Journal of Genetics, Genomics & Plant Breeding **10(2)70-83**
ISSN (Online): 2581-3293 DOI: 10.63551

RESEARCH ARTICLE

Designing and cleavage efficiency analysis of CRISPR/Cas9 guide RNA for long-juvenile trait in soybean

T. Tayalkar[1,2], A. Rani[1], A. Jajoo[2], A. K. Anshu[1], V. Kumar[1]

1 ICAR-National Soybean Research Institute, Khandwa Road, Indore, Madhya Pradesh-452001, India
2 Devi Ahilya Vishwavidyalaya, Indore, Madhya Pradesh-452001, India

Manuscript received: March 15, 2026
Decision on manuscript: March 28, 2026
Manuscript accepted: April 10, 2026

✉ **anitavks@yahoo.co.in**

Abstract

Effort was made to design CRISPER/ Cas9 guide RNA for long-juvenile trait in soybean. The long-juvenile trait is a crucial agronomic characteristic that extends the vegetative growth phase and delays flowering in soybean grown under short-day photoperiod conditions. *GmFT2a* (*Glyma.16g150700*) is a major gene responsible for early flowering under short-day conditions in soybean. CRISPR-Cas9-mediated genome editing is the most advanced and rapid process for genome editing, which requires the design and selection of unique guide RNA (gRNA) specific to the targeted gene. For CRISPR-Cas9-mediated editing of the *GmFT2a* gene, a total of 25 gRNAs targeting the first exon were designed using the Cas-Designer component of CRISPR RGEN tools. For cleavage efficiency analysis, a total of 9 gRNAs were selected based on minimum off-target sites (less than 30) and efficacy scores greater than 0.60. After cleavage efficiency analysis, three gRNAs were found to cleave their respective target sites more efficiently than the rest of the gRNAs. *In silico* off-target analysis further confirmed the high specificity and precision of the selected gRNAs for targeted editing of the *GmFT2a* locus. The optimized CRISPR-Cas9 construct developed in this study provides a promising strategy for generating long-juvenile soybean genotypes adaptable to tropical and subtropical agro-climatic conditions. Out of 3 gRNAs, one gRNA was inserted into the CRISPR/Cas9 expression vector "Cas9_MDC123" with the GmU6 promoter and scaffold and cloned into *Agrobacterium tumefaciens* strain EHA105 for editing of *GmFT2a* gene for the development of a soybean genotype with long-juvenile trait.

Keywords: Soybean, CRISPER/ Cas9, long-juvenile, gRNAs, *GmFT2a*

Introduction

Soybean is an important economic oilseed crop in India, contributing 20.71% of the country's total oilseed production. As a short-day photoperiod-sensitive crop, its productivity is significantly affected by changing photoperiods. Exposure to short-day photoperiod during the juvenile period triggers early flowering, leads to poor vegetative growth and reduced biomass, which ultimately affects the grain yield (Hu and Wiatrak, 2012; Bhatia *et al.,* 1999). Flowering time in soybean is regulated by a complex gene network that responds to photoperiodic changes. Among these genes, *GmFT2a* acts as a critical flowering activator and generates a signal that travels through phloem to the shoot apical meristem to promote the transition from vegetative to reproductive growth (Sun *et al.,* 2011). Under short-day conditions, the dominant form of *GmFT2a* gene is prominently expressed, inducing early flowering before completion of vegetative growth (Cai *et al.,* 2020; Nan *et al.,* 2014). In contrast, the recessive nonfunctional *Gmft2a* gene is responsible for the long-juvenile trait, which extends vegetative growth under short-day conditions (Zhao *et al.,* 2016). Within the photoperiodic pathway, *GmFT2a* functions at the extreme downstream, immediately prior to the activation of floral identity genes (Nan *et al.,* 2014).Completion of long-juvenile phase ensures sufficient vegetative growth to boost the crop productivity. Recently, Cai *et al.,* (2018, 2020) developed homozygous *Gmft2a* mutant soybean lines, which exhibited flowering delay of 6-7 days and produced an increased number of pods per plant in short-day conditions. Targeted mutagenesis of *GmFT2a* gene thus offers an effective solution to mitigate the problem of early flowering due to short-photoperiods. The present study represents a crucial initial step of the first attempt to target the *GmFT2a* gene using the CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9) technique to develop a long-juvenile soybean genotype specifically for Indian agriculture.

Over the years, CRISPR/Cas9 has emerged as a powerful and efficient genome-editing technology for crop improvement. Unlike traditional mutagenesis approaches, it is an efficient, simple and rapid technique. The system relies on two main components: a 20 bp synthetic gRNA, which aligns with the target DNA through complementary base pairing, and Cas9, a nuclease that acts as molecular scissors (Chen *et al.,* 2014). The gRNA with a scaffold bind to the Cas9 to form a gRNA-Cas9-Scaffold complex that scans the genome for a protospacer adjacent motif (PAM) next to the target sequence and creates a double-strand break (DSB) at the target site. These DSBs are subsequently repaired by the non-homologous end joining (NHEJ) repair (Wyman *et al.,* 2006). These repair mechanisms are error-prone, often introduces small deletion or insertion at the target site, disrupting the gene function. Recently, soybean genes for plant architecture, such as *GmSPL9*, GmLHY and GmAP1, were altered using CRISPR technique (Bao *et al.,* 2019; Cheng *et al.,* 2019; Chen *et al.,* 2020a). Designing of highly efficient gRNA is crucial for successful CRISPR/Cas9-mediated genome editing. Several factors such as GC content, target site location, nucleotide composition, secondary structure, proximity to the protospacer adjacent motif (PAM), and potential off-target influences gRNA efficiency. While computational tools such as CRISPR RGEN Tools (Bae *et al.,* 2014; Park *et al.,* 2015), CCTOP (Stemmer *et al.,* 2015), CRISPRP-P 2.0 (Liu *et al.,* 2017), CRISPOR (Haeussler *et al.,* 2016), E-CRISP

(Heigwer *et. al.*, 2014), CRISPR-PLANT (Xie *et al.*, 2014), CHOPCHOP (Labun *et al.*, 2019) and crisprBowtie (Hoberecht *et. al.*, 2022) have been developed to predict gRNA efficiency and off-target sites. Experimental cleavage efficiency analysis of gRNA designed using these computational tools remains essential, particularly in crop species with complex genomes, such as soybean. In the present study, we designed 25 gRNAs targeting the first exon of the *GmFT2a* gene in soybean. Based on their off-target and efficacy scores, 9 gRNAs were selected for *in vitro* cleavage efficiency analysis. A single high – efficiency gRNA was identified based on *in-vitro c*leavage efficiency analysis. A gRNA expression cassette – comprising the GmU6 promoter, gRNA sequence, and sgRNA scaffold – was designed for cloning into the CRISPR/Cas9 vector. The recombinant vector with a gRNA expression cassette was transformed into *Escherichia coli* strain DH5α and then mobilised into *Agrobacterium tumefaciens* strain EHA105.

