 (1999).

202. Boivin, K., Deu, M., Rami, J.F., Trouche, G. & Hamon, P. Towards a saturated sorghum map using RFLP and AFLP markers. *Theor. Appl. Genet.* **98**, 320-328 (1999).

203. Pereira, M.G., Lee, M., Bramel-Cox, P., Woodman, W., Doebley, J. *et al.* Construction of an RFLP map in sorghum and comparative mapping in maize. *Genome* **37**, 236-243 (1994).

204. Xu, G.W., Magill, C.W., Schertz, K.F. & Hart, G.E. A RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Theor. Appl. Genet.* **89**, 139-145 (1994).

205. Whitkus, R., Doebley, J. & Lee, M. Comparative genome mapping of sorghum and maize. *Genetics* **132**, 1119-1130 (1992).

206. Klein, R.R., Klein, P.E., Mullet, J.E., Minx, P., Rooney, W.L. *et al.* Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor. Appl. Genet.* **111**, 994-1012 (2005).

207. Kim, J.-S., Childs, K.L., Islam-Faridi, M.N., Menz, M.A., Klein, R.R. *et al.* Integrated karyotyping of sorghum by *in situ* hybridization of landed BACs. *Genome* **45**, 402-412 (2002).

208. Kim, J.-S., Klein, P.E., Klein, R.R., Price, H.J., Mullet, J.E. *et al.* Molecular cytogenetic maps of sorghum linkage groups 2 and 8. *Genetics* **169**, 955-965 (2005).

209. Bhattramakki, D., Dong, J., Chhabra, A.K. & Hart, G.E. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* **43**, 988-1002 (2000).

210. Kong, L., Dong, J. & Hart, G.E. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theor. Appl. Genet.* **101**, 438-448 (2000).

211. Peng, Y., Schertz, K.F., Cartinhour, S. & Hart, G.E. Comparative genome mapping of *Sorghum bicolor* (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. *Plant Breed.* **118**, 225-235 (1999).

212. Kosambi, D.D. The estimation of map distances from recombination values. *Ann. Eug.* **12**, 172-175 (1944).

213. Bedell, J.A., Budiman, M.A., Nunberg, A., Citek, R.W., Robbins, D. *et al.* Sorghum genome sequencing by methylation filtration. *P.L.O.S. Biol.*, 103-115 (2005).

214. Klein, P.E., Klein, R.R., Vrebalov, J. & Mullet, J.E. Sequence-based alignment of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement. *Plant J.* **34**, 605-621 (2003).

215. Brennan, C.A., Van Cleve, M.D. & Gumport, R.I. The effects of base analogue substitutions on the cleavage by the *Eco*RI restriction endonuclease of octadeoxyribonucleotides containing modified *Eco*RI recognition sequences. *J. Biol. Chem.* **261**, 7270-7278 (1986).

216. Huang, L.H., Farnet, C.M., Ehrlich, K.C. & Ehrlich, M. Digestion of highly modified bacteriophage DNA by restriction endonucleases. *Nucleic Acids Res.* **10**, 1579-1591 (1982).

217. Tasseron-de Jong, J.G., Aker, J. & Giphart-Gassler, M. The ability of the restriction endonuclease *Eco*RI to digest hemi-methylated versus fully cytosine-methylated DNA of the herpes *tk* promoter region. *Gene* **74**, 147-149 (1988).

218. Burn, J.E., Bagnall, D.J., Metzger, J.D., Dennis, E.S. & Peacock, W.J. DNA methylation, vernalization, and the initiation of flowering. *Proc. Natl. Acad. Sci. USA* **90**, 287-291 (1993).

219. Chan, S.W.L., Henderson, I.R. & Jacobsen, S.E. Gardening in the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* **6**, 351-360 (2005).

220. Hsieh, T.-F. & Fischer, R.L. Biology of chromatin dynamics. *Annu. Rev. Plant Biol.* **56**, 327-351 (2005).

221. Harushima, Y., Yano, M., Shomura, A., Sato, M., Shimano, T. *et al.* A high-density rice genetic linkage map with 2275 markers using a single F_2 population. *Genetics* **148**, 479-494 (1998).

