RESEARCH NOTE

Genetic analyses for deciphering the status and role of photoperiodic and maturity genes in major Indian soybean cultivars

SANJAY GUPTA[1]*, VIRENDER SINGH BHATIA[1], GIRIRAJ KUMAWAT[1], DEVSHREE THAKUR[2], GOURAV SINGH[1], RACHANA TRIPATHI[1,3], GYANESH SATPUTE[1], RAMGOPAL DEVADAS[4], SAYED MASROOR HUSAIN[1] and SURESH CHAND[3]

[1]*Indian Council of Agricultural Research - Indian Institute of Soybean Research, Khandwa Road, Indore 452 001, India*
[2]*Government Higher Secondary School, Jhalmala, Tehsil, Balod, Chhatisgarh 491 226, India*
[3]*Devi Ahilya Vishwavidyalaya, Indore 452 001, India*
[4]*Indian Council of Agricultural Research - National Research Centre for Orchids, Pakyong 737 106, India*

Abstract

Allelic combinations of major photoperiodic (*E1*, *E3*, *E4*) and maturity (*E2*) genes have extended the adaptation of quantitative photoperiod sensitive soybean crop from its origin (China ~35°N latitude) to both north (up to ~50°N) and south (up to 40°S) latitudes, but their allelic status and role in India (6–35°N) are unknown. Loss of function and hypoactive alleles of these genes are known to confer photoinsensitivity to long days and early maturity. Early maturity has helped to adapt soybean to short growing season of India. We had earlier found that all the Indian cultivars are sensitive to incandescent long day (ILD) and could identify six insensitive accessions through screening 2071 accessions under ILD. Available models for ILD insensitivity suggested that identified insensitive genotypes should be either *e3/e4* or *e1* (*e1-nl* or *e1-fs*) with either *e3* or *e4*. We found that one of the insensitive accessions (EC 390977) was of *e3/e4* genotype and hybridized it with four ILD sensitive cultivars JS 335, JS 95-60, JS 93-05, NRC 37 and an accession EC 538828. Inheritance studies and marker-based cosegregation analyses confirmed the segregation of *E3* and *E4* genes and identified JS 93-05 and NRC 37 as *E3E3E4E4* and EC 538828 as *e3e3E4E4*. Further, genotyping through sequencing, derived cleaved amplified polymorphic sequences (dCAPS) and cleaved amplified polymorphic sequences (CAPS) markers identified JS 95-60 with hypoactive *e1-as* and JS 335 with loss of function *e3-fs* alleles. Presence of photoperiodic recessive alleles in these two most popular Indian cultivars suggested for their role in conferring early flowering and maturity. This observation could be confirmed in F_2 population derived from the cross JS 95-60 × EC 390977, where individuals with *e1-as e1-as* and *e4e4* genotypes could flower 7 and 2.4 days earlier, respectively. Possibility of identification of new alleles or mechanism for ILD insensitivity and use of photoinsensitivity in Indian conditions have been discussed.

[Gupta S., Bhatia V. S., Kumawat G., Thakur D., Singh G., Tripathi R., Satpute G., Devadas R., Husain S. M. and Chand S. 2017 Genetic nalyses for deciphering the status and role of photoperiodic and maturity genes in major Indian soybean cultivars. *J. Genet.* **96**, xx–xx]

Introduction

Soybean (*Glycine max* (L.) Merr.) is a short-day plant and its different genotypes start to flower when the day length is less than their critical day length. It is grown worldwide from equator to 50°N and 35°S latitudes (Watanabe *et al.* 2012). The crop has adapted to such a wider range of latitudes through its natural variation in the major genes and quantitative trait loci (QTLs) controlling flowering time and maturity. To date, 10 major genes controlling flowering time and maturity have been indicated in soybean: *E1* and *E2* (Bernard 1971), *E3* (Buzzell 1971), *E4* (Buzzell and Voldeng 1980), *E5* (McBlain and Bernard 1987), *E6* (Bonato and Vello 1999), *E7* (Cober and Voldeng 2001), *E8* (Cober *et al.* 2010), *E9* (Kong *et al.* 2014) and *J* (Ray *et al.* 1995). Of these 10 genes, *E1*, *E3*, *E4* and *E7* have been reported as quantitative photoperiodic genes (Cober and Voldeng 2001) with dominant alleles conferring photosensitivity. While dominant alleles at *E1*, *E2*, *E3*, *E4*, *E5*, *E7* and *E8* loci delay time to flowering, recessive alleles at *E6*, *E9* and *J* loci delay flowering time to different extents, interacting with the environment and with genotypes at other loci (Watanabe *et al.* 2012; Kong *et al.* 2014).

The genes *E3* and *E4* were originally identified by different responses of flowering to long day conditions, which were

*For correspondence. E-mail: sanitaishu@gmail.com.

