Euphytica (2021)217:36
https://doi.org/10.1007/s10681-021-02775-2

QTL mapping for long juvenile trait in soybean accession AGS 25 identifies association between a functional allele of *FT2a* and delayed flowering

Rachana Tripathi · Nisha Agrawal · Giriraj Kumawat · Sanjay Gupta · Philips Varghese · Milind B. Ratnaparkhe · Virender S. Bhatia · Shivakumar Maranna · Gyanesh K. Satpute · Suresh Chand · Meeta Jain

Received: 17 October 2019 / Accepted: 20 January 2021

Abstract Soybean is a quantitative photoperiod sensitive short-day crop. In Asian countries, soybean is grown under rain-fed conditions. However, late planting due to delayed onset of the rainy season results in the reduced biomass leading to considerable yield losses. Long juvenility (LJ) trait in soybean delays flowering and helps in gaining biomass under short days. Our objective of this work was to understand the molecular basis of long juvenility in a genetic resource AGS 25. In this study, we used recombinant inbred lines (RILs) developed from an LJ genotype AGS 25 and conventional juvenile (CJ) genotype JS 93–05 and mapped a major quantitative trait locus (QTL) explaining upto 39.7% of the phenotypic variation. The identified QTL carried two candidate flowering genes *FT2a* and *FT2b,* in which genomic DNA sequencing of coding DNA sequence identified the former with a functional SNP variation but the latter was found monomorphic. The identified mutation in *FT2a* could be associated with LJ and validated in another segregating population (F_2 of SL 958 × AGS 25). The novel mutation caused a non-synonymous substitution of conserved glycine with aspartic acid (G169D) and the mechanism of LJ through this mutation is to be further divulged. This identified mis-sense mutation found associated with delayed flowering in two previous reports. The identified new functional SNP created differential *HinfI* (e9-CAPS) restriction sites and high-resolution melt (HRM) profiles (e9-HRM) for CJ and LJ parents and are associated with delayed flowering in RILs and F_2 population (SL 958 × AGS 25).

Keywords Long juvenility · Maturity · *E9, FT2a*, delayed flowering · Photoperiod

Rachana Tripathi and Nisha Agrawal contributed equally in the work

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10681-021-02775-2.

R. Tripathi · N. Agrawal · G. Kumawat · S. Gupta (✉) · M. B. Ratnaparkhe · V. S. Bhatia · S. Maranna · G. K. Satpute
ICAR- Indian Institute of Soybean Research, Indore, Madhya Pradesh, India
e-mail: sanitaishu@gmail.com

R. Tripathi · N. Agrawal · S. Chand · M. Jain
Devi Ahilya Vishwavidyalaya, Indore, Madhya Pradesh, India

P. Varghese
Agharkar Research Institute, Pune, Maharashtra, India

Introduction

Soybean (*Glycine max* (L.) Merrill) is the number one oilseed crop of the world with a contribution of 25% in world edible oil economy and 65% in animal feed concentrate. It is a quantitative photoperiodic short-

Published online: 10 February 2021

day crop and its genotypes adapt to a narrow latitudinal band. The crop has adapted from 50°N to 40°S latitudes mainly due to allelic variation in eleven major maturity genes. Among maturity genes, *E1* & *E2* (Bernard 1971), *E3* (Buzzell 1971), *E4* (Buzzell and Voldeng 1980), *E7* (Cober and Voldeng 2001) and *E8* (Cober et al. 2010) are photoperiodic genes and *E1* to *E4* has been characterized (Xia et al. 2012; Watanabe et al. 2009, 2011; Liu et al. 2008). Dominant alleles of *E1, E2, E3* and, *E4* genes confer photosensitivity, late flowering and maturity under long-day conditions. Recessive alleles of these genes make them insensitive to longer day lengths prevailing in higher latitudes and help them to adapt to these regions. Cultivated soybean has its origin in higher latitudes of China (32° to 40° N) from its wild ancestor (*Glycine soja* Sieb. & Zucc.) and has a long history of being a commercial crop in the higher latitudes by virtue of mutations in photoperiodic genes (Li et al. 2008). Mechanism of adaptation of the crop to higher latitudes has largely been elucidated and three modules viz PHYA–*E1*, *GI–CO* and miRNA-dependent have been identified (Cao et al. 2017).