222. Yuan, Q., Liang, F., Hsiao, J., Zismann, V., Benito, M.-I. *et al.* Anchoring of rice BAC clones to the rice genetic map *in silico*. *Nucleic Acids Res.* **28**, 3636-3641 (2000).

223. Zhong, X.B., Hans de Jong, J. & Zabel, P. Preparation of tomato meiotic pachytene and mitotic metaphase chromosomes suitable for fluorescence *in situ* hybridization (FISH). *Chromosome Res.* **4**, 24-28 (1996).

224. Lagercrantz, U. & Axelsson, T. Rapid evolution of the family of *CONSTANS LIKE* genes in plants. *Mol. Biol. Evol.* **17**, 1499-1507 (2000).

225. Kim, J.S., Islam-Faridi, M.N., Klein, P.E., Stelly, D.M., Price, H.J. *et al.* Comprehensive molecular cytogenetic analysis of sorghum genome architecture: distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. *Genetics* **171**, 1963-1976 (2005).

226. Chantereau, J., Trouche, G., Rami, J.F., Deu, M., Barro, C. *et al.* RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* **120**, 183-194 (2001).

227. Lin, Y.R., Schertz, K.F. & Paterson, A.H. Comparative analysis of QTLs affecting plant height and maturity across the poaceae, in reference to an interspecific sorghum population. *Genetics* **141**, 391-411 (1995).

APPENDIX

Table A1 · Primers for *Ma*$_5$ STS markers

Identifier	Forward sequence	Reverse sequence	Rice BAC alignment	Rice chr 7 location
CW067706gat8	GTACGGTGCTTCCATTCCAT	GGACAAGGAGGGCAGATACA	AP005737	53.4 cM
CW299433ag5ac5	CCTCGCGCCTTACTAACAAA	CGATGACGAATCGATGATTTT	AP005737	53.4 cM
CW299433atac26	AGCCGGTGATACGACAAAGT	TGGCTATGCATGAGTTCGAG	AP005737	53.4 cM
CW309106	TCAAATAAACACATTATATA	GGGAATGGACGAGAAAATCA	AP003808	55.9 cM
CW381389	ATCAGCCTGGACCATCCATA	GGTGTATGGTGTGTGGTGGA	AP005448	55.9 cM
CW453672	TCAAATAAACACATTATATA	GGGAATGGACGAGAAAATCA	AP003808	55.9 cM