Materials and methods

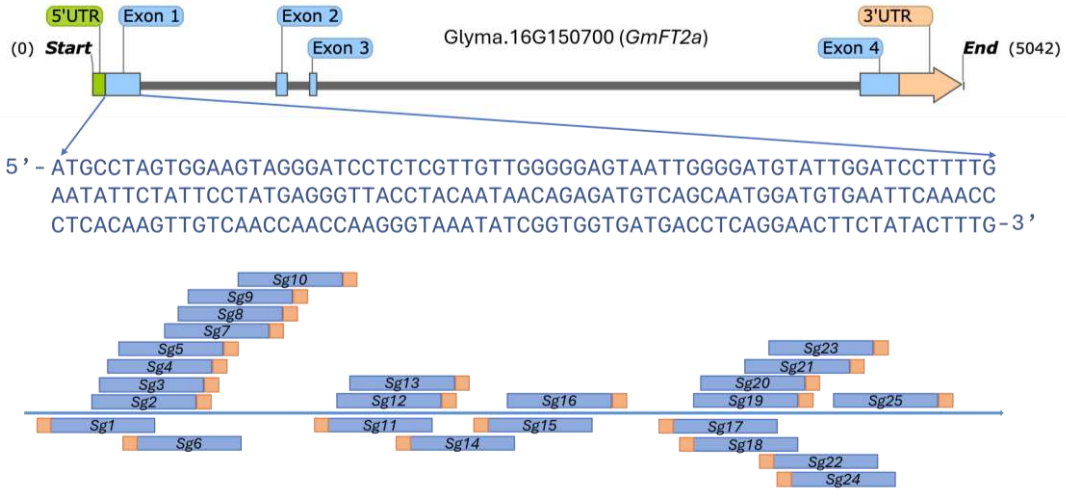
Target gene and sequence analysis

The complete gene sequence of *GmFT2a* (Glyma.16G150700) was retrieved from the soybean genome assembly *Glycine max* Wm82.a4.v1 (Phytozome genome ID: 508, NCBI taxonomy ID: 3847) in the Phytozome v14 database (https://phytozome.jgi.doe.gov/report/gene/Gmax_Wm82_a4_v1/Glyma.16G150700). The gene consists of four exons and three introns, with the coding sequence (CDS) spanning 531 base pairs encoding a 177 amino acid protein.

Designing of guide RNAs (gRNAs)

The first exon of the *GmFT2a* gene was selected as the primary target region for gRNA design to maximise the probability of generating frameshift mutations. Guide RNAs (gRNAs) were designed using the Cas-Designer component of CRISPR RGEN tools (http://www.rgenome.net/cas-designer/) by selecting the *Glycine max* Wm82.a4.v1 genome assembly, PAM type 5'-NGG-3' from *Streptococcus pyogenes* and gRNA length of 20 base pairs (Fig. 1).

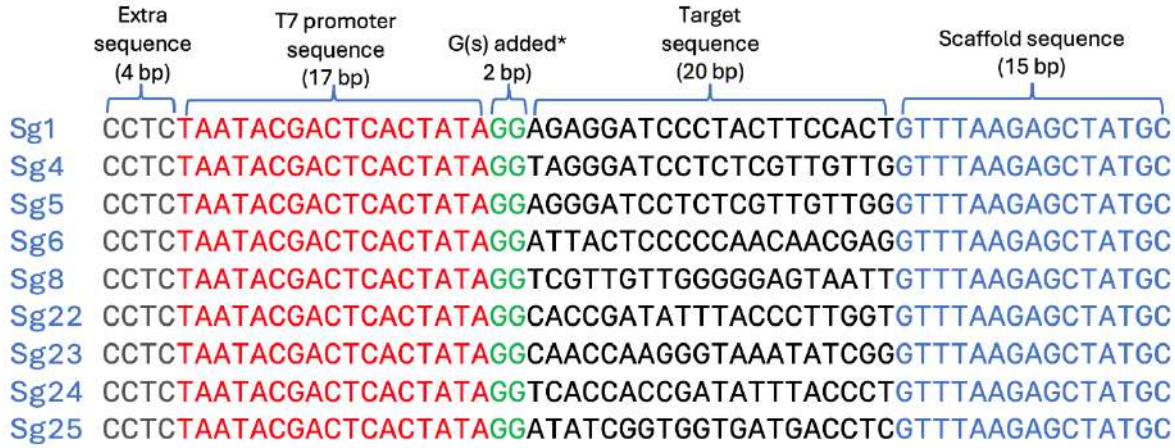
Fig. 1: Guide RNAs designed from the first exonic region of GmFT2a gene using the Cas-Designer component of CRISPR RGEN tools.