However, soybean adaptation and its commercial cultivation in short-day conditions of the lower latitudes (< 22°) became possible only after the identification of long juvenile trait (Hartwig 1970) which helped the crop gain sufficient biomass by delaying flowering even under short-day conditions. Delayed flowering in the long juvenile genotypes results in the gain of more height, nodes, grain yield and other agronomic characters (Lawn and James 2011a, b). Introduction of this trait has helped in expanding soybean cultivation to the lower latitudes and thus made Brazil the second-largest soybean producing country in the world (Gavioli 2013). Delayed flowering under short-day conditions was originally identified in PI 159925 (90 days) as against other varieties which flowered in ∼ 60 days (Hartwig and Kiihl 1979) and later the terms juvenile and long juveniles were used (Parvez and Gardner 1987; Sinclair and Hinson 1992) to describe this trait. Against the well-defined modules for photo-insensitivity, genetic and molecular bases of adaptation of soybean to short-day conditions are not well known. This trait has been shown to be governed by three recessive genes (Hartwig and Kiihl 1979). In contrast to photoperiodic genes, mutations at long juvenile loci result in delayed flowering under short-day conditions. A well-studied original long juvenile accession PI 159925 has been used to map and characterized long juvenile locus (*J*) (Ray et al. 1995; Lu et al. 2017). Later, a new loci *E6* was identified by using conventional juvenile (CJ) variety Parana and long juvenile (LJ) variety Paranagoiana (Bonato and Vello 1999). Genetic interaction between *E1* and *E6* has revealed that *E6* has a suppressive effect on *E1* and the function of *E6* depends on *E1*. *E6* mapped near to *J* loci on the linkage group (LG) C1 and is putatively either the allele of *J* or closely linked to it (Lu et al. 2017). Against *E6* and *J* which have been identified by involving LJ and CJ parents, another long juvenile locus *E9* was identified and mapped to J LG through QTL analysis of early flowering trait that was introduced from wild soybean accession (Kong et al. 2014).

Molecular bases of *J* and *E9* loci have been elucidated, and the former has been identified as the ortholog of *Arabidopsis thaliana EARLY FLOWERING 3* (*ELF3*) (Lu et al. 2017) and the latter as the florigen gene *FT2a* (Zhao et al. 2016). ELF3 is a conserved nuclear specific protein and is known to maintain circadian rhythms and controls flowering in many species (Hicks et al. 2001; Weller et al. 2012; Matsubara et al. 2012). J protein down-regulates *E1* by physically interacting with its promoter and promotes flowering (Lu et al. 2017). *Flowering locus T* (*FT*) has been well characterized in *Arabidopsis* and encodes a florigen protein which has a main role in different flowering pathways (Corbesier and Coupland 2006; Turck et al. 2008). Florigen protein moves to shoot apex through phloem and signals for floral initiation in *Arabidopsis* (Corbesier et al. 2007; Jaeger and Wigge 2007; Mathieu et al. 2007; Notaguchi et al. 2008). Over-expression of *FT* orthologs has been found to result in extremely early flowering in monocots and dicots (Böhlenius et al. 2006; Hsu et al. 2006; Lifschitz et al. 2006; Hayama et al. 2003; Izawa et al. 2002; Kojima et al. 2002; Yan et al. 2006). Some of the *FT* orthologs also induce premature flowering in *Arabidopsis* (Hsu et al. 2006; Lifschitz et al. 2006; Hayama et al. 2003). In soybean, Kong et al. (2010) identified ten *FT* homologs in five linked pairs in different homologous chromosomal regions. Among these ten *FT* homologs, *GmFT2a* and *GmFT5a* have been identified as flowering activators and integrators (Kong et al. 2010; Nan et al. 2014 and Sun et al. 2011) and *GmFT1a* and *GmFT4* as flowering

repressor (Liu et al. 2018; Zhai et al. 2014). In addition to *GmFT2a* and *GmFT5a*, four other *FT* homologs have been shown to promote flowering in the ectopic expression studies in *Arabidopsis* (Wang et al. 2015; Jiang et al. 2013). Chen et al. (2018) over-expressed *GmFT2b* and reported that flowering was hastened under long but not in short-day conditions. They further confirmed this result through CRISPR/Cas9 mediated mutagenesis and found that *GmFT2b* mutants delayed flowering significantly in long day but not in short-day conditions. *GmFT2a* and *GmFT5a* have been the most researched FT homologs, and their mechanism of action has been determined (Nan et al. 2014; Cai et al. 2018, 2019).

GmFT2a has been characterized as the gene for classical *E9* locus. Zhao et al. (2016) have reported the insertion of retroposon in the first intron which results in reduced transcript abundance of *FT2a* and associate it with delay flowering. Recently, (Jiang et al. 2019) sequenced ten *FT* family genes in 127 varieties and identified 34 SNPs and InDels in these genes. Among these variants, four were in exonic regions. SNP (s4645) in exon 4 of *FT2a* caused the conversion of glycine (G) to aspartic acid (D) and was found to be associated with flowering. Ogiso-Tanaka et al. (2019) detected novel mutants in flowering time-related genes by targeted resequencing using Ampli-Seq technology in soybean mini-core collection. One frameshift and two mis-sense variants were observed in *FT2a* and one of the mis-sense variants (Chr16_31114633) is same as reported by Jiang et al. (2019). In a more recent study, Sun et al. (2019) also reported a mis-sense mutation which was reported by Jiang et al. (2019) and Ogiso-Tanaka et al. (2019) to be associated with the late flowering habit of a Thai cultivar K3.