CW525955	ATCAGCCTGGACCATCCATA	GGTGTATGGTGTGTGGTGGA	AP005448	55.9 cM
AI724555	ACGATGCTGGTCACATGGTA	TATGCCTGGCCTAGCAATTC	AP005261	57.5 cM
CD207177-1	CAGACTCGACTCGTCAACCA	TAGAGAGGGGCGGGGAAG	AP005261	57.5 cM
CD207177-2	AGCAGTTCGAAACCAAAGGA	AGGTCCTCCACCGACGTG	AP005261	57.5 cM
CD423008	GCTACTCCTCCGGGTGCT	GGAGGGGTGGAGGTTGAG	AP005261	57.5 cM
CF488280	GCTACTCCTCCGGGTGCT	GGAGGGGTGGAGGTTGAG	AP005261	57.5 cM
CW271241	CAAGGGCACGAAATCTCTTC	CCCACATCCGCTATTCTTGT	AP004347	57.5 cM
CX607464	AGCTGCTCTCCATGAAATCG	AGGGATCAATGGTCGAGACA	AP004347	57.5 cM
35911.Contig1-1	GGCAACAAAATGGACCTGTT	TTTGTCCTTGTTGGCATTGA	AP005465	60.8 cM
35911.Contig1-2	CCATTTGCCAATGTGTGTGT	AGTGAACGTCGGTTTTCGTC	AP005465	60.8 cM
59L10ctg17	CGTGTTCACGTTAAGCTTTCA	GGGCAGCATCCACTGTTAAT	AP004299	60.8 cM
59L10ctg50	TGGACTAAACTCGCCAGGAG	GAACCTGGAGCTCGGGTAGT	AP004299	60.8 cM
59L10ctg50ssr	TGCATGCCCACTGTAATACG	CAGCAGCAACAGCAACAAAT	AP004299	60.8 cM
AW565589	AACCTCAACATGCAGACTTCG	ATCAAGGGATCAACAATGACTTG	AP003849	60.8 cM
AW671370	GAGGTGTTCGCATTCCCTAA	AATTGATGGCCCAGTCTCAC	AP005467	60.8 cM
34015.Contig1	AATCTTGCGTACACCCTGCT	TTGCATGACACATTAGATCACAA	AP005479	60.8 cM
CD211524	ACTACCTGCGATGCGAGACT	ACACCCCAGGTCTCACTGTC	AP004299	60.8 cM
CD211596	AGGTGCTGGTCTGGATGCTA	CAATCCAGAAGCAGATGCAC	AP005465	60.8 cM
CD236653	TAGCAGCAGCAGCATTTCAC	ATGCGAGTGGAGAAGTAGCC	AP004299	60.8 cM
CW271241ta18	CAAGGGCACGAAATCTCTTC	CCCACATCCGCTATTCTTGT	AP004347	60.8 cM
CW362404	AATCTTGCGTACACCCTGCT	TTGCATGACACATTAGATCACAA	AP005479	60.8 cM
BG462815	GGTGTTGCAGCCTTTGATTT	GACAGCCGCAAGACAAAGAT	AP004348	61.6 cM
BE356022	CAAAGCACCAACCCGATTAC	GGCGAGGCACAGGAGGTA	AP004266	61.9 cM
BF657641at18	GATCCCAAATCCCTTGAGGT	TTCAACGTAGCATTTCCACAA	AP005190	61.9 cM
BG463109	GGGGCATATGTATTTATTTCTTCA	ATTCGAAAGGCTTCATCACG	AP004051	61.9 cM
BI211204	GGAATCACCGACTGCAACTT	AAGACCCGACATCAAACCAG	AP003864	61.9 cM

