



Designing of MRPABCC5 specific CRISPR/Cas9 construct and its transient analysis using AGRODATE technique in soybean leaf discs

SMRUTIREKHA SAHU¹, ARCHANA SACHDEV², MONICA JOLLY*³, VEDA KRISHNAN⁴, ASHISH MARATHE⁵, JOSHNA JOSE⁶ and SHELLY PRAVEEN⁷

ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India

Received: 24 January 2020; Accepted: 24 January 2020

ABSTRACT

CRISPR/Cas9 has emerged as a tool of choice to precisely edit the genes pivotal for crop improvement programmes and gene function analysis. However, one of the puzzling aspects for utilization of this technique is the stability and efficiency of the construct. In the present study, we have designed the constructs dictated by certain parameters using web tools such as CRISPRscan, RNAfold webserver and OligoAnalyzer. The construct was transiently expressed in the soybean leaf discs using a vacuum based AGRODATE method (*Agrobacterium* mediated Transient Expression Assay) to edit the phytate transporter gene, i.e. *GmABCC5*. We observed an insertion of a nucleotide producing a SNP change in an analysed leaf disc sample depicting the efficacy of the construct.

Key words: CRISPR/Cas9, MRP ABC transporter, Phytic acid, Soybean

The global population stands at 7.7 billion that is expected to hit a mark of 8 billion by the end of the decade, which has shown a staggering rise by 1.08% from the preceding year. A major chunk of the growing population is significantly affected by an ailment called hidden hunger, due to their inaccessibility to nutrient-rich food. Therefore, to mitigate this silent epidemic, nutritionally enriched crops like soybean could be sought after which contains a plethora of bioactive components and nutraceuticals (Krishnan *et al.* 2018).

While soybean products have marked a niche in the consumer market, the presence of Anti-Nutritional Factors (ANFs) have significantly affected its potential consumption (Kumari *et al.* 2015). Amongst ANFs, Phytic Acid (PA) draws an unparalleled attention owing to the presence of negatively charged phosphate groups which chelates the divalent cations like Fe^{2+} , Zn^{2+} , Ca^{2+} and formation of insoluble phytate salts, severely hindering the bioavailability of these essential micronutrients. Proliferating occurrences of algal bloom and eutrophication collapsing the oxygen supply have been reported in water bodies due to the surface runoff of PA complexes which gets away undigested because of absence of phytase in the gut of monogastric animals. Taking into account these adverse impacts, numerous studies have been instigated amongst experimental groups across the world aimed at developing low PA mutants for soybean.

Despite its antinutritional attributes and ecological concerns, PA makes up approximately 75% of the total seed phosphorus content (Raboy *et al.* 1984) required for seed germination and other developmental processes, which necessitates its content balance for carrying out effective metabolic activities in plants. PA gets synthesised by two independent pathways, with several genes in the pathway serving as potential targets like *MIPS*, *IPKs*, *ITPK* and *ABCC5*, which were characterized for possible manipulation and development of low PA crops (Kumari *et al.* 2012; Krishnan *et al.* 2015; Basak *et al.* 2017; Punjabi *et al.* 2018; Pandey *et al.* 2018 and Marathe *et al.* 2018). Though experiments support the critical involvement of an early pathway *MIPS1* gene as key regulator of PA content, silencing of this gene has resulted in undesirable disturbance of inherent Pi and inositol homeostasis. Reports suggest a significant reduction in PA levels by utilising RNAi based silencing of late pathway gene (*IPK2*) and phytate transporter gene *MRP ABC* (Multidrug resistance-associated protein ATP Binding Cassette), thus maintaining the optimal phosphate inositol ratio crucial for survival of the plant. Post synthesis of the PA, it gets stored in the vacuoles actively channelled via a transporter protein encoded by *MRP* gene, which belongs to the ABC family. Role of these InsP6 transporters were confirmed by further studies in Arabidopsis *AtMRP5* and rice *OsMRP5*, and phylogenetic analysis of which confirmed their classification into ABC class of transporters as single copy genes. Two plausible reasons have been cited for the involvement of these transporters in biosynthesis of PA that are backed by experimental evidences (Shi *et al.* 2007) a) they might restrict the flow