Earlier, we identified a new long juvenile genetic resource (AGS 25) and characterized it for long juvenility across latitudes and sowing dates (communicated). In the present work, we mapped a major QTL for long juvenile trait in AGS 25, which corresponded to a functional mutation in *FT2a* recently reported by Ogiso-Tanaka et al.(2019), Jiang et al. (2019) and Sun et al. (2019). We validated this new variant in another F_2 population developed from LJ genotype AGS 25 and CJ genotype SL 958.

Material and methods

Plant material and field phenotyping

105 RILs ($F_{2:8}$) from the cross JS 93–05 × AGS 25 were developed by growing all individual F_2 plants and advancing these plants by growing individual plant progenies of a random plant of each line from F_3 to F_8 at Indore, Madhya Pradesh (22.7196°N, 75.8577°E, 553 m above mean sea level). F_2 population between SL 958 × AGS 25 was developed by the selfing of F_1s. RILs were evaluated in the sowings of 4 July 2018 (S1), 21 July 2018 (S2) in the field and 18 December 2018 (S3) in a glasshouse at Indore, temperature variation among different sowings dates till maturity was also recorded using zeal maximum and minimum thermometer (Supplementary file 1). RILs were also evaluated in a lower latitude location Pune, Maharashtra (18.5204°N) on 15 July 2018 sowing (S4). A F_2 population involving AGS 25 and conventional juvenile variety SL 958 was developed. F_2s were evaluated in S1 (200 plants) and S3 (176 plants). The soil of experimental plots of both locations and glasshouse was medium black and crop was grown under rain-fed conditions with irrigations under moisture stress conditions. Seed used for sowing was treated with fungicides (thiram & carbendazim) and *Rhizobium japonicum* culture in all of the experiments. Each line was grown in a three-meter row in field with a row to row distance of 45 cm and plant to plant distance of 5 cm. In the glasshouse, each line was grown in two pots with three plants per pot. Observations on days to R1 were recorded daily on five randomly selected plants. Observations on all of the six plants were recorded in the glasshouse. The opening of the first flower on any of the nodes on main stem was recorded as the date of flowering (R1) and days to R1 were calculated.

Genotypic characterization of AGS 25 for known flowering genes

The genotype of AGS 25 at *E1*, *E2, E3* and *E4* loci was determined using allele-specific markers (Xu et al. 2013; Zhao et al. 2016; Gupta et al. 2017) and at *E9* (SNP and Indel) using Agena Massarray (Kumawat et al. 2020). Later, after QTL analysis, AGS 25 was Sanger sequenced (Table 1) for coding DNA sequence (CDS) of *FT2a* and *FT2b.*

Table 1 List of primers used for Sanger sequencing of *E1* and *E9* CDS

Gene	Exon	Forward primer	Reverse primer
E1	Exon 1	CCCATCAAAGTTCACGACCC	AGGTTGAAGTACACGCTATTGC
E9	Exon 1	AGCGAAAGCATATCACTTCAAAC	TGGATGGTCAAAAACAATAACGTC
E9	Exon 2 And 3	TTCTTAGTGCAATCGAGGATCA	GCAAACATTGTGCGTTTGACA
E9	Exon 4	GGAAGAGAGGGAAAAGGACACA	AAAGAGTGTGGGAGATTGCC

Bulk segregant analysis and genotyping of RILs and F_2s

RILs were classified as CJ or LJ based on the mid-parent days to R1 which was calculated by averaging the days to the R1 of parents. RILs with a flowering duration equivalent to or less than that of mid-parent value were classified as CJ and those with more than mid-parent value were considered as LJ. The extraction of plant DNA from parents and RILs was carried out by CTAB (Cetyltrimethylammonium bromide) method (Doyle and Doyle 1990). The purified genomic DNA was quantified by the nanodrop (Denovix DS-11 + spectrophotometer). The quality of the DNA was checked by electrophoresis on 0.8% agarose. Polymorphism between JS 93–05 and AGS 25 was established using 872 BARCSOYSSR markers spread across 20 linkage groups. The BARCSOYSSR markers were randomly selected as well as a few taken from genomic regions of known flowering loci in soybean and homologs of *Arabidopsis* (Watanabe et al. 2012; Song et al. 2012). Amplification using SSR markers was carried out in thermocycler (Applied Biosystems, the USA) using following conditions: initial denaturation (94 °C) for 4 min, denaturation (94° C) for 50 s, annealing (55 °C) for 1 min, extension (72 °C) for 1 min and final extension (72 °C) for 10 min. The amplified products so obtained were resolved on 3.5% metaphor agarose, and SSR sizes were estimated using 50 bp DNA ladder.