Keywords. quantitative trait gene; inheritance; markers; molecular breeding; maturity genes.

induced by (i) light with a high red (R) to far-red (FR) quantum ratio generated by R-enriched fluorescent lamps and by (ii) light with a low R : FR ratio generated by FR-enriched incandescent lamps (Buzzell 1971; Buzzell and Voldeng 1980; Cober *et al.* 1996a). *E3* controls flowering under LD conditions with a high R : FR ratio, where *e3e3* recessive homozygous plants can initiate flowering under these conditions (Buzzell 1971). *E4* is involved in flowering under LD conditions with a low R : FR ratio and a recessive *e4* is necessary for plants homozygous for the *e3* to flower under these conditions (Buzzell and Voldeng 1980; Saindon *et al.* 1989a; Cober *et al.* 1996a, b). Two of the photoperiodic genes *E3* and *E4* have been identified as *Phytochrome A3* and *Phytochrome A2* genes, respectively (Liu *et al.* 2008; Watanabe *et al.* 2009). Allele-specific markers for these genes have been developed for genetic analysis and molecular breeding in soybean (Liu *et al.* 2008; Watanabe *et al.* 2009; Tsubokura *et al.* 2013; Xu *et al.* 2013). At *E3*, three mutations *e3-tr*, *e3-fs* and *e3-ns* have been reported (Watanabe *et al.* 2009 Xu *et al.* 2013). Allele *e3-tr* is a deletion mutation that lacks the 3′ region of the gene including exon 4, *e3-fs* is a frame-shift mutation that introduces a stop codon in exon 1 and *e3-ns* is a nonsense mutation in which a single nucleotide substitution in exon 3 creates a stop codon in place of a codon encoding glutamine. At *E4*, four single base-pair deletion mutations (*e4-kes*, *e4-kam*, *e4-tsu* and *e4-oto*) resulting in premature stop codons have been reported (Tsubokura *et al.* 2013).

E1 gene has also been reported to respond to ILD and it interacts with *E3* and *E4* for conferring ILD insensitivity in soybean (Cober *et al.* 1996a). *E1* has been characterized and found to encode legume-specific transcription factor that has a putative nuclear localization signal and B3, distantly related domain (Xia *et al.* 2012). Three mutations *e1-as*, *e1-nl* and *e1-fs* have been reported at this locus. *e1-as* is a hypoactive allele which causes loss of nuclear localization domain in E1 protein that results in its reduced activity. The *e1-fs* is a frame-shift mutation that causes 1-bp deletion in codon 17 and results in a premature stop codon. The *e1-nl*, a null allele, results from deletion of ~130 kb (including the entire *E1* gene) region (Xia *et al.* 2012). Molecular basis of photoperiodic locus *E7* is still unknown. *E2* is an orthologue of *Arabidopsis GIGANTEA* gene (*GI*) (Watanabe *et al.* 2011). Recessive genotype at this locus causes early flowering by inducing the expression of soybean florigen homologue, *FT2a*, whereas the effect of *E2* allele on flowering under different environments remains stable (Watanabe *et al.* 2011).

Different soybeans adapting to a narrow latitudinal band have evolved with diverse combination of these genes and QTLs that control flowering behaviour (Tsubokura *et al.* 2013). Precision breeding for developing varieties for a specific area would involve identification of combinations of these genes suitable for that area and their incorporation during breeding process (Saindon *et al.* 1989b; Tsubokura *et al.* 2014). Soybean has originated in China which shares its boundaries with India. Soybean is known to be sporadically grown in Indian states, neighbouring China since ages (Piper and Morse 1910; Hymowitz 1970). However, the feasibility of commercial soybean cultivation in India could be demonstrated through introduction of varieties like Bragg, Clark 63, Lee, Improved Pelican, Davis, Hardee from USA in early sixties. These introduced varieties have served as founder stock by becoming parents in many of the present day Indian soybean cultivars. Soybean is grown as rainfed crop in India and key to its adaptation has been the development of early maturing (~85–100 days) cultivars which complete their life cycle in short growing season. Understanding the genetic basis of earliness in Indian cultivars would help to develop the targeted breeding programmes.

Presently, India is the fifth largest soybean producing country after US, Brazil, Argentina and China (FAOSTAT 2013, http://faostat3.fao.org/browse/rankings/countries_by_commodity/E; Statista 2016, http://www.statista.com/statistics/267270/production-of-soybeans-by-countries-since-2008/). In spite of the characterization of major photoperiodic genes, their status and role have remained unestablished in Indian cultivars, largely due to their nonavailability of genetic stocks of these genes and their similar phenotypic effects on flowering and maturity. In our earlier work, we had identified six ILD insensitive accessions (Singh *et al.* 2008), which showed little effect of ILD at Indore and over various locations spread over large latitudinal range (15.27° to 29°N) for days to R1. We observed that all the Indian cultivars were ILD-sensitive. For genetic analysis of photoperiod trait in identified photoinsensitive lines, the available literature suggested that two genotypic groups: (i) *e3e4* (Buzzell 1971; Buzzell and Voldeng 1980; Saindon *et al.* 1989a; Cober *et al.* 1996a; Abe *et al.* 2003) and (ii) *e1* (*e1-nl* or *e1-fs*) with either *e3* or *e4* confer photoperiod insensitivity to ILD (Xu *et al.* 2013). With the availability of markers and sequences for these genes, we proceeded with genetic analysis of these genes and report their status and putative role in major Indian cultivars.