*Corresponding author e-mail : monicajolly28@gmail.com

Target sequence	Primer sequence
GCTGTTCTGGTGGCTACGTTTGG	Forward Primer: 5' <u>GATTGCTGTTCTGGTGGCTACGTT</u> 3' Reverse Primer: 5' <u>AAACAACGTAGCCACCAGAACAGC</u> 3'

of phosphate to the substrate, *myo*-inositol essential for InsP-kinase to carry out the biosynthesis, b) the negative feedback mechanism exercised by these transporters, implying decreased biosynthesis due to accumulation of PA in cytosol. Therefore, targeting the *MRP ABC* transporter gene could be helpful in creating low PA mutants without adversely affecting the phenotypic attributes of the plant.

Successful incidents of *lpa* mutant generation have been reported utilising classical breeding approaches and reverse genetics showing a significant reduction of PA content with augmented Pi levels. However, events of dismal agronomic performance, poor germination rate, off-target mRNA silencing have notoriously discredited the former strategies. Furthermore, RNAi impacts post transcriptional level of gene expression and does not disrupt the gene function therefore, unable to elucidate the particular gene function analysis. Taking into account the abovesaid factors, an easier, fast paced and precise gene editing tool have surfaced to the rescue, i.e. Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated system (CRISPR/Cas9).

CRISPR/Cas9 belongs to the advanced era of Genome Editing Tools (GETs) which requires a well-designed single guide RNA (sgRNA) sequence to drive the enzyme machinery after recognition of Protospacer Adjacent Motif (PAM) sequence, that differs from its counterparts, i.e. Zinc Finger Nucleases (ZFNs), Transcription activator-Like Effector Nucleases (TALENs) which carry a protein-based endonuclease called Fok-I and requires its dimerization to carry out Double Strand Breaks (DSBs) in the target gene knockout thus making it difficult to programme the protein machinery and dimerization for individual knockout experiment. Thus the ease of target design and requirement of single-protein effectors renders the CRISPR/Cas system a more user-friendly approach in creation of gene-edited organisms. With several reports pouring in for crop improvement using CRISPR/Cas9, this technique could be responsible to dramatize the face of agriculture in the coming years.

Even though this technique has been on the forefront, incidences of off-target effects shown by the gRNAs have made the exact mechanism of this technique to remain elusive. Recent times have seen an upsurge of *in silico* tools for predicting the sgRNA efficiency and differentiate them on the basis of their probable activities inside the plant system. This could potentially fast pace the choice of target sequences required to edit the gene of interest and decline the time window of generating the edited crop plants.

Transient transformation assays enable for rapid and easy assessment of the sgRNA efficiency before proceeding

for stable transformation (Guidarelli *et al.* 2015). In the present study, CRISPR/Cas9 has been employed to edit the PA transporter gene, *GmABCC5* with a unique *Agrobacterium* mediated Transient Expression Assay (AGRODATE) in soybean leaf discs (Krishnan *et al.* 2019) overcoming the recalcitrance of soybean towards stable transformation (Hada *et al.* 2018).

MATERIALS AND METHODS

Plant materials

Soybean (*Glycine max*) seeds of var. DS-9712 were sown in pots and maintained in the National Phytotron Facility (NPF) at 28°C, 10:14 hr of photoperiod condition for a duration of 6 weeks.

