2001

Genetic analysis of reproductive traits in soybean [Glycine max (L) Merr]

I. Made Tasma

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, and the Molecular Biology Commons

Recommended Citation


https://lib.dr.iastate.edu/rtd/680

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
Genetic analysis of reproductive traits in soybean \( \textit{Glycine max} \) (L.) Merr.

By

I. Made Tasma

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Randy C. Shoemaker, Major Professor
E. Charles Brummer
Paul Scott
James T. Colbert
Roger Wise

Iowa State University

Ames, Iowa

2001
This is to certify that the Doctoral dissertation of

I. Made Tasma

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College
# TABLE OF CONTENTS

## ABSTRACT

### CHAPTER 1. GENERAL INTRODUCTION

- Introduction
- Dissertation Organization
- Literature Review

### CHAPTER 2. INHERITANCE OF GENES CONTROLLING PHOTOPERIOD INSENSITIVITY IN SOYBEAN

- Abstract
- Introduction
- Materials and Methods
- Results and Discussion
- Acknowledgements
- References

### CHAPTER 3. MAPPING GENETIC LOCI FOR FLOWERING TIME, MATURITY, AND PHOTOPERIOD INSENSITIVITY IN SOYBEAN

- Abstract
- Introduction
- Materials and Methods
- Results
- Discussion
- Acknowledgements
- References
ABSTRACT

Traits that affect reproduction (e.g., time of flowering, maturity, and photoperiod insensitivity) in soybean are important agronomic characters. These characters are important for developing soybean cultivars with a wider geographical adaptation. The objectives of this study were to: (1) estimate the number of genes controlling photoperiod insensitivity in soybean; (2) map quantitative trait loci (QTL) controlling flowering time, maturity, and photoperiod insensitivity in soybean, and determine if these traits are controlled by the same or different loci in the soybean genome; (3) map homologous and cloned flowering time gene sequences in soybean; and (4) correlate these sequences with maturity ($E$) loci by means of near isogenic lines (NILs). We tested a hypothesis that these flowering time gene homologs may be candidate genes controlling the QTL mapped in the study of objective two.

Objectives 1 and 2 have used two independent $F_6\times F_7$ recombinant inbred (RI) lines of 101 and 100 progeny and objective 3 also included $F_{2:4} G. max X G. soja$ population (60 progeny) in addition to the two sets of RI lines. At least three genes were proposed to control photoperiod insensitivity in soybean and thus making this an acceptable target for QTL analyses. A large-effect QTL for days to R1 (the day when 50% of the plants in a plot have an open flower at one of the top nodes with a fully expanded leaf), R3 (the number of days after emergence when 50% of the plants in a plot had presented the first 5 mm pod at one of the top four nodes with a fully expanded leaf), and R7 (the number of days after emergence when 50% of pods in a plot had mature pod color), and photoperiod insensitivity was found at the same location on linkage group (LG) C2 in both populations. This QTL explained as much as 47% of the total phenotypic variance. This result suggests that photoperiod
insensitivity, flowering time, and maturity may be controlled by the same gene(s) or by

tightly clustered genes in the same chromosomal region. In addition to the large effect QTL,
minor QTL were also detected controlling the four traits in both populations. Minor QTL
account for as much as 17.8% and 12.1% of phenotypic variance in populations IX132 and
IX136, respectively. Thus, time of flowering, maturity, and photoperiod insensitivity in
these soybean populations are proposed to be controlled by a major QTL with a large effect
and modified by several minor QTL. Eighteen soybean cDNA clones, identified by BLAST
to have high similarities with 18 previously cloned flowering time genes, were used as
probes in this study. Ten of the 18 cDNA clones now have been mapped. The homologous
sequences were mapped onto LGs A2 (CRY2), B1 and H (COL1), A1 and B2 (PHYA), C1
(DETI and LD), D2 (AP2), E and K (PHYB), F (COL2), L (FCA), and Q (CCA1). None of
these cDNA sequences have been found to be directly associated with previously mapped
QTL for flowering time. However, analyses of these candidate genes using NILs show that
the homologous gene sequence FCA was associated with maturity locus E3. The map
position and phenotypic data support the hypothesis that homologous gene sequence FCA is
a strong candidate gene for maturity locus E3. Analyses of NILs suggest that PHYB
homolog may be associated with maturity locus E1. However, current data show E1 and
PHYB mapped in different LGs.
CHAPTER 1. GENERAL INTRODUCTION

Introduction

Traits that affect reproduction (e.g., time of flowering, maturity, and photoperiod insensitivity) in soybean are characters of agronomic interest. These characters are important for developing soybean cultivars with a wider geographical adaptation. Soybean cultivars have been classified into 13 maturity groups (MGs) according to their relative time of maturity at different latitudes (Fehr 1987). Individual soybean cultivars are adapted to a narrow range of latitudes (most are restricted to about 4 degrees of latitude or about 480 km of their adapted area). Outside this range, plants mature too early in the south and consequently have low yields, or fail to mature before frost in the north (Summerfield and Wien 1979). The species as a whole, however, has been grown widely from latitudes 0 to 50°.

The transition from vegetative to reproductive development is the outcome of the activation of genes responsible for inflorescence and floral organ formation. These genes, which control the identity of the apex and the morphogenesis of the floral organ, are strictly regulated since their improper expression would result in abnormal flowers and inflorescences (Veit et al. 1993: Levy and Dean 1998; Okamuro et al. 1993). The initial activation of these genes is generally the result of environmental cues indicating the appropriate time to flower. The mechanisms by which environmental factors activate inflorescence and floral organ production are complex and many genes are known to be involved in the transduction of environmental signals that regulate flowering (Levy and Dean 1998; Koornneef et al. 1998; Bernier et al. 1993; Coupland et al. 1995).
Among the environmental factors that are sensed by plants, day length may be the most important in inducing flowering. Another important factor is temperature. The role of photoperiod, or day length, in flowering was demonstrated by Garner and Allard in the 1920s in their classical studies using the tobacco mutant Maryland Mammoth and the soybean variety Biloxi (Thomas and Vince-Prue 1997).

Studies from the model plant *Arabidopsis thaliana* show that flowering (the early stage of plant reproduction) is the result of a sequential action of two groups of genes: namely, floral meristem identity, also called floral initiation process (FLIP) genes, and organ identity genes. FLIP genes are those that switch the fate of meristems from a vegetative phase to a floral phase. Organ identity genes are those that direct the formation of various flower parts (Levy and Dean 1998; Piñiero and Coupland 1998; Koornneef et al. 1998). Genes that control flowering time interact with FLIP genes to initiate flowering. Flowering time genes are those that display their major effects on the duration of vegetative development. These genes act before the actions of FLIP genes. They may activate or repress floral identity genes under different environmental conditions. Mutations of the FLIP genes cause primordia that would normally develop as flowers in the wild-type plants to form structures with shoot-like characteristics (Haughn et al. 1994).

There are four cloned and well characterized FLIP genes in *Arabidopsis*: *LEAFY (LFY)*, *APETALA1 (API)*, *APETALA2 (AP2)*, and *CAULIFLOWER (CAL)* (Haughn et al. 1994; Piñiero and Coupland 1998). All four genes probably encode transcription factors. Another FLIP gene that is not as well characterized is *UNUSUAL FLORAL ORGANS (UFO)*.

Many genes that control flowering time have been identified. Most were identified from studies of a few different plant species such as *Arabidopsis* and pea (Koornneef et al.
From \textit{Arabidopsis} alone, at least 80 loci currently have been reported to affect the timing of flowering (Levy and Dean, 1998). In principle, these genes could act at any level in the regulatory system, from the perception of environmental signals to the activation of floral meristem identity genes in the apex. The genes were identified through the analyses of natural variation of different \textit{Arabidopsis} ecotypes and through characterization of induced mutations (Coupland 1995; Coupland 1997).

Some of these genes (about 25) have been cloned. Table 1 presents examples of the genes cloned from \textit{A. thaliana}. In addition, many genes that were initially studied for their roles in other aspects of plant development – such as light perception and hormone metabolism – also play roles in the regulation of flowering time (Koornneef et al. 1998). At least a dozen of such genes have been cloned.

Candidate gene approaches using homologous gene sequences from \textit{Arabidopsis} have been reported in several crop species. In barley, three homologous sequences [\textit{CONSTANTS (CO), TERMINAL FLOWER I (TFL1)}, and \textit{GIGANTEA (GI)}] from \textit{Arabidopsis} were found to be associated with the previously identified flowering time QTL (Christodolou et al. 2001). In rice, the homologous flowering time gene sequence \textit{PNZIP} from \textit{Pharbitis nil} and homologous sequence \textit{GIGANTEA (GI)} from \textit{Arabidopsis} were found to be associated with the previously detected flowering time QTL (Thompson et al. 2001). The authors concluded that even though the homologous gene sequences do not correspond to known flowering time loci, their presence in the rice and barley genomes suggest there is a considerable conservation of the flowering time genetic pathways and provide support for the continued use of this candidate gene approach. Lagercrantz et al. (1996) also reported that the \textit{CO}
Table 1. Example of cloned *Arabidopsis* genes that affect flowering time (Levy and Dean 1998; Koornneef et al. 1998)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence similarity and probable function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADG-1</strong></td>
<td>-ADP glucose pyrophosphorylase involved in starch metabolism</td>
</tr>
<tr>
<td><strong>CO</strong></td>
<td>-Putative transcription factors with two zinc fingers</td>
</tr>
<tr>
<td><strong>CRY2</strong></td>
<td>-Cryptochrome 2, a flavin-containing blue light photoreceptor</td>
</tr>
<tr>
<td><strong>FCA</strong></td>
<td>-RNA binding protein with a protein-protein interaction domain</td>
</tr>
<tr>
<td><strong>FPF-1</strong></td>
<td>-Novel protein that may involve in response to gibberellic acid (GA)</td>
</tr>
<tr>
<td><strong>FT</strong></td>
<td>-TFL1 homolog</td>
</tr>
<tr>
<td><strong>GA1</strong></td>
<td>-ent-kaurene synthase A an enzyme involved in GA biosynthesis</td>
</tr>
<tr>
<td><strong>GI</strong></td>
<td>-Novel protein with putative membrane-spanning regions</td>
</tr>
<tr>
<td><strong>LD</strong></td>
<td>-Glutamine-rich homeobox transcription factor</td>
</tr>
<tr>
<td><strong>PGM</strong></td>
<td>-Phosphoglucomutase, involved in starch metabolism</td>
</tr>
<tr>
<td><strong>PHYA</strong></td>
<td>-Light-labile Red-Far Red (R-FR) light photoreceptor</td>
</tr>
<tr>
<td><strong>CCA-1</strong></td>
<td>-MYB-related transcription factor; <em>LHY</em> homolog</td>
</tr>
<tr>
<td><strong>CLF</strong></td>
<td>-Homology to Enhancer of Zeste, a <em>Drosophila</em> polycomb gene</td>
</tr>
<tr>
<td><strong>ELF3</strong></td>
<td>-Novel protein</td>
</tr>
<tr>
<td><strong>ESD4</strong></td>
<td>-Novel protein</td>
</tr>
<tr>
<td><strong>LHY</strong></td>
<td>-MYB-related transcription factor; <em>CCA-1</em> homolog</td>
</tr>
<tr>
<td><strong>PHYB</strong></td>
<td>-Light stable R-FR light photoreceptor</td>
</tr>
<tr>
<td><strong>SPY</strong></td>
<td>-O-linked N-acetylglucosamine transferase, involved in modification protein</td>
</tr>
</tbody>
</table>
(CONSTANS) gene isolated from Arabidopsis, a gene that promotes flowering time under the long day photoperiodic pathway (Levy and Dean 1998), is a putative candidate gene for two QTL controlling flowering time in black mustard, *Brassica nigra*.

The objectives of this study were to: (1) estimate the number of genes controlling photoperiod insensitivity in soybean; (2) map quantitative trait loci (QTL) controlling flowering time, maturity, and photoperiod insensitivity in soybean, and determine if these traits are controlled by the same or different loci in the soybean genome; (3) map homologous and cloned flowering time genes in soybean; and (4) correlate these sequences with maturity (*E*) loci by means of near isogenic lines (NILs). We tested a hypothesis that these homologous gene sequences may be candidate genes controlling the QTL mapped in the study of objective two. Studies of objectives 1 and 2 have used two independent *F*₆ recriment inbred (RI) lines of size 101 and 100 progeny each and the study for objective 3 also included *F*₂₄ *G. max X G. soja* population (60 progeny) in addition to the two RI lines.

**Dissertation Organization**

This dissertation consists of a general introduction, three papers, and a general conclusion. The first paper describes the inheritance of genes controlling photoperiod insensitivity in soybean and has been published in Soybean Genetics Newsletter. The second paper deals with mapping QTL for flowering time, maturity, and photoperiod insensitivity in soybean. This manuscript has been accepted for publication in Molecular Breeding. The third paper describes mapping of homologous and cloned flowering time genes in soybean and their association with maturity (*E*) loci. This final manuscript will be submitted for publication to Crop Science. Each of the manuscripts was written in a paper format by
adopting the writing styles of the journal to which each manuscript has been or will be
submitted. The papers are followed by a General Conclusion. Literature cited in the general
introduction and conclusion are listed following the general conclusion.

Literature Review

Flowering time

Flowering time of different plant varieties often varies widely depending upon the
geographical locations in which they are grown. One good example of this is the soybean.
Different soybean cultivars are grown at different latitudes to obtain timing of flowering and
maturity necessary for optimal commercial production. Soybean cultivars have been
classified into 13 maturity groups (MGs) according to their relative time of maturity at
different latitudes (Fehr 1987). Individual cultivars are adapted to a narrow range of latitudes
(most are restricted to within about 4 degrees of latitude). The species as a whole, however,
has been grown widely from latitudes 0 to 50°.

Flowering time of most plant species has been reported to be the outcome of the
interaction between two main factors, environmental cues and endogenous cues (Levy and
Dean 1998; Koornneef et al. 1998). The most important environmental signals that regulate
flowering time are day length, including light intensity and quality, and temperature.

The phenomenon whereby day length regulates flowering is referred to as
photoperiodism (Vince-Prue 1975). Photoperiodic control of flowering allows plants to
coordinate their reproduction with the environment and with other members of their species.
An understanding of the effect of day length on reproductive development has agronomic
importance because the ability to alter flowering time allows for the cultivation of a species in environments that may differ greatly from the one in which it originally evolved.

Day length sensitive plant species have been classified into two classes: short day and long day plants (Vince-Prue 1975). Short day plants are those that flower, or flower earlier, when day lengths are less than a critical period (longer dark period). Long day plants are those that flower, or flower earlier, when day lengths are longer than a critical period (shorter dark period). Those in which flowering times are not affected by photoperiod are known as photoperiod insensitive or day neutral plants.

Vernalization, an extended period of low temperature, promotes flowering in some plant species, especially those from northern latitudes, such as Arabidopsis, barley, wheat, and alfalfa (Koomneef et al. 1998).

The sensitivity of plants to environmental stimuli, especially day length, increases with the age of the plant. This suggests that internal factors that change with plant development, in addition to the environmental factors, are also important in determining floral initiation (Mosley and Thomas 1995). The timing of flowering, therefore, is the result of the interactions between environmental factors, which signal the conditions favorable for the success of reproductive development, and the endogenous developmental competence of the plants.

Physiological studies show that the perception of day length is largely confined to young expanded leaves (Vince-Prue 1975). Other reports indicate that buds can also be sites of photoperiod perception (Jacobs and Sutters 1974), but neither root nor stem tissues appear to be capable of functioning in photoperiod induction. The perception site is in the leaf or bud, but the flowering response occurs at the shoot apical or lateral meristems. Exposure of
leaves to appropriate photoperiods must, therefore, lead to the formation or release of a transmissible stimulus, which is capable of evoking floral morphogenesis.

After inductive day lengths for floral induction occurs in the leaves, a graft-transmissible substance is then transported to the apex where it triggers flowering (Lang et al. 1977). It was proposed that this transmissible substance was ‘florigen’, a flower-promoting hormone produced in leaves under favorable photoperiods and transported to the shoot apex in the phloem (Thomas and Vince-Prue 1997). There was also evidence that in the leaves of plants grown under non-inductive photoperiods, substances are formed that inhibit flowering (Lang et al. 1977). This inhibitor was proposed as an ‘antiflorigen’ (Thomas and Vince-Prue 1997). Many research years were spent on trying to isolate these substances. The failure to identify an individual substance led to the proposal that a complex mixture of substances may be responsible for flower initiation. This is a multifactorial control model which proposed that a number of promoters and inhibitors (including phytohormones) are involved in controlling the developmental transition (Bernier 1988).

Perception of day length by plants involves two photoreceptors, phytochromes and cryptochromes (Kendrick and Kronenberg 1994). Phytochromes are the red/far-red light receptors and are photochromic proteins that exist in two isomeric forms: the red-light absorbing form (Pr) and the far-red-light absorbing form (Pfr). Cryptochromes are the blue/UV-A light receptors and are flavoproteins that share amino acid sequence similarity with DNA photolyases, which catalyze blue/UV-A light-dependent DNA repairing (Cashmore et al. 1999). Blue light (approximately 400-500 nm) and red light (approximately 600-700 nm) are the two spectra of solar radiation that are most effectively absorbed and utilized by the photosynthetic system of the plants. Plants possess multiple.
discrete molecular species of phytochromes, the apoproteins of which are encoded by a small family of divergent genes (Furuya 1993). In *A. thaliana*, for example, there are five genes encoding phytochrome apoproteins: *PHYA, PHYB, PHYC, PHYD*, and *PHYE* (Quail et al. 1995). *Arabidopsis* has at least two cryptochrome genes, *CRY1* and *CRY2* (Levy and Dean 1998).

Another aspect involved in photoperiodism is the endogenous circadian rhythm. Physiological studies show that the endogenous circadian rhythm provides the timer that enables plants to distinguish between LDs and SDs (Thomas and Vince-Prue 1997). The transition from light to dark sets the phase of the rhythm. Detection of this point enables the plants to measure the duration of photoperiod. In soybean, the earliest evidence supporting this aspect of control was reported by Snyder (1940) with Biloxi soybean. Maximum floral inhibition was obtained when short cycle treatments consisting of 3 minutes of light and 3 h of darkness were applied in cycles of 12 or 36 h. The response of the plants treated with an increasing number of short cycles was dependent on the time of the application. Allard and Garner (1941) further reported that Biloxi and Peking soybean were dependent on the ratio of light and darkness and upon the length of the inductive cycle.

**Genetics of flowering time**

**Early studies**

Genetic approaches have been used to study the flowering differences between varieties of the same species since the early decades of the twentieth century. Typically, this was conducted by making crosses between two varieties showing different flowering time and then following the segregation of flowering time among progeny of the cross. Between
1910 and 1920, for example, this approach was used to analyze flowering time in peas (Keble and Peelew 1910), rice (Hoshino 1915), cotton (Leake 1911), wheat (Thompson 1918), and tobacco (Allard 1919). Two genetic mechanisms were reported to control flowering time between a pair of varieties. The first was a single-gene model showing differences in flowering time between two varieties were due to a single gene (Allard 1919). More often, however, differences were shown to be the result of interactions between multiple loci (Goodwin 1944). This latter result supported the previous physiological studies showing that flowering time was under a multifactorial control (Thomas and Vince-Prue 1997).

Induced mutation studies, either with chemicals or radiation, were also reported in these early studies and were found to be useful because the differences between the mutants and the progenitors were caused by a single gene, not multiple loci. Alteration in the vernalization response of a spring barley to create a line with the flowering behavior of a winter variety was one example of induced mutation studies in crop plants (Stubbe 1959). Another example was the mutations of Arabidopsis, each of which delayed flowering time (Redei 1962). These Arabidopsis mutants have been used extensively by later scientists in studies to better understand the genetic control of flowering time. The most recent advances of our understanding on the genetic mechanism that controls this important trait has come from the model plant A. thaliana.

Genetics of flowering time from model plant A. thaliana

A. thaliana is classified as a facultative long day plant that needs about 16 h of light exposure to accelerate flowering. Flowering is also accelerated by extended exposure (3-6
weeks) to low temperature after germination (Koomneef et al. 1991). The early flowering ecotypes commonly used in the laboratory, such as Lansberg *erecta* (*Ler*) and Columbia, flower in about three weeks if grown under inductive long days (16 h light/8 h dark) and form an average of four rosette leaves prior to bud appearance. When grown under noninductive, short days (10 h light/14 h dark), however, they flower at least six weeks and form an average of 19 rosette leaves prior to bud appearance (Coupland 1995).