CJ and LJ bulks were formed by combining the equal quantity of genomic DNA from ten extreme type plants of each bulk type and bulk segregant analysis was conducted by genotyping of identified polymorphic SSRs (Michelmore et al. 1991). Polymorphic SSR markers identified by bulk segregant analysis were used for genotyping of the RIL population (n = 105). For validation of the identified SNP, polymorphism between AGS 25 and SL 958 was established for flanking SSR markers. Polymorphic markers were used for genotyping of F_2 plants (n = 200) in S1 and 176 plants in S3 as per the procedure described for the RIL population.

High-resolution melting (HRM) assay

Forward primer sequence 5′AGAATTTGCTGAACT TTACAACCT3′ and reverse primer 5′TTTCTTAG TATAACCTCCTTCCACC3′ for HRM analysis were designed from the flanking regions of identified SNP using primer3 software (Untergasser et al. 2012). Real-time PCR based HRM reactions were prepared using an HRM ultrafast reaction mixture containing SYBR Green (Agilent technologies, the USA). Each 20-µl real-time PCR reaction contained 2X HRM mix, 0.5-µmol of each primer and 20 ng of DNA. The reactions were performed using AriaMx real-time PCR systems thermal cycler (Agilent Technologies, the USA) with the following amplification and melting conditions: one cycle of hot start denaturation at 95 °C for 3 min, 40 cycles at 95 °C for 5 s and 60 °C for 10 s with an endpoint detection of fluorescence intensity in each cycle. The thermal profile to generate high-resolution melt curve follows one cycle of denaturation from 65–95 °C for 30 s and fluorescence data were acquired at 0.2 °C increments every 10 s (Lochlainn et al. 2011).

Cleaved amplified polymorphic sequence (CAPS) marker assay

Web-based tool dCAPS Finder 2.0 was used to predict restriction sites in *FT2a* CDS (Neff et al. 2002). The fourth exon of *FT2a* was amplified using PCR primers that were used for sequencing. The genomic sequence of the fourth exon from AGS 25 and JS 93–05 was amplified using PCR and was restriction

digested with *HinfI* restriction enzyme. *HinfI* digested PCR product was separated on 3.5% metaphor agarose gel for the detection of a new allele.

Data Analysis

QTL analysis was performed using QTL IciMapping Software (Meng et al. 2015). Single marker analysis (SMA) of variance and inclusive composite interval mapping (ICIM) method was used for QTL detection. LOD threshold for declaring a QTL was identified using 1000 permutations at $P = 0.05$ for type I error. DNA sequence alignment was performed by Clustal Omega tool (https://www.ebi.ac.uk/Tools/msa/clustalo).

Results

RIL phenotyping and bulk segregant analysis

A continuous variation for days to R1 was observed in RILs and both of the parents showed wide differences in days to R1 in all the growing environments (Fig. 1). Soybean accession AGS 25 flowered in 56 & 53 days while cultivar JS 93–05 flowered in 40 & 37 days in S1 and S2 sowings, respectively. The mean flowering duration of RILs was 1.7 days more in S2 than in S1.

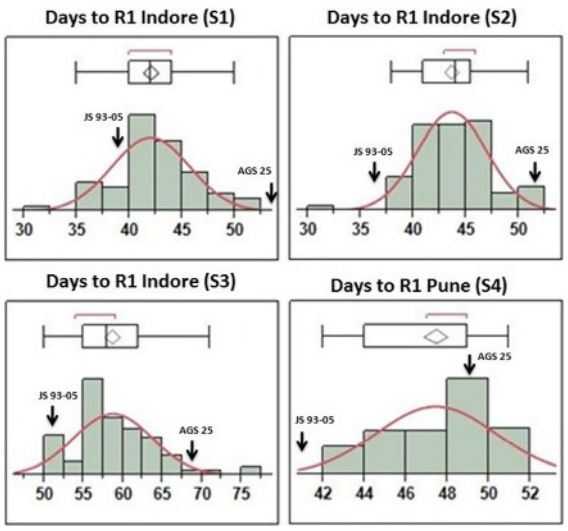

Fig. 1 Distribution of days to R1 in RILs. **a** days to R1 Indore (S1) sown on 4 July 2018; **b** days to R1 Indore (S2) sown on 21 July 2018; **c** days to R1 Indore (S3) sown on 18 December 2018; **d** days to R1 Pune (S4), sown on 15 July 2018

In both of these sowings, transgressive segregants were observed in RILs for CJ parent. During S3, AGS 25 and JS 93–05 flowered in 69 and 52 days and transgressive segregants for flowering were observed for both parents. At Pune (S4), AGS 25 flowered in 50 days and JS 93–05 flowered in 35 days with a population mean of 49.77 days and transgressive segregants were observed for LJ parent (Additional Fig. 1).