Off-target prediction and cleavage efficiency analysis

Off-targets and efficacy scores of designed gRNAs were determined using the Cas-OFFinder component of CRISPR RGEN tools (http://www.rgenome.net/cas-offinder/) and the CRISPRater tool of CCTop-CRISPR/Cas9 target online predictor (https://cctop.cos.uni-heidelberg.de:8043/), respectively. Based on the number of off-target sites less than 30 and efficacy scores greater than 0.60, nine gRNAs were selected for cleavage efficiency analysis using the manufacturer's protocol of Guide-it Complete sgRNA Screening Kit (Takara Bio USA, Inc.; Cat. No. 632636). For generating gRNA-encoding template using PCR, forward primers were designed for each selected gRNAs, which contain a T7 promoter sequence (17 bp), gRNA sequence (20 bp) and scaffold sequence (15 bp) (Fig. 2). These gRNA templates were in vitro transcribed using components provided in the kit. For cleavage assay, genomic DNA was extracted from young leaf of NRC150 soybean variety following the cetyl trimethyl ammonium bromide method (Doyle and Doyle, 1990). For PCR amplification of the first exon of *GmFT2a* gene, primers were designed using the Primer3 website (https://primer3.ut.ee/). For Cas9-gRNA cleavage assay, the amplified PCR product and each in vitro transcribed gRNA template were incubated with the Guide-it-recombinant Cas9 nuclease in individual reaction. The reaction products were analysed by capillary gel electrophoresis.

Fig. 2: Forward primer sequences designed for in vitro expression analysis of selected gRNAs



CRISPR/Cas9 expression vector

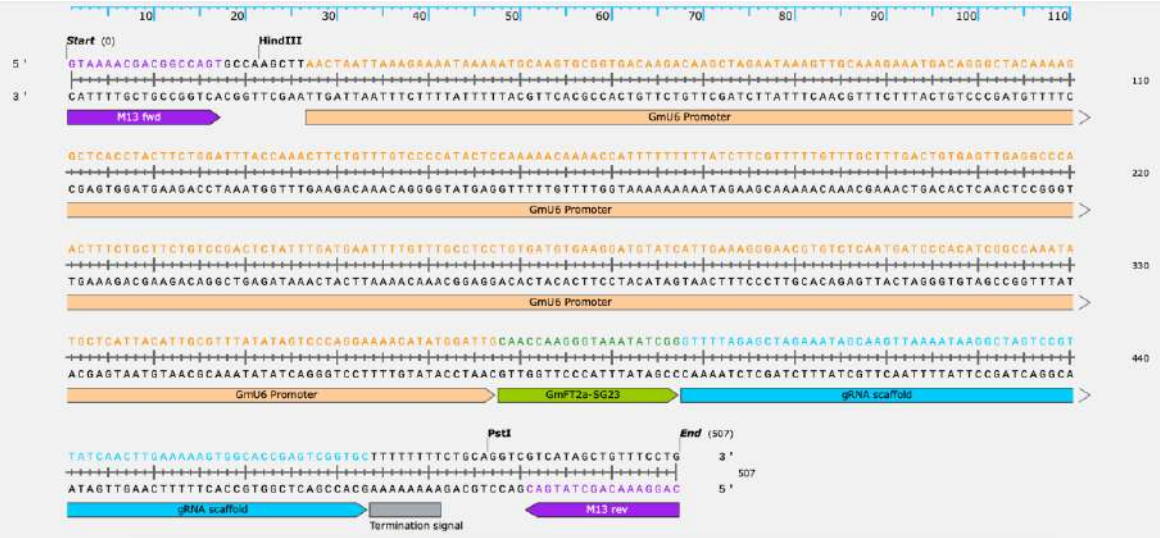
The CRISPR/Cas9 expression vector Cas9_MDC123 (Addgene plasmid #59184) was generously provided by Robert Stupar through the Addgene repository (RRID: Addgene_59184). The plasmid was received in the *Escherichia coli* (*E. coli*) DH5α strain as an agar stab and was revived on Luria–Bertani (LB) agar medium supplemented with 50 µg/ml kanamycin. The Cas9_MDC123 vector bacterial selection and Basta (Blpr) resistance for selection of transformed plants. For backbone contains a *Streptococcus pyogenes* Cas9 gene codon-optimised for *Glycine max* and includes a C-terminal SV40 nuclear localisation signal to facilitate efficient nuclear import of the Cas9 protein. Cas9 expression is regulated by a dual Cauliflower Mosaic Virus 35S (2x35S) promoter and terminated by the nopaline synthase (NOS) sequence. The vector has kanamycin resistance for

insertion of gRNA expression cassettes, the vector offers unique HindIII and PstI restriction sites.

Designing of gRNA expression cassette and transfer in Cas9_MDC123 expression vector

The guide RNA (gRNA) expression cassette (507 bp), comprising of GmU6 promoter (351 bp), the target-specific sequence (20 bp), the gRNA scaffold (76 bp), and a termination signal (8 bp), was designed using SnapGene software. For efficient cloning into the vector, the cassette was flanked with M13 forward primer sequence (17 bp) and a HindIII restriction site at the 5' end and with PstI restriction site followed by the M13 reverse primer sequence (17 bp) at the 3' end (Fig. 3). Cloning for insertion of gRNA expression cassette into Cas9_MDC123 expression vector was outsourced to Eurofins Genomics India using restriction enzyme-mediated cloning strategy.

Fig. 3: Schematic representation of the gRNA expression cassette



Bacterial transformation and primer designing for plasmid verification

The synthesised CRISPR/Cas9 expression vector *GmFT2a*_Cas9_MDC123 was transformed into *E. coli* DH5α. Competent cell preparation and transformation were performed using bacterial transformation kit (Thermo Scientific, Cat. No. K2711). Following transformation, cells were plated on LB agar supplemented with kanamycin (50 µg/ml) for antibiotic selection, and incubated overnight at 37°C. For screening of transformed colony, primers were designed using primer3 web tool to amplify the *Glycine max* codon-optimised *Cas9*, *Blpr* gene, and gRNA specific sequence for selection of gRNA cassette in the T-DNA region within the Cas9_MDC123 backbone (Table 1). Plasmid DNA was isolated from selected PCR-positive colonies and verified by Sanger sequencing.