Selection of target-specific sgRNA

Using the FASTA sequence (KF183505.1) as input into the Stupar's CRISPR website (<http://stuparlab.cfans.umn.edu/protocols/crisprcas9-glycine-max>), it yielded five potential sgRNA sequences specific for *GmABCC5*. With respect to low off-target activity shown by the targets using CasOFFinder (<http://www.rgenome.net/cas-offinder/>), CRISPRscan (<http://www.crisprscan.org/>) and NEB cutter (<http://nc2.neb.com/NEBcutter2/>), it fetched us three favourable sgRNAs containing 5' NGG 3' PAM (Protospacer Adjacent Motif) required for the activity of *Streptomyces pyogenes*. These sgRNAs were then examined for their efficiency using RNAfold Webserver program (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) and OligoAnalyzer (<https://www.idtdna.com/pages/tools/oligoanalyzer>) to characterize the secondary structure, Minimum Free energy (MFE) and GC content of the target sequences.

Construction of the vectors

The construction framework was kept in accordance with the assembly protocol developed by Stupar's lab. Forward and reverse primers (Table 1) containing the BbsI overhangs were used for designing the construct subsequent to annealing. The annealed target oligos were then ligated into pBlu/gRNA containing the 557 bp gRNA cassette and ampicillin resistance. DH5α strain of *E. coli* was transformed with the recombinant vector and cloning was confirmed through restriction digestion with EcoRI (Conc. 10000 units/mL). Recombinant clones were thereafter screened through colony PCR using gRNA specific primers and confirmed through Sanger sequencing. Plasmids from PCR positive clones were isolated using QIAprep Spin Miniprep Kit (250). The Cas9/gRNA-*GmABCC5* construct

was transformed into the competent *Agrobacterium* EHA105 cells, spread on LA-Kan plates and incubated at 28°C. Randomly picked colonies were subjected to colony PCR to identify the positively cloned cells.

Transient expression assay of gene constructs in soybean leaf discs

In order to analyse the efficiency of gene constructs in a rapid and simple manner, transient expression of the theoretically validated construct was followed using AGRODATE technique designed specifically for soybean leaves (Krishnan *et al.* 2019). Soybean leaves of 8-week old plants were collected, leaf discs were cut and were treated with 200 ml of 70% ethanol for 2 min and, later with 200 ml of 30% sodium hypochlorite for 20 min. The leaf discs were then incubated in an infiltration medium containing agrobacterium cells harbouring Cas9/gRNA-*GmABCC5* suspended in a medium containing 10mM 2-(N morpholino) ethane sulphonic acid sodium salt, 10mM MgCl₂ (pH 5.4), 0.2M acetosyringone, 400mg/l cysteine, 0.01% Tween 20 and 0.5 mM DTT subjected to a final OD of 0.6. A 20 ml disposable syringe was used to draw up to 10 ml of buffer and leaf discs were put into it. To remove any air gap, tip was sealed with a parafilm, the plunger was pulled back to create vacuum, shaken vigorously for 2 min and finally, the plunger was released. The leaf discs were removed out, washed in sterile water and placed on MS-media containing petri plates which were incubated in the growth chamber having 80% humidity, at 26±2°C and 16:8 hr photoperiod for 4 days.

Genomic DNA isolation and detection of mutations

After a span of 4 days, genomic DNA was isolated from the leaves using CTAB method. PCR amplification of target genes were carried out with gene specific primers to identify mutation in soybean leaf discs. The PCR product was purified using ThermoFisher Scientific Gel Extraction Kit and was outsourced for sequencing. The sequencing results were analysed through NCBI-BLAST.

Statistical analysis

Data was represented as mean ± SE (n=3). *P* value <

0.05 was considered statistically significant.

RESULTS AND DISCUSSION

In silico analysis of the sgRNA sequences

To obtain the desired target sequences for soybean genome, *Glycine max* specific web-tool was used to design the sgRNA sequences along with their forward and reverse primers and corresponding restriction enzyme sites (Michno *et al.* 2015). It yielded five oligonucleotides out of which three have been selected on the basis of their off-target activity using CasOFFinder. The remaining protospacer sequences were checked for their efficacy by utilising several web-tools like CRISPRscan score, Minimum Free Energy (MFE) (Fig 1a, 1b) and GC content of the target sequences for *GmABCC5*. Out of the three targets examined for the above parameters (Table 2), target 3 showed the highest CRISPRscan score of 59, -1.5 kcal/mol of Minimum Free Energy depicting stable secondary structure for target 3 and higher GC content of 56.5% making it a better suited target to proceed for designing the CRISPR construct.