Identification of QTL for long juvenility

Out of 872 SSR markers screened for polymorphism covering 20 linkage groups, 254 markers were polymorphic between the two parental lines AGS 25 and JS 93–05. The bulk segregant analysis identified two polymorphic SSR markers flanking *E9* locus (BARCSOYSSR_16_1014 and BARCSOYSSR_16_1035) on J LG. The polymorphic SSR markers identified between two contrasting bulks were further used to genotype RIL population. In single marker analysis (SMA) of QTLs BARCSOYSSR_16_1014 and BARCSOYSSR_16_1035 were found associated in S1 (LOD > 6.0), S2 (LOD > 2.9), S3 (LOD > 6.3) and S4 (LOD > 3.1) and explained up to 14.26% total phenotypic variance (data not shown). We further genotyped the genomic region identified on chromosome 16 (LG J) with nine new polymorphic SSR markers to delimit the confidence interval of QTL. QTL analysis using ICIM identified a major QTL between BARCSOYSSR_16_1012 and BARCSOYSSR_16_1014 in all of the sowings (Table 2) with LOD range of 3.53 (S2) to 12.43 (S3). This QTL explained 14.68 (S2) to 26.51% (S1) of phenotypic variance. AGS 25 contributed the late maturing QTL with the additive effect of 1.28 days in S2 to 3.6 days in S4. One more QTL was identified between BARCSOYSSR_16_1010 and BARCSOYSSR_16_1012 in the winter sowing environment (S3). Candidate gene analysis of the flanking genomic region between BARCSOYSSR_16_1010 and BARCSOYSSR_16_1014 identified two candidate genes Glyma.16g150700 (*FT2a*) and Glyma.16g151000 (*FT2b*), which are homologs of *FT* having a role in flowering regulation.

Table 2 QTL analysis results of inclusive composite interval mapping (ICIM) in RILs population

Sowing environments	Left marker name	Right marker name	LOD value	% Phenotypic variance	Additive effect
S1	BARCSOYSSR_16_1012	BARCSOYSSR_16_1014	6.722	26.512	2.18
S2	BARCSOYSSR_16_1012	BARCSOYSSR_16_1014	3.533	14.688	1.287
S3	BARCSOYSSR_16_1010	BARCSOYSSR_16_1012	11.960	18.617	3.530
S3	BARCSOYSSR_16_1012	BARCSOYSSR_16_1014	12.435	19.500	3.622
S4	BARCSOYSSR_16_1012	BARCSOYSSR_16_1014	4.004	16.259	1.349

Sequence variation in CDS regions of candidate flowering genes and identification of a novel functional allele

AGS 25 was genotyped using known gene-specific markers for all of the maturity alleles including three markers (one SNP, one InDel and one *SORE-1* retrotransposon in the first intron) of *E9*. It did not differ from JS 93–05 at any of these loci and the genotype of both of the parents was *E1/E2/E3/E4/E9*. Since a major QTL was detected between flanking markers of *E9* but no previously reported alleles were detected in the long juvenile accession AGS 25, we sequenced the entire CDS of two candidate genes (*GmFT2a & GmFT2b)*. While both of the parents were monomorphic for *GmFT2b* (Supplementary file 2), the long juvenile parent AGS 25 had a single base mutation at 506th position (Supplementary file 3) in *GmFT2a* as compared to Williams 82. This SNP resulted in the non-synonymous substitution of glycine with aspartic acid at 169th position in the fourth exon (Fig. 2). The restriction site prediction in CDS of *FT2a* identified two restriction sites for the *HinfI* enzyme in the fourth exon of Williams 82 but the non-synonymous substitution in AGS 25 created an additional site in the fourth exon. This information could be used in the development of new CAPS marker (e9-CAPS). This mutation also reduced the melting temperature of polymerase chain reaction amplified DNA containing identified mutation (e9-HRM).

QTL analysis was conducted again in the RIL population using genotyping data of e9-HRM (Fig. 3). Only single QTL between BARCSOYSSR_16_1014 and e9-HRM could be detected in S1 (7.87 LOD), S2 (4.20 LOD), S3 (11.941 LOD) and S4 (4.97 LOD) and explained 31.9, 18.7, 9.7 and 21.6% of total phenotypic variation in S1, S2, S3 and S4, respectively (Table 3 and Fig. 4). Differential HRM profiles could also be observed between AGS 25 and another conventional juvenile genotype SL 958. The identified QTL was further validated in a F_2 population derived from these parents in S1 and S3 using the same set of nine polymorphic markers and e9-HRM. The QTL was identified between e9-HRM and BARCSOYSSR_16_1012 (LOD = 12.87) in S1 and e9-HRM and BARCSOYSSR_16_1014 (LOD = 4.25) in S3. These QTLs explained 26.67 and 11.94% of total phenotypic variance in S1 and S3, respectively (Table 4) and like RILs AGS 25 contributed the delayed flowering QTL. The newly identified delayed flowering allele of *FT2a* has been named as *e9-aa*.