A systematic genetic approach to identifying genes involved in the flowering transition has been conducted with *Arabidopsis* (Koornneef et al. 1998) and, in a less intensive study, with pea (Weller et al. 1997). In *Arabidopsis*, genes that promote flowering were identified as mutations that delay flowering time. Genetic variation causing similar phenotypic effects was recovered by crossing different ecotypes. Alleles causing late flowering extend the duration of vegetative growth and, therefore, increase the number of leaves formed before the development of flowers.

Two types of mutants were reported in *Arabidopsis*. late flowering and early flowering mutants (Coupland 1997; Koornneef et al. 1998). This classification was based on their phenotypes in response to the environmental factors, day length, and vernalization (Figure 1). These are known as flowering time mutations and the corresponding genes are known as flowering time genes. Mutants that remain in the vegetative phase indefinitely and never undergo the transition to flowering have not been identified, but mutations that disrupt flowering responses in distinct ways have been described. This suggests some degree of redundancy exists between genes that promote flowering. Therefore, inactivation of a single gene is partially compensated for by other genes (Levy and Dean 1998; Piñeiro and Coupland 1998).
Figure 1. Types of mutants (late and early flowering mutants) recovered from *A. thaliana* (redrawn from Coupland 1997). Not all mutants within each group are shown in this picture.
Late flowering mutants were the first group to be identified and analyzed in detail (Redei 1962; Koornneef et al. 1991). Some of these genes act to promote flowering and others to repress it. Some appear to interact with environmental factors (e.g., photoperiod and temperature) while others appear to act in an autonomous way. Thus, based on these observations, together with data from double mutant studies and transgenic plant analyses, it was proposed that at least four different genetic pathways may control flowering time in *Arabidopsis* (Levy and Dean, 1998; Koornneef et al. 1998). An example of the current model of the genetic pathways is shown in Figure 2 (Levy and Dean 1998).

The first group of late flowering mutants includes *co*, *jha*, and *gi*, which show little or no response to vernalization, and flowering is delayed under long day (LD) but not under short day (SD) (Figure 1). This suggests the wild-type plants promote flowering under LD and this group of genes, therefore, is proposed to act through a pathway that promotes flowering under LD conditions (Levy and Dean 1998; Koornneef et al. 1998). The second group of late flowering mutants includes *fca*, *fe*, *fpa*, *fve*, *fy*, *ld*, and *fri*, which flower later than the wild-type plants both under LD and SD conditions and show a decrease in flowering time in response to vernalization (Figure 1). These genes are proposed to act in an autonomous pathway that promotes flowering independent of environmental conditions (Coupland 1997; Koornneef et al. 1998; Levy and Dean 1998) (Figure 2). The third group includes *gal* and *gi*, which have been identified as affecting synthesis or responses to the plant hormone gibberelic acid (GA). These mutants have slight effects on flowering time under LDs, but severe effects on SDs (mutants flower similar to wild-type plants under LDs but flower much later under SDs) (Figure 1). This suggests the hormone is more important for flowering under short-day conditions (Wilson et al. 1992; Levy and Dean 1998).
Figure 2. Proposed genetic pathways controlling flowering time in *A. thaliana* (redrawn from Levy and Dean 1998). GA = Gibberellic acid
These mutants were proposed to act in the GA pathway (Levy and Dean 1998) (Figure 2). The fourth group of mutants are those that were categorized as early flowering mutants (Coupland 1997; Levy and Dean 1998).

The early flowering mutants were identified more recently than the late flowering mutants. Most early flowering mutants have been categorized by their response to photoperiod. Since the mutants flower early regardless of day length, the products of the corresponding wild-type genes are thought to act in repression of flowering and, therefore, act in the repression pathway (Levy and Dean 1998) (Figure 2). Mutants in this category include clf, elf1, elf2, emf1, emf2, esd4, flc, pef1, pef2, pef3, phyB, speedy, tf11, tf12, and wlc (Figures 1 and 2). The EMF genes have been considered to play a major role in repressing flowering. Strong mutant alleles of emf1 and emf2 flower with essentially no preceding vegetative phase (the mutants flower directly after emergence without forming any rosette leaves) (Sung et al. 1992).

Based on mutant characteristics described above, it is clear there are several ways for plants to be photoperiod insensitive, depending on the type of genetic route a particular mutant has. For example, the response of particular mutants (e.g., gal and gai) to the hormone GA in the GA pathway delays flowering under the non-inductive short day length. The wild-type plants, therefore, flower early regardless of day length and thus are photoperiod insensitive.

A few mutants were identified through analyses of natural variation from different ecotypes (late flowering ecotypes). These analyses found several mutants that are typically dominant. FRI, FLC, FKR, and JUV are examples of mutants identified through this method. All the quantitative trait loci (QTL) controlling this trait were identified by this method.
Most well studied mutants were identified by characterizing induced mutations using both chemicals [e.g., ethyl methane sulfonate (EMS)] and radiation. These latter mutants were generally recessives and were derived from three rapid progenitor ecotypes (Landsberg erecta, Columbia, and Wassilewskija).

Among the 80 loci recognized so far that affect flowering times in Arabidopsis, at least 25 have been cloned (Table 1). In addition, many genes that were initially studied for their roles in other aspects of plant development – such as light perception and hormone metabolism – also play roles in the regulation of flowering time and are accordingly called flowering time genes (Koomneef et al. 1998).

**Genetics of flowering time and maturity in soybean**

Soybean [Glycine max (L.) Merr.] is classified as a short day species in which flower buds are induced under short day conditions (Garner and Allard 1920; Kenworthy et al. 1989). Agronomists have long recognized the response to day length in soybean as a principal factor in determining the time of flowering and maturity and, hence, the geographical area of adaptation for an individual cultivar. As previously stated, soybean cultivars are divided into 13 MGs based on their relative time of maturity and are adapted to a narrow range of latitudes. Their adaptation variation is equivalent to variation in the timing of reproductive events. These events are caused by genotypic interactions with seasonal changing environmental factors of photoperiod and temperature, which act as cues.

Criswell and Hume (1972) studied 111 MG 00 soybean strains and found the number of days to flowering for about 70% of them was unaffected by photoperiod treatments ranging from 12 to 24 h. A total of 1,978 strains from MG 00 to MG VIII were grown under
16-h photoperiod and also under natural day length at the Asian Vegetable Research and Development Center in Taiwan (AVRDC, 1975). Eleven percent (212) of the strains were delayed in flowering less than five days by extended photoperiod and all 212 were from MG IV or earlier. Nissly et al. (1981) tested 515 strains of MG III in Urbana, Illinois, and in Issabela, Puerto Rico, under natural day length and under photoperiod extended with artificial light in addition to natural day length. The strains tested exhibited a wide range of photoperiod sensitivity. Thirty-two strains were identified as having low photoperiod sensitivity and one of these, PI 317.334B ('Kitami Shiro'), exhibited almost no photoperiod sensitivity at either location. PI 317.334B was then classified as a day-neutral strain. In another study, PI 317.336 ('Shinsei') was also reported as a day-neutral genotype (Guthrie 1972). The study was based on field conditions extended with incandescent light. PI317.336 delayed flowering only 3 days under continuous light compared to that grown in 12 h light based on a 24-h cycle.

Studies by Johnson et al. (1960) and Byth (1968) have shown that, in general, early maturing soybean cultivars are less affected by changes in photoperiod than are later ones. Thus, the search for day-neutrality has been most successful in the earliest maturing soybean germplasm. The identification of strains insensitive to photoperiod would provide germplasm that could be used to develop cultivars with wide areas of adaptation.

A minimum number of inductive nights is needed for floral induction and flowering in soybean. Wilkerson et al. (1989) reported a minimum of 5-6 long nights (8 h light/16 h darkness) was needed to cause flower expression. Shanmugasundaram and Tsou (1978) tested sensitive and insensitive soybeans with 10 h and 16 h photoperiod and reported that, for soybean photoperiod-sensitive lines, the induction period was 27 short days (10 h) and
that anthesis occurred 10 days after the completion of induction. The critical time to begin
induction was 9 days after emergence. The insensitive lines were not affected by
photoperiod.

Five genes have been reported to affect flowering time and maturity in soybean
(Cober et al. 1996a). The genes are known as the E-series: $E_1$ and $E_2$ (Bernard 1971), $E_3$
(Buzell 1971), $E_4$ (Buzell and Voldeng 1980), and $E_5$ (McBlain and Bernard 1987).
McBlain et al. (1987) reported these loci interact with photoperiod in the control of
flowering. Three of the genes ($E_1$, $E_2$, and $E_3$) have been mapped on LGs C2, O, and L,
respectively, based on the consensus soybean genetic map (Cregan et al. 1999).

Under natural day length, the dominant soybean maturity locus $E_1$ was reported to
delay flowering 16-23 days and maturity 15-18 days compared to its respective recessive
allele $e_1$ (Bernard 1971): $E_2$ delayed flowering 7-14 days and maturity 14-17 days (Bernard
1971); $E_3$ delayed maturity 6-8 days (Buzzell 1971; Buzzell and Bernard 1975); $E_4$ delayed
flowering 1-6 days and maturity 8-20 days (Saindon et al. 1989); and $E_5$ delayed flowering
and maturity similar to $E_2$ (McBlain and Bernard 1987). The effects of these maturity loci
were also tested under natural day length extended to 20 h with incandescent light, which is
called incandescent long day length (ILD) (Saindon et al. 1989; Cober et al. 1996a). The
insensitivity to ILD was reported to be controlled by the $E_3$ and $E_4$ loci (Buzzell and
Voldeng 1980; Saindon et al. 1989) in which only $e_3e_3e_4e_4$ genotype is ILD-insensitive. It
was also concluded that $E_3$ is the major locus conferring long day length insensitivity in
soybean (Saindon et al. 1989), but for breeding purposes, both $E_3$ and $E_4$ loci should be
considered when breeding for insensitivity to long daylength using ILD.
The response of the maturity \( E \) genes on light quality has also been studied in soybean (Cober et al. 1996b). The use of different lamp types emitting lights of different qualities changed flowering responses to long day length. ‘Harosoy’ near isogenic lines (NILs) were grown under 20-h long days of different light qualities as measured by the bichromatic ratio of red to far-red quanta (R:FR). Sensing of the R:FR ratio is the function of phytochromes in the light grown plants. Photoperiod response was greater (i.e., later flowering) with the decrease of R:FR ratio. The \( E1 \) allele was found to be most sensitive to light quality and required an R:FR approximating that of natural daylight for response to long days. The \( E3 \) allele showed the least sensitivity and the \( E4 \) allele showed intermediate sensitivity to long day length. The genotype \( ele3e4 \), previously reported to be insensitive to long day length, was found to be sensitive to long day length under low R:FR ratio. Long days of high R:FR were not effective in delaying flowering for some genotypes. Based on the differential responses of the three loci studied on the changes in R:FR ratio, the authors speculated either a close relationship between the soybean \( E \) alleles and phytochromes or the possibility that some photoperiod sensitive loci may be part of the phytochrome gene family.

Determinate (\( dtl \)) and indeterminate (\( Dtl \)) growth habits have also been reported to affect flowering time and maturity in soybean. Foley et al. (1986) reported that determinate lines flower one day earlier and mature 3.5 days earlier than indeterminate lines. Previously, Bernard (1972) reported a similar result showing that determinate, near-isogenic lines matured 2 to 3 days earlier than indeterminate lines.

Soybean lines lacking “short-day requirements” for one or all stages of development have been identified (Criswell and Hume 1972; Polson 1972; and Nissly et al. 1981). Flowering and maturity responses to long day length are known to be controlled by major
genes at two loci, \( E3 \) (Buzzell 1971; Kilen and Hartwig 1971; Buzzell and Voldeng 1980) and \( E4 \) (Buzzell and Voldeng, 1980). Both \( E3 \) and \( E4 \) control responses to long day length and should be of major importance in breeding genotypes adapted to the very long day length (>16h) at the higher latitudes.

Several QTL associated with flowering time and maturity have been previously mapped in soybean. Keim et al. (1990) reported that five markers on linkage groups C1, C2, and D1 (Shoemaker and Specht 1995) were found to be associated with maturity (including time of flowering) in an \( F_2 \) population derived from a cross between \( G. \ max \) and \( G. \ soja \). The observed QTL explained 17-23% of total phenotypic variance. Using another \( F_2 \) population of ‘Minsoy’ X ‘Noir 1’. Mansur et al. (1996) also reported a major QTL for flowering time on LG C2 and minor QTL for maturity on LG L and M (Cregan et al. 1999). Lee et al. (1996) reported that QTL for maturity traits were found on LG K. The study was based on an \( F_2 \) population derived from a cross between PI 97100 and ‘Coker 237’. The observed QTL explained 26.2-31.2% of phenotypic variance. These results suggested the putative QTL for maturity traits may be population-specific as indicated by different genomic regions (LG) that control the same trait(s) in different mapping populations. Yamanaka et al. (2000) reported that three QTL (\( FT1 \), \( FT2 \), and \( FT3 \)) were found to control flowering time in an \( F_2 \) population derived from a cross between two varieties, ‘Mizuzudaizu’ and ‘Moshidou Gong 503’. A major QTL (\( FT1 \)) mapped on LG 3-2 (LG C2, Cregan et al. 1999) accounted for approximately 70% of the total variation and appeared to correspond to the maturity locus \( E1 \), the locus for flowering time and maturity of the classical genetic map. The minor effect QTL were observed on other LGs (LGs 25 and 16). The authors could not correlate these LGs to the consensus genetic map because of the limited number of genetic markers.
Molecular markers

The use of genetic markers has been reported since the early decades of the twentieth century. Many classical (morphological) markers such as dwarfism, chlorophyll deficiency, and leaf color have been used in earlier genetic studies. Sax (1923) detected variation of seed size, a quantitatively inherited trait in *Phaseolus vulgaris*, using a morphological marker seed coat color. A linkage between a major gene controlling flower color and a quantitative trait flowering time was also reported in peas (Rasmusson 1935). The number of morphological markers, however, is limited and their expression is strongly affected by the environment. This limitation has been largely solved by the discovery of molecular markers that include isozyme and DNA markers.

Isozyme markers are based on the coding sequences detected by different allele mobilities at the protein level in gel electrophoresis. The isozyme marker can be used to detect more polymorphisms than morphological markers. Isozyme loci are also commonly scored as co-dominant markers. Since only the coding regions of the genome influence these markers and their expression is often strongly affected by the environment and stages of plant development, the number of polymorphisms detected by this marker is often limited. In addition, the specificity of the enzyme system used to detect the marker loci further limits the utility of this marker (Kepart 1990).

The limitation of morphological and isozyme markers has been largely solved by the introduction of DNA markers (Smith and Smith 1992). This is mainly because almost all DNA sequences in the genome potentially could be used as markers. The markers may either be the coding or the non-coding regions of the genome. DNA markers are also not affected
inter-allelic (epistatic) interactions. These features of DNA markers facilitate the
development of genetic maps for many important crop plants and domesticated animals.

A DNA marker is typically a DNA sequence showing polymorphism between two
individuals of a species. Many types of DNA markers have been described lately. Among
those are restriction fragment length polymorphisms (RFLPs), random amplified
polymorphic DNA (RAPD), simple sequence repeats (SSRs) or microsatellites, and
amplified fragment length polymorphism (AFLP).

RFLPs were the first DNA markers identified (Botstein et al. 1980; Soller and
Beckmann 1983). This marker is based on different lengths of DNA fragments of a
particular chromosomal locus created when DNA of different individuals, from the same or
different species, is digested with restriction enzymes. The polymorphisms are due to base
sequence changes or DNA rearrangements that are naturally occurring and are simply
inherited as Mendelian characters. RFLPs are also ubiquitous throughout plant tissues and
throughout the plant genome's coding and non-coding sequences, demonstrate relatively high
polymorphics, and appear to be selectively neutral and co-dominantly inherited (Helentjaris
and Burr 1989). Many genetic maps of important crops have been constructed mainly based
on RFLP markers, including soybean (Shoemaker et al. 1993; Shoemaker and Specht 1995;
Mansur et al. 1993; Cregan et al. 1999). However, this technique is relatively laborious,
expensive, and few loci can be detected per assay. Because of the use of radioactive
materials, this technique may be less attractive compared to more recently developed
markers.

The development of the polymerase chain reaction (PCR) (Saiki et al. 1988) has
expanded the repertoire and efficiency of available DNA marker systems. PCR-based
markers include random amplified polymorphic DNA (RAPD), microsatellites or simple sequence repeats (SSRs), and amplified fragment length polymorphism (AFLP).

RAPD technique (Williams et al. 1990) is considered to be more straightforward compared to RFLP analysis and requires only a small amount (nanogram) of genomic DNA. RAPD markers provide a quick and simple method to gather information on genetic variability in a wide range of organisms for taxonomic purposes and also for genetic mapping (Reiter et al. 1992). RAPD uses a single, arbitrarily short oligonucleotide primer (8 to 10 nucleotides in length) to amplify genome segments flanked by two complementary primer-binding sites in inverted orientation (Williams et al. 1990). This marker is mostly scored as a dominant marker and therefore is less powerful as a genetic marker than the co-dominant markers such as RFLPs. This technique is also sensitive to subtle changes in reaction conditions and, hence, difficult to reproduce (Jones et al. 1997).

SSRs, also called microsatellites, are locus-specific, abundant, multiallelic, co-dominant PCR-based markers. They are also uniformly distributed throughout the genome of numerous species, including soybean (Akkaya et al. 1995). SSRs are short segments of DNA consisting of a small number of repeated nucleotide sequences such as \((\text{CA})_n\), \((\text{AAT})_n\), \((\text{AGAT})_n\), \((\text{AT})_n\), and \((\text{ATT})_n\) (Akkaya et al. 1992, Saghai Maroof et al. 1994). These repeated sequences are flanked by conserved DNA sequences allowing them to be amplified by PCR using primers designed from the conserved sequences. Variation in the number of tandem repeats results in the length differences of the PCR products. The development of the oligonucleotide primers of this marker, however, is time-consuming, expensive, and relatively few loci can be identified per assay. Once such primers are available, however, they are very easy to use because of their simplicity. SSRs were reported to exhibit high
levels of polymorphisms in many plant species such as soybean, barley, and rice (Akkaya et al. 1992; Margente-Oliveri 1993; Wu and Tanksley 1993; Saghai Maroof et al. 1994). In a sample of 20 rice genotypes, Wu and Tanksley (1993) reported there were 5 to 11 alleles at each of 8 microsatellite loci. Studies from barley sequences were reported by Saghai Maroof et al. (1994). They observed the number of alleles per SSR locus ranged from three to thirty-seven. Cregan et al. (1995) further reported, SSR loci having as many as 23 alleles were found in soybean. A consensus soybean genetic map containing more than 300 SSR markers has been constructed (Cregan et al. 1999). Because of their high degree of polymorphism and informativeness, SSRs are very useful in many genetic applications.

The recently developed AFLP (Vos et al. 1995) appears to be a promising genetic marker that may be useful in particular genetic applications. These include genetic mapping, varietal fingerprinting, and genetic diversity studies. The major advantage of this technique is its power to identify large numbers of potentially polymorphic loci. A large number of amplified products are generated in a single assay (Powell et al. 1996).

AFLP markers are a modification of RFLPs, but PCR-based and a priori sequence information is not required to develop this marker. The genetic basis of polymorphisms results from a change in the restriction site, as happens in RFLPs. Genomic DNAs are restricted with two different restriction endonucleases (typically, one is a six-base cutter and the other is a four-base cutter: e.g., EcoRI and MseI). Adaptors are then used to tag the restriction fragments. Universal primers are designed based on the adaptors’ sequences and are used in the PCR amplifications. A two- to three-selected base extension is attached to the primers and acts as a selective binding of the primers on the target DNA templates (Vos et al. 1995). The PCR products are then separated by a polyacrylamide gel electrophoresis.
AFLPs are classified as dominant markers since only the presence and the absence of a particular DNA band is scored. Their power, especially on genetic mapping purposes, therefore, is less than that of co-dominant markers. In soybean, Keim et al. (1997) constructed an AFLP-based genetic linkage map. It contains 650 AFLP markers which are distributed across 28 linkage groups (LGs). The map covers 3441 cM with an average distance of 4 cM between any two adjacent markers.

**Mapping quantitative trait loci (QTL)**

In the last decade, the advent of molecular markers has greatly facilitated the systematic dissection of QTL into their underlying Mendelian factors (Lee 1995). Being very numerous and phenotypically neutral, molecular markers have allowed the construction of high resolution genetic maps for many plant species. These maps have facilitated the detection of QTL for many agronomically important traits.

Dissection of QTL using such a Mendelian factor was first cited by Sax in 1923 who reported the association of a quantitatively inherited trait, seed size, with a monogenic, classical marker, seed-coat color in *Phaseolus vulgaris* (Sax 1923). A similar study was conducted in peas (*Pisum sativum*), reporting observed linkages between a quantitative trait for flowering time and a major gene controlling flower color (Rasmusson 1935).