The stability of sgRNA is an essential feature which is suggested to have an impact on the consistency of Cas9 activity. CRISPRscan uses a unique algorithm to rank the target sequences on the basis of predicted values by utilising yeast as a model organism with the highest score given to the gRNA targets with correctly oriented structure (Moreno-Mateos *et al.* 2015). The proper folding of the secondary structure of 21 bp protospacer sequence and their corresponding sgRNA structure yielded MFE value which was obtained from RNAfold webserver (Gruber *et al.* 2008), with more negative values showing enhanced stability of the protospacer as well as the sgRNAs. Protospacer sequences showing circled configuration for their secondary structure were found to bear MFE value of 0 kcal/mol which were theoretically unstable (Chu *et al.* 2016). GC content were shown to affect sgRNA secondary (Doench *et al.*, 2014) which had an impact on their effective attachment to Cas9 endonuclease and a range of 40-70% have been devised to be optimum for effective editing of the gene (Ren *et al.* 2014). GC content were obtained by putting the target as an input into OligoAnalyzer 3.0.

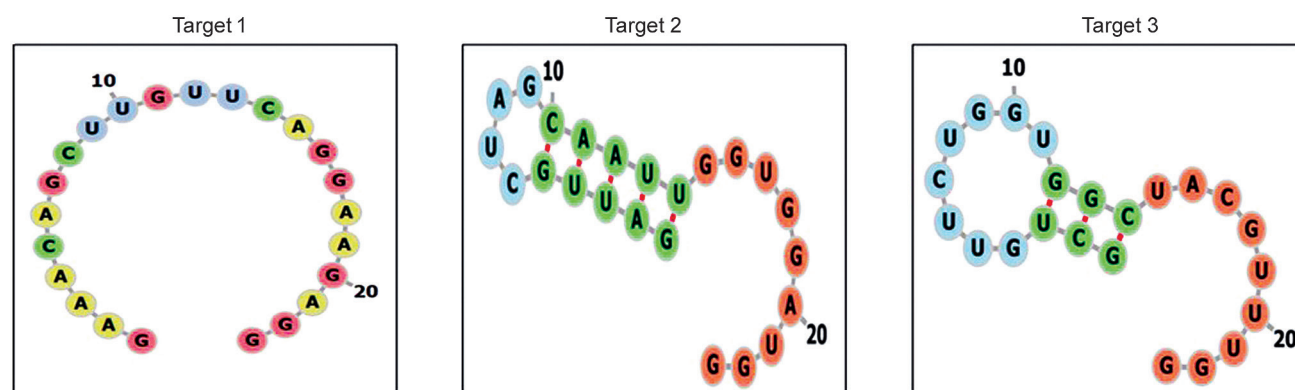


Fig 1 a - Secondary structure of the protospacer sequences of target 1, target 2, target 3.