Material and methods

Genetic analysis at E3 and E4 loci and assessment of effect of photoperiodic genes for days to R1

Six earlier identified photoinsensitive accessions (MACS 330, EC 325097, EC 333897, EC 34101, EC 325118 and EC 390977), four widely cultivated ILD-sensitive cultivars (JS 335, JS 95-60, JS 93-05 and NRC 37) and one ILD-sensitive exotic accession (EC 538828) were genotyped using fragment length polymorphic (FLP) markers for *E3* and *E4* genes (Liu *et al.* 2008; Watanabe *et al.* 2009). Sources of these accessions and pedigree of the cultivars are given in table 1. A double recessive (*e3e3e4e4*) photoperiod insensitive accession EC 390977 was hybridized with these cultivars and accession EC 538828. F_1s and F_2s were evaluated under 20-h ILD at experimental farm of ICAR-Indian Institute of Soybean Research, Indore (22.4°N),

Table 1. Details of seven soybean accessions and four cultivars used for *E1* gene sequencing with GenBank accession numbers of their *E1* coding DNA sequence.

Accession/ cultivar name	Origin/ source	Pedigree	GenBank accession no. for *E1* gene sequence
MACS 330	India	Indigenous collection	KU312397
EC 325097	Hungary	Exotic collection	KU312398
EC 333897	USA	Exotic collection	No amplification
EC 34101	Hungary	Exotic collection	KM386867
EC 325118	Hungary	Exotic collection	KM386868
EC 390977	Taiwan	Exotic collection	KM386869
JS 335	India	JS 78-77 × JS 71-5	KU312396
JS 93-05	India	Secondary selection from PS 73-22	KM386863
JS 95-60	India	Selection from PS 73-22	KM386864
NRC 37	India	Gaurav × Punjab 1	KM386871
EC 538828	Japan	Exotic collection	KM386858

India, during 2011–12. Extended photoperiod was created by providing lighting with 40 W incandescent bulbs at a height of 3 feet above the crop canopy and the bulbs were connected to an automatic timer. Planting was done in a single row along the row of bulbs and 10 cm distance between plants was maintained. Recommended package of practices were followed. Data for days to R1 were recorded daily. The F_2 ratios of 3 : 1 and 15 : 1, for photoperiod sensitive to insensitive plants were tested for segregation of one and two genes, respectively. For assessing the effect of *E1*, *E3* and *E4* loci on days to R1, the F_2 population of JS 95-60 × EC 390977 was phenotyped under 15 h of ILD during 2012–13.

Genomic DNA from parents, F_1 and F_2 of three crosses namely NRC 37 × EC 390977, JS 93-05 × EC 390977 and EC 538828 × EC 390977 were isolated by cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1990). Genotyping at *E3* and *E4* loci was done through allele-specific (Fragment Length Polymorphic markers) (Liu *et al.* 2008; Watanabe *et al.* 2009). Since, allele specific marker for *E3* could amplify *e3* allele only, therefore, polymorphism between photoinsensitive and photosensitive parents was established using SSR markers present near *E3* locus (table 2). For assessing the effect of *E1*, *E3* and *E4* genes on days to flower, F_2 population from cross JS 95-60 × EC 390977 was genotyped using SSR markers for *E1* and *E3* and FLP marker for *E4*. For SSR amplification, PCR reactions were carried out using the following cycling parameters: initial denaturation at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 55°C for 2 min, 72°C for 2 min and finally, a primer extension cycle of 10 min at 72°C. The amplification products were separated on 3% metaphor agarose gels. Gels were run for 3 h at 120 V in 1× TBE buffer. The size of the fragments was estimated using a 50-bp DNA ladder (Thermo Fisher Scientific, Bengaluru, India).

Allelic characterization at E1, E2, E3 and E4 loci

PCR amplified DNA fragments of these 11 accessions and cultivars, from primers 5′TGCCTTCACTTCCATTTCACA3′ and 5′AGGTTGAAGTACACGCTATTGC3′, were outsourced (Scigenom, Bengaluru, India) for Sanger sequencing of coding region of *E1*. Multiple sequence alignment tool ClustalW2 was used to identify different alleles of *E1*.

Genotyping at *E2*, *E3* and *E4* loci was further done using *e2* (dCAPS: Derived Cleaved Amplified Polymorphic Sequences), *e3-fs* Cleaved Amplified Polymorphic Sequences, *e3-ns* (dCAPS), *e4-oto*, *e4-kam*, *e4-kes* (CAPS) and *e4-tsu* (dCAPS) markers (Watanabe *et al.* 2011; Tsubokura *et al.* 2013; Xu *et al.* 2013).