The main principle of QTL detection is to obtain correlation between the trait value under examination and the genotype of every marker (Tanksley 1993). A significant correlation test means that at least one QTL may exist near the marker locus. To be able to detect QTL using the marker loci, a linkage disequilibrium must exist between alleles at the marker loci and alleles of the linked QTL (Tanksley 1993).
The identification of QTL has been performed mainly by one of two methods, single-point analysis of variance or interval mapping (Tanksley 1993). Single-point analysis (Soller et al. 1976; Edward et al. 1987) is done by analyzing the association of one marker at a time with the trait. This is the simplest method of detecting QTL since no complete genetic linkage maps are required prior to the analysis. The main weakness of this method, however, is that the QTL effects tend to be underestimated and may be confounded with recombination between the marker and the QTL (Tanksley et al. 1993). The QTL effect will decrease considerably if the recombination frequency between the marker and the QTL increases (Edward et al. 1987). Confounding of the QTL effect and recombination frequency also causes difficulty in estimating the effect of small QTL and the effect of the recombination of distance markers and QTL (Lander and Botstein 1989). To detect QTL with small to moderate effects, this method requires a larger population size compared to the interval mapping method (Soller et al. 1976). This method, however, is powerful if the QTL are located near a marker.

Interval mapping (Jensen 1989; Lander and Botstein 1989) depends on the availability of genetic maps, preferably with markers that cover the entire genome. This method is mainly based on maximum likelihood estimation and allows testing of many markers simultaneously to find ones that may link with the trait. The log of odd (LOD) score is used to estimate the QTL location and effect. The presence of a QTL, basically, is estimated from the trait distribution within each marker genotype class and the mean differences between the genotype class of flanking markers (Lander and Botstein 1989). Interval mapping procedures would solve the problems faced by the single factor analysis (Tanksley 1993). This method, compared to the others, will be very beneficial in detecting
QTL if the distance between linked markers is greater than 20 cM, so that many crossovers are likely between the marker loci and the QTL.

Another method of QTL detections includes multiple linear regression (Wright and Mower 1994), used to detect QTL with additive, dominant, and epistatic effects in different crosses and generations (Moreno-Gonzalez 1992). The computation, however, is complex, especially if the number of QTL is large. In addition, this method is still debatable, mainly concerning the genetic model used (Dudley 1993).

The most recent method in QTL detection is the composite interval mapping (Zeng 1994). This method is a blend of the maximum likelihood method and the multiple linear regression. The basic concept of this technique is an interval test where statistical testing of the interval is free from the interference of QTL located outside the interval (Jansen 1993; Zeng 1994). This method, according to Zeng (1994), increases the precision of the QTL location and increases the efficiency of QTL mapping compared to other methods.

Candidate genes

The candidate-gene (CG) approach has emerged as a promising method of merging QTL analysis with the extensive data available from cloning and characterization of genes involved in many different traits, including those controlling time of flowering. In this method, genes involved in biochemical pathways leading to trait expression are employed as markers for QTL analysis. This approach, also called the 'positional' candidate gene approach, relies on QTL mapping and on examination of known function genes or mutations which map in the same region (co-localization between the gene(s) or mutations and the
QTL, the effect of which may be related to the trait (Prioul et al. 1999). This approach has
been applied successfully in several plant species (Goldman et al. 1993; Causse et al. 1995).

Byrne et al. (1996) reported the most compelling case for linking candidate genes
involved in the flavone synthesis pathway of maize with the host defense response (DR)
phenotype associated with QTL resistance to corn earworm. In this study, the pI locus,
encoding a transcriptional activator, together with three other candidate genes accounted for
75.9% of the phenotypic resistance variance. The application of the CG approach to QTL of
disease resistance was also reported by Faris et al. (1999). Several candidate genes,
including chitinase, oxalate oxidase, peroxidase, superoxide dismutase, and thaumatin,
mapped within the previously identified resistant QTL and explained a greater amount of the
phenotypic variation. The authors also concluded the candidate-gene approach can be an
efficient method for QTL identification.

Pnueli et al. (1998) reported the SELF-PRUNING (SP) gene in tomato was found to
be the functional ortholog of CENTRORADIALS (CEN) of Anthirrhinum and of TERMINAL
FLOWER1 (TFL1) of Arabidopsis. The SP gene controls the indeterminate growth habit of
tomato plants. The homozygous recessive allele of the SP gene confers accelerated
termination of sympodial units by the inflorescence, resulting in a limited growth of the
shoot, a bushy, compact constitution, and nearly homogeneous fruit setting (Atherton and
Harris 1986). A principle assumption of this study was that Arabidopsis, Antirrhinum, and
tomato employ similar genes to regulate their growth habits. These genes, however, operate
in different meristematic contexts: monopodial, indeterminate, and photoperiod-sensitive in
the former two species and sympodial, determinate, and day-neutral in the latter species
(Parnis et al. 1997). Lagercrantz et al. (1996) also reported the CO (CONSTANS) gene
isolated from *Arabidopsis*, which promotes flowering under the long day photoperiodic pathway (Levy and Dean 1998), is a putative candidate gene for two QTL controlling flowering time in black mustard, *Brassica nigra*. The QTL were on LGs 2 and 8, explaining 53 and 12% of the total phenotypic variances, respectively.

The most advanced CG studies for plant QTL characterization concerned several complex traits in maize. Physiological studies had shown that mutations in the *dwarf3* gene had an effect on plant height (Fujioka et al. 1988) and that overexpression of the *phytochrome B (PhyB)* gene led to a dwarf phenotype. These two genes co-localized with a plant-height QTL on chromosome 9 (Touzet et al. 1995). Further analyses of these CGs showed *dwarf3* was the best candidate because only this CG showed polymorphism between near-isogenic lines (NILs) containing contrasting alleles in the QTL region.

QTL for protein and starch contents in maize kernels were mapped. The *Sh2* gene encodes one of the ADP-Glucose pyrophosphorylase subunits, an enzyme of the starch biosynthesis pathway, was found to be located in the vicinity of one of the detected QTL on chromosome 3 (Goldman et al. 1993). Using another segregating population, the *Sh2* locus was also colocalized with a QTL responsible for controlling amylose content in kernels (Prioul et al. 1999).

Kianian et al. (1999) reported a map colocalization between a cDNA for plastidic acetyl-CoA carboxylase (ACCase) and a major groat oil content QTL of oat on LG11 using two RI lines. The QTL linked to ACCase accounted for up to 48% of phenotypic variance for groat oil content. These results support the hypothesis that ACCase has a major role in determining the groat oil content in oat. The ACCase is involved in the first step of *de novo* fatty acid synthesis.
Key genes controlling the acclimation to water stress in maize have been identified. This was done by mapping QTL that controlled key enzyme activities in the control and the stress leaves (Pelleschi et al. 1999). It was previously reported the early response to water shortage was a two- to three-fold increase in the vacuolar invertase activity. The Ivr2, a gene encoding vacuolar invertase, colocalised with two QTL, one for the control and one for the stressed plants, on chromosome five. Ivr2 was proposed to be a likely CG for regulation of invertase activity under control conditions but not under stressed conditions. Most identified stress-specific regions were not colocalized with any CGs, suggesting that the regions may be controlled by other unknown regulatory genes.

Near-isogenic lines and their uses in genetic mapping

Near-isogenic lines (NILs) or isogenic lines are produced by repeated backcrossing of characters under monogenic control from a donor parent (DP) into a standard recurrent parent (RP) line (Fehr 1987). This results in a series of lines differing from RP at only one locus or, more precisely, by a short length of chromosome adjacent to that locus. After repeated backcrossing, most of the donor genome would be eliminated from the RP except for the segment surrounding the introgressed gene of interest. Muehlbauer et al. (1988) reported that a soybean NIL developed after six backcrosses would have, on average, 2.1% of its genome originated from DP. Sixty-five percent of the donor genome contribution in the NIL would be located in the linkage block surrounding the monogenic locus backcrossed into the NIL. If the source of the gene (DP) and the RP are sufficiently polymorphic with respect to one another, the introgressed segment can be used as a target to determine whether a given genomic clone is located near the gene of interest. Markers that have high
probability of being linked to the introgressed gene(s), therefore, could be found by
surveying markers that are polymorphic between the RP and the NIL. A linkage test using a
standard segregating population could be conducted to verify the marker is linked with the
introgressed gene.

Many NILs containing different traits have been developed in soybean (Bernard et al.
1991). These include NILs for loci controlling flowering time and maturity. The NILs are
readily used for mapping particular traits of interest.
CHAPTER 2. INHERITANCE OF GENES CONTROLLING PHOTOPERIOD INSENSITIVITY IN SOYBEAN

A paper published in Soybean Genetics Newsletter

I. M. Tasma, L. L. Lorenzen, D. E. Green, and R. C. Shoemaker

Abstract

The objective of this research was to study the inheritance of genes controlling photoperiod insensitivity in soybean. Two single-cross populations, IX132 (PI 317.336 X ‘Corsoy’) and IX136 (PI 317.334B X ‘Corsoy’), were developed for this purpose. The populations were inbred to obtain 101 and 100 F_6 lines, respectively, using a modified single seed descent. Flowering time (days to R1) of the RI lines from each population was observed in the growth chamber at 12 and 20 h photoperiods using fluorescent and incandescent lamps. Results show the RI lines have dramatically different responses to

---


2 Department of Agronomy, Iowa State University, Ames, Iowa 50011. USA

3 Office of Biotechnology, Iowa State University, Ames, Iowa 50011. USA

4 USDA-ARS, Corn Insects and Crop Genetics Research Group, Iowa State University, Ames, Iowa 50011. USA: Author for correspondence.
day length. A normal distribution of flowering times was observed when the lines were grown in the growth chamber with a 12 h photoperiod. When the lines were grown in the growth chamber with a 20 h photoperiod, however, a discontinuous distribution was observed. This suggested that the insensitivity of the RI lines on long day length may be controlled by a few major genes. The time of flowering was delayed in almost all lines when grown in growth chambers with a 20 h photoperiod compared to those with a 12 h photoperiod. The flowering delays were 5 to 75 days in population IX132 and 0 to 75 days in population IX136. Chi-square tests show the segregation data fit a 1:6:1 ratio in populations IX132 and IX136. Based on these tests, a minimum of three genes are proposed to control photoperiod insensitivity in both populations.

Introduction

Soybean [*Glycine max* (L.) Merr.] is recognized as a short day plant (Garner and Allard 1920; Kenworthy et al. 1989). Most soybean genotypes require short day exposure to initiate flowering. Shanmugasundaram and Tsou (1978) reported that for photoperiod sensitive genotypes, 27 short days (10 h photoperiod) were required for flowering induction: anthesis was observed 9 days after the completion of the induction. They also reported the critical time of short day exposure was 9 days after emergence.

Photoperiod insensitivity has also been reported in soybean (Yoshida 1952; Criswell and Hume 1972; Guthrie et al. 1972; Nissly et al. 1981). Two insensitive lines include PI 317.336 (‘Sinshei’) and PI 317.334B (‘Kitami-Shiro’), two genotypes introduced from Japan. In addition, it has also been reported that the early-maturing genotypes are less affected by
changes in photoperiod than later-maturing genotypes (Johnson et al. 1960; Byth 1968; Criswell and Hume 1972; Kenworthy et al. 1989).

At least five genes have been reported to control flowering time and maturity in soybean (Cober et al. 1996a). The genes are known as the $E$-series: $E_1$, $E_2$, $E_3$, $E_4$, and $E_5$ (Bernard 1971; Buzzell 1971; Buzell and Voldeng 1980; McBlain and Bernard 1987). McBlain et al. (1987) reported that these loci interact with photoperiod in the control of flowering. Under natural daylength, the dominant $E$-genes tend to delay time of flowering and maturity but the magnitude of effect of each gene can be different (Cober et al. 1996a).

Flowering and maturity response of soybean to long day length also have been studied. The studies were based on near isogenic lines (NILs) containing specific alleles of $E$-genes tested mainly under field conditions with natural light extended to 20 h with incandescent light, also refered to incandescent long day length (ILD). The NILs were developed by backcrossing to transfer particular alleles of $E$-genes and alleles of a gene controlling growth habit ($D_{tl}$) from donor parents into the recurrent parent ‘Harosoy 63’ ($e_1e_2E_3E_4e_5D_{tl}$) in the Cober et al. (1996a) study, and into recurrent parents ‘Harosoy-e3’ ($e_3e_3E_4E_4$), ‘Evans-e3’ ($e_3e_3E_4E_4$), and ‘Maple Arrow’ ($E_3E_4e_4e_4$) in the Saindon et al. (1989) study. From the Saindon et al. (1989) study, only plants with the genotypes $e_3e_3e_4e_4$ were ILD-insensitive. They also proposed a two-gene model ($E_3$ and $E_4$) each with two alleles ($E_3e_3E_4e_4$) plus epistasis of $E_3$ to $e_4$ to explain the ILD-insensitivity. It was concluded that $E_3$ is the major locus that controls long day insensitivity in soybean, but both $E_3$ and $E_4$ must be considered when breeding for insensitivity to long day length using ILD. A later study by Cober et al. (1996a) reported that under natural daylength, the $E_3$ and $E_4$ alleles each delayed flowering 5 days and maturity 15 days compared with the alternative
early-maturation alleles. The allele $E_4$ was found to be epistatic to $e_3$. The $E_3$ and $E_4$ alleles each delayed flowering 10 days under ILD compared with natural day length. Thus, based on these studies, the insensitivity to long day length in soybean is controlled by two loci, $E_3$ and $E_4$.

The study reported herein was based on segregating progeny of two single-cross populations grown in a growth chamber. The parents of the cross had widely different responses to photoperiod. The PI parents (PI 317.336 and PI 317.334B) were reported as day-neutral genotypes and hence insensitive to long day length with regard to flowering time and maturity (Guthrie 1972; Nissly et al. 1981; Metz et al. 1985). The ‘Corsoy’ parent, on the other hand, is sensitive. Recombinant inbred (RI) lines were developed from each population. Using this strategy, we tested the hypothesis that the insensitivity of soybean in response to long day length is controlled by two genes.

**Materials and Methods**

**Population development**

Two single-cross populations were developed. The first population, IX132, was developed by crossing PI 317.336 (‘Shinsei’) and ‘Corsoy’. The second population, IX136, was developed by crossing PI 317.334B (‘Kitami-Shiro’) and ‘Corsoy’. Both PI parents were reported to be photoperiod insensitive with regard to flowering time and maturity and are classified as maturity group 0 (MG 0) (PI 317.336) and MG III (PI 317.334B) (Guthrie 1972; Nissly et al. 1981; Metz et al. 1985). Both PI parents also display strongly determinate
stem morphology (Metz et al. 1985). Corsoy, on the other hand, is classified as MG II, has indeterminate stem morphology, and is photoperiod sensitive.

The populations were advanced separately by a modified single seed descent. Between 10 and 20 F\textsubscript{1} seeds were made for each cross. F\textsubscript{1} plants were bulk harvested to obtain F\textsubscript{2} seed. The F\textsubscript{2} seed bulk was divided into a portion to reserve in cold storage and a portion to plant in bulk. The populations were advanced to the F\textsubscript{6} generation by pod bulking (Fehr 1987a, b). A two-to-three-seeded pod was harvested from each plant in the F\textsubscript{2} through F\textsubscript{5} generations, and the seeds were bulked. At each generation, the seed bulk was divided into a portion to reserve in cold storage and a portion to plant in bulk. Between 120 and 140 randomly selected F\textsubscript{6} plants from each population were harvested individually. The F\textsubscript{6-7} seeds from these plants were grown for evaluation in growth chamber. A total of 101 lines from population IX132 and 100 lines from population IX136 were evaluated for days to first flower (days to R1).

**Growth chamber observation**

The parents and progeny of each cross were grown in growth chamber (GC) under two different day lengths, 12 h and 20 h using cool white lamps and 40-W incandescent lamps that provided a total photosynthetic photon flux of 230 \( \mu \text{mol m}^{-2} \text{s}^{-1} \). A one-gallon pot containing soil (2 loam : 1 peat : 1 sand) was used. Three seeds were planted in each pot then thinned to one plant at the two open leaf stage. Experiments were arranged in a randomized complete block design. Each line consisted of three pots and was replicated twice. Plants were fertilized with N-P-K (20-4.7-16.6) once a week applied together with irrigation water. Each population was grown in one chamber. Temperature was set at 27°C
during the day and 18°C at night. The R1 stage (Fehr and Caviness 1977) was observed as
the number of days after emergence when an open flower at one of the top nodes with a fully
expended leaf appeared on a plant. Flowering time (days to R1) was recorded separately for
each day length treatment. The difference in flowering times between 12 h and 20 h day
length was used to determine the number of days flowering time was delayed by treatment
of a 20 h photoperiod. These data together with that of the 20 h photoperiod were used to
predict the number of genes controlling photoperiod insensitivity.

Data analysis

The flowering delay for each R1 line due to long day treatment was determined by
subtracting the number of days to R1 when grown in growth chamber with 20 h photoperiod
(gc20), from the number of days to R1 when grown in a growth chamber with 12 h
photoperiod (gc12). The population means and standard deviations for each day length
treatment were calculated for each population. Normality tests of flowering time distribution
frequencies were conducted using the W-test (Shapiro and Wilk 1965). The t-tests (Steel and
Torrè 1980) were carried out to test if the transgressive segregants observed were
significantly different from the low or high performant parents. We classified the phenotypes
based on 'natural' break points. 'Natural' break points are those that separate the extreme
and the moderate phenotypes. For example, 'natural' break points of days to R1 in gc20 are
on days 35 and 90 (Figure 1 b). Using natural break points the phenotypes can be divided
into three classes (Tables 2). A Chi-square test was then performed to test the goodness of
fit of each proposed segregation ratio.
Results and Discussion

The distribution frequencies for days to R1 in gc20 and days that flowering was delayed are shown in Figure 1. Flowering delay of each R1 line was determined by subtracting the number of days to R1 when grown in the gc20, from the number of days to R1 when grown in gc12. Normality tests (Wilk and Shapiro 1965) show that a normal distribution was observed for R1 lines grown in gc12. A non-normal (discontinuous) distribution, however, was observed for the lines grown in gc20 and number of days flowering time was delayed due to long day treatment (Table 1 and Figure 1).

The lines grown in gc20 clearly group into distinct phenotypic classes. For example, in population IX132 there are no plants observed in days 35 and 90 (Figure 1a). These are 'natural' break points that separate the extreme sensitive and insensitive phenotypes from those of moderate insensitivity. Break points were also observed for number of days to R1 delayed (Figure 1b), with no plants observed in days 15 and 65, thus again separating the three phenotypes. Natural break points are also observed in population IX136 (Figures 1c and 1d).

The number of genes that control photoperiod insensitivity was predicted using the segregation data for days to R1 from R1 lines grown in gc20 and the data for days to R1 delayed due to long day treatment. Chi-square tests of the segregation data show that photoperiod insensitivity fits very well to a 1:6:1 segregation ratio (P=0.84) in population IX132 based on the gc20 data. In this population a 1:6:1 segregation ratio (P=0.91) based on the flowering delayed data (Table 2) is also acceptable. In population IX136, photoperiod insensitivity also fits a 1:6:1 segregation ratio (P=0.15) based on gc20 data as well as a 1:6:1 segregation ratio (P=0.07) based on the flowering delayed data (Table 2). A Chi-square test
rejected the two-gene model segregation ratio of 1:2:1 for population IX132 (Table 2) for both gc20 data and data for flowering delay. A Chi-square test also rejected the two-gene model segregation ratio of 1:2:1 for the gc20 and flowering delay data for population IX136. These data suggest that photoperiod insensitivity is controlled by a minimum of three genes in populations IX132 and IX136.

Results of our experiment reject the hypothesis that insensitivity to long day length is controlled by two genes. The main reason may be due to the differences in light quality employed in both experiments. Cober et al. (1996b) reported that soybean maturity loci response differently on changes in light quality based on the ratio of red to far-red light. Our experiment used a combination of fluorescent and incandescent light and the lines were grown in growth chamber. Previous studies (Saindon et al. 1989; Cober et al. 1996a) used ILD and the lines were grown in field conditions. Data of our study suggest that an addition of fluorescent light on the long day treatment enabled us to detect another locus controlling photoperiod insensitivity. Only two loci were detected using the combination of natural and incandescent light (Saindon et al. 1989; Cober et al. 1996a). An addition of fluorescent light on the long day treatment, therefore, may allow detection of more maturity loci that control photoperiod insensitivity in soybean. Therefore, based on this study, we find it plausible to accept that three or more genes control insensitivity of soybean to long day length, thus making this an acceptable target for QTL analyses.

Acknowledgements

Joint publication of the USDA-ARS-Corn Insect and Crop Genetics Research Unit and Journal Paper no. J 18489 of the Iowa Agriculture and Home Economics Experiment
Station, Ames, Iowa (project no. 3236). We thank the government of Indonesia for financial support for IMT.