Table 1: Oligonucleotide sequence of primer used for PCR screening

Name	Forward (5' to 3')	Reverse (5' to 3')
Cas9	CATGATCAAGTTCAGAGGACAC	GTGGTGCTCATCGTATCTCT
Blpr	CAACCACTACATCGAGACAAGCAC	CTGAAGTCCAGCTGCCAGAAA
sgRNA-specific primer	GTGCTGCAAGGCGATTAAGTTGG	TGTCGTGCTCCACCATGTTGACC
GmFT2a-specific primer	TCAAACACAATGGAATCGAGGC	ACTTGACCTTCCCTTAAACAC

The confirmed plasmid DNA was mobilised into *Agrobacterium tumefaciens* strain EHA105. Competent cells of EHA105 were prepared using the freeze-thaw method described by Hofgen and Willmitzer (1988), with minor modifications. Briefly, 1 ml of an overnight-grown primary culture of EHA105 was inoculated into 200 ml of LB broth supplemented with rifampicin (50 µg/ml) and incubated at 28°C with shaking for 4–5 h. Bacterial cells were harvested by centrifugation at 5000 rpm for 10 min, washed once with TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 7.5), and resuspended in the same buffer. For transformation, 500 µL of competent EHA105 cells were mixed with 1 µg of GmFT2a_Cas9_MDC123 plasmid DNA and incubated on ice for 5 minutes. The competent cell-DNA mixture was then frozen in liquid nitrogen for 5 min, followed by incubation at 37°C for 5 min. Thereafter, 500 µL of LB broth was added, and the culture was incubated at 28°C for 2 h. Transformed cells were plated on LB agar medium supplemented with kanamycin (50 µg/ml) and rifampicin (50 µg/ml) as selection markers and incubated at 28°C for 48 h. Putative *A. tumefaciens* transformants were screened by colony PCR using Cas9, Blpr and sgRNA-specific primers.

Results and discussion

Successful transfer of T-DNA into plants mainly depends on the efficiency of *Agrobacterium tumefaciens*–mediated transformation. In soybean, *A. tumefaciens*–mediated transformation efficiency is very low (Li *et. al.*, 2017); thus, generating large numbers of T-DNA insertion mutants requires considerable effort. To achieve high transformation efficiency, steps such as designing and selecting a highly effective gRNA and gRNA expression cassette, choosing a high-efficiency expression vector, and efficiently transferring the vector into an appropriate *Agrobacterium* strain are crucial.

Confirmation of *GmFT2a* target region sequence in soybean variety NRC150

To confirm the target sequence in soybean variety NRC150, the first exon of the *GmFT2a* gene was PCR amplified using the designed flanking primers (Table 1). The amplified product was subjected to sanger sequencing, and the obtained sequence was aligned with the reference *GmFT2a* sequence retrieved from the Phytozome database using Clustal W and UGENE software. The alignment confirmed 100% sequence similarity between both sequences.

Designing and cleavage efficiency analysis of guide RNAs (gRNAs)

A total of 25 gRNAs were designed to target the first exon of GmFT2a (Table 2). Targeting the first exon is advantageous, as mutations introduced at this position are likely to generate premature stop codons, ensuring a functional gene knockout (Cai *et al.*, 2018). These gRNAs were designed using the Cas-Designer module of the CRISPR RGEN tools, which was previously used by Song *et. al.*, (2022) to design gRNAs targeting the *GmIPK1* gene in soybean. The designed gRNAs were evaluated for off-targets and efficacy scores using the Cas-OFFinder component of CRISPR RGEN Tools and the CCTop-CRISPR/Cas9 target predictor, respectively. Do *et. al.*, (2019) utilised these tools to target the *GmFAD2–1A* and *GmFAD2–1B* genes in soybean. The designed gRNAs exhibited efficacy scores ranging from 0.54 to 0.8, with off-target counts between 11 and 382 (Table 2). To evaluate double-stranded break formation at the target locus, nine gRNAs (Sg1, Sg4, Sg5, Sg6, Sg8, Sg22, Sg23, Sg24 and Sg25) were selected for cleavage efficiency analysis based on low off-target sites (<30) and high efficacy scores (>0.60).

Table 2: List of gRNAs designed from first exon of *GmFT2a* gene

gRNA Name and gRNA Sequence (5' → 3')	Protospacer adjacent motif	1st Exon Position	Physical position in *Glycine max* Wm82.a4.v1	Direction	GC %	Efficacy score by CRISPRater	No. of off targets	Mismatches (throughout the whole genome up to 3 mismatches)			
								0	**1**	**2**	**3**
Sg1:AGAGGATCCC TACTTCCACT	AGG	4	31303265	–	50	0.76	14	0	1	1	2
Sg2:AGTAGGGATC CTCTCGTTGT	TGG	13	31303274	+	50	0.75	382	0	1	1	2
Sg3:GTAGGGATCC TCTCGTTGTT	GGG	14	31303275	+	50	0.75	202	0	1	1	6
Sg4:TAGGGATCCT CTCGTTGTTG	GGG	15	31303276	+	50	0.73	28	0	0	0	12
Sg5:AGGGATCCTC TCGTTGTTGG	GGG	16	31303277	+	55	0.74	11	0	0	0	2
Sg6:ATTACTCCCC CAACAACGAG	AGG	22	31303283	–	50	0.63	15	0	0	1	1
Sg7:CTCGTTGTTG GGGGAGTAAT	TGG	25	31303286	+	50	0.57	15	0	0	1	1
Sg8:TCGTTGTTGG GGGAGTAATT	GGG	26	31303287	+	45	0.70	29	0	0	1	4
Sg9:CGTTGTTGGG GGAGTAATTG	GGG	27	31303288	+	50	0.76	114	0	0	1	5
Sg10:GGGAGTAAT TGGGGATGTAT	TGG	36	31303297	+	45	0.54	84	0	0	1	38
Sg11:ATAGGAATA GAATATTCAAA	AGG	61	31303322	–	20	0.72	201	0	1	1	17
Sg12:TTGAATATT CTATTCCTATG	AGG	65	31303326	+	25	0.60	51	0	1	1	10
Sg13:TGAATATTC TATTCCTATGA	GGG	66	31303327	+	25	0.66	43	0	1	3	2
Sg14:TTATTGTAG GTAACCCTCAT	AGG	79	31303340	–	35	0.74	28	0	1	1	0
Sg15:GCTGACATC TCTGTTATTGT	AGG	92	31303353	–	40	0.69	23	1	1	0	1
Sg16:AATAACAGA GATGTCAGCAA	TGG	97	31303358	+	35	0.63	34	1	1	1	2
Sg17:GGTTGGTTG ACAACTTGTGA	GGG	133	31303394	–	45	0.76	17	2	0	1	5
Sg18:TGGTTGGTT GACAACTTGTG	AGG	134	31303395	–	45	0.70	20	2	0	0	6
Sg19:CACAAGTTG TCAACCAACCA	AGG	137	31303398	+	45	0.63	26	1	0	0	1
Sg20:ACAAGTTGT CAACCAACCAA	GGG	138	31303399	+	40	0.65	39	0	0	0	1
Sg21:AACCAACCA AGGGTAAATAT	CGG	148	31303409	+	35	0.62	40	0	0	2	6
Sg22:CACCGATAT TTACCCTTGGT	TGG	150	31303411	–	45	0.68	9	0	0	2	0
Sg23:CAACCAAGG GTAAATATCGG	TGG	151	31303412	+	45	0.80	11	0	0	2	0
Sg24:TCACCACCG ATATTTACCCT	TGG	154	31303415	–	45	0.73	2	0	0	2	0
Sg25:ATATCGGTG GTGATGACCTC	AGG	164	31303425	+	50	0.72	15	0	1	1	3