Discussion

AGS 25 is an adapted long juvenile resource in India. In extreme short-day conditions of winter where all of the genotypes take longer time to flower (Zhao et al. 2016), it maintains the flowering duration difference from others and further delays flowering. Our previous inheritance studies have identified a single recessive gene for this trait in AGS 25 (Anonymous 2014). The objective of the present research was to identify the molecular basis of long juvenility in this resource.

No polymorphism was detected between AGS 25 and other conventional juvenile parents for known photoperiodic and long juvenile markers. QTL analysis in a population from AGS 25 and JS 93–05 could locate a QTL near *FT* homologue pair (*Ft2a* and *Ft2b*). This QTL was consistently detected in all of the four locations and explained up to 39.7 of variation for days to R1 with AGS 25 contributing the positive allele. Since *GmFT2a* has

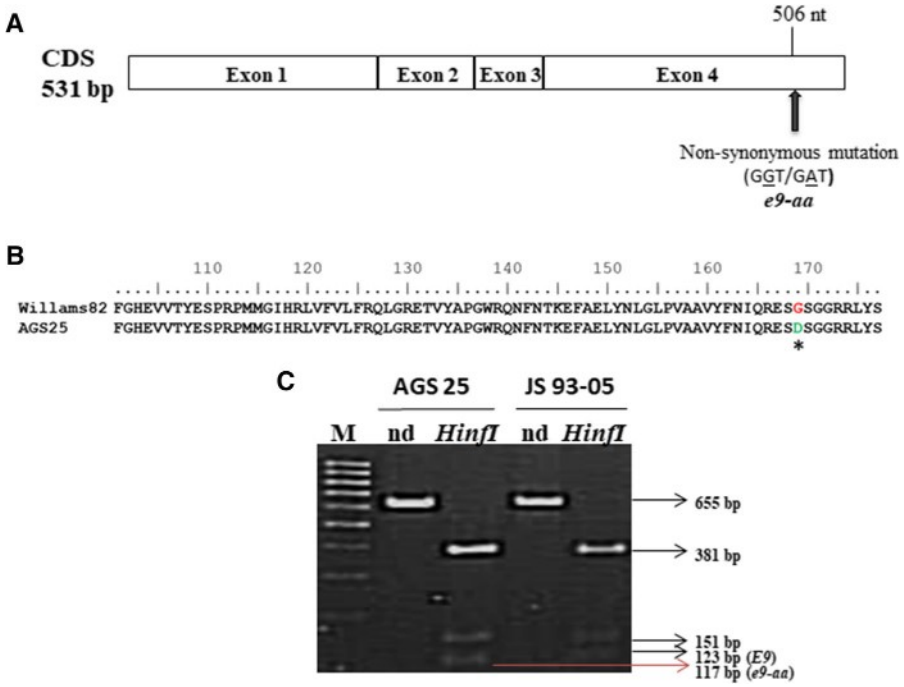

Fig. 2 Novel allele (*e9-aa*) of *E9* locus in long juvenile soybean genotype AGS 25. **a** Position of *e9-aa*, **b** predicted changes in amino acid sequence, asterisks indicate substituted amino acid. **c** CAPS marker distinguishing recessive allele of long juvenile genotype AGS 25 (*e9-aa*) from conventional juvenile genotype JS 93–05 (*E9*), nd-not digested