References


Table 1. Time of flowering of the parents and the F_{6.7} RI lines of both populations grown in growth chamber with 12 h (GC12) and 20 h (GC20) photoperiods

<table>
<thead>
<tr>
<th>Trait</th>
<th>Parents</th>
<th>Population</th>
<th></th>
<th></th>
<th></th>
<th>Normality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P.I.s&lt;sup&gt;a&lt;/sup&gt;</td>
<td>'Corsoy'</td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Min/Max values&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GC12</td>
<td>19.8</td>
<td>23.4</td>
<td>1.90</td>
<td>19.1/27.5</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>GC20</td>
<td>26.1</td>
<td>58.1</td>
<td>20.21</td>
<td>24.3/95.0**</td>
<td>NN</td>
</tr>
<tr>
<td></td>
<td>GC20-GC12</td>
<td>6.3</td>
<td>34.7</td>
<td>19.13</td>
<td>0.9/70.0**</td>
<td>NN</td>
</tr>
</tbody>
</table>

Population IX132

|            | GC12    | 19.9       | 24.7        | 2.80          | 20.2/32.2     | N         |
|            | GC20    | 26.3       | 59.3        | 20.91         | 27.0/95.0**   | NN        |
|            | GC20-GC12 | 6.4        | 35.0        | 19.50         | 0.0/71.0**    | NN        |

Population IX136

<sup>a</sup>PI317.336 for population IX132 and PI317.334B for population IX136

<sup>b</sup>Min = minimum. Max = maximum values for each respective trait

<sup>c</sup>Normality tests were based on the W-test method (Wilk and Shapiro 1965): The null hypothesis of the test is that data distribute normally: N = fail to reject the null hypothesis; NN = reject the null hypothesis.

**Significantly different from high parent phenotype at P=0.01 based on t-tests (Steel and Torrie 1980).
Table 2. Chi-square analyses of segregation ratios for flowering time and number of days that flowering was delayed among RI lines from populations IX132 and IX136. Phenotypic classes were assigned based upon 'natural' break-points in the distribution frequencies within each population. Two-gene and three-gene models were tested.

<table>
<thead>
<tr>
<th>Flowering time (days to RI)</th>
<th>Total progeny tested</th>
<th>Ratio&lt;sup&gt;b&lt;/sup&gt;</th>
<th>$X^2$&lt;sup&gt;c&lt;/sup&gt;</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (Ins.)</td>
<td>Intermediate (Sens.)</td>
<td>Late</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-38&lt;sup&gt;d&lt;/sup&gt;</td>
<td>39-80&lt;sup&gt;d&lt;/sup&gt;</td>
<td>&gt;80&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-15&lt;sup&gt;e&lt;/sup&gt;</td>
<td>16-55&lt;sup&gt;e&lt;/sup&gt;</td>
<td>&gt;55&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Population IX132**

**Based on gc20 data**

| Observed | 14 | 76 | 11 | 101 | 1:6:1 | 0.36 | 0.84 |
| Expected | 12.6 | 75.8 | 12.6 | 101 |
| Observed | 14 | 76 | 11 | 101 | 1:2:1 | 26.14 | <0.001 |
| Expected | 25.3 | 50.4 | 25.3 |

**Based on flowering delayed data**

| Observed | 14 | 75 | 12 | 101 | 1:6:1 | 0.20 | 0.91 |
| Expected | 12.6 | 75.6 | 12.6 |
| Observed | 14 | 75 | 12 | 101 | 1:2:1 | 24.10 | <0.001 |
| Expected | 25.3 | 50.4 | 25.3 |
Table 2. continued

<table>
<thead>
<tr>
<th>Population IX136</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Based on gc20 data</strong></td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td><strong>Based on flowering delayed data</strong></td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
<tr>
<td>Observed</td>
</tr>
<tr>
<td>Expected</td>
</tr>
</tbody>
</table>

*Ins=Insensitive to long day length; Sens=Sensitive to long day length; Interm.=Intermediate phenotype.

b Eight (three-gene model) and four (two-gene model) genotypic classes were tested in both populations.

c The null hypothesis of the test is that the progeny segregate in the ratios tested.

d The range values of days to RI of the RI lines grown in gc20 accepted for each phenotypic class.

e The range values of flowering delay accepted for each phenotypic class.
Figure 1. Distribution frequency of flowering time (days to R1) of the F<sub>6:7</sub> RI lines in population IX132 and IX136 when grown in growth chamber with 20 h photoperiod (a and c) and number of days flowering time was delayed due to 20 h photoperiod (b and d). * = The upper class limit.
CHAPTER 3. MAPPING GENETIC LOCI FOR FLOWERING TIME, MATURITY, AND PHOTOPERIOD INSENSITIVITY IN SOYBEAN

A paper accepted by Molecular Breeding

I. M. Tasma\textsuperscript{2}, L. L. Lorenzen\textsuperscript{3}, D. E. Green\textsuperscript{2} & R. C. Shoemaker\textsuperscript{3,4}

Abstract

Time of flowering and maturity in soybean \textit{[Glycine max (L.) Merr]} are important reproductive characters of agronomic interest. These traits are useful for developing soybean cultivars with a wider geographical adaptation. The objective of this research was to use molecular markers to identify chromosomal regions that control traits for flowering time, maturity and photoperiod insensitivity in soybean. Two single-cross populations, IX132 (PI 317.336 X 'Corsoy') consisting of 101 progeny, and IX136 (PI 317.334B X 'Corsoy') consisting of 100 progeny were used. Days to R1 (the day when 50\% of the plants in a plot have an open flower at one of the top nodes with a fully expanded leaf) was observed among

\textsuperscript{1}Reprinted with permission of Journal Molecular Breeding, 2001. USA

\textsuperscript{2}Department of Agronomy. Iowa State University. Ames, Iowa 50011. USA

\textsuperscript{3}Office of Biotechnology. Iowa State University. Ames, Iowa 50011. USA

\textsuperscript{4}USDA-ARS, Corn Insects and Crop Genetics Research Group. Iowa State University. Ames, Iowa 50011. USA: Author for correspondence
F_{6.7} RI lines in the field during 1991 and 1992 and in the growth chamber at 12 h and 20 h photoperiods using fluorescent and incandescent lamps. Days to R3 (the number of days after emergence when 50% of the plants in a plot had presented the first 5 mm pod at one of the top four nodes with a fully expanded leaf) was observed in the field during 1991 and in the growth chamber with 12 h photoperiod. Days to R7 (the number of days after emergence when 50% of pods in a plot had mature pod color) was observed in the field in 1991. A total of 139 markers (88 RFLPs and 51 SSRs) in the IX132 population and 125 markers (73 RFLPs and 52 SSRs) in the IX136 population were used to map quantitative trait loci (QTL) affecting these traits. Results show that a large-effect QTL for days to R1, R3, and R7, and photoperiod insensitivity was found at the same location on linkage group (LG) C2 in both populations. This result suggests that photoperiod insensitivity, flowering time, and maturity may be controlled by the same gene(s) or by tightly clustered genes in the same chromosomal region. In addition to the large effect QTL, minor QTL were also detected controlling the four traits in both populations. Minor QTL account for as much as 17.8% and 12.1% of phenotypic variance in populations IX132 and IX136, respectively. Thus, time of flowering, maturity, and photoperiod insensitivity in these soybean populations are proposed to be controlled by a major QTL with a large effect and modified by several minor QTL.

**Introduction**

A consensus soybean genetic map has been constructed [12]. The map is composed primarily of RFLP and SSR markers in addition to a few RAPD, and classical genetic markers. Many QTL studies for agronomically important traits have been reported. These include traits for reproduction and morphology [24, 32, 27], hard seededness [23], seed
protein and oil content [15, 6], plant height and lodging [27, 32, 24], iron deficiency chlorosis [29], yield [37], and resistance to sudden death syndrome [21]; soybean cyst nematode [11]; and brown stem rot [28].

Time of flowering and maturity in soybean [Glycine max (L.) Merr] are important reproductive characters of agronomic interest. These traits are useful for developing soybean cultivars with a wider geographical adaptation. Expression of the traits is known to be a function of daylength, temperature, and plant genotype [49]. A soybean cultivar's narrow area of adaptation is due mainly to differences in daylength perception that affects the length of time required for reproductive periods [9].

Soybean genotypes that are day-neutral (photoperiod insensitive) have been reported [19, 36, 35]. These include PI 317.336 ('Sinshei') and PI 317.334B ('Kitami-Shiro'), two genotypes that were introduced from Japan. Use of these genotypes may help to facilitate a better understanding of the genetic mechanism controlling flowering time and maturity in soybean.

Conventional genetic studies indicate that five genes influence flowering time and maturity in soybean [9]. The genes are known as the E-series: E1, E2, E3, E4, and E5 [4, 7, 8, 33]. McBlain et al. [34] reported that these loci interact with photoperiod in the control of flowering. Under natural day length, the dominant alleles tend to delay flowering time and maturity but the magnitude of effect of each gene can be different [9].

Several QTL associated with flowering time and maturity have been previously mapped in soybean. Keim et al. [24] reported that five markers on linkage groups C1, C2, and D1 [43] were found to be associated with maturity (including time of flowering) in an F2 population derived from a cross between G. max and G. soja. The observed QTL explained
17-23% of total phenotypic variance. Using another F₂ population of ‘Minsoy’ X ‘Noir 1’, Mansur et al. [32] also reported a major QTL for flowering time on LG C2 and minor QTL for maturity on LG L and M [12]. Lee et al. [27] reported that QTL for maturity traits (including time of flowering) were found on LG K. The study was based on an F₂ population derived from a cross between PI 97100 and ‘Coker 237’. The observed QTL explained 26.2-31.2% of phenotypic variance. These results suggested that the putative QTL for maturity traits (including those for flowering time) may be population-specific as indicated by different genomic regions (LG) that control the same trait(s) in different mapping populations. No QTL studies controlling photoperiod insensitivity have been reported.

Studies from other crop species such as peas, rice, cotton, wheat, and tobacco show that two genetic mechanisms can control flowering time: a single gene mechanism [2] and a multigene mechanism [18]. Yano et al. [50] reported that two major- and three minor-effect QTL controlled flowering time among progeny of a cross between cultivated rice Indica and Javanica. Furthermore, Lin et al. [30] found five minor QTL controlling flowering time among BC₁F₃ lines from the rice recurrent parent ‘Nipponbare’ and the donor parent ‘Kasalath’.

The objective of this study was to map the chromosomal regions that control flowering time (days to R1) [17] and maturity (days to R3 and days to R7) [17], and photoperiod insensitivity in soybean. Flowering time (days to R1) and maturity (days to R8) in soybean have been reported to be highly correlated traits [32, 37]. No evidence, however, has yet been presented to support that they are also correlated with photoperiod insensitivity. They may all be controlled by the same gene or cluster of genes. By analyzing two populations, each segregating for all three of these traits, we hope to correlate directly the
locations of QTL for each trait.

Materials and methods

Plant materials

Two single-cross populations were used in this study, IX132 (PI 317.336 X ‘Corsoy’) and IX136 (PI 317.334B X ‘Corsoy’). Each of the P.I.s is known to be photoperiod insensitive [19, 36, 35]. The populations were advanced to the $F_6$ generation using a modified single seed descent method. The development of these RI lines has been described previously [47]. The $F_{6:7}$ seeds from each population were grown for evaluation in field and growth chamber. A total of 101 lines from population IX132 and 100 lines from population IX136 were used in this study.

Field observation

The trait data of the parents and progeny from each single-cross population were recorded in the field in 1991 and 1992. The experiments were conducted at the Bruner Farm, Iowa State University Agronomy Field Research Station, Ames, IA. The $F_{6:7}$ lines were grown in plots 1.5 meters long planted with 50 seeds. Plots were separated by 1.2 meters (within rows) with 1-meter spacing between rows. Each entry was replicated twice in a randomized complete block design. A border row was planted along each side of a block. The R1 stage [17] was observed as the number of days after emergence when 50% of the plants in a plot had an open flower at one of the top nodes with a fully expanded leaf. In addition to the R1 data, the days to the stages R3, and R7 were recorded in the 1991
experiment to test if these traits are controlled by the same or different loci in the soybean genome. The R3 stage [17] was observed as the number of days after emergence when 50% of the plants in a plot had presented the first 5-mm-long pod at one of the top four nodes with a fully expanded leaf. The R7 stage was observed as number of days after emergence when 50% of pods in a plot had mature pod color [17].

*Growth chamber observation*

The parents and RI lines of each cross were grown in a growth chamber (GC) in 1992 under two different day lengths, 12 h and 20 h, using fluorescent and incandescent lights. The experiments were arranged in a randomized complete block design. Three seeds were planted in each pot then thinned to one plant at the two-open leaves stage. Each line consisted of three pots and was replicated twice. Flowering time (days to R1) was observed as the number of days after emergence when an open flower at one of the top nodes with a fully expended leaf appeared on a plant [17]. The difference in flowering time between treatment (12 h and 20 h) daylengths was used to determine the number of days flowering time was delayed by the larger photoperiod treatment, 20 h. These data, together with that of the 20 h photoperiod, were used to detect QTL for photoperiod insensitivity. Days to R3 was observed as in the field study.

*Pubescent color observation*

Pubescence color was scored in plantings conducted in Puerto Rico during Fall of 2000 under the natural conditions of the island. Twenty seeds of each genotype were planted in 2 ft. row plots, at a seeding rate of 10 seeds per ft., leaving a 1 ft. space between genotypes
planted in the same row. Adjacent rows were 24 inches apart. For each genotype, seedlings emerged approximately four days after planting. Standard practices used at the research site in Puerto Rico (weeding, irrigation, and application of herbicides, insecticides, and fungicides) were used in this planting.

Pubescence color of each genotype was visually observed and recorded during the growing season when plants had reached vegetative stage V3 [17], and confirmed by a second scoring done at maturity at stage R7 [17]. No segregation for pubescence color was observed among individual plants of each of the genotypes.

Chi square tests were then conducted for each segregating population with the null hypothesis that the progeny segregate in a 1:1 ratio.

**Molecular data**

DNA was extracted from leaves of the parents and the RI lines as described previously [22]. To identify the RFLP probe-enzyme combinations revealing polymorphism between the parents, the parental DNAs (PI317.336, PI 317.334B, and ‘Corsoy’) were digested with each of five restriction enzymes (DraI, EcoRI, EcoRV, HindIII, and TaqI) according to the manufacturer (Gibco BRL, Gaithersburg, Maryland). Digested DNA was then electrophoresed in 0.8 % Ultrapure agarose gel (Gibco BRL, Gaithersburg, Maryland) in 1X TAE buffer [39] at 22-30 V for 16-20 h. The DNAs were then transferred [44] onto Zeta- Probe* GT nylon membranes (Bio Rad, California), probed, and autoradiographed as described previously [23]. Two hundred ninety-two RFLP probes were tested. The probes were chosen so that markers were no more than 20 cM apart [12]. This marker density, in combination with the use of interval QTL mapping, should allow the detection of major QTL
anywhere in the genome [46]. The polymorphic probes were then hybridized to membranes containing digested DNA of the RI lines. We compared RFLP patterns to those of 
*Glycine max* breeding line A81-356022 and *Glycine soja* PI 468.916, parents used to develop the USDA/ISU soybean public map [42], to identify 'anchored' RFLP markers.

SSRs were selected to fill in the gaps for chromosomal regions not covered with RFLPs. SSR primers were synthesized by the Iowa State University Nucleic Acid Facility using sequences obtained from the Soybean database (SoyBase) or were a gift of Dr. Perry Cregan (USDA/ARS, Maryland). PCR reactions were as described previously [1], except that the cycling time was 30 seconds each for DNA denaturation (94°C), primer annealing (47°C), and DNA extension (68°C), and the reactions were carried out for 45 cycles in a 10-μl reaction. A 60-ng DNA template was used in the reaction instead of 30 ng. PCR-cycling was performed in a MJ PTC-100 Programmable Thermal Controller. All PCR products were then electrophoresed in a 3-4 % Ultrapure agarose (Gibco BRL, Gaithersburg, Maryland) gels containing ethidium bromide. The percentage of the agarose gel depended on the mobilities of the two alleles. The gels were run in 1X TAE buffer [39] at 150 V for 2-3 h. A 20-base-pair ladder was used as a DNA marker. The gels were then photographed with an AlphaImager 2000. One hundred and twenty SSR primers were chosen based on their distribution in the ISU/USDA public map. The polymorphic SSRs were then used to amplify DNA from the lines in each mapping population.

*Linkage mapping*

Placement of markers was done using the program Mapmaker [25]. For grouping the markers, a minimum LOD score of 3.0 and a maximum distance of 40 cM were used as a
threshold value to declare linkage in the pair-wise loci analysis. The linkage map was constructed using the Haldane map function [20]. The gene orders were assigned using the ‘compare’, ‘try’, and ‘ripple’ (minimum LOD score of 3.0) commands.

**QTL analysis and mapping**

Two steps of QTL analyses were conducted. First, a single factor analysis of variance (ANOVA) was carried out for all pair-wise combinations of marker loci and quantitative traits [16]. QTL analyses were performed on the line mean data for single years and combined across years of field observations and for different photoperiod treatments of growth chamber observations. A $P \leq 0.01$ level of significance was used as evidence that there was linkage between a marker locus and a QTL. The data were then analyzed using the Mapmaker-QTL interval mapping program [31]. This program was found to be appropriate for the analysis of RI line data in soybean and other crop species although it was not designed for RI lines [11, 21, 26, 48]. QTL analyses also were performed on the line mean data of single years and combined across years of field observations and for different photoperiod treatments of growth chamber observations. A LOD score of 2.5 was used as a minimum to declare the presence of a QTL in a particular genomic region. The LOD score peaks were used to estimate the most likely position of QTL on the linkage map. The amount of variation explained by a marker was determined using the coefficient of determination ($R^2$) value (SAS-GLM) [40]. In this study, a QTL that explains more than 20% of total phenotypic variance is considered one that has a large effect.
Results

Trait data

Table 1 presents trait data for the parents and RI lines of both populations in field and growth chamber conditions. Normality tests of trait distribution frequencies based on the W-test method [41] show that almost all trait and environment combinations show normal distribution (Table 1). In both populations, only distribution frequencies for trait data in growth chamber with 20h photoperiod (GC20) and those of flowering delay data (GC20-GC12) did not follow normal distribution (Table 1). Using t-tests [45], it was shown that none of the minimum transgressive segregants of either population was significantly different from the low parent phenotypes (Table 1). Most of the maximum transgressive segregants, on the other hand, were significantly different from the high parent phenotypes (Table 1).

Marker segregation and map construction

A total of 292 RFLP probes and 120 SSR primer pairs were chosen from the soybean linkage map [12]. Of the 292 RFLP markers tested, 91 (31.2%) were found to be polymorphic with one or more restriction enzymes in population IX132 and 85 (29.1%) were polymorphic with one or more restriction enzymes in population IX136. Of the 120 SSR primers tested, 58 (48.3%) were polymorphic in population IX132 and 60 (50.0%) were polymorphic in population IX136.

Of the 139 markers segregating in population IX132, 15 markers did not follow the expected 1:1 segregation ratio (P ≥ 0.05). Of these, 10 were skewed toward the ‘Corsoy’ parent and five were skewed toward the PI 317.336 parent. Of the 125 segregating markers
in population 1X136, 11 markers did not follow the expected 1:1 segregation ratio. Of these, six were skewed toward the 'Corsoy' allele and five were skewed toward the PI 317.334B allele.

The segregating markers were distributed across all 20 linkage groups (LGs) in both populations and covered 986.1 cM in population 1X132 with an average distance between adjacent markers of 11.9 cM, and covered 696.4 cM in population 1X136 with the average distance between adjacent marker loci of 9.5 cM. This map coverage does not include the unlinked anchored-markers. Three LGs (B1, F, and N) in population 1X132 and four (A1, B1, B2, and N) in population 1X136 contained only one marker. Seven markers (A136V, A186T, A124H, A176V, K011T, K003H, and L1851) and nine markers (A059I, A333H, A176V, A725V, A132T, L1851, A064D, Bng132V, and A638I) in population 1X132 and 1X136, respectively, remained unlinked. In both populations, the 't' locus mapped closely to Satt205 on LG C2 (Figures 1a and 1b).