Table 2. SSR markers used for detecting polymorphism at *E3* locus between parents of cross NRC 37 × EC 390977 and JS 93-05 × EC 390977.

SSR name	Forward primer sequence (5′ → 3′)	Reverse primer sequence (5′ → 3′)
BARCSOYSSR_19_1430	AAGCGCCCTTTCAGTTTATG	TGCAGACAAAACGATAGCAAA
BARCSOYSSR_19_1447	CACAATATAATTGAGAGACACTTTCAT	CCCAAGTTTTCATTGTCTCAA
BARCSOYSSR_19_1462	CATAACTTCATTACAATTTTTACACCA	TGGATAAACTAGGTTTTTGGCTT
BARCSOYSSR_19_1478	GTTTGGCTGGAAGGATGTGGT	TCTCTTTCCAACAAGAAGTCGTC
BARCSOYSSR_19_1499	CATCAATTTTATCGATATTCTACACC	TTTTGGAAATGGAAGAACTACTTAAA
BARCSOYSSR_19_1513	CCCTCTCCCTCTTTGAATCC	TTGCCACCAAGGTTGATGTA
BARCSOYSSR_19_1527	TTTCCTCTAATAAACATAATGTCGAG	AAATTGTGAGATTAATGGGAATG
Satt 229	TGGCAGCACACCTGCTAAGGGAATAAA	GCGAGGTGGTCTAAAATTATTACCTAT
Satt 006	CAATGTGATTAGTTTTGGAAA	GGGTTAATGTTGTTTTTTATA
Satt 664	GCGTAGATGCTCAACATCAACACTAATCTG	GCGGACGATGAAGAAATATACTATTACGA
Satt 373	TCCGCGAGATAAATTCGTAAAAT	GGCCAGATACCCAAGTTGTACTTGT

Table 3. Genotypes and phenotypes at *E3* and *E4* loci using FLP markers.

Genotype at *E3* and *E4* loci	Phenotype of cultivar	Accession
e3e3E4E4	ILD sensitive	EC 538828
	ILD insensitive	MACS 330, EC 325097
–e4e4	ILD sensitive	–
	ILD insensitive	EC 333897
–E4E4	ILD sensitive	JS 335, JS 95-60, JS 93-05, NRC 37
	ILD insensitive	EC 34101, EC 325118
e3e3e4e4	ILD sensitive	–
	ILD insensitive	EC 390977

–, no amplification at *E3*.

Table 4. Segregation for photoinsensitivity in F_2 populations of crosses developed from photosensitive and photoinsensitive parents.

Parental combination of cross	F_2 segregation (sensitive : insensitive)	Segregation tested	χ^2 probability
NRC 37 × EC 390977	144 : 6	15 : 1	0.2549
JS 93-05 × EC 390977	64 : 4	15 : 1	0.9003
JS 335 × EC 390977	88 : 5	15 : 1	0.7278
EC 390977 × JS 95-60	41 : 4	15 : 1	0.4645
EC 390977 × EC 538828	24 : 11	3 : 1	0.3797

Table 5. The χ^2 testing of segregation (1 : 2 : 1) of alleles at *E3* and *E4* loci in F_2 populations.

Cross name	Genotype	Observed ratio	χ^2 probability
JS 93-05 × EC 390977	*E4E4* : *E4e4* : *e4e4*	19 : 30 : 36	0.7183
	E3E3 : *E3e3* : *e3e3*	14 : 32 : 19	0.6755
NRC 37 × EC 390977	*E4E4* : *E4e4* : *e4e4*	28 : 44 : 33	0.1990
	E3E3 : *E3e3* : *e3e3*	39 : 61 : 27	0.2916
EC 390977 × EC 538828	*E4E4* : *E4e4* : *e4e4*	5 : 7 : 11	0.0359

Results

Identification of genotype of Indian cultivars at E3 and E4 loci

Recessive *e3e3* was observed in ILD-insensitive MACS 330, EC 325097, EC 390977 and sensitive EC 538828 accessions (table 3). Genotype of rest of the accessions and cultivars could not be determined at *E3* locus because of nonamplification of the product. Recessive *e4e4* was present in two ILD insensitives (EC 333897 and EC 390977) only and rest of the accessions and cultivars had *E4E4* genotype. Double recessive photoinsensitive accession EC 390977 (*e3e3e4e4*) was hybridized with photosensitive cultivars and accession (EC 538828). All the F_1s were sensitive to ILD. In F_2, phenotypic segregation ratio of 15 sensitive to 1 insensitive was observed in crosses involving cultivars and that of 3 : 1, in EC 538828 × EC 390977 (table 4).