Additional explanations

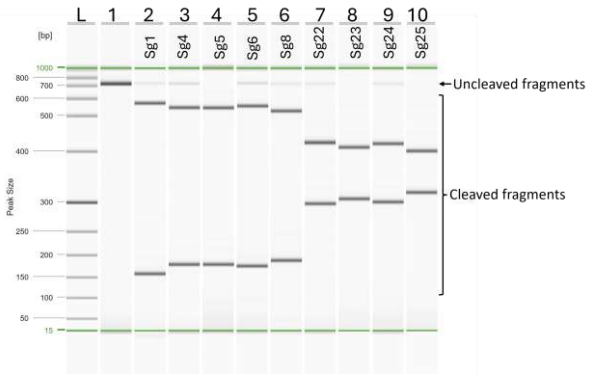
Position of first base of gRNA in 1st exon of Glyma.16G150700 (*GmFT2a*) gene
Physical position of first base gRNA on chromosome 16 in *Glycine max* Wm82.a4.v1 genome assembly
Position of gRNA in upper strand (+) or complementary lower strand (-) of DNA
Efficacy score calculated by CRISPRater tool of CCTop – CRISPR/Cas9 target online predictor website
The number of off-targets is determined using the Cas-OFFinder component of CRISPR RGEN tools.

To evaluate double-stranded break formation at the target locus, nine gRNAs (Sg1, Sg4, Sg5, Sg6, Sg8, Sg22, Sg23, Sg24 and Sg25) were selected for cleavage efficiency analysis based on low off-target sites (<30) and high efficacy scores (>0.60). An *in vitro* Cas9 cleavage assay was performed for each of the nine *in vitro* transcribed gRNA templates using a 720 bp PCR amplicon of the *GmFT2a* target region from the NRC150 variety and Guide-it recombinant Cas9 nuclease. Among these, three gRNAs, Sg5, Sg23, and Sg25, were identified as highly efficient. The 720 bp target DNA fragment was completely cleaved by the Sg5-Cas9 complex into 175 bp and 547 bp fragments, whereas the Sg23-Cas9 complex produced 308 bp and 412 bp fragments, and the Sg25-Cas9 complex generated 320 bp and 400 bp fragments (Table 3). In contrast, the remaining six gRNAs showed both cleaved and uncleaved fragments (Fig. 4). This in vitro analysis kit, previously used for cleavage efficiency analysis in rice and sweet potato (Bahariah *et. al.,* 2021; Brown *et. al.,* 2026), significantly reduces the labour and time required for *in vivo* studies. Our results demonstrate its high efficiency and reliability for soybean research.

Table 3: DNA fragments generated after *in vitro* Cas9 cleavage using designed gRNA

Lane No.	gRNA	Fragment size (bp)	
1	Untreated	720	0
2	Sg1	152	568
3	Sg4	172	548
4	Sg5	173	547
5	Sg6	170	550
6	Sg8	183	537
7	Sg22	298	422
8	Sg23	308	412
9	Sg24	302	418
10	Sg25	320	400

Fig. 4: Analysis of *in vitro Cas9*-gRNA cleavage by capillary gel electrophoresis. Lane L: 50 bp size marker; lane 1: untreated target DNA. Lanes 2, 3, 5, 6, 7 and 9 represent uncleaved fragments, and lanes 4, 8 and 10 represent completely cleaved fragments.



Three efficient gRNAs, Sg23 (5′-CAACCAAGGGTAAATATCGG-3′ with a TGG PAM) was selected for cloning. The target site for Sg23 is located 151 bp downstream of the start codon within the first exon of the *GmFT2a* gene. Although the sequence of Sg5 is identical to the GmFT2a-sp3 gRNA used by Cai *et al.,* (2018), Sg23 was preferred due to its higher efficacy score (0.80 compared to 0.74 for Sg5). Sg5 and Sg25 exhibited 9 and 12 off-targets in exonic regions, respectively, whereas Sg23 had only 3 off-target sites in exonic regions. Two of these off-targets of Sg23, *GmFT2b* (*Glyma.16G151000*) and *GmFT2c* (*Glyma.02G069200*), are members of the phosphatidylethanolamine-binding protein family and are homologs of the *GmFT2a* gene (Wu *et al.,* 2017). *GmFT2b* promotes flowering under long-day conditions and is not expressed under short-day conditions (Chen *et. al.,* 2020b), while *GmFT2c* is non-functional in cultivated soybean (Wu *et al.,* 2017). Therefore, even if these two homologs are targeted, no significant phenotypic effect is expected.