**QTL for days to R1**

A large effect QTL was detected by the same markers on LG C2 in both populations (Figures 1a and 1b). This QTL accounts for as much as 40.7% of total phenotypic variance in population 1X132 and as much as 47.4% in population 1X136, depending on the environment (Tables 2 and 3). Minor-effect QTL were also observed in both populations (Figures 1a and 1b). In population 1X132, the minor effect QTL for days to R1 were on LGs G and J (Figure 1a). In population 1X136, the minor effect QTL was on LG G (Figure 1b). These QTL individually account for as much as 15.07% of total phenotypic variance in population 1X132 and 11.4% in population 1X136, depending on the environment (Tables 2
The allelic effect for days to R1 for the loci corresponding to the peak in the QTL intervals (Tables 2 and 3) was determined based on mean phenotypic differences between the two genotypic classes (see the legends of Tables 2 and 3 for details). The allele effects with field data were as high as 16.3 days in population IX132 (Table 2) and 15.1 days in population IX136 (Table 3). In both field and growth chamber environments, the alleles derived from the PI parents in both populations were found to increase flowering time compared to those derived from the 'Corsoy' parent (Tables 2 and 3).

**QTL for days to R3**

A major effect QTL for days to R3 was also observed on LG C2 in both populations. The QTL accounts for as much as 36.5% total phenotypic variance in population IX132, and 44.4% of total phenotypic variance in population IX136 and was detected with the same markers as with the trait days to R1. Minor effect QTL were also observed in both populations. In population IX132, the minor effect QTL are on LGs A2 and J (Figure 1a and Table 2). In population IX136, the minor effect QTL are on LGs L and Satt276 (Figure 1b and Table 3). These QTL account for 6.9-17.8% and 5.1-12.1% of total phenotypic variances in populations IX132 and IX136, respectively (Tables 2 and 3). The allele effects of the minor QTL for this trait ranged from 1.9 to 6.8 days in IX132 and from 1.3 to 6.9 days in IX136 (Tables 2 and 3).

**QTL for days to R7**

A large effect QTL for days to R7 was also observed on LG C2 in both populations. These QTL account for 24.95% of total phenotypic variance in population IX132 and 21.7%
phenotypic variance in population IX136. In population IX132, the minor effect QTL are on LGs A2 and J (Figure 1a). In population IX136, a minor effect QTL is on LG A1 (Table 3). In populations IX132 and IX136, the minor QTL individually account for a maximum of 10.6% and 9.8% of total phenotypic variance, respectively. The allele effects for trait days to R7 range from 6.6 to 10.5 days in IX132 and from 6.4 to 7.0 days in IX136 (Tables 2 and 3).

QTL for photoperiod insensitivity

This trait was scored as days to R1 in gc20 minus days to R1 in gc12 (Tables 2 and 3, and Figures 1a and 1b). As with the other traits, a large effect QTL was observed on LG C2 in both populations. The QTL accounts for 22.3% total phenotypic variance in population IX132, and 20.8% of total phenotypic variance in population IX136. A minor QTL was detected in population IX136 only. The minor QTL is on LG G (Table 3, and Figure 1b). The allele effects for this trait is 17.7 days in IX132 and have ranges from 6.6 to 18.3 days in IX136 (Tables 2 and 3). Photoperiod insensitivity increases with alleles from the 'Corsoy' parent (Tables 2 and 3).

Discussion

In this study, we mapped QTL for four major reproductive, maturity, and photoperiod insensitivity traits. This is the first time these traits have been mapped concurrently in independent populations. We determined that the major QTL controlling all four of these traits maps to the same location on LG C2.

The consistency of the location of the major QTL controlling the traits days to R1, R3, R7, and photoperiod insensitivity on LG C2 strongly suggest that all four traits may be
controlled by the same gene or cluster of genes. The LG C2 location of the major QTL
detected in this study is consistent with the results reported by Keim et al. [24] and Mansur et
al. [32] for the number of days to R1 and R7. In addition to the major QTL detected on LG
C2, Keim et al. [24] also resolved QTL for R8 on LG C1 (associated with marker K472) and
on LG D1 (associated with marker R013-2) [43]. In another study, Lee et al. [27] identified
two independent QTL for days to R8 on LG K. Our study did not find QTL on LGs C1, D1,
or K. The lack of detection of a major QTL for maturity traits on LG C2 by Lee et al. [27]
could be due to the lack of segregation of maturity alleles at this locus in their population.
Their study involved southern soybean germplasm whereas the other studies involved
northern germplasm. These gene pools represent broadly different maturity groups. Also,
because of gaps in our maps, some major and minor QTL might remain undetected.

Minor QTL controlling traits days to R1, R3, R7, and photoperiod insensitivity.
however, were observed on LGs A2, G, and J in population 1X132 and minor QTL for days
to R1, R3, R7, and photoperiod insensitivity were observed on LGs G, L, and A1 in
population 1X136.

QTL for the different traits were found within the same chromosomal intervals across
environments. For example, on LG G of population 1X132, QTL for days to R1 in two
environments (field 1991 and combined data of field 1991/1992) are located in
approximately the same region as the QTL for photoperiod insensitivity (based on gc20 data)
(Figure 1a). A similar condition was observed on LGs J and A2. This was also observed on
LG G of population 1X136 (Figure 1b). This observation is not surprising given that these
traits are known to be highly correlated [32, 37]. The traits may all be controlled by the
same gene, or by a cluster of genes. Another possible explanation is that a single QTL
present in each region has a pleiotropic effect on all observed traits in our populations.

We mapped the pubescent color locus, 't', that is known to be linked with maturity locus *El* on LG C2 [12]. The 't' locus mapped closely to Satt205 in both populations (Figures 1a and 1b) and Satt205 is tightly linked with the major QTL on LG C2 in both populations. The obvious interpretation is that the major QTL on LG C2 may be *El*. Cober et al. [10], however, reported that another maturity locus may also link to the pubescent color locus. This was based on the observation that a tawney pubescent parent of early maturing soybean germplasm (*elel*) produced progeny in which maturity was not associated with pubescent color. Given these results, we can not rule out the possibility that although *El* may reside in close proximity to the major QTL on LG C2, many of the effects of this QTL may be the result of actions of a second maturity gene.

Previously, other maturity genes *E3* and *E2* were mapped [12] to LG L and LG O, respectively. None of the QTL we identified corresponded to either of these positions. However, simulations have shown that not all QTL will be identified in small populations (100 to 200 F2 progeny) [3]. Therefore, QTL may exist at these locations but we failed to identify them. The modeling studies also demonstrated that small populations will result in overestimating the effect of a QTL [3]; thus, the distinction in this study between a large effect QTL and small effect QTL may be biased.

We also observed transgressive progeny. Transgressive segregants are commonly observed in soybean [27, 32]. Rick and Smith [38] proposed three possible causes for the occurrence of transgression: *de novo* mutation, complementary actions of genes from the two parents, and unmasking recessive genes due to inbreeding. Complementary gene actions, however, are the most commonly proposed and are mainly based on evidence from the QTL
mapping studies [46]. Transgressive segregation can occur among the progeny of both intra- and interspecific crosses. In the interspecific hybridization of tomato (*Lycopersicon esculentum* *X* *Lycopersicon pennellii*), DeVicente and Tanksley [14] reported that eight out of 11 quantitative traits observed in the F₂ progeny of the cross show transgressive segregations. They found that the presence of the transgressives were directly linked to the presence of complementary QTL alleles in the two parents. Thirty-six percent of the detected QTL were reported to have allelic effects opposite of those predicted by the parents. A similar mechanism may explain the observed transgressive segregants observed in our experiment. The populations used in this study are crosses between Plant Introductions and the cultivar ‘Corsoy’. A complementary-gene action between alleles of the two parents may be involved in the initiation of the transgressive progeny. Since our populations are RI lines, the gene actions involving overdominant and dominant effects cannot be observed. The transgression observed in this study may be due to the complementation of dominant loci between the two parents. This may explain why almost all the transgressive progenies observed from this study have late-flowering phenotypes. In addition, lateness is associated with dominant alleles of the *E* loci, both under natural daylength or daylength extended to 20 h with incandescent or fluorescent light. The *E₁* allele delays flowering and maturity the most, followed by *E₂*, *E₃*, and *E₄* [9], compared to their respective recessive alleles. *E₅* delays flowering and maturity similar to *E₂* [33]. Both P.I.s (Japanese cultivars) have *E₁*. 'Corsoy' has *e₁* [5], and all three genotypes may also have other dominant *E* genes. Progeny of these crosses may have different combinations of dominant *E* genes and act additively (i.e. more dominant *E* loci show later flowering and maturity) resulting in transgressive progeny toward the late phenotypes.
Acknowledgements

Joint publication of the USDA-ARS-Corn Insect and Crop Genetics Research Unit and Journal Paper no. J19263 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa (project no. 3236). Names are necessary to report factually on the available data; however, the USDA neither guarantees nor warrants the standard of the product to the exclusion of others that may also be available. We thank Dr. Perry Cregan for the gift of SSR primers, Dr. Arla Bush for the initial assays of SSR markers, Dr. Kimberly Lewers for the help of initial data analysis. Dr. Silvia Cianzio for collecting phenotypic data and the government of Indonesia for financial support of IMT.

References


genomes and eight quantitative trait loci controlling flowering time in a winter x spring barley

27. Lee SH, Bailey MA, Mian MAR, Shipe ER, Ashley DA, Parrott WA, Hussey
RS, and Boerma HR: Identification of quantitative trait loci for plant height, lodging, and
maturity in a soybean population segregating for growth habit. Theor Appl Genet 92: 516-

28. Lewers KS, Crane EH, Bronson CR, Schupp JM, Keim P, and Shoemaker RC:
Detection of linked QTL for soybean brown stem rot resistance in 'BSR 101' as expressed in


30. Lin SY, Sasaki T, and Yano M: Mapping quantitative trait loci controlling seed
dormancy and heading date in rice, Oryza sativa L., using backcross inbred lines. Theor


32. Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, and Lark KG: Genetic mapping of
agronomic traits using recombinant inbred lines of soybean. Crop Sci. 36: 1327-1336
(1996). 33. McBlain BA, and Bernard RL: A new gene affecting the time of flowering and

34. McBlain BA, Bernard RL, Cremeens CR, and Korczak JF: A procedure to identify


45. Steel RGD, and Torrie JH: Principles and procedures of statistics: A biometrical


Table 1. Phenotypic performance of the parents and the F₆.ₑ RI lines of populations IX132 and IX136 grown in field and growth chamber conditions

<table>
<thead>
<tr>
<th>Trait</th>
<th>Parents</th>
<th>Population</th>
<th>Trait</th>
<th>Parents</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P.I.s  b 'Corsoy'</td>
<td>Mean</td>
<td>Standard deviation</td>
<td>Min/Max values c</td>
<td>Normality tests d</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R1 (F91)</td>
<td>42.7</td>
<td>32.5</td>
<td>51.1</td>
<td>11.05</td>
<td>29.0/68.5**</td>
</tr>
<tr>
<td>R1 (F92)</td>
<td>45.2</td>
<td>33.1</td>
<td>51.6</td>
<td>11.63</td>
<td>33.0/69.5**</td>
</tr>
<tr>
<td>R1 (F91/92)</td>
<td>43.9</td>
<td>32.8</td>
<td>51.4</td>
<td>10.74</td>
<td>31.3/69.8**</td>
</tr>
<tr>
<td>R1 (GC12)</td>
<td>19.8</td>
<td>25.0</td>
<td>23.4</td>
<td>1.90</td>
<td>19.1/27.5</td>
</tr>
<tr>
<td>R1 (GC20)</td>
<td>26.1</td>
<td>57.9</td>
<td>58.1</td>
<td>20.21</td>
<td>24.3/95.0**</td>
</tr>
<tr>
<td>R1 (GC20-GC12)</td>
<td>6.3</td>
<td>32.9</td>
<td>34.7</td>
<td>19.13</td>
<td>0.9/70.0**</td>
</tr>
<tr>
<td>R3 (F91)</td>
<td>58.8</td>
<td>47.5</td>
<td>60.5</td>
<td>10.72</td>
<td>41.5/81.5*</td>
</tr>
<tr>
<td>R3 (GC12)</td>
<td>27.0</td>
<td>36.1</td>
<td>28.2</td>
<td>2.33</td>
<td>23.6/34.7</td>
</tr>
<tr>
<td>R7 (F91)</td>
<td>116.2</td>
<td>101.7</td>
<td>111.1</td>
<td>11.34</td>
<td>84.5/132.5</td>
</tr>
<tr>
<td>R1 (F91)</td>
<td>42.3</td>
<td>32.5</td>
<td>53.1</td>
<td>10.76</td>
<td>30.0/75.0**</td>
</tr>
<tr>
<td>R1 (F92)</td>
<td>50.1</td>
<td>33.1</td>
<td>54.9</td>
<td>11.26</td>
<td>35.5/80.5**</td>
</tr>
<tr>
<td>R1 (F91/92)</td>
<td>46.2</td>
<td>32.8</td>
<td>53.9</td>
<td>10.93</td>
<td>31.8/77.8**</td>
</tr>
<tr>
<td>R1 (GC12)</td>
<td>19.9</td>
<td>25.0</td>
<td>24.7</td>
<td>2.80</td>
<td>20.2/32.2</td>
</tr>
<tr>
<td>R1 (GC20)</td>
<td>26.3</td>
<td>57.9</td>
<td>59.3</td>
<td>20.91</td>
<td>27.0/95.0**</td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th></th>
<th>GC20-GC12</th>
<th>F91</th>
<th>GC12</th>
<th>F91</th>
<th>GC12</th>
<th>GC12</th>
<th>GC12</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1 (GC20-GC12)</td>
<td>6.4</td>
<td>57.9</td>
<td>25.8</td>
<td>25.8</td>
<td>117.5</td>
<td>117.5</td>
<td>117.5</td>
</tr>
<tr>
<td>R3 (F91)</td>
<td>32.9</td>
<td>47.5</td>
<td>36.1</td>
<td>47.5</td>
<td>101.7</td>
<td>101.7</td>
<td>101.7</td>
</tr>
<tr>
<td>R3 (GC12)</td>
<td>35.0</td>
<td>62.3</td>
<td>28.9</td>
<td>62.3</td>
<td>112.7</td>
<td>112.7</td>
<td>112.7</td>
</tr>
<tr>
<td>R7 (F91)</td>
<td>19.50</td>
<td>9.41</td>
<td>2.93</td>
<td>9.41</td>
<td>9.69</td>
<td>9.69</td>
<td>9.69</td>
</tr>
<tr>
<td></td>
<td>0.0/71.0**</td>
<td>43.5/84.5**</td>
<td>23.5/37.7</td>
<td>90.0/135.0</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

\(^a\)R1 = days to R1; R3 = days to R3; R7 = days to R7; F = field; GC = growth chamber

\(^b\)PI317.336 for population IX132 and PI317.334B for population IX136

\(^c\)Min = minimum, Max = maximum values for each respective trait

\(^d\)Normality tests were based on the W-test method [41]; The null hypothesis of the test is that data distribute normally: N = fail to reject the null hypothesis; NN = reject the null hypothesis

*Significantly different from high parent phenotypes at P=0.05 based on t-tests [45]

**Significantly different from high parent phenotypes at P=0.01 based on t-tests [45].
Table 2. Chromosomal intervals significantly associated with variation for days to R1, R3, and R7 of soybean recombinant inbred lines in population IX132 during 1991 and 1992.

<table>
<thead>
<tr>
<th>Chromosomal interval</th>
<th>Linkage group</th>
<th>$R^2$ (%)</th>
<th>LOD score</th>
<th>Allele effect (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satt205-Satt286</td>
<td>C2</td>
<td>36.07</td>
<td>8.86</td>
<td>15.8 (P36)</td>
</tr>
<tr>
<td>A112T-A378H</td>
<td>G</td>
<td>8.28</td>
<td>3.83</td>
<td>2.1 (P36)</td>
</tr>
<tr>
<td>Satt380-RGA1T</td>
<td>J</td>
<td>7.76</td>
<td>3.25</td>
<td>6.1 (P36)</td>
</tr>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(field 1991)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satt205-P029D_2</td>
<td>C2</td>
<td>40.73</td>
<td>10.44</td>
<td>16.3 (P36)</td>
</tr>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(field 1992)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satt205-P029D_2</td>
<td>C2</td>
<td>40.10</td>
<td>10.40</td>
<td>16.1 (P36)</td>
</tr>
<tr>
<td>A112T-A378H</td>
<td>G</td>
<td>7.32</td>
<td>2.47</td>
<td>1.4 (P36)</td>
</tr>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(field 1991 and 1992)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satt205-P029D_2</td>
<td>C2</td>
<td>38.85</td>
<td>9.54</td>
<td>2.9 (P36)</td>
</tr>
<tr>
<td>Satt380-RGA2V</td>
<td>J</td>
<td>15.07</td>
<td>4.04</td>
<td>1.5 (P36)</td>
</tr>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(growth chamber 12 h, gc12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satt205-P029D_2</td>
<td>C2</td>
<td>27.55</td>
<td>6.35</td>
<td>20.6 (P36)</td>
</tr>
<tr>
<td>A112T-A378H</td>
<td>G</td>
<td>5.57</td>
<td>2.67</td>
<td>9.6 (P36)</td>
</tr>
<tr>
<td><strong>Days to R1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(growth chamber 20 h, gc20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. (continued)

<table>
<thead>
<tr>
<th>Flowering delay</th>
<th>(differences between gc20 and gc12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
</tr>
<tr>
<td></td>
<td>22.27</td>
</tr>
<tr>
<td></td>
<td>5.12</td>
</tr>
<tr>
<td></td>
<td>17.7 (P36)</td>
</tr>
</tbody>
</table>

*Days R3* *(field 1991)*

<table>
<thead>
<tr>
<th>Satt205-PO29D_2</th>
<th>C2</th>
<th>27.79</th>
<th>7.58</th>
<th>13.3 (P36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt380-RGA1T</td>
<td>J</td>
<td>10.63</td>
<td>2.93</td>
<td>6.8 (P36)</td>
</tr>
<tr>
<td>Satt378-Satt228</td>
<td>A2</td>
<td>6.89</td>
<td>3.15</td>
<td>5.6 (P36)</td>
</tr>
</tbody>
</table>

*Days to R3* *(gc12)*

<table>
<thead>
<tr>
<th>Satt205-PO29D_2</th>
<th>C2</th>
<th>36.52</th>
<th>8.90</th>
<th>3.4 (P36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGA2V-K005V</td>
<td>J</td>
<td>17.78</td>
<td>5.48</td>
<td>1.9 (Cor)</td>
</tr>
</tbody>
</table>

*Days to R7* *(field 1991)*

<table>
<thead>
<tr>
<th>Satt205-PO29D_2</th>
<th>C2</th>
<th>24.95</th>
<th>4.65</th>
<th>10.5 (P36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt380-RGA1T</td>
<td>J</td>
<td>8.66</td>
<td>3.30</td>
<td>6.6 (P36)</td>
</tr>
<tr>
<td>Satt378-Satt228</td>
<td>A2</td>
<td>10.59</td>
<td>3.40</td>
<td>7.7 (P36)</td>
</tr>
</tbody>
</table>

*a Based on USDA/ISU soybean public map [42]:

*b Coefficient of determination and allele effects of the first marker locus in the interval that is significantly associated with the trait locus (p ≤ 0.01), based on a single-factor ANOVA of SAS-GLM;

*c Based on Mapmaker-QTL analyses of the indicated chromosomal interval:

*d Mean differences of the respective traits between the two genotypic classes carrying PI317.336 (P36) and Corsoy (Cor) alleles of the first marker in each the interval.