To confirm that the segregating genes are *E3* and *E4*, three F_2 populations from NRC 37, JS 93-05 and EC 538828 were genotyped for *E3* and *E4*. At *E3* locus, SSR markers BARCSOYSSR_19_1447 and BARCSOYSSR_19_1527 were identified as polymorphic between parents and the former was used for genotyping. The validity of this marker was further established in F_2 population of JS 93-05 × EC 390977 by correlating its genotype with that of FLP marker. Plants with *–e3* genotype through FLP marker were either *E3e3* or *e3e3* through SSR marker. This suggested for the close location and robustness of the marker for its use in our population. Events of crossing over for estimation of distance between SSR locus and *E3* gene (FLP marker) could not be identified as presence of *e3* through FLP meant either *E3e3* or *e3e3* genotype. However, if the absence of band through fragment length *E3* marker is considered as the presence of *E3E3* (dominant) genotype, we identified that five plants were of this genotype, but were identified as *E3e3* by SSR marker. These plants should have arisen due to crossing over. Of the 65 total F_2 plants, five have been identified as result of crossing over. This suggests that SSR marker BARCSOYSSR_19_1447 is at least 7.7 cM away from the *E3* and may be used for genotyping purpose.

Table 6. Summary of marker genotypes and ILD phenotypes in F_2 of crosses NRC 37 × EC 390977 and JS 93-05 × EC 390977.

	NRC 37 × EC 390977		JS 93-05 × EC 390977	
Genotype	Number of plants	Phenotype	Number of plants	Phenotype
E3E3E4E4	9	Photosensitive	5	Photosensitive
E3E3E4e4	7	Photosensitive	6	Photosensitive
E3e3E4E4	10	Photosensitive	8	Photosensitive
E3e3E4e4	23	Photosensitive	16	Photosensitive
e3e3E4E4	9	Photosensitive	6	Photosensitive
e3e3E4e4	6	Photosensitive	8	Photosensitive
E3E3e4e4	11	Photosensitive	3	Photosensitive
E3e3e4e4	12	Photosensitive	9	Photosensitive
e3e3e4e4	5	Photoinsensitive	4	Photoinsensitive
e3e3e4e4	1*	Photosensitive	–	–

*Photosensitive phenotype identified as insensitive by genotyping.

Molecular markers for *E3* and *E4* segregated in expected 1 : 2 : 1 ratio (table 5) in crosses involving EC 390977 with JS 93-05 and NRC 37, but segregation distortion was observed in EC 538828 × EC 390977, probably because of small population size. In all the three F_2 populations, photoperiod insensitive individuals had the *e3e3e4e4* genotype and all the photoperiod sensitive individuals had one of the other possible seven genotypes. One photoperiod sensitive plant from NRC 37 × EC 390977 was identified as insensitive through molecular marker (table 6).

Allelic characterization of genotypes at E1, E2, E3 and E4 loci

Sequencing of *E1* coding region revealed that all the accessions and cultivars were *E1E1* except for JS 95-60 and EC 325097 which have *e1-as e1-as* and EC 333897, where amplification product was not obtained for sequencing (table 7). CAPS and dCAPS markers identified recessive genotype for *E2* in four photoperiod insensitive (MACS 330, EC 325097, EC 333897 and EC 325118) and one sensitive (EC 538828) accessions. All the insensitive accessions had recessive *e3-tr* genotype at *E3* locus except for EC 333897. Among sensitive cultivars, only JS 335 had recessive (*e3-fs e3-fs*) genotype. Of all the cultivars and accessions, only two insensitive accessions EC 325097 and EC 325118 had recessive *e4-kes* and *e4-kam* alleles, respectively, at *E4* locus.

Effect of E1, E3 and E4 loci on flowering under ILD

EC 390977 and JS 95-60 flowered in 37 and 58 days, respectively. Days to R1 in F_2 ranged from 33 to 80 days with a mean of 55.4 days. Two SSR markers for *E1* locus identified a mean delaying effect of seven days in *E1E1* genotype as compared to recessive genotypes (*e1-as e1-as*) (figure 1). Mean of heterozygotes was at par with that of homozygous dominant genotype at this locus. *E4E4* genotype caused a mean delay of only 2.4 days over *e4e4*.

Discussion

Maturity gene-specific stocks were not available with us, but we had identified six ILD insensitive accessions from

Table 7. Genotypes at *E1*, *E2*, *E3* and *E4* loci through sequencing and CAPS and dCAPS markers.