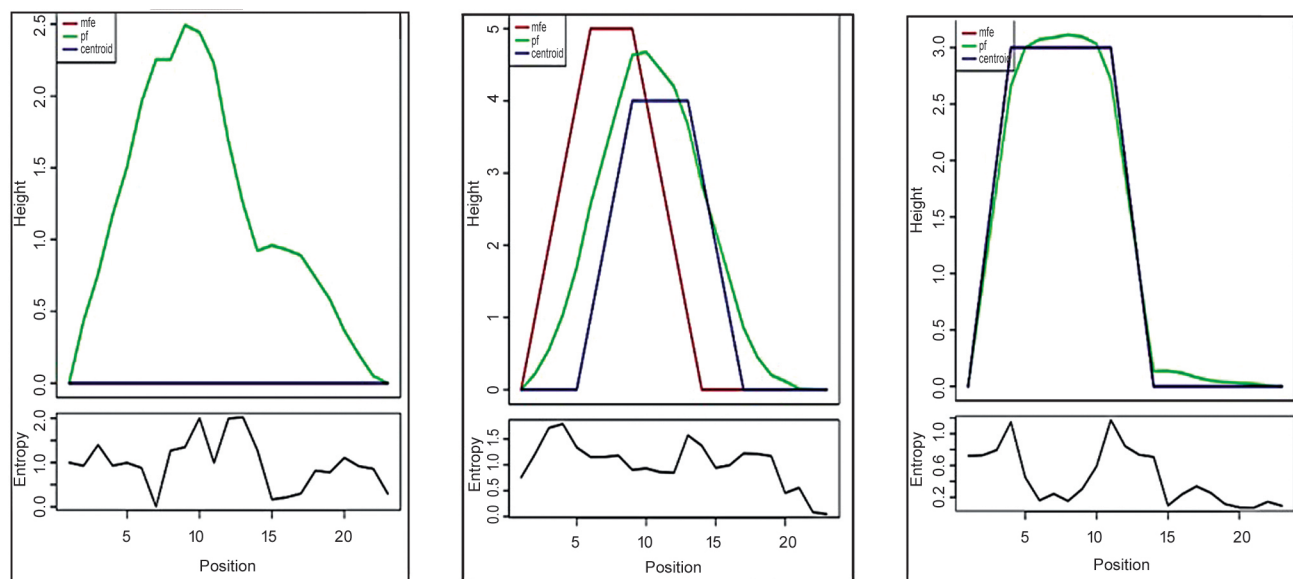


Fig 1 b - Mountain plots corresponding to the MFE (Minimum Free Energy) of the target 1, target 2 and target 3

Table 2 *In silico* tool-based sgRNA efficiency prediction values

Target no.	CRISPRscan score	MFE of target sequence (kcal/mol)	GC%
1.	40	0	47.8
2.	36	-1.10	47.8
3.	59	-1.5	56.5

Designing the construct with target sgRNAs

The target oligos, devoid of any off-target effects were annealed into the shuttle vector, pBlu/gRNA shuttle vector containing ampicillin resistance. DH5 α strains of *E. coli* were transformed with pBlu/gRNA-*GmABCC5*. The cloning was confirmed by restriction digestion with EcoRI showing a release of 557 bp containing pBlu/gRNA-*GmABCC5* run on 1% agarose gel and sequencing confirmed the insertion of the target oligos in the shuttle vector (Fig 2 a, b).

The gRNA cassettes from positive clones were further sub cloned into destination vector, Cas9MDC123

through restriction enzyme-based cloning using EcoRI. The recombinant vector was mobilized into DH5 α and positive colonies were confirmed through colony PCR and sequencing (Fig 3 a, b).

Transient transformation

The positive clones containing the Cas9/gRNA-*GmABCC5* were further mobilized into EHA105 strain of *Agrobacterium tumefaciens* for transient transformation in the soybean leaf discs.

Plant tissue culture of dicots like soybean pose certain impediments owing to the tedious amount of labour and lower efficiency of stable transformation. Thus, it would be beneficial to lessen the time window and screen the efficiency of the gene constructs through transient expression assays (Sparkes *et al.* 2006). Out of several assays developed, we have used a unique AGRODATE technique for examining the transient expression of Cas9/gRNA-*GmABCC5* constructs in the leaf discs from 8-week old soybean plants