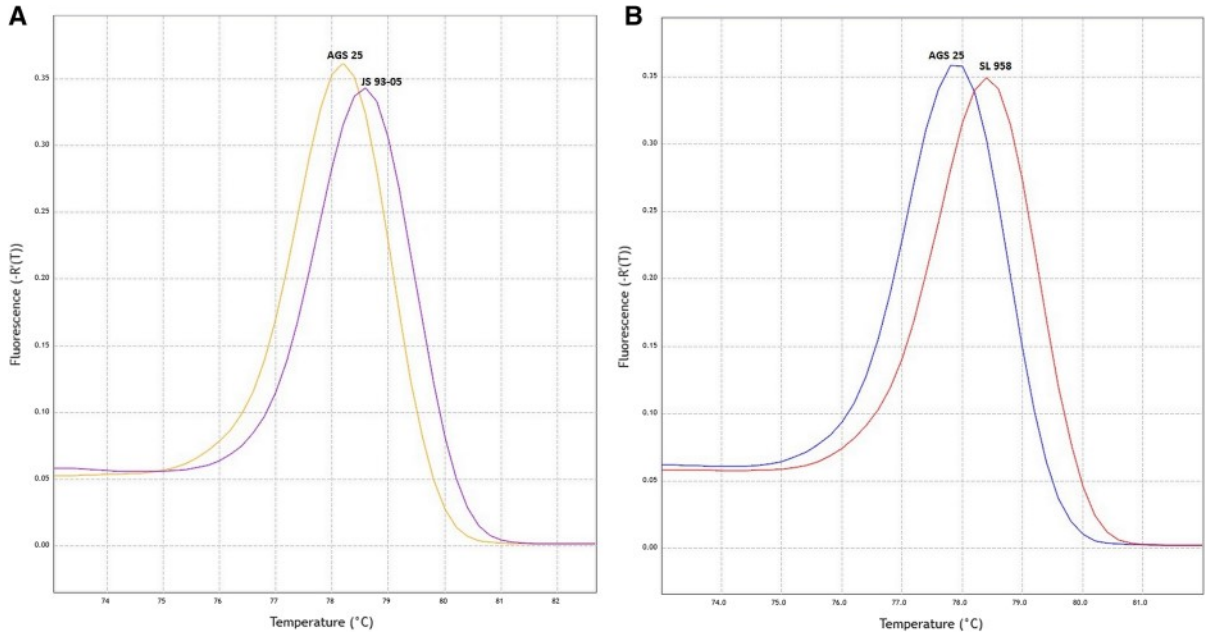

Fig. 3 HRM profile of (**a**) AGS 25 and JS 93–05 and (**b**) AGS 25 and SL 958

Table 3 QTL analysis results of inclusive composite interval mapping (ICIM) in RILs population after including newly developed gene-specific marker e9-HRM

Traits	Linkage group	Position (cM)*	Left marker name	Right marker name	LOD value	% Phenotypic variance	Additive effect
S1	J	0.00	BARCSOYSSR_16_1014	e9-HRM	7.8750	31.9177	2.1102
S2	J	4.00	BARCSOYSSR_16_1014	e9-HRM	4.2046	18.7313	1.4163
S3	J	3.00	BARCSOYSSR_16_1014	e9-HRM	11.9441	39.7671	3.6121
S4	J	5.00	BARCSOYSSR_16_1014	e9-HRM	4.9778	21.6431	1.3639

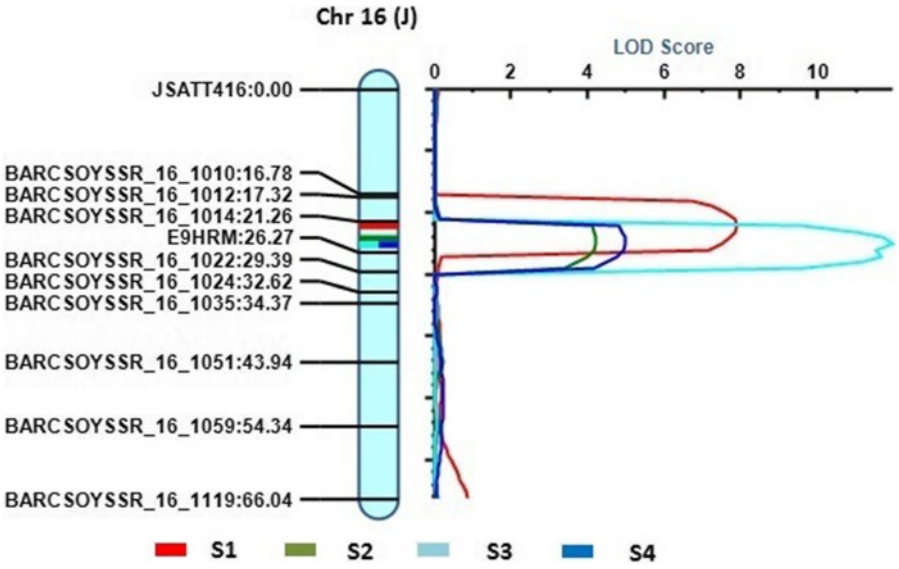

Fig. 4 Detection of QTL for long juvenility determined by SSR markers on LG J using JS 93–05 × AGS 25 RILs. Days to R1(S1) sown on 4 July 2018 at Indore; (S2) sown on 21 July 2018 at Indore; (S3) sown on 18 December 2018 at Indore and (S4) sown on July 2018 at Pune. Genetic linkage map, marker names and genetic distances (cM) are displayed on the left side of the graph

Table 4 QTL analysis results of inclusive composite interval mapping (ICIM) in SL 958 × AGS 25 population after including newly developed gene specific marker e9-HRM

Traits	Linkage group	Left marker name	Right marker name	LOD value	% Phenotypic variance	Additive effect
S1	J	e9-HRM	BARCSOYSSR_16_1012	12.877	26.672	−2.021
S3	J	e9-HRM	BARCSOYSSR_16_1014	4.2524	11.942	−2.423

been characterized as classical long juvenile locus *E9* and has been reported to delay flowering under short-day conditions (Cober 2011) it appeared to be a putative candidate for long juvenility in AGS 25. The role of *GmFT2b* was further obviated as it hastens flowering in long-day conditions only (Chen et al. 2018) and such conditions are not present in India. Association of *FT2a* and not of *FT2b* with long juvenility in AGS 25 could be putatively attributed to the former because of parental polymorphism detected through sequencing. This polymorphism, present in the exon 4, resulted in the non-synonymous substitution of glycine with aspartic acid at 169th position in the fourth exon (Fig. 2) and has recently been detected in three other studies (Jiang et al. 2019;