Accession/ cultivar	*E1* coding region sequencing	CAPS and dCAPS genotype at loci			Final genotype with sequencing and all markers
		E2	*E3*	*E4*	
Photoinsensitive accessions					
MACS 330	*E1E1*	*e2e2*	*e3-tr e3-tr*	*E4E4*	*E1E1 / e2e2 / e3e3 / E4E4*
EC 325097	*e1-as e1-as*	*e2e2*	*e3-tr e3-tr*	*e4-kes e4-kes*	*e1e1 / e2e2 / e3e3 / e4e4*
EC 333897	—*	*e2e2*	*E3E3*	*E4E4*	*– / e2e2 / E3E3 / e4e4*
EC 34101	*E1E1*	*E2E2*	*e3-tr e3-tr*	*E4E4*	*E1E1 / E2E2 / e3e3 / E4E4*
EC 325118	*E1E1*	*e2e2*	*e3-tr e3-tr*	*e4-kam e4-kam*	*E1E1 / e2e2 / e3e3 / e4e4*
EC 390977	*E1E1*	*E2E2*	*e3-tr e3-tr*	*E4E4*	*E1E1 / E2E2 / e3e3 / e4e4*
Photosensitive accessions					
JS-335	*E1E1*	*E2E2*	*e3-fs e3-fs*	*E4E4*	*E1E1 / E2E2 / e3e3 / E4E4*
JS 95-60	*e1-as e1-as*	*E2E2*	*E3E3*	*E4E4*	*e1e1 / E2E2 / E3E3 / E4E4*
JS 93-05	*E1E1*	*E2E2*	*E3E3*	*E4E4*	*E1E1 / E2E2 / E3E3 / E4E4*
NRC-37	*E1E1*	*E2E2*	*E3E3*	*E4E4*	*E1E1 / E2E2 / E3E3 / E4E4*
EC 538828	*E1E1*	*e2e2*	*E3E3*	*E4E4*	*E1E1 / e2e2 / e3e3 / E4E4*

*No amplification of product for sequencing.

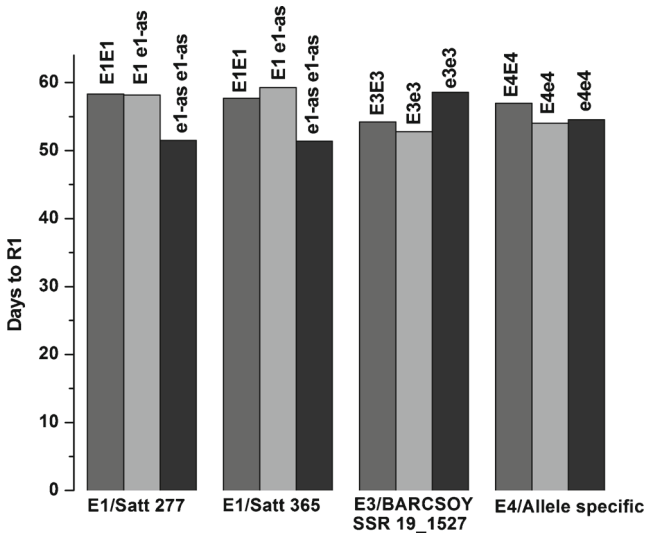

Figure 1. Effect of photoperiodic genes on days to R1 in F_2 from JS 95-60 × EC 390977.

our earlier screening of 2071 soybean accessions from 2002 to 2007. These accessions had exhibited no effect of ILD extended day length at Indore and latitudes in multilocation testing. Indian cultivars were highly responsive to these conditions and expressed delayed flowering in higher latitudes and ILD extended day length. Available information up to year 2007 suggested that genotypes of these accessions should be *e3e3e4e4*. With the accessibility of the allele-specific markers for *E3* and *E4* (Liu *et al.* 2008; Watanabe *et al.* 2009), we could determine that one of the insensitive accessions (EC 390977) has this combination and used it to identify the alleles at *E3* and *E4* loci in popular Indian cultivars. These markers identified the presence of *E4E4* in all the cultivars, but genotype at *E3* could not be detected because of amplification of only *e3* allele. Since *E4* FLP marker scores only a single mutation event and genotype for *E3* was unknown, to determine their status in Indian soybean cultivars, we undertook inheritance and cosegregation studies involving double recessive genotype EC 390977. Inheritance studies identified the segregation of two genes in all the crosses involving cultivars and that of one gene in cross involving EC 538828. Cosegregation analysis involving *E4* FLP and *E3* SSR markers (BARCSOYSSR_19_1447) could confirm that segregating genes were indeed *E3* and *E4*. From this study, we could confirm that genotype of JS 93-05 and NRC 37 is *E3E3E4E4* and that of EC 538828 is *e3e3E4E4*. Although F_2 involving JS 335 and JS 95-60 also segregated in 15 sensitive : 1 insensitive, cosegregation studies or sequencing is required to confirm the involvement of *E3* and *E4*. Although, soybean genome has been sequenced, yet the polymorphism information on BARCSOYSSR_19_1447 would be useful to breeders, while using these accessions and cultivars in their breeding programmes.

In our later genotyping, we found that JS 335 has recessive *e3-fs* allele and hence, its F_2 should have segregated into 3 : 1 instead of observed 15 : 1. JS 335 has been a mega soybean cultivar of India and adapts to a wide range of latitudes in India. This finding calls for further research to find if yet another ILD sensitivity mechanism is present in JS 335. In our insensitive accessions, we could find EC 325097, EC 325118 and EC 390977 to fit in the *e3e4* insensitivity model, but all of them were insensitive to ILD and latitudes. These remaining genotypes also offer the possibility for identification of another insensitivity mechanism or new alleles at these loci.