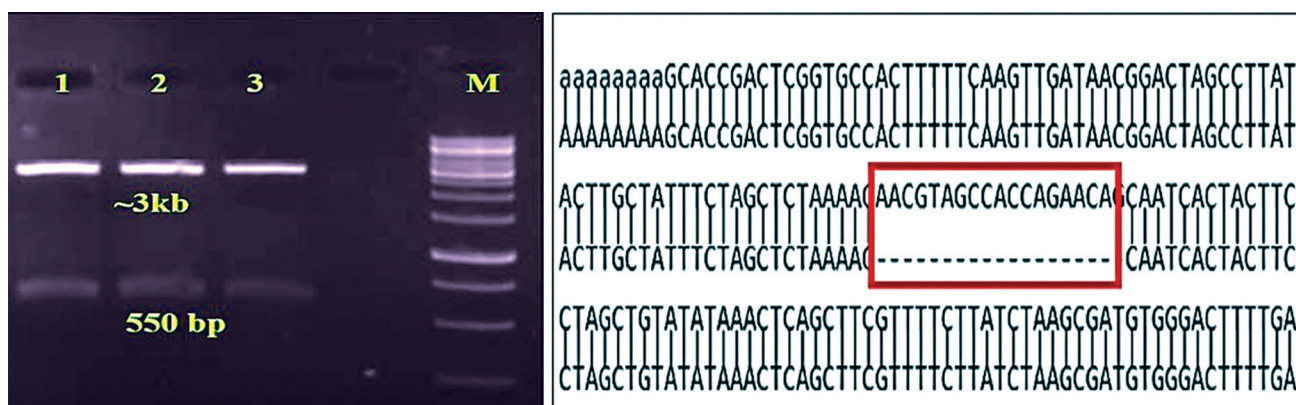


Fig 2 a) Restriction analysis of gRNA-*GmABCC5* fragment from pBlu/gRNA vector with EcoRI (Lane M: 1Kb ladder, Lane 1,2,3: Restriction of pBlu/gRNA-*GmABCC5* with EcoRI); b) Sequence confirmation of the target via BLAST alignment with vector backbone of pBlu/gRNA.

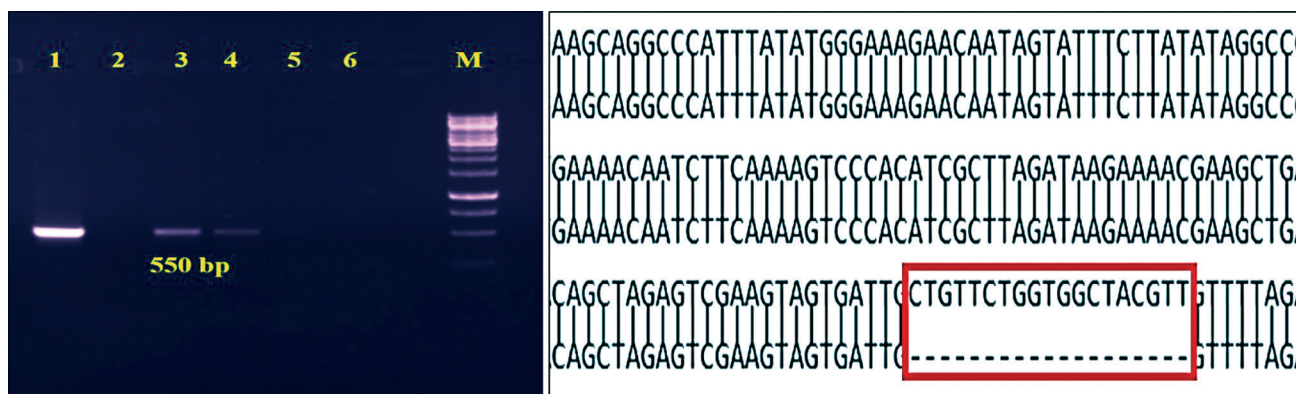


Fig 3 a) PCR amplification of random colonies of Cas9/gRNA-GmMRPABCC5 with the gRNA specific primers (Lane 1,3,4 – DNA amplicon of 550bp, Lane 6 – no amplification of negative, i.e. Cas9 gene with gRNA specific primers, Lane M – 1 kb DNA marker); b) Sequence confirmation of the target via BLAST alignment with vector backbone of pBlu/gRNA.