Transfer of gRNA expression cassette in Cas9_MDC123 vector and selection of transformed colony

The gRNA expression cassette containing the GmU6 promoter, Sg23 gRNA, and gRNA scaffold was designed (Fig. 3). The GmU6 promoter was specifically selected to enhance the efficiency of CRISPR/Cas9-mediated mutagenesis in soybean (Di *et al.,* 2019). Similarly, the Cas9_MDC123 vector used for cloning is highly specialised for soybean applications. The *Glycine max* codon-optimised Cas9 (*GmCas9*) in this vector was specifically modified for enhanced expression. Its efficacy has previously been validated by targeting glutamine synthase (*Glyma.18g041100*), chalcone-flavanone isomerase (*Glyma.20g241500*), and GmSWEET15 genes in soybean (Michno *et al.,* 2015; Wang *et al.,* 2019). The nucleotide sequences of *Arabidopsis* codon-optimised *Cas9* (*AtCas9*) and *GmCas9* are similar, approximately 11% of the bases have been changed to match soybean codon preference. In addition, *GmCas9* lacks the IV2 intron and carries only a single NLS at the C-terminus, whereas *AtCas9* contains NLSs at both the N- and C-termini (Michno *et. al.,* 2015). The vector also contains the Basta resistance (Blpr) gene for robust selection of genome-edited lines. The gRNA expression cassette was cloned into the Cas9_MDC123 vector (Fig. 5a) using a restriction enzyme-based double-digestion strategy. The resulting construct (Fig. 5b) was transformed into DH5α strain of *E. coli*, and putative transformants were selected on LB agar plates supplemented with kanamycin (50 µg/ml) after overnight incubation at 37 °C. To identify positive clones, kanamycin-resistant colonies were screened via PCR using Cas9-, bar- (Blpr), and sgRNA-specific primers. The amplification of 544 bp, 452 bp, and 650 bp fragments with the Cas9-, Blpr-, and sgRNA-specific primers, (Fig. 6), confirmed the successful integration of the GmFT2a_Cas9_MDC123 construct. Further, the correct orientation and sequence of the sgRNA cassette within the vector were confirmed by Sanger sequencing (Fig. 7).

Mobilization of GmFT2a_Cas9_MDC123 construct into *A. tumefaciens* EHA105

For mobilisation into *A.tumefaciens,* plasmid DNA was isolated from confirmed E. coli DH5α clones containing the GmFT2a_Cas9_MDC123 vector. The construct was then transferred into *A. tumefaciens* strain EHA105, a disarmed and hypervirulent strain with high transformation efficiency (Hood *et al.,* 1993). After incubation at 28°C for 48 h, two colonies were obtained on LB agar supplemented with kanamycin (50 µg/ml) and rifampicin (50 µg/ml).

Fig. 5: (a) Cas9_MDC123 vector

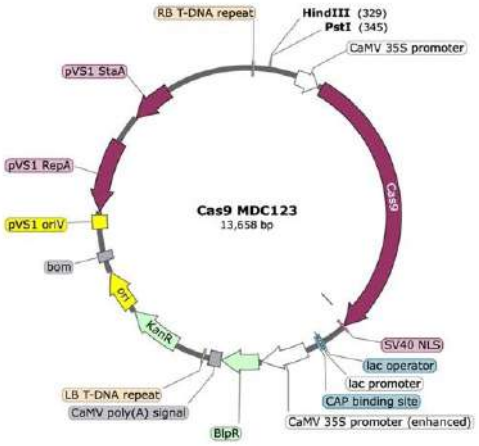


Fig. 5: (b) Cloned GmFT2a_Cas9_MDC123 vector

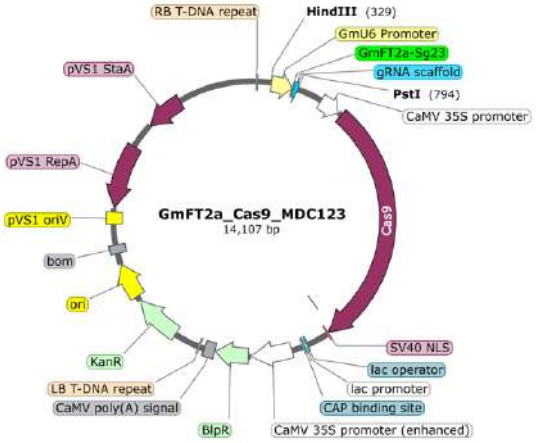
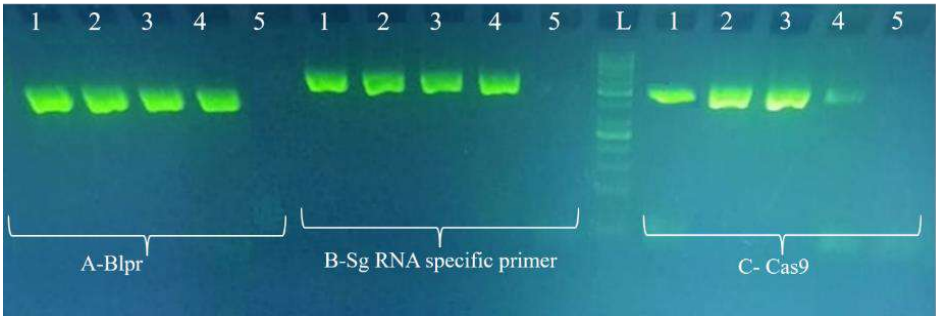


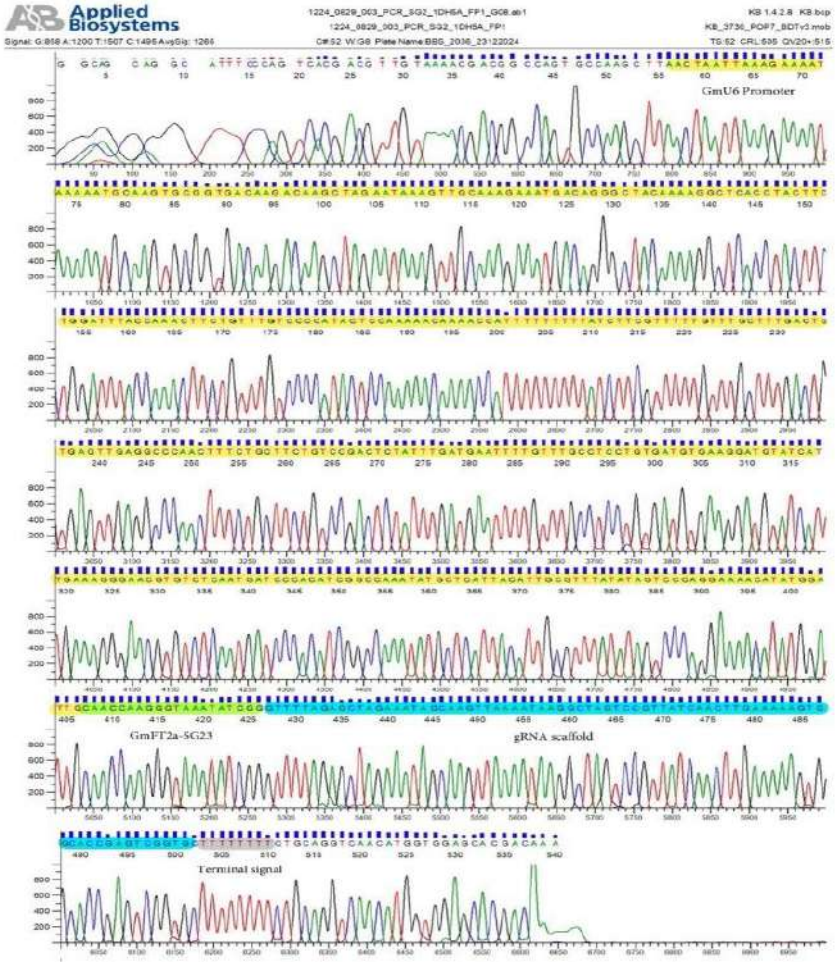
Fig. 6: PCR confirmation of transformed *E. Coli* DH5α colonies on 2% agarose gel using