Parentheses by each number is the parent whose marker increases the value of the trait.
Table 3. Chromosomal intervals significantly associated with variation for days to R1, R3, and R7 of soybean recombinant inbred lines in population IX136 during 1991 and 1992

<table>
<thead>
<tr>
<th>Chromosomal interval</th>
<th>Linkage(^a) group</th>
<th>R(^{2b}) (%)</th>
<th>LOD(^c) score</th>
<th>Allele effect(^{bd}) (days)</th>
</tr>
</thead>
</table>

**Days to R1**  
(*field 1991*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
<td>36.56</td>
<td>9.17</td>
<td>13.4 (P34B)</td>
</tr>
<tr>
<td>Satt276-</td>
<td>A1</td>
<td>11.46</td>
<td>4.12</td>
<td>7.6 (P34B)</td>
</tr>
</tbody>
</table>

**Days to R1**  
(*field 1992*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
<td>42.94</td>
<td>11.01</td>
<td>15.1 (P34B)</td>
</tr>
<tr>
<td>A073H-K443H</td>
<td>G</td>
<td>9.16</td>
<td>4.01</td>
<td>6.6 (P34B)</td>
</tr>
</tbody>
</table>

**Days to R1**  
(*field 1991 and 1992*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
<td>39.94</td>
<td>10.13</td>
<td>14.2 (P34B)</td>
</tr>
<tr>
<td>A073H-K443H</td>
<td>G</td>
<td>10.21</td>
<td>4.15</td>
<td>6.8 (P34B)</td>
</tr>
</tbody>
</table>

**Days to R1**  
(*growth chamber 12 h, gc12*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
<td>47.39</td>
<td>12.43</td>
<td>3.9 (P34B)</td>
</tr>
</tbody>
</table>

**Days to R1**  
(*growth chamber 20 h, gc20*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-Satt286</td>
<td>C2</td>
<td>22.71</td>
<td>4.14</td>
<td>18.3 (P34B)</td>
</tr>
<tr>
<td>A073H-K443H</td>
<td>G</td>
<td>5.24</td>
<td>3.47</td>
<td>6.6 (P34B)</td>
</tr>
</tbody>
</table>

**Flowering delay**  
(*differences between gc20 and gc12*)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Linkage group</th>
<th>R(^{2b})</th>
<th>LOD(^c)</th>
<th>Allele effect(^{bd})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-Satt286</td>
<td>C2</td>
<td>20.77</td>
<td>3.36</td>
<td>15.1 (P34B)</td>
</tr>
<tr>
<td>A073H-K443H</td>
<td>G</td>
<td>6.75</td>
<td>2.52</td>
<td>5.9 (P34B)</td>
</tr>
</tbody>
</table>
Table 3. (continued)

<table>
<thead>
<tr>
<th>Days to R3</th>
<th>(field 1991)</th>
<th>Days to R3 (growth chamber 12 h)</th>
<th>Days to R7 (field 1991)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt205-Satt286</td>
<td>C2</td>
<td>28.65</td>
<td>6.94</td>
</tr>
<tr>
<td>Satt276-</td>
<td>A1</td>
<td>12.13</td>
<td>3.79</td>
</tr>
<tr>
<td>Satt205-PO29D_2</td>
<td>C2</td>
<td>44.42</td>
<td>11.61</td>
</tr>
<tr>
<td>Satt182-P029D_1</td>
<td>L</td>
<td>5.10</td>
<td>3.40</td>
</tr>
<tr>
<td>Satt205-Satt286</td>
<td>C2</td>
<td>21.72</td>
<td>5.50</td>
</tr>
<tr>
<td>Satt276-</td>
<td>A1</td>
<td>9.83</td>
<td>2.99</td>
</tr>
</tbody>
</table>

* Based on USDA/ISU soybean public map [42]:

Coefficients of determination and allele effects of the first marker locus in the interval that is significantly associated with the trait locus (p ≤ 0.01), based on a single-factor ANOVA of SAS-GLM:

Based on Mapmaker-QTL analyses of the indicated chromosomal interval:

Mean differences for the trait between the two genotypic groups carrying PI 317.334B (P34B) and 'Corsoy' (Cor) alleles of the first marker in each the interval. The parental allele that is associated with an increase in trait value is given in parentheses.
Figure 1 a. Location of QTL in population IX132. Thin lines on left and right sides of each linkage group indicate the confidence intervals of each detected QTL. 1 = days to R1 (field 1991); 2 = days to R1 (field 1992); 3 = days to R1 (field 1991/1992); 4 = days to R1 (growth chamber with 12 h photoperiod, gc12); 5 = days to R1 (growth chamber with 20 h photoperiod, gc20); 6 = flowering delay (differences between gc20 and gc12); 7 = days to R3 (field 1991); 8 = days to R3 (gc12); 9 = days to R7 (field 1991). Arrows indicate the putative location of each detected QTL.
**Figure 1 b.** Location of QTL in population IX136. Thin lines on left and right sides of each linkage group indicate the confidence intervals of each detected QTL. 1 = days to R1 (field 1991); 2 = days to R1 (field 1992); 3 = days to R1 (field 1991/1992); 4 = days to R1 (growth chamber with 12 h photoperiod, gc12); 5 = days to R1 (growth chamber with 20 h photoperiod, gc20); 6 = flowering delay (differences between gc20 and gc12); 7 = days to R3 (field 1991); 8 = days to R3 (gc12); 9 = days to R7 (field 1991). Arrows indicate the putative location of each detected QTL.
CHAPTER 4. MAPPING FLOWERING TIME GENE HOMOLOGS
IN SOYBEAN AND THEIR ASSOCIATION
WITH MATURITY (E) LOCI

A paper to be submitted to Crop Science


ABSTRACT

Time of flowering is a quantitatively inherited character and is an important reproductive trait of agronomic interest. This trait is important for developing soybean cultivars with a wider area of adaptation. The genetics of flowering time has been extensively studied in the model plant Arabidopsis thaliana. Many genes have been identified and some have been cloned and characterized. The first objective of this study was to map onto the soybean genetic map orthologous genes known to be involved in photoperiod recognition and time of flowering. The second objective was to correlate the mapped homologous genes with maturity (E) loci by means of near isogenic lines (NILs). Three

1Department of Agronomy, Iowa State University, Ames, IA 50011
2USDA-ARS-Corn Insect and Crop Genetics Research Group and Department of Agronomy, Iowa State University, Ames, IA 50011.
3Curator, USDA Soybean Germplasm Collection/University of Illinois, Urbana-Champaign
4Plant Research Center, Agric. Agri-Food Canada, Ottawa, ON, Canada, K1A OC6.

*Author for correspondence
single-cross populations consisting of two recombinant inbred (RI) lines and an F₂ population were used in the mapping study and 41 NILs and two recurrent parents (RP) were used in the correlation study. One RI population, IX132, was developed by crossing PI317.336 and ‘Corsoy’; the second, IX136, by crossing PI317.334B and ‘Corsoy’. Both P.I.s have been reported to be photoperiod insensitive. The ‘Corsoy’ parent has been reported to be a photoperiod-sensitive cultivar. The RI populations were previously used to map QTL for flowering time, maturity, and photoperiod insensitivity in soybean. The F₂ population was developed from an interspecific cross between G. max (breeding line A81-356022) and G. soja (PI 468.916) and was previously used to construct the USDA/ISU public map and the soybean consensus map. Eighteen soybean cDNA clones, identified by BLAST to have high similarities with 18 previously cloned flowering time genes, were used as probes in this study. Ten out of the 18 cDNA clones were mapped. The homologous sequences were mapped onto linkage groups (LGs) A2 (CRY2), B1 and H (COLI), A1 and B2 (PHYA), C1 (DET1 and LD), D2 (AP2), E and K (PHYB), F (COL2), L (FCA), and Q (CCA1). None of these cDNA sequences has been found to be directly associated with previously mapped QTL for flowering time. However, analyses of these candidate genes using NILs show that the homologous gene sequence FCA was associated with maturity locus E3. That FCA is a strong gene candidate for maturity locus E3 is further supported by map position and phenotypic data. Analyses of NILs suggest that PHYB homolog may be associated with maturity locus E1. However, current data show they mapped in different LGs.
INTRODUCTION

Flowering time of most plant species has been reported to be controlled by interaction between environmental and endogenous cues (Levy and Dean 1998). The environmental cues are those such as light duration and quality and temperature which signal the onset of conditions favorable to the success of reproductive development. The endogenous cues are those that control the developmental competence of the plant.

Studies from the model plant A. thaliana show that flowering time is the result of a sequential action of two groups of genes: flowering time genes and floral identity or floral initiation process (FLIP) genes (Levy and Dean 1998; Piñiero and Coupland 1998; Koornneef et al. 1998). Flowering time genes are those that display their major effects on the duration of vegetative development. These genes act before the actions of FLIP genes and they may activate or repress floral identity genes under different environmental conditions. FLIP genes, on the other hand, are those that switch the fate of meristems from the vegetative to the floral phase. Mutations of these genes cause primordia which would normally develop as flowers in the wild-type plants to instead form structures with shoot-like characteristics.

Many genes that control flowering time have been identified. Gene discovery has been based on studies from a few different plant species, such as Arabidopsis and pea (Koornneef et al. 1998; Piñiero and Coupland, 1998, Weller et al. 1997). From Arabidopsis alone, at least 80 loci currently have been reported to affect the timing of flowering (Levy and Dean 1998). The genes were identified through the analyses of natural variation of different Arabidopsis ecotypes and through characterization of induced mutations (Coupland 1995; Coupland 1997). Some of these genes act to promote flowering while others to repress it. Some appear to interact with environmental factors (e.g., photoperiod and temperature).
while others appear to act in an autonomous way. Based on these observations, together with data from double mutant studies and transgenic plant analyses, it was proposed at least four different genetic pathways may control flowering time in *Arabidopsis* (Levy and Dean 1998; Koornneef et al. 1998).

Among the 80 *Arabidopsis* flowering time genes reported so far, at least 25 have been cloned and characterized (Levy and Dean 1998; Koornneef et al. 1998). In addition, several genes from studies of other aspects of plant development – such as light perception and hormone metabolism – were cloned and were found to play roles in the regulation of flowering time and, therefore, are also called flowering time genes.

Cloned flowering time genes have been mapped in several different plant species. In barley, three homologous sequences from *Arabidopsis* [*CONSTANTS (CO)*, *TERMINAL FLOWER 1 (TFL1)*, and *GIGANTEA (GI)*] were found to map closely to previously identified flowering time QTL (Christodolou et al. 2001). They, however, did not correspond precisely to the QTL peak. Another study in rice reported the homologous flowering time genes *GIGANTEA (GI)* from *Arabidopsis* and the homologous sequence *PNZIP* from morning glory (*Pharbitis nil*) were associated with the previously detected flowering time QTL (Thompson et al. 2001). The authors concluded that even though the homologous gene sequences do not correspond to known flowering time loci, their presence in the rice and barley genomes suggest that there is a considerable conservation of the flowering time genetic pathways and provide support for the continued use of this candidate gene approach.

In soybean [*Glycine max* (L.) Merr.], at least five genes have been reported to affect flowering time and maturity (Cober et al. 1996a). The genes are known as the *E*-series: *EI*,...
Several QTL associated with flowering time and maturity have also been previously mapped in soybean. A large effect QTL controlling flowering time, maturity, and photoperiod insensitivity was mapped on LG C2 (Tasma et al. 2001). The QTL was found to be located in the same location in two populations IX132 and IX136. This QTL explained as much as 47% of total phenotypic variance and was also found to closely map to the pubescent color locus T (Tasma et al. 2001). Smaller effect QTL were found to be located on other LGs (A2, G. J. and L) in both populations. These later QTL individually explained as much as 17.8% of total phenotypic variance. Keim et al. (1990) reported that five markers on linkage groups C1, C2, and D1 (Shoemaker and Specht 1995) were found to be associated with maturity (including time of flowering) in an F$_2$ population derived from a cross between *G. max* and *G. soja*. The observed QTL explained 17-23% of total phenotypic variance. Using another F$_2$ population of 'Minsoy' X 'Noir 1', Mansur et al. (1996) also reported a major QTL for flowering time on LG C2 and minor QTL for maturity on LG L and M (Shoemaker and Specht 1995). Lee et al. (1996) reported that QTL for maturity traits were found on LG K, based on an F$_2$ population derived from a cross between PI 97100 and 'Coker 237'. The observed QTL explained 26-31% of phenotypic variance.

No studies have yet been reported that map flowering time gene homologs in soybean. In this study we mapped candidate genes for these traits. In addition, we compared this map location with flowering time QTL previously identified in the same two
RI soybean populations (Tasma et al. 2001). Finally, we determined the relationship between homologous gene sequences and the known maturity ($E$) loci.

Correlating these homologous gene sequences with the known soybean maturity ($E$) loci would give us insight to further analyze this important trait at the molecular level. NILs have been used as a genetic resource to identify linkages between molecular and classical markers. Many NILs with different allele combinations of the $E$ loci have been developed (Bernard et al. 1991). The rationale for using NILs for mapping the molecular and classical genetic markers has been reported (Muehlbauer et al. 1988; Young et al. 1988; Muehlbauer et al. 1989; Muehlbauer et al. 1991).

**MATERIALS AND METHODS**

**Plant Materials**

**Mapping populations**

Three single-cross populations, two recombinant inbred (RI) lines and an $F_2$, were used in this study. The RI population, IX132, was developed by crossing PI317.336 and ‘Corsoy’. The RI population, IX136, was developed by crossing PI317.334B and ‘Corsoy’. Both P.I. parents have been reported to be photoperiod insensitive (Gutrie 1972; Nissly et al. 1981, Metz et al. 1985). The ‘Corsoy’ parent, on the other hand, has been reported to be a photoperiod sensitive cultivar. These RI populations were used to map QTL for flowering time, maturity, and photoperiod insensitivity in soybean (Tasma et al. 2001).

An $F_2$ population was developed from an interspecific cross between *G. max* (breeding line A81-356022) and *G. soja* (PI 468.916). This latter population was used to
construct the USDA/ISU public map (Shoemaker and Olson 1993: Shoemaker and Specht 1995) and the soybean consensus map (Cregan et al. 1999). The development of the populations was reported previously (Tasma et al. 2000: Keim et al. 1990). The size of each population was 101 (IX132), 100 (IX136), and 60 progeny (G. max X G. soja).

Near isogenic lines

Two sets of NILs were used. The first set comprised the L-lines developed at the University of Illinois, Urbana-Champaign (Bernard et al. 1991), and the second set comprised the OT-lines developed at the Plant Research Center, Ottawa, Ontario, Canada. The L-lines were developed using recurrent parent cultivars ‘Clark’ and ‘Harosoy’ and were obtained from Dr. Randall Nelson (USDA-ARS/University of Illinois, Urbana-Champaign). The OT-lines were developed using recurrent parent cultivar ‘Harosoy’ and were obtained from Dr. Elroy Cober (Eastern Cereal and Oilseed Research Center, Ontario, Canada). The NILs contain different allele combinations of the £ loci (Table 1). The NILs also contain different alleles of growth habit locus Dtl.

Identification of Cloned Homologous Flowering Time Genes

From literature review, we identified cloned flowering time genes to be used in this study (e.g., PHYA, PHYB, CRY2, CO, and FCA) (Table 2). Using a NCBI (National Center for Biotechnology Information) nucleotide search menu, we searched for cloned flowering time genes (e.g., by searching for soybean AND PHYA to obtain soybean ESTs encoding PHYA). Soybean clones were then selected based upon similarity of the soybean EST sequences to those of the cloned flowering time genes in the GeneBank databases as
determined by BLAST. The EST sequences were blasted using either the nucleotide sequence (Blastn) or the protein sequence (Blastp) programs. The BLAST results are shown in Table 4. Characteristics of the cloned flowering time genes used in this study are shown in Table 2 and the cDNA clones used as probes in this study are shown in Table 3.

**PCR Amplification of cDNA Clones**

The cDNA clones were obtained from Incyte Genomics (St. Louis, Missouri). The clones were cultured in a LB (Luria Broth) plate medium containing antibiotic Ampicillin (0.1 µg/µl). A single colony was then picked and cultured in a 5-ml LB medium containing Ampicillin (0.1 µg/µl) and was incubated in a shaker (250 rpm) at 37°C overnight. The cDNA insert of each clone was amplified by PCR. We used either a single colony or an overnight bacterial-LB culture containing a particular cDNA clone as template in the PCR amplification. Generally, we used T3 and T7 promoter primers (Gibco BRL, Maryland) to amplify the inserts since most cDNA clones used were constructed in the pBluescript vector (Gibco BRL, Maryland). When cDNA clones were constructed in a pSPORT1 vector (Gibco BRL, Maryland), the primer pairs used were SP6 and T7 (Gibco BRL, Maryland). All the primers were synthesized at the Iowa State University DNA facility. A 100-µl total PCR reaction was set up containing 1.5 mM MgCl₂, 0.2 mM each of T3 and T7 or SP6 and T7 primers, 100 µM of each nucleotide, 1x PCR buffer (containing 50 mM of KCl, 10 mM of Tris-HCl pH 8.3), 2 units of Taq DNA polymerase (Gibco BRL, Maryland), and a single bacterial colony or a 10-µl overnight bacterial culture containing the respective cDNA clone. An initial 4-minute treatment at 94°C was conducted before the cycling began. Cycling consisted of 45 s denaturation at 94°C, one minute primer annealing at 52°C, and two
minutes DNA extension at 72°C for 30 cycles on a MJ Research PTC-100 thermocycler. At the end of the cycles, an additional five-minute DNA extension at 72°C was carried out. The PCR product was then diluted with sterile water to make a concentration of 25 ng per µl. The amplified cDNA inserts were then used as probes in the RFLP analyses.

**DNA Extraction, Restriction Digestion, Electrophoresis, and Southern Analysis**

DNA was extracted from leaves as described previously (Keim et al. 1988). To identify the RFLP probe-enzyme combinations revealing polymorphism between the parents of the mapping populations, a parental screen analysis was conducted. All parental DNAs [PI317.336, PI 317.334B, 'Corsoy', *G. max* breeding line A81-356022 and *G. soja* (PI 468.916)] were digested with each of nine restriction enzymes (DraI, EcoRI, EcoRV, HhaI, HindIII, PstI, Styl, TaqI and *XhoI*) according to the manufacturer (Gibco BRL, Maryland). The recurrent parent and the NIL DNAs were also digested with the same restriction enzymes as those used to digest the parental and mapping progeny DNAs. Digested DNAs were then electrophoresed in 0.8 % Ultrapure agarose gel (Gibco BRL, Maryland) in 1X TAE buffer (Sambrook et al. 1989) at 22-30 V for 16-20 h. The DNAs were then transferred (Southern, 1975) onto Zeta-Probe® GT nylon membranes (Bio Rad, California), probed, and autoradiographed as described previously (Keim et al. 1990). In the parental screen analyses, the polymorphic probes were then used to hybridize the progeny DNA of each population. A Chi-square test was then carried out to determine segregation ratio of the homologous gene sequences in each population. The null hypothesis of the tests was that progeny segregate in a 1:1 ratio of the opposite alleles for populations IX132 and IX136 and a 1:2:1 ratio for the
G. max X G. soja population. Only the mapped homologous-cloned flowering time gene sequences were used to probe the NIL DNAs and their respective recurrent parents.

**Linkage Mapping**

Markers used in this study were those reported by Tasma et al. (2001) for the RI populations, IX132 and IX136, and those reported by Cregan et al. (1999), Shoemaker and Olson (1993), and Shoemaker and Specht (1995) for the G. max X G. soja population. Placement of markers was done using the Mapmaker mapping program (Lander et al. 1987). For grouping the markers, a minimum LOD score of 3.0 and a maximum distance of 40 cM were used as a threshold value to declare linkage in the pairwise loci analysis. The linkage map was constructed using the Haldane map function (Haldane, 1919). The gene orders were assigned using the 'compare', 'try', and 'ripple' (minimum LOD score of 3.0) commands.

**RESULTS**

**Homologous Gene Segregation in the Progeny of Each Population**

A total of 14, 11, and 17 homologous gene and restriction enzyme combinations representing 8, 7, and 10 distinct homologous genes, respectively, were found to be segregated in populations IX132, IX136, and G. max X G. soja, respectively (Tables 5, 6, and 7). Chi-square tests showed almost all the segregating restriction fragments followed the null hypotheses of 1:1 segregation ratios in populations IX132 and IX136, and a 1:2:1 ratio in
G. max X G. soja population (Tables 5, 6, and 7). Only two (FCA-I and FCA_Sty) of the segregating restriction fragments did not follow the expected segregation ratios in populations IX132. Only one (AP2_H1) of the markers did not follow the expected segregation ratio in G. max X G. soja population.

Location of Cloned Flowering Time Genes in the Soybean Genome

The mapped candidate genes consist of four photoperiod recognition sequences (PHYA, PHYB, CRY2, and CCA1), five flowering time genes (FCA, DET1, COL1, COL2, and LD), and one floral identity gene sequence (AP2). The AP2, CCA1, COL2, DET1, FCA, FUS5, PHYA, and PHYB were mapped in population IX132 (Figure 1). PHYB mapped to two different LGs, E and K. The others were mapped on one LG only (Figure 1). The COL2, DET1, FCA, FUS5, PHYA, and PHYB were mapped in population IX136 (Figure 2). As in population IX132, PHYB was also mapped in two different LGs (E and K) in this population. They were also found to be linked with the same markers in both populations (Figure 2). The COL1, COL2, CRY2, FCA, LD, PHYA, and PHYB were mapped in the G. max X G. soja population (Figure 3). Consistent with data from populations IX132 and IX136, PHYB was again mapped on two LGs (E and K). In this population, PHYA and COL1 also mapped in two LGs (A1 and B2 for PHYA; B1 and H for COL1) (Figure 3).

The map locations of homologous sequences were found to be consistent across populations (Figures 1, 2, and 3). The homologous gene sequences that were mapped in all three populations were PHYA (LG B2), PHYB (LGs E and K), and FCA (LG L). All were found to be linked with the same markers. Those that were mapped in two populations (IX132 and IX136) were DET1 (LG C1), FUS5 (LG E), and COL2 (LG F). The homologous
sequences mapped in only one population include $AP2$ (LG D2 of IX132), $CCAl$ (LG Q of IX132), $COLl$ (LGs B1 and H of $G.\ max \times G.\ soja$), $CRY2$ (LG A2 of $G.\ max \times G.\ soja$), and $LD$ (LG C1 of $G.\ max \times G.\ soja$).