Hypoactive *e1-as* allele was identified in cultivar JS 95-60 which flowers in 25–27 days after sowing and matures in 85–90 days. In central India, where cultivar JS 95-60 rules (35% of total Indian soybean breeder seed requisition), farmers require cultivar of early maturity to accommodate potato and wheat/chickpea after the harvest of soybean. Presence of *e1-as e1-as* genotype in this cultivar may be one of the reasons for its early flowering. Our recent finding that accession IC 15089, with almost similar flowering and maturity duration as that of JS 95-60, has *e1-as* allele also supports the role of this recessive allele in conferring earliness (unpublished data). *E1* is a major photoperiodic gene, controlling days to flower and delaying effect up to 25 days, has been reported for its dominant allele under ILD conditions (Bernard 1971; Cober *et al.* 1996a). Observed mean difference of 7 days in flowering for *E1E1* genotype from homozygous recessives in both the SSR markers may be because of *e1-as* allele. The allele *e1-as*, encodes a protein that is dysfunctional in its nuclear localization because of a point mutation in the putative nuclear localization signal (Xia *et al.* 2012). These researchers found that *e1-as* is a leaky allele and may retain partial *E1* function.

Soybean originated in China at ~35°N and its further spread more towards northern latitudes, with longer day conditions, became possible through loss of function mutations in photoperiodic genes (Cober *et al.* 2010; Jiang *et al.* 2014; Tsubokura *et al.* 2014). Soybeans belonging to 000 maturity groups have recessive alleles at all these loci. Much of the inheritance work on photoinsensitivity has been conducted in the countries of higher latitude. Theoretically, soybeans adapting to lower latitudes, like India, should have photosensitive alleles and those adapted to tropical zone have long juvenile character as well. This explains the suitability of maturity groups V and later to adapt to India and that is probably the reason for Bragg (MG VII), Lee (MG VI), Clark 63 (MG IV), improved Pelican (MG VIII), Davis (MG VI) and Hardee (MG VIII) becoming the founder soybean cultivars in India. Long juvenile character which is not affected by photoperiod, delays flowering under short day conditions and thus promotes plant growth which results in good yield when the crop is grown near tropics. Lawn and James (2011a, b) have shown the importance of photoinsensitivity trait in Australia whose latitudes (9–44°S) almost correspond to that of India (6–35°N). They introduced elite US photoinsensitive cultivars Charleston and Sprite 87 (MG III) and backcrossed them as recurrent parent with donors of long juvenile trait. Developed backcross lines with both long juvenility and

photoinsensitivity traits could adapt to a very large range of latitudes and sowing dates. They could attribute the importance of photoinsensitivity in fixing the days to flower over a range of latitudes and sowing dates in presence of long juvenility trait. Presence of recessive alleles in the most popular cultivars, JS 335 and JS 95-60, also suggests for the utility of photoinsensitivity in India but a definite conclusion can only be obtained by developing near isogenic lines for these genes and evaluating them over various latitudes and sowing dates in India.

Identified recessive alleles offer a number of opportunities for their use in Indian soybean molecular breeding programmes. In India, based on the observation on 320 accessions evaluated during eight years in All India Coordinated Soybean Research Project, we have found that average days to flower in northern plain zone (average latitude 28.75°), central zone (average latitude 22.61°) and southern zone (average latitude 16.13°) decrease from 51 to 41, to 38, respectively. We have found that average plant height for these zones decreased from 66 in northern plain zone to 56 cm in central zone, and to 42 cm in southern zone. Hence, one of the utilities of identified recessive alleles would be in imparting a small degree of photoinsensitivity in cultivars adapting to northern plain zone, where flowering is too late, plants grow longer and lodge. Their another usefulness would be in adaptation of soybean accessions from lower latitudes to higher latitudes of northern India by making them insensitive to relatively longer day conditions of these latitudes. This would allow the use of elite germplasm of southern latitudes possible for developing desirable early maturing cultivars in central India. Further, presence of hypoactive allele in the most popular cultivar JS 95-60 and its quite earliness suggest for converting popular long duration (>100 days) cultivars of central India with photoinsensitive genes for faster development of new genotypes with desired maturity in adaptive genetic background. This is the first report on status of photoperiodic genes in Indian soybean gene pool and identified alleles and their molecular markers would be of great value in Indian soybean breeding programmes.

Acknowledgement

We thank the Director, ICAR-Indian Institute of Soybean Research for providing all the required facilities for conducting the work.

References

Abe J., Xu D. H., Miyano A., Komatsu K., Kanazawa A. and Shimamoto Y. 2003 Photoperiod-insensitive Japanese soybean landraces differ at two maturity loci. *Crop Sci.* **43**, 1300–1304.

Bernard R. L. 1971 Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* **11**, 242–244.

Bonato E. R. and Vello N. A. 1999 *E6* a dominant gene conditioning early flowering and maturity in soybeans. *Genet. Mol. Biol.* **22**, 229–232.

Buzzell R. I. 1971 Inheritance of a soybean flowering response to fluorescent day length conditions. *Can. J. Genet. Cytol.* **13**, 703–707.