Ogiso-Tanaka et al. 2019; Sun et al. 2019). Based on this SNP, two types of markers e9-HRM and e9-CAPS - were developed and validated. QTL analysis using e9-HRM limited the associated region between this marker and other SSR markers in all environments in both of the populations. FT is a member of the ancient phosphatidylethanolamine-binding protein (PEBP) (Wang et al. 2015) and has critical role in induction of flowering. It shares high degree of amino acid similarity with another family member TFL1 which is a growth habit gene and has repressing effect on flowering (Ahn et al. 2006). Both of these genes have four exons and 4th exon was identified as the most critical exon for diversion of functions of *FT* and *TFL* (Ahn et al. 2006). In the 4th exon, also glycine is conserved in flowering inducer FT-like phosphatidylethanolamine-binding proteins, whereas aspartic acid at this position is a conserved feature of floral repressor TFL1-like proteins (Ahn et al. 2006; Wang et al. 2015). Notably, the flowering inhibitor *GmFT1a* also has aspartic acid in the same location, indicating that this amino acid substitution might have an important role in the function of *GmFT2a* (Jiang et al. 2019). Therefore, it is to be divulged in the further study that how the mis-sense mutation at 169th position is affecting the functionality of *Gmft2a*.

In the present investigation, the molecular basis of long juvenility in AGS 25 could be attributed to the genomic region carrying the novel functional mutation in *ft2a* and CAPS & HRM markers are developed. AGS 25 is a valuable genetic resource for India for developing soybean varieties with wider sowing window in rainy season. In India, soybean is cultivated only during rainy seasons and there is scope to breed soybean for winter season in areas with mild temperatures. In such areas, early flowering due to shorter winter day lengths results in reduced plant growth. AGS 25 offers the opportunity for developing soybean for alternate season. Identified markers would be useful for rapid conversion of existing varieties to their long juvenile version.

Acknowledgements Authors are thankful to Director, ICAR-IISR, Indore for providing all the required facilities for the experiments.

Author contributions SG identified AGS 25, developed RILs, planned all experiments and coordinated the work; RT and NA performed phenotyping and genotyping of RILs; GK and MR analyzed molecular data and conducted MassArray of soybean germplasm; VB, SM, GKS, SC and MJ contributed in inheritance studies and phenotyping; PV conducted RIL phenotyping in Pune.

Funding The research was supported by grant from Department of Biotechnology, Ministry of science and technology, Government of India through Department of Biotechnology project number "BT/PR15159/AGIII/103/919/2016".

Availability of data and materials All data generated or analyzed during this study are included in this published article and its supplementary information files.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Ahn JH, Miller D, Winter VJ, Banfield MJ, Lee JH et al (2006) A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. EMBO J 25:605–614

Anonymous: Annual Report 2013–2014. Directorate of Soybean Research Indore

Bernard RL (1971) Two major genes for time of flowering and in soybeans. Crop Sci 11:242–244. https://doi.org/10.2135/cropsci.0011183X001100020022x

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science 312:1040–1043

Bonato ER, Vello NA (1999) *E6* a dominant gene conditioning early flowering and in soybeans. Genet Mol Biol 22:229–232. https://doi.org/10.1590/S1415-47571999000200016

Buzzell RI (1971) Inheritance of a soybean flowering response to fluorescent day length conditions. Can J Genet Cytol 13:703–707. https://doi.org/10.1139/g71-100

Buzzell RI, Voldeng HD (1980) Inheritance of insensitivity to long day length. SoybeanGenet Newsl 7:26–29. https://doi.org/10.1093/jhered/esp113

Cai Y, Chen L, Liu X, Guo C, Sun S, Wu C, Jiang B, Han T, Hou W (2018) CRISPR/Cas9-mediated targeted mutagenesis of *GmFT2a* delays flowering time in soya bean. Plant Biotechnol J 16:176–185

Cai Y, Wang L, Chen L, Wu T, Liu L, Sun S, Wu C, Yao W, Jiang B, Yuan S, Han T, Hou W (2019) Mutagenesis of GmFT2a and GmFT5a mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. Plant Biotechnol J. https://doi.org/10.1111/pbi.13199

Cao Y, Li S, He X, Chang F, Kong J, Gai J, Zhao T (2017) Mapping QTLs for plant height and flowering time in a Chinese summer planting soybean RIL population. Euphytica 213:39. https://doi.org/10.1007/s10681-016-1834-8