A. Blpr primer (452 bp), B. SgRNA-specific primer (650 bp) and C. Cas9 primer (544 bp). Lanes 1-4 represent amplification of the respective primer in transformed colonies. Lane 5 represents non-template control for each set, showing no amplification. Lane L: 50 bp size marker



Although the number of transformed colonies obtained was relatively low, this is frequently observed with large vectors. In such experiments, maintaining the structural integrity of the vector in transformed cells is critical, as even a single verified colony is sufficient for plant transformation. PCR amplification using sgRNA cassette-, Cas9-, and Blpr-specific primers in the obtained colonies confirmed the complete mobilisation of the functional T-DNA region. This is an important step, as large vectors can sometimes be damaged during transfer from *E. coli* to *Agrobacterium* (Liu *et al.*, 2023). *Agrobacterium* mediated genetic transformation is one of the most reliable methods for soybean transformation. In future studies, the confirmed *Agrobacterium* clones developed in the present work will be used for transformation and editing of the *GmFT2a* gene to improve soybean productivity and performance.

Fig. 7: Sanger sequencing chromatogram of cloned GmFT2a_Cas9_MDC123 vector showing the presence of the GmU6 promoter, GmFT2a-Sg23 and gRNA scaffold sequence.



Hence, in conclusion, the present study designed and evaluated 25 CRISPR/Cas9 gRNAs for targeted editing of the soybean *GmFT2a* gene to develop the long juvenile genotype. Based on *in silico* screening and in vitro cleavage efficiency analysis, the Sg23 gRNA was identified as the most effective candidate, as it demonstrated a high efficacy score of 0.80, a minimal number of exonic off-targets, and superior in vitro cleavage performance. The Sg23 gRNA was successfully cloned into the Cas9_MDC123 vector under the control of the GmU6 promoter. The recombinant construct was verified and mobilised into *A. tumefaciens* EHA105, providing a functional platform for the genome editing of *GmFT2a* gene. Overall, this work establishes a robust foundation for CRISPR/Cas9-mediated *GmFT2a* mutagenesis and the development of soybean lines with improved adaptation and productivity under short-day conditions.

Acknowledgements

This work was supported by the WISE-PhD fellowship (Sanction No. DST/WISE-PhD/LS/2024/302), Department of Science and Technology, Ministry of Science and Technology, Government of India.

References

1. Bae, S., Park, J. and Kim, J.S. 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics, 30(10):1473-1475.
2. Bahariah, B., Masani, M. Y. A., Rasid, O. A. and Parveez, G. K. A. 2021. Multiplex CRISPR/Cas9-mediated genome editing of the FAD2 gene in rice: A model genome editing system for oil palm. J. Genet. Eng. Biotechnol., 19(1):86.
3. Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., Qiu, D., Shan, Z., Yang, Z., Yuan, S., Zhang, C., Zhang, X., Liu, B., Kong, F., Li, X., Zhou, X., Tran, L.P. and Cao, D. 2019. CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9 genes alters plant architecture in soybean. BMC Plant Biol., 19(1):131.
4. Bhatia, V.S., Tiwari, S.P. and Joshi, O.P. 1999. Yield and its attributes as affected by planting dates in soybean (*Glycine max*) varieties. Indian J. Agric. Sci., 69:696-699.
5. Brown, A. P. A., Egnin, M., Bukari, F., Ritte, I. P. and Bernard, G. C. 2026. Establishment of efficient CRISPR-Cas9 PEG-mediated DNA-free genome editing through ribonucleoproteins method in hexaploid sweetpotato (*Ipomoea batatas* L. (Lam)) targeting the EIF-4E genes. Plants, 15(3):447.
6. Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B., Han, T. and Hou, W. 2018. CRISPR/Cas9-mediated targeted mutagenesis of *GmFT2a* delays flowering time in soya bean. Plant Biotechnol. J., 16:176-185.
7. Cai, Y., Wang, L., Chen, L., Wu, T., Liu, L., Sun, S., Wu, C., Yao, W., Jiang, B., Yuan, S., Han, T. and Hou, W. 2020. Mutagenesis of *GmFT2a* and GmFT5a mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. Plant Biotechnol. J., 18(1):298-309.
8. Chen, K. and Gao, C. 2014. Targeted genome modification technologies and their applications in crop improvements. Plant Cell Rep., 33:575–583.
9. Chen, L., Nan, H., Kong, L., Yue, L., Yang, H., Zhao, Q., Fang, C., Li, H., Cheng, Q., Lu, S., Kong, F., Liu, B. and Dong, L. 2020a. Soybean AP1 homologs control flowering time and plant height. J. Integr. Plant Biol., 62:1868-1879.
10. Chen, L., Cai, Y., Qu, M., Wang, L., Sun, H., Jiang, B., Wu, T., Liu, L., Sun, S., Wu, C., Yao, W., Yuan, S., Han, T. and Hou, W. 2020b. Soybean adaption to high-latitude regions is associated with natural variations of *GmFT2b*, an ortholog of FLOWERING LOCUS T. Plant Cell Environ., 43:934-944.
11. Cheng, Q., Dong, L., Su, T., Li, T., Gan, Z., Nan, H., Lu, S., Fang, C., Kong, L., Li, H., Hou, Z., Kou, K., Tang, Y., Lin, X., Zhao, X., Chen, L., Liu, B. and Kong F. 2019. CRISPR/Cas9-mediated targeted mutagenesis of *GmLHY* genes alters plant height and internode length in soybean. BMC Plant Biol., 19(1):562.
12. Di, Y.H., Sun, X.J., Hu, Z., Jiang, Q.Y., Song, GH., Zhang, B., Zhao, S.S. and Zhang, H. 2019. Enhancing the CRISPR/Cas9 system based on multiple GmU6 promoters in soybean. Biochem. Biophys. Res. Commun., 519(4):819-823.
13. Do, P.T., Nguyen, C.X., Bui, H.T., Tran, L.T.N., Stacey, G., Gillman, J.D., Zhang, Z.J. and Stacey, M.G. 2019. Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous *GmFAD2-1A* and *GmFAD2-1B* genes to yield a high oleic, low linoleic and α-linolenic acid phenotype in soybean. BMC Plant Biol., 19(1):311.