**Correlation of the Flowering Time Gene Homologs and the Maturity ($E$) Loci**

Among ten cloned flowering time genes used as probes and the 41 NILs and RPs tested, only two candidate genes were found to be polymorphic in eight NILs. These were homologs of $FCA$ and $PHYB$. An example of a polymorphism observed from this study is shown in Figure 4. The $FCA$ probe was polymorphic on NILs L71-920, L74-441, L80-5914, L63-2404, L63-3270, and L80-5879 with respect to their RP, 'Clark' (Table 8 and Figure 4). Restriction endonucleases showing polymorphisms in these NILs were $Styl$, $TaqI$, and $HindIII$. Genotypes of each NIL and the respective RP are shown in Table 8. The pattern of polymorphisms among the various NILs is consistent with the association of $FCA$ with maturity locus $E3$. $FCA$ was mapped on LG L and was detected in all three populations (IX132, IX136, and $G.\ max \times G.\ soja$). $E3$ was also mapped on LG L based on the consensus map (Cregan et al. 1999) in an approximately similar location to where $FCA$ was mapped. Thus, $FCA$ may correspond to the maturity locus $E3$.

The homologous gene sequence $PHYB$ was polymorphic on NIL L66-531 with respect to its RP 'Clark’ and on NILs L67-2324 and L71-1116 with respect to their RP 'Harosoy’ (Table 8). Restriction endonucleases showing polymorphisms were $Hhal$ and $HindIII$. The analysis of NILs provides compelling evidence of an association between candidate gene $PHYB$ and maturity locus $E1$. However, $PHYB$ was mapped on LGs E and K in all three populations. $E1$, on the other hand, was mapped on LG C2 (Cregan et al. 1999).
All other homologous-cloned flowering time gene sequences were found to be monomorphic across all NILs used in this study. No direct association between map position of candidate genes and flowering QTL were observed. However, indirect associations can be inferred. These will be discussed in a later section.

**DISCUSSION**

In this study we mapped ten flowering time gene homologs in soybean. The sequences were found to be distributed on nine LGs of the soybean genetic map (Cregan et al. 1999). Seven of the homologous sequences are unique and were detected in only one LG. Three, however, were detected in two LGs, suggesting that they represent duplicated loci. Results of this study further support the notion that the cloned flowering time gene sequences are conserved across plant genera. The candidate gene sequence *FCA* was found to be associated with maturity locus *E3*. The candidate gene *PHYB* was found to be polymorphic in three NILs and seems to be associated with maturity locus *E1*, but their map positions are on different LGs.

None of the homologous gene sequences were found to be directly associated with the previously identified flowering time QTL (Tasma et al. 2001). One possible explanation for this is that we have failed to detect QTL that in fact might be associated with candidate genes. Simulation studies show not all QTL will be detected in a small population (100 to 200 progeny) (Beavis 1997). In addition, due to gaps in our maps, some small effect QTL may not be detected. Thus, some of the homologous sequences mapped in this study may be responsible for a particular QTL, but we failed to detect the QTL and no association of the traits and these sequences was found.
Using \textit{FCA} as a probe, six NILs (L71-920, L74-441, L80-5914, L63-3270, L63-2404, and L80-5879) were found to be polymorphic with their respective recurrent parents 'Clark' (Table 8 and Figure 4). The alleles transferred into the RP 'Clark' were \textit{e2} and \textit{e3} for the L71-920; \textit{E1}, \textit{e2}, and \textit{e3} for the L80-5914; \textit{E1} and \textit{e3} for the L74-441; \textit{dtl} and \textit{e3} for the L63-3270; \textit{e3} for the L63-2404; and \textit{dtl}, \textit{E1}, \textit{e2}, and \textit{e3} for the L80-5879. A polymorphism was found between all six L-lines and 'Clark' (Figure 4), suggesting the polymorphisms are due to genetic variation in the genome region including the contrasting alleles \textit{e3} and \textit{E3} of the six L-lines and 'Clark'. Thus, from this analysis we conclude that \textit{FCA} may be associated with maturity locus \textit{E3}.

\textit{FCA} was mapped on LG L and was detected in the same locations in all three populations used in this study (Figures 1, 2, and 3). \textit{E3} was also previously mapped on LG L based on the soybean consensus map (Cregan et al. 1999). \textit{FCA} and \textit{E3} are both located near SSR marker Satt006. No flowering time (days to R1) QTL were detected on LG L from our previous study using these RI populations (Tasma et al. 2001). However, we mapped a small effect QTL for days to R3 (days to first pod) based on data in a growth chamber with a 12 h photoperiod on LG L, detected in population LX136 only (Tasma et al. 2001). Its location, however, is distant from the \textit{FCA}. Mansur et al. (1996) discovered minor effect QTL for flowering time (days to R1) and maturity (days to R8) on LG L linked with SSR marker Satt006. Yamanaka et al. (2001) found a major QTL (\textit{FT3}) for flowering time on LG L linked very tightly with SSR marker Satt373. \textit{FCA} was mapped closely to Satt373 in all three populations used in this study (Figures 1, 2, and 3). Some effects of the minor QTL observed by Mansur et al. (1996) and the major QTL \textit{FT3} observed by Yamanaka et al. (2001), therefore, may be due to the \textit{E3} locus.
Thus, a possible reason for our inability to detect association between the flowering time QTL and the mapped location of FCA was that we were unsuccessful in finding particular QTL that in reality might be linked to or corresponded with the mapped homologous gene sequences.

A degree of common functionality has been demonstrated between FCA of Arabidopsis and the maturity locus E3 of soybean with regard to their response to inductive and non-inductive day lengths. In Arabidopsis, FCA promotes flowering independent of photoperiod (photoperiod insensitive) and this gene was placed in the autonomous or constitutive signal transduction pathway (Levy and Dean 1998; Piñiero and Coupland 1998). E3 was also reported as a major locus conferring photoperiod insensitivity in soybean (Saindon et al. 1989). This study was based on natural day length extended to 20 h with incandescent light also called incandescent long day length (ILD) (Saindon et al. 1989; Cober et al. 1996a). In addition, under natural daylength, the dominant (late maturing) allele of maturity locus E3 was reported to delay maturity the least compared to the effects of dominant alleles of other maturity loci with respect to the recessive (early maturing) alleles, supporting the insensitive nature of the E3 locus. Cober et al. (1996b) also reported the E3 allele was the least sensitive to changes of light quality in response to long day length compared to other E loci. Thus, based upon their responses to the inductive and non-inductive photoperiods, there are characteristics in common between the effect locus FCA has in Arabidopsis and the maturity locus E3 shows in soybean. This, and the similar map locations of FCA and E3 on LG L, make the soybean FCA homolog a strong candidate for the soybean maturity locus E3.
Analysis of NILs strongly indicated that *PHYB* is associated with maturity locus *El*. The map position of *El* (LG C2) has previously been shown to correlate with a QTL for flowering time (Tasma et al. 2001). However, *PHYB* mapped to two LGs (E and K) (this study) and *El* was mapped on LG C2 (Cregan et al. 1999). This suggests that we may have mapped duplicate loci of *PHYB* paralogs. Soybean has been recognized as an ancient polyploid and has been reported to be rich in duplicated loci (Shoemaker et al. 1996). Common RFLP markers are detected in these regions indicating the two mapped *PHYB* are duplicated genes. A report by Shoemaker et al. (1996) that LG E and LG K contain homoeologous regions (i.e., RFLP markers in common) further supports that the two mapped *PHYB* are duplicated loci. A similar situation was observed between homoeologous regions on LGs B1 and H (Shoemaker et al. 1996) to which duplicate copies of the *COLI* homolog were mapped.

There may be more than two copies of *PHYB* homologous sequences in the soybean genome. It is possible, therefore, that we have failed to map the *PHYB* homolog on LG C2. Another possibility is that another gene (possibly an unknown maturity locus) has been backcrossed during the NIL development, along with *El*. Therefore, we identified polymorphisms around the other gene and it is the other gene that corresponds to *PHYB*. Another possibility is that *El* may be a duplicated gene and another copy may be on LG E and/or K where *PHYB* has been mapped. However, this would mean that only one gene at a time was segregating or previous researchers would have seen a segregation ratio indicative of duplicate genetic factors. Further investigations, however, would clarify all the above hypotheses.
In *Arabidopsis* (a LD plant), *PHYB* inhibits floral initiation (Coupland 1997). The *PHYB* mutant is classified as an early flowering mutant and flowers earlier than the wild-type plants under both LD and SD conditions. However, the early phenotype of the mutant is more pronounced in SD conditions (Goto et al. 1991; Mockler et al. 1999). *PHYB* has also been placed in the repression genetic pathway (Levy and Dean 1998). *PHYB* mutants in sorghum (*Ma3*<sup>R</sup>, Childs et al. 1997) and pea (*lvi-1*, Weller and Reid 1993) both showed early flowering and decreased photoperiod sensitivities. Sorghum is classified as a short day plant and pea is a long day plant. The *PHYB* mutation in pea flowers early in SD but not in LD (Weller and Reid 1993) and *PhtYB* mutations in sorghum flowers early in LD but not in SD (Pao and Morgan 1986). Thus, wild-type *PHYB* inhibits flowering both in LD and SD plants. The flowering inhibition of *PHYB*, however, appears more apparent in the photoperiod that normally suppresses flowering in the respective plant.

In soybean, *E1* also inhibits flowering. Under natural day length *E1* delayed flowering 16-23 days and maturity 15-18 days (Bernard 1971) compared to its contrasted allele *e1*. This flowering delay is greater than the effects of other dominant maturity loci (Cober et al. 1996a).

Light quality studies showed that the *E1* allele was most sensitive to light quality and required a R:FR ratio approximating that of natural daylight for response to long day length (Cober et al. 1996b). The *E3* and *E4* alleles were reported to have the least and intermediate sensitivities to changes in light quality, respectively. The *Ele3e4* isolate showed a strong response to long days compared to the *ele3e4* isolate when a R:FR range from two to one (a R:FR ratio equivalent to natural light) but showed very small response with R:FR ratio above two. The *Ele3e4* isolate flowered 5 days later than the *ele3e4* isolate at a R:FR ratio of two.
but 15 days later than the ele3e4 at R:FR ratio of one. This flowering delay is characteristic of the effect of allele $E1$ under field conditions (Bernard 1971; McBlain et al. 1987). Sensing the R:FR ratio is the function of phytochrome in light grown plants.

A study in *Arabidopsis* showed that *PHYB* and at least one other phytochrome mediate light quality responses where long day low R:FR ratio demonstrated accelerated flowering (Halliday et al. 1994). In soybean, on the other hand, long day low R:FR conditions resulted in delayed flowering (Cober et al. 1996b). Cober et al. (1996b) speculated that long day low R:FR conditions may simply elicit more extreme photoperiodic responses, i.e., earlier flowering in long day plants and delayed flowering in short day plants. Thus, based on this evidence, there are common characteristics in phenotypes between *PHYB* in *Arabidopsis* and maturity locus $E1$ in soybean.

Even though phenotypic data show some common features between the candidate gene *PHYB* and the maturity locus $E1$, and analysis of NILs indicate an association, current map positions of these two genes are different. However, we cannot rule out *PHYB* as a candidate gene for the maturity locus $E1$. Further investigation should clarify this discrepancy.

In this study, ten flowering time gene homologs have been mapped in the soybean genome, suggesting they are conserved across plant genera. We have also mapped homologs of *PHYB* and *COL1* to homoeologous genome regions. This further supports the notion that the soybean genome is rich in duplicated regions. The homologous gene sequence *FCA* is a strong candidate for maturity locus $E3$. The *PHYB* homologous gene sequence may associate with maturity locus $E1$, but current evidence shows they map in different LGs. The results of this study will facilitate further studies of these maturity loci at the molecular level.
ACKNOWLEDGEMENTS

The authors thank Dr. Detroy Green for the RI populations used in this study. Dr. John Erpelding for technical advise. Dr. David Grant for helpful discussion. Dr. M. Paz and Mr. Thad Wunder for the excellent technical assistance. and the government of Indonesia for financial support of IMT. Names are necessary to report factually on the available data; however, the USDA neither guarantees nor warrants the standard of the product to the exclusion of others that may also be available.

REFERENCES


Plant Physiol. 113: 611-619.


Haldane, J.B.S. 1919. The combination of linkage values, and the calculation of distances between the loci of linked factors. J. Genet. 8: 299-309.

other phytochrome mediate the accelerated flowering response of *Arabidopsis thaliana* L. to low red/far-red ratio. Plant Physiol. 104: 1311-1315.


Shoemaker, R.C., K. Polzin, J. Labate, J. Specht, E.C. Brummer, T. Olson, N. Young, V.


<table>
<thead>
<tr>
<th>Name</th>
<th>Genotype</th>
<th>Background (RP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>'Clark'</td>
<td>Dtle1E2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>'Harosoy'</td>
<td>Dtle1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L71-920</td>
<td>Dtle1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L62-667</td>
<td>Dtle1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L63-3117</td>
<td>Dtle1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L65-778</td>
<td>Dtle1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L67-153</td>
<td>Dtle1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L63-2404</td>
<td>Dtle1E2e3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L84-307</td>
<td>Dtle1E2e3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L63-3270</td>
<td>Dtle1E2e3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L64-4584</td>
<td>Dtle1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L63-3016</td>
<td>Dtle1E2E3e4e5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L80-5914</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L71-802</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L80-5879</td>
<td>Dtle1e2E3e4e5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L74-102</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L66-432</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L67-2324</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L66-531</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L71-1116</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L74-441</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L76-865</td>
<td>Dtle1E1e2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L65-3366</td>
<td>Dtle1E1E2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L71L-3004</td>
<td>Dtle1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th>Variety</th>
<th>Genotype</th>
<th>Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>L66-546</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L84-337</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L94-1110</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L64-4830</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L92-1195</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Clark'</td>
</tr>
<tr>
<td>L74-66</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>L71L-3015</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT89-5</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT89-6</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT93-26</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT93-28</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-37</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-39</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-41</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-43</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-47</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-49</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT94-51</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
<tr>
<td>OT99-17</td>
<td>D1E1E2E3E4E5E7</td>
<td>'Harosoy'</td>
</tr>
</tbody>
</table>

RP = recurrent parent.
Table 2. Characteristics of flowering time genes used in this study

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein encoded</th>
<th>Function</th>
<th>Species(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHYA (Phytochrome A)</td>
<td>Apoprotein</td>
<td>Light-labile R-FR light photoreceptor</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>PHYB (Phytochrome B)</td>
<td>Apoprotein</td>
<td>Light-stable R-FR light photoreceptor</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>PHYE (Phytochrome E)</td>
<td>Apoprotein</td>
<td>Photoreceptor</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>CRY2 (Cryptochrome 2)</td>
<td>Flavoprotein</td>
<td>Blue-light photoreceptor</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>COPI</td>
<td>Not yet classified</td>
<td>Light regulation</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>DET1 (Deetiolated 1)</td>
<td>Nuclear localized protein</td>
<td>Light regulation</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>GT2</td>
<td>Not yet classified</td>
<td>Light regulation</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>FCA</td>
<td>RNA binding protein with protein-protein interaction domains</td>
<td>Promote flowering independent of photoperiod</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>CO (Constans)</td>
<td>Zinc finger protein</td>
<td>Promote flowering in response to inductive photoperiod</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>LD (Luminidependens)</td>
<td>Nuclear localized glutamine-rich domains</td>
<td>Promote flowering independent of photoperiod</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>AG (Agamous)</td>
<td>MADS box protein domains</td>
<td>Flower formation</td>
<td><em>A. thaliana</em></td>
</tr>
<tr>
<td>Gene</td>
<td>Protein encoded</td>
<td>Function</td>
<td>Species&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>---------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>AP2</td>
<td>AP2 protein</td>
<td>Flower formation</td>
<td>A. thaliana</td>
</tr>
<tr>
<td>(Apetala 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Quite similar to CO</td>
<td>Not yet determined</td>
<td>A. thaliana</td>
</tr>
<tr>
<td>(Constans like 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Quite similar to CO</td>
<td>Not yet determined</td>
<td>A. thaliana</td>
</tr>
<tr>
<td>(Constans like 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCA1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>MYB-related transcription factor</td>
<td>Circadian clock regulation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>A. thaliana</td>
</tr>
<tr>
<td>(Circadian Clock Associated 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FPFL</td>
<td>Novel protein</td>
<td>May involve in signaling or</td>
<td>A. thaliana</td>
</tr>
<tr>
<td></td>
<td></td>
<td>response to GAs&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>GAI</td>
<td>Member of novel family with putative transcription factor</td>
<td>Promote flowering in response to non-inductive photoperiod</td>
<td>A. thaliana</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPI&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Not yet classified</td>
<td>Seed ripening</td>
<td>A. thaliana</td>
</tr>
<tr>
<td>(Viviparous 1)</td>
<td></td>
<td></td>
<td>Zea mays</td>
</tr>
</tbody>
</table>

<sup>a</sup>Two homologs of CO gene: they are quite similar in structure with CO gene (Kooimeef et al. 1998)

<sup>b</sup>Overexpression of CCA1 abolished circadian clock (Levy and Dean 1998)

<sup>c</sup>GAs = gibberelic acids

<sup>d</sup>Species from which the genes were originally isolated
Table 3. Soybean cDNA clones used as probes in this study

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Incyte Genomics I.D.</th>
<th>bp mRNA sequence</th>
<th>cDNA insert size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL2</td>
<td>Gm-c1012-422</td>
<td>553</td>
<td>1100</td>
</tr>
<tr>
<td>CRY2</td>
<td>Gm-c1019-3678</td>
<td>615</td>
<td>1000</td>
</tr>
<tr>
<td>PHYB</td>
<td>Gm-c1027-285</td>
<td>700</td>
<td>1200</td>
</tr>
<tr>
<td>PHYE a</td>
<td>Gm-c1019-2148</td>
<td>573</td>
<td>900</td>
</tr>
<tr>
<td>PHYA</td>
<td>Gm-c1016-1937</td>
<td>389</td>
<td>1100</td>
</tr>
<tr>
<td>AP2</td>
<td>Gm-c1014-291</td>
<td>554</td>
<td>1100</td>
</tr>
<tr>
<td>VP1 a</td>
<td>Gm-c1012-166</td>
<td>488</td>
<td>1300</td>
</tr>
<tr>
<td>LD</td>
<td>Gm-c1004-3823</td>
<td>184</td>
<td>1300</td>
</tr>
<tr>
<td>GT2 a</td>
<td>Gm-c1012-1002</td>
<td>484</td>
<td>1900</td>
</tr>
<tr>
<td>COP1 a</td>
<td>Gm-c1019-2052</td>
<td>531</td>
<td>1350</td>
</tr>
<tr>
<td>AG a</td>
<td>Gm-c1019-25</td>
<td>407</td>
<td>1400</td>
</tr>
<tr>
<td>FCA</td>
<td>Gm-c1015-116</td>
<td>440</td>
<td>1100</td>
</tr>
<tr>
<td>COL1</td>
<td>Gm-c1012-1126</td>
<td>537</td>
<td>1000</td>
</tr>
<tr>
<td>DET1</td>
<td>Gm-c1020-1634</td>
<td>445</td>
<td>900</td>
</tr>
<tr>
<td>CCA1</td>
<td>Gm-c1004-7468</td>
<td>422</td>
<td>1400</td>
</tr>
<tr>
<td>FUS5</td>
<td>Gm-c1027-1042</td>
<td>710</td>
<td>1500</td>
</tr>
<tr>
<td>FPF1 a</td>
<td>Gm-c1018-799</td>
<td>394</td>
<td>800</td>
</tr>
<tr>
<td>GAI a</td>
<td>Gm-c1012-2128</td>
<td>538</td>
<td>700</td>
</tr>
</tbody>
</table>

aMonomorphics in all three populations used in this study
Table 4. BLAST results of the cDNA sequences used as probes in this study