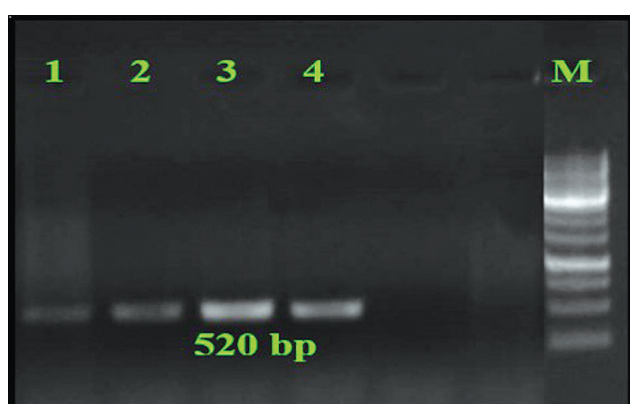


Fig 4 Gel picture showing positive PCR amplicon of genomic DNA with ABCC5 specific primers (Lane 1,2,3,4 : Positive amplicons, Lane M : 1kb marker).

(Krishnan *et al.* 2019; Nekrasov *et al.* 2013).

Mutation analysis

Post transformation, genomic DNA was isolated from the transformed leaf discs, amplified using gene specific (*GmABCC5*) primers (Table 3) (Fig 4) which were then sequenced for mutation analysis. Sequencing results confirmed the presence of a single base insertion at 4 bp upstream of Protospacer Adjacent Motif (PAM) sequence in one of the four samples transformed (Fig 5).

Amongst the several variants of CRISPR/Cas9, type II is more preferable owing to the use of a single Cas9 protein to generate breaks. Cells have inherent repair machinery which gets induced by Double Stranded Breaks (DSBs) and

Table 3 *MRPABCC5* gene specific primers for colony PCR in soybean

Gene specific cassette primers	T _m (°C)
Forward primer: 5' ATGAAGCCATAGAGGCCATGGATATCCCTAC 3'	64
Reverse primer: 5' TAGCTTCTGTGCAGCTGCAAGACCAAAC 3'	65

through Non-Homologous End Joining (NHEJ), an error prone pathway, Indels get introduced causing mutation of the gene (Bortesi and Fischer 2015).

Modest quantity of studies been carried at to target several genes using CRISPR/Cas9 technique in soybean over the recent years. Efficient genome targeting necessitates the use of a single binary vector carrying the sgRNA and Cas9 cassettes with their suitable delivery into the plant system and, in order to fast-pace the experiment *Agrobacterium* mediated transient expression assay of the constructs was carried out. Here, we illustrated about editing of the genome using CRISPR/Cas9 which could lead to anticipated mutations of the gene owing to the observed results of nucleotide substitution and insertion.

Conclusion

The *in-silico* analysis of the constructs depicted efficiency of the target sequences pertaining to their intrinsic parameters such as GC content, sequence composition, sgRNA folding which in turn affect their efficient binding



Fig 5 PCR product analysis of the positive amplicon for GmABCC5, showing successful cloning and transient expression of soybean leaf discs. Sequence analysis showed an insertion of a nucleotide (red letter) in one of the leaf disc sample.

to the Cas9 endonuclease thus, affecting the editing machinery. With the targeted edits shown by the transient expression of CRISPR/Cas9 construct for *GmABCC5* gene, the above-mentioned method could cater as an effective precheck for their validation before moving forward for stable transformation in crops like soybean. The gene constructs showing appreciable results could further be carried forward for stable transformation for generation of low phytate soybean mutants and their downstream functional characterisation.