Table A1 · Continued

Identifier	Forward sequence	Reverse sequence	Rice BAC alignment	Rice chr 7 location
BM325521	CACCGGATCATATCATGCAC	GGCTTCCCCAAAAATGAAAT	AP003864	61.9 cM
CD204472	TCTCACACGTCCCATCCAT	GACATCGATCTCGTAAAAACAGG	AP004051	61.9 cM
CD229207	CTTCTCCGAGCTCCTCACC	CGTTGGAAACGTCCAACTCT	AP004051	61.9 cM
CD463372	GAAGCTCAGGGACATCATGC	GCGCAGTTGAGAAGAACCTT	AP005190	61.9 cM
Txp431	TGAAAAAGCCCTCCAACTTC	TTCTTAAACTCGCTTTCTAAATTATCA	AP004259	62.4 cM
224b4cgc5	GTTCCCCATTTGCCTCCT	ATAAACCCGCCCAAAAACAG	AP004259	62.4 cM
224b4ttg7	TAGGTCGCCACCTGACTTCT	CATTCAGCTCATCGTTCCAA	AP004259	62.4 cM
42115.Contig1	GTTGTCCGCGGAAATACACT	GATGACGACGATGACACACC	AP003995	62.4 cM
42115.Contig1R2		GACGATGACGACGATGACAC	AP003995	62.4 cM
AW678868-1	GATCAGATCGACCCAGCATT	TGCCACCAATTAACCAGTCA	AP003956	62.4 cM
AW678868-2	ACGAGGGAAATGATGTGACC	TCAGCCTTCTCCAGGTCAGT	AP003956	62.4 cM
BE360675	ACTTCGTCACTGGGCACTTT	TTCTTCTCCACGCGAAGTTT	AP004006	62.4 cM
BG947782	TTGGGACATGAAGTTGAGCA	CAGTTTTTCCAGTGCCAGGT	AP003956	62.4 cM
CB925377	CCACGATTCTTGGTGGGTAG	GTAGTACGCCATGCTCGTCA	AP005186	62.4 cM
CB926798	CTTACACCGACGGTTGTTCC	GAGCAGGGTGATGGTGAAGT	AP005186	62.4 cM
CD212538	AGAAGCACAGAGGGTCCTGA	GTGGATGGACAAATGGAACC	AP004259	62.4 cM
CD230864	TGTTTCGGATGGACAGATCA	GCTGGTAGCTCTCGTTCAGG	AP004259	62.4 cM
CD234380	ACCGAGTCAGCTTCATGCTT	CACCTCACCATGTCCATGTC	AP004259	62.4 cM
CD462208	TGTGGAATTTGGTTCCATGA	ATTGGCCTTGGGTAAGATCC	AP004259	62.4 cM
CD463104	ATGCCACTGATGGGACTAGG	CTCACAGCTTCACACCAGGA	AP005177	62.4 cM
CF074323-1	GATGGGTTGGTGGTGGAC	CCTCCGCCTGTAGCATCC	AP004259	62.4 cM
CF074323-2	CTTCGTGGTCCGGCATGG	GCTCAGAGACGGTTTCCAGA	AP004259	62.4 cM
CF430085	AAATTGCTGCTGCACTTCCT	TGTGTTCACTGGCTGAGAGG	AP003956	62.4 cM
CF485892	TGCTCACCCTTCAACAACTG	CCAGCTTCCAGCAAAAACTC	AP005177	62.4 cM
Txp428	CACTGGCCAAGGTTTCACTT	CATGGAATGCAACATAGCAA	AP004259	62.4 cM
CW247848	TATAAGCGAGTGGCACCA	GACGCAAGGCAATGTCCTAC	AP005127	67 cM
Txp429	CACTGCCGTTGGAATCCTAT	ATGCGCTGCAGCTTTATCTT	AP003815	67 cM
BE599905-276	GTTCACAGAATCAGCCTACCAGAA	TCCACAGGTTGGTCCTTTGG	AP003815	67 cM
CD236027	AATCCTTCCAACCCATTTCC	GTGGAGAGGTGGGAGCAC	AP005196	67 cM
7663.Contig1	ATGCTGCACCCAATACACAA	GGATTGTCGGTGTCCTACCT	AP005103	69.2 cM
CD233373-117	TCATTTTTCTTTCCCTATGGGAAA	GAAACCGTGATCGAGAATTTGAT	AP005103	69.2 cM
2432.Contig1	TTGCCTCCAAAGGTCAAAAT	AGCGATCGACCCTAGTGTGT	AP003825	105.7 cM