14. Doyle, J. J. and Doyle, J. L. 1990. Isolation of plant DNA from fresh tissue. Focus, 12(1):13-15.
15. Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A., Mianne, J., Renaud, J. B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly, J. S. and Concordet, J. P. 2016. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Bio., 17(1):148.
16. Heigwer, F., Kerr, G. and Boutros, M. 2014. E-CRISP: fast CRISPR target site identification. Nat. Methods, 11(2):122-123.
17. Hoberecht, L., Perampalam, P., Lun, A. and Fortin, J. 2022. A comprehensive bioconductor ecosystem for the design of CRISPR guide RNAs across nucleases and technologies. Nat. Commun., 13(1):6568.
18. Hofgen, R. and Willmitzer, L. 1988. Storage of competent cells for *Agrobacterium transformation*. Nucleic Acids Res., 16(20):9877.
19. Hu, M. and Wiatrak, P. 2012. Effect of Planting date on soybean growth, yield, and grain quality: Review. Agron. J., 104:785-790.
20. Labun, K., Montague, TG., Krause, M., Torres Cleuren, YN., Tjeldnes, H. and Valen, E. 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res., 47(W1):W171-W174.
21. Li, S., Cong, Y., Liu, Y., Wang, T., Shuai, Q., Chen, N., Gai, J. and Li, Y. 2017. Optimization of *Agrobacterium*-mediated transformation in soybean. Front. Plant Sci., 8:246.
22. Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K. and Chen, L. L. 2017. CRISPR-P 2.0: An improved CRISPR-Cas9 tool for genome editing in pants. Mol. Plant., 10(3):530-532.
23. Liu, X., Miceli, J.F., III, Patton, S., Murray, M., Evans, J., Wei, X. and Wang, P. 2023. Agrobacterial transformation enhancement by improved competent cell preparation and optimized electroporation. Life, 13(11):2217.
24. Michno, J. M., Wang, X., Liu, J., Curtin, S. J., Kono, T. J. Y. and Stupar, R. M. 2015. CRISPR/Cas mutagenesis of soybean and *Medicago truncatula* using a new web-tool and a modified Cas9 enzyme. GM Crops Food., 6(4):243-252.
25. Nan, H., Cao, D., Zhang, D., Li, Y., Lu, S., Tang, L., Yuan, X., Liu, B. and Kong, F. 2014. *GmFT2*a and *GmFT5a* redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor GmFDL19 in soybean. PLOS ONE, 9(5):e97669.
26. Park, J., Bae, S. and Kim, J.-S. 2015. Cas-Designer: A web-based tool for choice of CRISPR-Cas9 target sites. Bioinformatics, 31(23):4014-4016.
27. Song, J.H., Shin, G., Kim, H.J., Lee, B.S., Moon, J.Y., Jeong, J.C., Choi, H.K., Kim, I.A., Song, H.J., Kim, C.Y. and Chung, Y.S. 2022. Mutation of *GmIPK1* gene using crispr/cas9 reduced phytic acid content in soybean seeds. Int. J. Mol. Sci., 23(18):10583.
28. Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J. and Mateo, J.L. 2015. CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PLOS ONE, 10(4):e0124633.
29. Sun, H., Jia, Z., Cao, D., Jiang, B., Wu, C., Hou, W., Liu, Y., Fei, Z., Zhao, D. and Han, T. 2011. *GmFT2a*, a soybean homolog of FLOWERING LOCUS T, is involved in flowering transition and maintenance. PLOS ONE, 6(12): e26231.
30. Wang, S., Yokosho, K., Guo, R., Whelan, J., Ruan, Y.-L., Ma, J.F. and Shou, H. 2019. The soybean sugar transporter *GmSWEET15* mediates sucrose export from endosperm to early embryo. Plant Physiol., 180(4): 2133-2141.

31. Wu, F., Sedivy, E. J., Price, W. B., Haider, W. and Hanzawa, Y. 2017. Evolutionary trajectories of duplicated FT homologues and their roles in soybean domestication. Plant J., 90(5): 941-953.
32. Wyman, C. and Kanaar, R. 2006. DNA double-strand break repair: all's well that ends well. Annu. Rev. Genet., 40: 363-383.
33. Xie, K., Zhang, J. and Yang, Y. 2014. Genome-wide prediction of highly specific guide RNA spacers for the CRISPR-Cas9 mediated genome editing in model plants and major crops. Mol. Plant, 7(5): 923-926.
34. Zhao, C., Takeshima, R., Zhu, J., Xu, M., Sato, M., Watanabe, S., Kanazawa, A., Liu, B., Kong, F., Yamada, T. and Abe, J. 2016. A recessive allele for delayed flowering at the soybean maturity locus E9 is a leaky allele of FT2a, a FLOWERING LOCUS T ortholog. BMC Plant Bio., 16: 20.

.