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Incyte Genomics I.D.</th>
<th>cDNA sequences similar to</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL1</td>
<td>Gm-c1012-1126</td>
<td><em>Malus domestica CONSTANS</em>-like protein 1 (<em>COL1</em>)</td>
<td>2e-43</td>
</tr>
<tr>
<td>COL2</td>
<td>Gm-c1012-422</td>
<td><em>A. thaliana CONSTANS</em>-like protein 2 (<em>COL2</em>)</td>
<td>2e-19</td>
</tr>
<tr>
<td>CRY2</td>
<td>Gm-1019-3678</td>
<td>*A. thaliana CRYPTOCHROME 2 mRNA, complete codes</td>
<td>2e-45</td>
</tr>
<tr>
<td>PHYB</td>
<td>Gm-1027-285</td>
<td>*Glycine max PHYB gene exons 1-5, complete codes</td>
<td>0.0</td>
</tr>
<tr>
<td>PHYE †</td>
<td>Gm-c1019-2148</td>
<td>*Lycopersicon esculentum PHYTOCHROME E (<em>PHYE</em>) gene, complete code</td>
<td>3e-04</td>
</tr>
<tr>
<td>PHYA</td>
<td>Gm-c1016-1937</td>
<td>*Phaseolus vulgaris PHYTOCHROME A (<em>PHYA</em>) gene</td>
<td>1e-48</td>
</tr>
<tr>
<td>GT2 †</td>
<td>Gm-c1012-1002</td>
<td>*A. thaliana GT2 gene</td>
<td>2e-04</td>
</tr>
<tr>
<td>FCA</td>
<td>Gm-c1015-116</td>
<td>*A. thaliana FLOWERING TIME PROTEIN FCA</td>
<td>1e-04</td>
</tr>
</tbody>
</table>
Table 4. (continued)

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Incyte Genomics I.D.</th>
<th>cDNA sequence similar to</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET1</td>
<td>Gm-c1020-1634</td>
<td><em>A. thaliana DET1</em> gene, complete codes</td>
<td>1e-12</td>
</tr>
<tr>
<td>CCA1</td>
<td>Gm-c1004-7468</td>
<td><em>A. thaliana CCA1</em> gene</td>
<td>0.013</td>
</tr>
<tr>
<td>FUS5</td>
<td>Gm-c1027-1042</td>
<td><em>A. thaliana FUS5</em> mRNA, complete codes</td>
<td>4e-04</td>
</tr>
<tr>
<td>COP1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Gm-c1019-2051</td>
<td><em>A. thaliana COP1</em> protein</td>
<td>1e-17</td>
</tr>
<tr>
<td>AP2</td>
<td>Gm-c1014-291</td>
<td><em>A. thaliana</em> floral homeotic protein <em>APETALA2 (AP2)</em></td>
<td>8e-54</td>
</tr>
<tr>
<td>FPF1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Gm-c1018-799</td>
<td><em>A. thaliana FPF1</em> protein</td>
<td>6e-08</td>
</tr>
<tr>
<td>AG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Gm-c1019-25</td>
<td><em>A. thaliana AGAMOUS (AG)</em> protein</td>
<td>1e-16</td>
</tr>
<tr>
<td>GAI&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Gm-c1012-2128</td>
<td><em>A. thaliana GAI</em> gene</td>
<td>1e-51</td>
</tr>
<tr>
<td>LD</td>
<td>Gm-r1030-3560</td>
<td><em>A. thaliana LUMINDEPENDENS (LD)</em> protein</td>
<td>4e-34</td>
</tr>
<tr>
<td>VPI&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Gm-c1012-166</td>
<td><em>A. thaliana VPI</em>-like protein</td>
<td>4e-09</td>
</tr>
</tbody>
</table>

<sup>a</sup>Monomorphics in all three populations used in this study
Table 5. Chi-square tests of homologous cloned flowering time gene segregation in population IX132

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Observed ratio</th>
<th>Expected ratio</th>
<th>$X^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>$FCA_J$</td>
<td>35</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>$FCA_{Sty}$</td>
<td>34</td>
<td>58</td>
<td>46</td>
</tr>
<tr>
<td>$FCA_{Hha}$</td>
<td>45</td>
<td>51</td>
<td>48</td>
</tr>
<tr>
<td>$PHYB_{Hha1}$</td>
<td>31</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>$PHYB_{Hha2}$</td>
<td>34</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>$COL2_{Hha}$</td>
<td>44</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>$CCA1_{H}$</td>
<td>51</td>
<td>57</td>
<td>49</td>
</tr>
<tr>
<td>$CCA1_{Pst}$</td>
<td>44</td>
<td>45</td>
<td>44.5</td>
</tr>
<tr>
<td>$AP2_{I1}$</td>
<td>45</td>
<td>48</td>
<td>46.5</td>
</tr>
<tr>
<td>$AP2_{I2}$</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>$AP2_{H}$</td>
<td>51</td>
<td>46</td>
<td>48.5</td>
</tr>
<tr>
<td>$PHYA_{H}$</td>
<td>44</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td>$DETI_{D}$</td>
<td>39</td>
<td>37</td>
<td>38</td>
</tr>
<tr>
<td>$PUS5_{V}$</td>
<td>41</td>
<td>59</td>
<td>50</td>
</tr>
</tbody>
</table>

$^a$D=DraI; H=HindIII; Hha=HhaI; I=EcoRI; V=EcoRV; Pst=PstI; T=TaqI

$^b$A=the allele from the PI 317.336 parent; B=the allele from the 'Corsoy' parent

$^c$The null hypothesis of the test is that the alleles segregate in a 1:1 ratio

$^*$Significantly different from 1:1 at $P=0.05$
**Table 6.** Chi-square tests of homologous cloned flowering time gene segregation in population 1X136

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Observed ratio</th>
<th>Expected ratio</th>
<th>$X^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>$FCA_J$</td>
<td>37</td>
<td>54</td>
<td>45.5</td>
</tr>
<tr>
<td>$FCA_{Sty}$</td>
<td>35</td>
<td>49</td>
<td>42</td>
</tr>
<tr>
<td>$FCA_H$</td>
<td>37</td>
<td>54</td>
<td>45.5</td>
</tr>
<tr>
<td>$FCA_T$</td>
<td>40</td>
<td>51</td>
<td>45.5</td>
</tr>
<tr>
<td>$PHYB_{Hha1}$</td>
<td>31</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>$PHYB_{Hha2}$</td>
<td>33</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>$COL2_{Hha}$</td>
<td>37</td>
<td>27</td>
<td>32</td>
</tr>
<tr>
<td>$CCAI_J$</td>
<td>38</td>
<td>55</td>
<td>46.5</td>
</tr>
<tr>
<td>$PHYA_H$</td>
<td>31</td>
<td>35</td>
<td>33</td>
</tr>
<tr>
<td>$DET1_D$</td>
<td>39</td>
<td>45</td>
<td>42</td>
</tr>
<tr>
<td>$PUS5_V$</td>
<td>47</td>
<td>46</td>
<td>46.5</td>
</tr>
</tbody>
</table>

$D=DraI; H=HindIII; Hha=HhaI; I=EcoRI; V=EcoRV; Pst=PstI; T=TaqI$

$A=$ the allele from the PI 317.334B parent ; $B=$ the allele from the ‘Corsoy’ parent

The null hypothesis of the test is that the alleles segregate in a 1:1 ratio
### Table 7. Chi-square tests of homologous cloned flowering time gene segregation in the F\(_2\)A4 G. max × G. soja population

<table>
<thead>
<tr>
<th>Soybean homolog</th>
<th>Observed ratio</th>
<th>Expected ratio</th>
<th>(X^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>H</td>
<td>B</td>
</tr>
<tr>
<td><strong>FCA_Hha</strong></td>
<td>19</td>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td><strong>PHYB_Hha</strong></td>
<td>11</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td><strong>PHYB_H</strong></td>
<td>13</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td><strong>COL1_D1</strong></td>
<td>18</td>
<td>29</td>
<td>11</td>
</tr>
<tr>
<td><strong>COL1_D2</strong></td>
<td>17</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td><strong>COL1_I</strong></td>
<td>13</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td><strong>COL2_V</strong></td>
<td>14</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td><strong>AP2_D</strong></td>
<td>13</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td><strong>AP2_H1</strong></td>
<td>5</td>
<td>36</td>
<td>10</td>
</tr>
<tr>
<td><strong>AP2_H2</strong></td>
<td>10</td>
<td>27</td>
<td>13</td>
</tr>
<tr>
<td><strong>AP2_I</strong></td>
<td>33</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td><strong>PHYA_H1</strong></td>
<td>6</td>
<td>25</td>
<td>17</td>
</tr>
<tr>
<td><strong>PHYA_H2</strong></td>
<td>23</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td><strong>PHYA_I</strong></td>
<td>13</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td><strong>CRY2_Pst</strong></td>
<td>16</td>
<td>24</td>
<td>13</td>
</tr>
<tr>
<td><strong>LD_Sty</strong></td>
<td>13</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td><strong>PUS5_V</strong></td>
<td>20</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>

\(^a^D=\text{DraI}; \ H=\text{HindIII}; \ Hha=\text{HhaI}; \ I=\text{EcoRI}; \ V=\text{EcoRV}; \ Pst=\text{PstI}; \ T=\text{TaqI}\)
\[ \text{A} = \text{the allele from the } G. \text{ max parent} : \text{B} = \text{the allele from the } G. \text{ soja parent} : \text{H} = \text{heterozygous progeny} \]

\[ \text{The null hypothesis of the test is that the alleles segregate in a 1:2:1 (A:H:B) ratio} \]

\[ \text{Dominant gene segregation} \]

\[ \text{Significantly different from 1:2:1 at } P=0.01 \]
Table 8. Homologous-cloned flowering time gene sequences showing polymorphisms between NILs and their respective RPs and their genotypes

<table>
<thead>
<tr>
<th>Homologous sequences</th>
<th>R.E.</th>
<th>Polymorphic types</th>
<th>NILs' genotypes</th>
<th>RP</th>
<th>RP genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FCA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Syr</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L71-920</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L74-441</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L80-5914</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L63-2404</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L63-3270</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L80-5879</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td><strong>Taql</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L71-920</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L74-441</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L80-5914</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L63-2404</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L63-3270</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L80-5879</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td><strong>HindIII</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L63-3270</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td>L80-5879</td>
<td></td>
<td>DtelE2e1E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE2E3E4e5E7</td>
</tr>
<tr>
<td><strong>PHYB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Hhal</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L66-531</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE1E2E3E4e5E7</td>
</tr>
<tr>
<td><strong>HindIII</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L66-531</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
<td>‘Clark’</td>
<td></td>
<td>DtelE1E2E3E4e5E7</td>
</tr>
<tr>
<td>L67-2324</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
<td>‘Harosoy’</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
</tr>
<tr>
<td>L71-1116</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
<td>‘Harosoy’</td>
<td></td>
<td>DtelE1e2E3E4e5E7</td>
</tr>
</tbody>
</table>

*R.E.=restriction endonuclease; polymorph.=polymorphism; NIL = near isogenic line; RP = recurrent parent.
Figure 1. Location of homologous-cloned flowering time genes on the soybean genetic map based on data from population 1X132 (PI317.336 X ‘Corsoy’). *FCA_I, FCA_T, and FCA_H were also mapped on this site. D = DraI, I = EcoRI, V = EcoRV, Hha = HhaI, H = HindIII, Pst = PstI, Sty = StyI, and T = TaqI.
Figure 2. Location of homologous-cloned flowering time genes on the soybean genetic map based on data from population IX136 (PI317.334B X 'Corsoy'). *FCA_1, and FCA_T were also mapped in this site. D = Dral, I = EcoRI, V = EcoRV, Hha = Hha1, H = HindIII, Pst = PstI, Sty = Sty1, and T = Taq1.
Figure 3. Location of homologous-cloned flowering time genes on the soybean genetic map based on data from G. max X G. soja population. D=DrAI, I=EcoRI, V=EcoRV, Hha=HhaI, H=HindIII, Pst=PstI, Sty=StyI, T=TaqI. Only a portion of the map in each linkage group (LG) is shown. Other markers in each LG are not shown.
Figure 4. Polymorphisms observed between NILs (L71-920 and L74-441) and their respective RP 'Clark' probed with the homologous-gene sequence FCA. Upon analyzing the data it was found that FCA is associated with the maturity locus E3. DNAs were digested with a restriction endonuclease Styl. Blue arrows indicate polymorphic bands. Numbers on the left side of the picture indicate the lambda marker size in kb.
CHAPTER 5. GENERAL CONCLUSION

Traits that affect reproduction (e.g., time of flowering, maturity, and photoperiod insensitivity) in soybean are characters of agronomic interest. These characters are important for developing soybean cultivars with a wider geographical adaptation. The objectives of this study were to: (1) estimate the number of genes controlling photoperiod insensitivity in soybean; (2) map quantitative trait loci (QTL) controlling flowering time, maturity, and photoperiod insensitivity in soybean, and determine if these traits are controlled by the same or different loci in the soybean genome; (3) map homologous and cloned candidate flowering time genes in soybean; and (4) correlate these sequences with maturity ($E$) loci by means of near isogenic lines (NILs). We tested a hypothesis that these homologous gene sequences may be candidate genes controlling the QTL mapped in the study of objective two. Some of these candidate genes may be associated with the known soybean maturity loci.

In the first paper, we estimate the number of genes controlling photoperiod insensitivity in soybean. Two single-cross populations, IX132 (PI 317.336 X 'Corsoy'), and IX136 (PI 317.334B X 'Corsoy') were developed and were used in this study. The populations were inbred to obtain 101 and 100 $F_{6:7}$ lines using a modified single seed descent. Flowering time (days to R1) of the RI lines from each population was observed in the growth chamber at 12 and 20 h photoperiods using a combination of fluorescent and incandescent lamps. The RI lines show dramatically different responses to day length. A normal distribution of flowering times was observed when the lines were grown in the growth chamber with a 12 h photoperiod. When the lines were grown in the growth chamber with a 20 h photoperiod, however, a discontinuous distribution was observed. This suggested
the insensitivity of the RI lines to long day length might be controlled by a few major genes. The time of flowering was delayed in almost all lines when grown in the growth chamber with a 20 h photoperiod compared to those grown in the growth chamber with a 12 h photoperiod. The flowering delays were 5 to 75 days in population IX132 and 0 to 75 days in population IX136. From analyses of the segregation data, we proposed a minimum of three genes control photoperiod insensitivity in both populations and thus makes this an acceptable target for QTL analyses.

In the second paper, the two RI lines used in the first study were used to map QTL controlling time of flowering, maturity, and photoperiod insensitivity in soybean. Days to R1, R3, and R7 (Fehr and Caviness 1977) were observed among the F$_6$;7 RI lines in the field during 1991 and 1992 and in the growth chamber at 12 and 20 h photoperiods using fluorescent and incandescent lamps. A total of 139 markers (88 RFLPs and 51 SSRs) in the IX132 population and 125 markers (73 RFLPs and 52 SSRs) in the IX136 population were used to map QTL affecting these traits. A large-effect QTL for days to R1, R3, R7, and photoperiod insensitivity was found at the same location on linkage group (LG) C2 in both populations. This QTL explained as much as 47% of total phenotypic variance. This result suggests that photoperiod insensitivity, flowering time, and maturity may be controlled by the same gene(s) or by tightly clustered genes in the same chromosomal region. In addition to the large effect QTL, minor QTL were also detected that controlled the four traits in both populations. Minor QTL account for as much as 17.8% and 12.1% of phenotypic variance in populations IX132 and IX136, respectively. Thus, time of flowering, maturity, and photoperiod insensitivity in these soybean populations are proposed to be controlled by a major QTL with a large effect and modified by several minor QTL. The large effect QTL
found on LG C2 should facilitate marker-assisted selection (MAS) of these traits. Satt205 links tightly with the QTL and should be useful in the MAS.

In the third paper, we mapped onto the soybean genetic map genes known to be involved in photoperiod recognition and time of flowering and then correlated the mapped homologous gene sequences with the maturity (E) loci by means of NILs. Three single-cross populations consisting of the two F0,7 RI lines used in the first and second studies and an F2 population of an interspecific cross between G. max and G. soja were used in this study. In addition, 41 NILs and 2 RPs were used in the correlation study. Eighteen soybean expressed sequence tags (ESTs) identified by BLAST to have high similarities with 18 previously cloned heterologous flowering time genes were used as probes in this study. Ten of the 18 cDNA clones now have been mapped. The homologous sequences were mapped onto LGs A2 (CRY2), B1 and H (COL1), A1 and B2 (PHYA), C1 (DETI and LD), D2 (AP2), E and K (PHYB), F (COL2), L (FCA), and Q (CCA1). None of these cDNA homologous sequences have been found to be directly associated with previously mapped QTL for flowering time (Tasma et al. 2001) and, therefore, we reject the hypothesis that these homologous gene sequences may be candidate genes of the mapped flowering time QTL. However, analyses of these sequences using NILs show an association between the homologous gene sequence FCA with maturity locus E3. The map positions and phenotypic data support that homologous gene sequence FCA is a strong candidate for maturity locus E3. Analyses of NILs suggest that PHYB homolog may be associated with maturity locus E1. However, current data show they mapped in different LGs.

Interestingly, these soybean maturity loci are the two most important loci in terms of their effects in delaying flowering and maturity in response to inductive and non-inductive
day lengths. *E3* was reported to be a locus controlling photoperiod insensitivity in response incandescent long day length (ILD) (Saindon et al. 1989; Cober et al. 1996a). ILD is a natural day length extended to 20 h with incandescent light. Under natural day length, *E1* was reported to delay flowering and maturity the most, compared to the effects of other maturity loci (Bernard 1971). In addition, *E1* was reported to be the most sensitive maturity locus to changes in light quality, mainly in response to low red:far-red (R:FR) ratio, a light quality similar to natural light. *E3*, on the other hand, was reported to be the most insensitive to changes in light quality (Cober et al. 1996b). Because of their significant importance for breeding soybean cultivars with a wider adaptation, further investigations on these two maturity loci need to be done. This is to determine the nature of these associations and to explore these maturity loci further at the molecular level.

Analysis of NILs strongly indicated that *PHYB* is associated with maturity locus *E1*. The map position of *E1* (LG C2) has previously been shown to correlate with a QTL for flowering time (Tasma et al. 2001). However, *PHYB* was mapped on two LGs (E and K) (this study) and *E1* was mapped on LG C2 (Cregan et al. 1999). This suggests that we may have mapped duplicate loci of *PHYB* paralogs. Soybean has been recognized as an ancient polyploid and has been reported to be rich in duplicated loci (Shoemaker et al. 1996). Common RFLP markers are detected in these regions indicating the two mapped *PHYB* are duplicated genes. A report by Shoemaker et al. (1996) that LG E and LG K contain homoeologous regions (i.e., RFLP markers in common) further supports the hypothesis that the two mapped *PHYB* are duplicated loci. A similar situation was observed between homoeologous regions on LGs B1 and H (Shoemaker et al. 1996) to which duplicate copies of the *COL1* homolog were mapped. There may be more than two copies of *PHYB*
homologous sequences in the soybean genome. It is possible, therefore, that we may have failed to map the \textit{PHYB} homolog on LG C2. Another possibility is that another gene (possibly an unknown maturity locus) has been backcrossed during the NIL development, along with \textit{EI}. Therefore, we identified polymorphisms around the other gene and it is the other gene that corresponds to \textit{PHYB}.

In this study, a minimum of three genes were proposed to control photoperiod insensitivity in soybean. Time of flowering, maturity, and photoperiod insensitivity in these soybean populations are proposed to be controlled by a major QTL with a large effect and modified by several minor QTL. Ten flowering time gene homologs have been mapped in the soybean genome, suggesting they are conserved across plant genera. We have also mapped homologs of \textit{PHYB} and \textit{COL1} to homoeologous genome regions. This further supports the notion that the soybean genome is rich in duplicated regions. The homologous gene sequence \textit{FCA} is a strong candidate for maturity locus \textit{E3}. The \textit{PHYB} homologous gene sequence may associate with maturity locus \textit{EI}, but current evidence shows they map in different LGs.
REFERENCES CITED


signals that induce flowering. Plant Cell 5: 1147-1155.


protein and starch concentration in the Illinois Long Term selection maize strains.


polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18: 6531-6535.


ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professor Dr. Randy Shoemaker for the opportunity given to me to be one of his lab members. I have learned a lot of things in the laboratory. I also highly appreciate his guidance, patience and time allocation. He almost always provides time for me to discuss research problems and other academic matters as long as my Ph.D. program. I also thank all of my program of study (POS) committee members: Dr. Charlie Brummer, Dr. Paul Scott, Dr. James Colbert, and Dr. Roger Wise for serving as committee and the constructive critiques of this dissertation. I thank all the Shoemaker Lab members for their friendship and cooperation as long as my Ph.D. investigation.

I thank the government of Indonesia through the Agency for Agricultural Research and Development, Department of Agriculture for financial support. I thank Dr. Joko Budianto, Dr. Djoko Damardjati, Dr. Endang S. Thohari, Dr. Mappaona, Ir. Sumarsini, and Ir. Hetty Herawati for their help in managing my study. In addition I thank Dr. Pasril Wahid, Ir. Hasnam, Ph.D., Dr. Azmi Dhalimi, and Dr. Nurliani Bermawie for their permission and encouragement to attend a Ph.D. program at Iowa State University.

Finally I thank my parents, brothers and sisters for their love and pray throughout my study. I also appreciate numerous friends here in Ames for their friendships.