Table A1 · Continued

Identifier	Forward sequence	Reverse sequence	Rice BAC alignment	Rice chr 7 location
BQ656077F	CGCGGTTTTAAAAGGGAAA	GCAATCCTCCTTGGTGTTGT	NA	NA
BQ656077F2	ACTCGATTGCGTTCCTGCT	CCCAACATCCTCGAAATCAT	NA	NA
BQ656077R3		ATTGCACGGACGGTGTTACT	NA	NA
BQ656077R4		AACAGAACATCATCACCCCC	NA	NA
BQ656077R5		GCTTGGGGGCAACATACTT	NA	NA

Table A2 · Primers for *Ma*$_6$ STS markers

Identifier	Forward sequence	Reverse sequence	Rice BAC alignment	Rice chr 4 location
35146.Contig1	TCCAGACATTTACAGCAGCTT	GCATGTAGCTAGCGCGATTT	AL662977	41.7 cM
17352.Contig1	CCACCGATGACTTGTGTACG	GGAGTTTGCAAAGGTCCAGA	AL662997	49.7 cM
34413.Contig1	GGATTGGAGGACGAATCAGA	GCACCATGAGGGAGCTAAGT	AL731591	49.7 cM
34413.Contig2	AGAAAAGGCTCGGGAACAAT	ATTTCTGGGTGCACAAAAGC	AL731591	49.7 cM
41716.Contig1	TGAGATCTACCTCGGCCATC	TGACAAGGGTAAGGCCAAGA	AL662997	49.7 cM
45276.Contig1	ACCGCGAGGTCTACGACA	CGTCCTCAGACGAGGAGAAG	AL731591	49.7 cM
BE357713	CGGCGACTACAAGAAGATCA	CTCCTCCTTGCCCTCCTC	AL731591	49.7 cM
BF585499	GAAGCAGGTGGGCGGGTGCAC	TACCTGCACTGCGCGCTCACAAC	AL731591	49.7 cM
BG051187	AGATCGTCTCCGTTTCCGTCAAC	TATCCGAGTGGGCACGTAAGT	AL731591	49.7 cM
CC616682	GCACTACCGAGGGGTGAG	AAGTTGAGCTTGGCCTTGTG	AL731591	49.7 cM
CD208734-1	GATTCGGTGTTGCGATTCC	CTGCAGCTGAAAGCAAGCTA	AL606621	49.7 cM
CD208734-2	TGGGACCGACTATTCTTCTCAT	CAATGCAGCTTTTTCAAGCA	AL606621	49.7 cM
CD208734-3	CCTCTCCCCACCTCTCCTAC	CGGTGGTCCTTCCTCTCC	AL606621	49.7 cM
CD226594	AGGAAGGGATGCTTGAGGTT	CTGGTGACAATGTGTGATCCTT	AL731591	49.7 cM
13124.Contig1	GGCATTGGGAGAAACAAAAA	CTTGGCAAACACATGTCACC	AL731642	52.6 cM
BE598359	AGCGAGGGGTGGTGTACCTGATG	CGCCTCAAATCTTGGATGGGTAG	AL606610	52.6 cM
BG357895	TCGCAGCTCTACCACCAG	GAAGCCTCTCCTCCAGCTC	AL606610	52.6 cM
BG412843	CTTCTGCGGCAGCTTCAC	CCAAAAACCGGTACTGGTAAA	AL606610	52.6 cM
BM317777	CCTGGGCACACAACAACAGTCTG	GCCGAAAGCAAGATGGTTCTC	AL731462	52.6 cM
CD221096(gca)7	CCTCACCTCCCTCTTTCTCC	CCTGACGCCATTTTTAGTCG	AL731642	52.6 cM
CD221096(gca)7-2	GCTCTGCAATTCCATTCCAT	TCTTCTCTCCTGACGCCATT	AL731642	52.6 cM
AW677166	TGAACGGCTATGTATGTCTTGG	ATGGTGGCCTTCAAGATCAG	AL607005	56.1 cM
AW677340	CTTTCACAAACTTGGCTGCGTAA	GCGAACACAATGTAAGGGCTATG	AL607005	56.1 cM
BE594647(cca)9-1	AGCTGATGGCGTCCAACTAC	CTCCTCCGCATCGTCTTG	AL731639	56.1 cM
BE594647(cca)9-2	CAGAGAACCAGGAGGAGCTG	AAGGAGCCTGCGTGAGTG	AL731639	56.1 cM
BG355728	GGACCTCCAAAGATTTTTACTGA	ATAAGGACACTAGCGATCACAG	AL731639	56.1 cM
BG462875	GGAGATCCTCGGCATCGTGTAC	AAGGCCACGCCACAACTACATAC	AL731639	56.1 cM
CW053469	TCCAGGGACAGGAAAGTGAG	CTCCTCCGCATCGTCTTG	AL731639	56.1 cM
12255.Contig1	GAGAGAGAGCGCGATGAGAC	ATCCATCGCAAACCGATAAA	AL731641	57.5 cM
22744.Contig1	AAGGTAAGTGAAGCCCAAGG	TGAAGCGAGGAAGAGAAGGA	AL731641	57.5 cM
AW678663-1	CGGTGGAGGATGATCTTGAC	CTTCGAAAGCCTCTTCATGG	AL731641	57.5 cM
AW678863-2	GGAGTCTGTGAGCCTGAAGC	TGTCATCTTGCTCCTCGTTG	AL731641	57.5 cM