Buzzell R. I. and Voldeng H. D. 1980 Inheritance of insensitivity to long day length. *Soybean Genet. Newsl.* **7**, 26–29.

Cober E. R. and Voldeng H. D. 2001 A new soybean maturity and photoperiod-sensitivity locus linked to *E1* and *T*. *Crop Sci.* **41**, 698–701.

Cober E. R., Tanner J. W. and Voldeng H. D. 1996a Genetic control of photoperiod response in early-maturing near-isogenic soybean lines. *Crop Sci.* **36**, 601–605.

Cober E. R., Tanner J. W. and Voldeng H. D. 1996b Soybean photoperiod-sensitivity loci respond differentially to light quality. *Crop Sci.* **36**, 606–610.

Cober E. R., Molnar S. J., Charette M. and Voldeng H. D. 2010 A new locus for early maturity in soybean. *Crop Sci.* **50**, 524–527.

Doyle J. J. and Doyle J. L. 1990 Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.

Hymowitz T. 1970 On the domestication of the soybean. *Econ. Bot.* **24**, 408–421.

Jiang B., Nan H., Gao Y., Tang L., Yue Y., Lu S. *et al.* 2014 Allelic combinations of soybean maturilty loci *E1*, *E2*, *E3* and *E4* result in diversity of maturity and adaptation to different latitudes. *PLoS One* **9**, e106042.

Kong F., Nan H., Cao D., Li Y., Wu F., Wang J. *et al.* 2014 A new dominant gene *E9* conditions early flowering and maturity in soybean. *Crop Sci.* **54**, 2529–2535.

Lawn R. J. and James A. T. 2011a Application of physiological understanding in soybean improvement. I. Understanding phenological constraints to adaptation and yield potential. *Crop Pasture Sci.* **62**, 1–11.

Lawn R. J. and James A. T. 2011b Application of physiological understanding in soybean improvement. II. Broadening phenological adaptation across regions and sowing dates. *Crop Pasture Sci.* **62**, 12–24.

Liu B., Kanazawa A., Matsumura H., Takahashi R., Harada K. and Abe J. 2008 Genetic redundancy in soybean photoresponses associated with duplication of *phytochrome A* gene. *Genetics* **180**, 996–1007.

McBlain B. A. and Bernard R. L. 1987 A new gene affecting the time of flowering and maturity in soybeans. *J. Hered.* **78**, 160–162.

Piper C. V. and Morse W. J. 1910 The soybean; history, varieties, and field studies. *Bull. U.S. Dep. Agric. Bur. Plant Ind.* **197**, 84.

Ray J. D., Hinson K., Mankono E. B. and Malo F. M. 1995 Genetic control of a long-juvenile trait in soybean. *Crop Sci.* **35**, 1001–1006.

Saindon G., Voldeng H. D., Beversdorf W. D. and Buzzell R. I. 1989a Genetic control of long day length response in soybean. *Crop Sci.* **29**, 1436–1439.

Saindon G., Beversdorf W. D. and Voldeng H. D. 1989b Adjusting of the soybean phenology using the *E4* loci. *Crop Sci.* **29**, 1361–1365.

Singh R. K., Bhatia V. S., Yadav S., Athale R., Lakshmi N., Guruprasad K. N. *et al.* 2008 Identification of genetically diverse genotypes for photoperiod insensitivity in soybean using RAPD markers. *Physiol. Mol. Biol. Plants* **14**, 369–375.

Tsubokura Y., Matsumura H., Xu M., Liu B., Nakashima H., Anai T. *et al.* 2013 Genetic variation in soybean at the maturity locus *E4* is involved in adaptation to long days at high latitudes. *Agronomy* **3**, 117–134.

Tsubokura Y., Watanabe S., Xia Z., Kanamori H., Yamagata H., Kaga A. *et al.* 2014 Natural variation in the genes responsible for maturity loci *E1*, *E2*, *E3* and *E4* in soybean. *Ann. Bot.* **113**, 429–441.

Watanabe S., Hideshima R., Xia Z., Tsubokura Y., Sato S., Nakamoto Y. *et al.* 2009 Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* **182**, 1251–1262.

Watanabe S., Xia Z., Hideshima R., Tsubokura Y., Sato S. and Harada K. 2011 A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* **188**, 395–407.

Watanabe S., Harada K. and Abe J. 2012 Genetic and molecular bases of photoperiod responses of flowering in soybean. *Breed. Sci.* **61**, 531–543.

Xia Z., Watanabe S., Yamada T., Tsubokura Y., Nakashima H., Zhai H. *et al.* 2012 Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci.* **109**, E2155–E2164.

Xu M., Xu Z., Liu B., Kong F., Tsubokura Y., Watanabe S. *et al.* 2013 Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biol.* **13**, 91.

Received 16 February 2016, in final revised 3 May 2016; accepted 6 May 2016
Unedited version published online: 9 May 2016
Final version published online: 9 January 2017

Corresponding editor: UMESH C. LAVANIA