Table A2 · Continued

Identifier	Forward sequence	Reverse sequence	Rice BAC alignment	Rice chr 4 location
AW678863-3	ACTGGCTTCACCCTCACCCTCAG	GCAGCGATCACCACCCAGATG	AL731641	57.5 cM
BG050332	GACGTGGCCGTTGATGGAGTAC	AACAGCCGAGGTGGAGAGGTAGC	AL731641	57.5 cM
BG102021	CTCATCCACCACCATTTCCT	CGGGTTAAAGTGAACCCAAA	AL731641	57.5 cM
BG649498(gca)6	AACGCTACAAGGTGGAGTGC	GAGGACCAGTGCTGGAAAGA	AL731641	57.5 cM
BG050402(Txp434)	CGAGGTCCAGGAGTACACG	CGGCCTCCATGAGGAGTAAT	AL731641	57.5 cM
BI099358	CAATTTGAACAGTAAGACCTATCTCA	TCAGGATCCAATTCATCTTCG	AL731641	57.5 cM
BM323660	CCGCTCCCCAGATCACATAC	AGCTTGCCTCCTTGTTGTTAAAG	AL731641	57.5 cM
CD423046	AGATTGACCCAATGCTGGAA	CTCCAGGAGCCCATTCTCTA	AL731641	57.5 cM
AW284270	TGTTTGGCTTTGGGGTCTCGTA	AGCCTCTCATTGTGGGGAAAGTG	AL662947	58.6 cM
BE598024	TGCAAGTTCGAGGCCACCGTCAC	CCCGGCCAGAGGTATTCACAT	AL662945	58.6 cM
BG605968	AAGGGCGATTCTACTCCGATCTG	TTTCCGGCGATTGCTACCAC	AL662945	58.6 cM
BI075348	GCGTCGCCGTCACTCCGTTCT	CGGCGCAGTTCCAGGACCAG	AL662947	58.6 cM
BM325368	CTGGGCGTTTACCTGTTGTC	CTGTGTGGGATGTGCTTGAG	AL662947	58.6 cM
BE593589	CCGGTCATCACCAGCCATATA	TTTTTCATGACATTTCCGAACTG	AL606453	58.9 cM
BE595056	AAGTTCCGGTCTTTAAGTCAA	TTCCGTGTATCAGCCCAGTC	AL606598	58.9 cM
BG158604	CTCAGGGATCTCGGGTTC	CGGCAGTATCTGGAGTTACTT	AL606598	58.9 cM
BG947398	CCCACCCGTCCATCGTTTG	CGGCTGGAGGAAGGTCTCGTA	AL606453	58.9 cM
BM326197	CTTCTTGCGAGTCCTCACTT	CATCATGGGAAACGCTGGACTG	AL606453	58.9 cM
CF759033	TGCGTGACCAAGAAATCAAG	GGAGGACCAAGATGATCCAA	AL606453	58.9 cM
AW563373	AACAGATCCAGTGTGGCATTATC	AATCACCAATGGCAGAATCAAC	AL662944	60.2 cM
BE358270	GCGCTATCAGGTGGGAACA	AAAGCCTCTTGACCAGCCTTATC	AL606618	60.2 cM
BI211826	CGGCTGCAAATAAGAACGATGA	TTTACGGCAGTTGGAGACGAATC	AL662944	60.2 cM
BG355669	CTTACATCATCTTTGGCGTGTGA	ATCCTTCGATCTTAGCGGTGTG	AL606632	62.1 cM
BG560161	AAGCATCTCAAATAAGCCAATTC	CAACCAGAAGGGCAATGAC	AL606458	62.6 cM
BM323488	CTGTAATTCGCATCACTTCACT	ACATCAGGAGAGATGCCTCTG	AL606626	62.6 cM
BM328686	CAAGCGATCTGCGAGGGAATGAA	CGGCTGGGAAGAGGATGAGACAC	AL606458	62.6 cM
BE355764	AGAGCTTTGAAACGGCAACTAGA	GGCGGATCACCATCTCAGAGTAC	AL606452	65 cM
BE357397	CTTCTCTTCCATGGGCGTGTG	CAACAGCTTCAAGGCGCAGAT	AL606452	65 cM
BM326325	CTCAGCAGTCATCAACCCCTGTG	CGCCCATACACCACGATCA	AL606452	65 cM
BF587114	CCTCGAGAGCCTTCTTGCCACTG	GGGGCTGGTCTCCGTGTTC	AL731593	66.4 cM
BI139518	GCCGTGAAAATGGTGATGAGTCT	TTCCGGTCTTCATTGCTAGTCT	AL731593	66.4 cM

VITA

Name:	Jeffrey Alan Brady
Address:	Texas A&M Research and Extension Center 1229 N U.S. Hwy 281 Stephenville, Texas 76401
Phone:	254-968-4144
Email Address:	j-brady@tamu.edu
Education:	B.A., English, Tarleton State University, August 1993 M.S., Biology, Tarleton State University, December 1998 Ph.D., Genetics, Texas A&M University, May 2006