REFERENCES

- Bortesi L and Fischer R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* **33**(1): 41-52.
- Chu V T, Graf R, Wirtz T, Weber T, Favret J, Li X and Kühn R. 2016. Efficient CRISPR-mediated mutagenesis in primary immune cells using CrispRGold and a C57BL/6 Cas9 transgenic mouse line. *Proceedings of the National Academy of Science* **113**(44): 12514-12519.
- Doench J G, Hartenian E, Graham D B, Tothova Z, Hegde M, Smith I and Root D E. 2014. Rational design of highly active sgRNAs for CRISPR/Cas9-mediated gene inactivation. *Nature Biotechnology* **32**(12): 1262.
- Gruber A R, Lorenz R, Bernhart S H, Neuböck R and Hofacker I L. 2008. The vienna RNA websuite. *Nucleic Acids Research* **36**(suppl_2): W70-W74.
- Hada A, Krishnan V, Jaabir M M, Kumari A, Jolly M, Praveen S and Sachdev A. 2018. Improved *Agrobacterium tumefaciens*-mediated transformation of soybean [*Glycine max* (L.) Merr.] following optimization of culture conditions and mechanical techniques. *In Vitro Cellular & Developmental Biology-Plant* **54**(6): 672-688.
- Guidarelli M and Baraldi E. 2015. Transient transformation meets gene function discovery: the strawberry fruit case. *Frontiers in Plant Science* **6**: 444.
- Krishnan V, Jose J, Jolly M, Vinutha T, Kumar R, Manickavasagam M and Sachdev A. 2019. 'AGRODATE': a rapid *Agrobacterium*-mediated transient expression tool for gene function analysis in leaf discs. *Journal of Plant Biochemistry and Biotechnology* 1-11.
- Kumar A, Kumar V, Krishnan V, Hada A, Marathe A, Parameswaran C and Sachdev A. 2019. Seed targeted RNAi-mediated silencing of *GmMIPS1* limits phytate accumulation and improves mineral bioavailability in soybean. *Scientific Reports* **9**(1): 7744.
- Kumari S, Jolly M, Krishnan V, Dahuja A, and Sachdev A. 2012. Spatial and temporal expression analysis of D-myo-inositol 3-phosphate synthase (MIPS) gene family in *Glycine max*. *African Journal of Biotechnology* **11**(98): 16443-16454.
- Marathe A, Krishnan V, Vinutha T, Dahuja A, Jolly M and Sachdev A. 2018. Exploring the role of Inositol 1, 3, 4-trisphosphate 5/6 kinase-2 (*GmITPK2*) as a dehydration and salinity stress regulator in *Glycine max* (L.) Merr. through heterologous expression in *E. coli*. *Plant Physiology and Biochemistry* **123**: 331-341.
- Michno J M, Wang X, Liu J, Curtin S J, Kono T J, 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 and Food* **6**(4): 243-252.
- Moreno-Mateos M A, Vejnar C E, Beaudoin J D, Fernandez J P, Mis E K, Khokha M K and Giraldez A J. 2015. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nature Methods* **12**(10): 982.
- Nekrasov V, Staskawicz B, Weigel D, Jones J D and Kamoun S. 2013. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature Biotechnology* **31**(8): 691.
- Pandey V, Krishnan V, Basak N, Marathe A, Thimmegowda V, Dahuja A and Sachdev A. 2018. Molecular modeling and in silico characterization of GmABCC5: a phytate transporter and potential target for low-phytate crops. *3 Biotech* **8**(1): 54.
- Punjabi M, Bharadvaja N, Jolly M, Dahuja A and Sachdev A. 2018. Development and evaluation of low phytic acid soybean by siRNA triggered seed specific silencing of inositol polyphosphate 6-/3-/5-kinase gene. *Frontiers in Plant Science*: 9.
- Raboy V, Dickinson D B and Below F E. 1984. Variation in seed total phosphorus, phytic acid, zinc, calcium, magnesium, and protein among lines of *Glycine max* and *G. soja*. *Crop Science* **24**(3): 431-434.
- Raboy V, Gerbasi P F, Young K A, Stoneberg S D, Pickett S G, Bauman A T and Ertl D S. 2000. Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiology* **124**(1): 355-368.
- Ren X, Yang Z, Xu J, Sun J, Mao D, Hu Y and Deng P. 2014. Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in *Drosophila*. *Cell Reports* **9**(3): 1151-1162.
- Shi J, Wang H, Schellin K, Li B, Faller M, Stoop J M and Glassman K. 2007. Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. *Nature Biotechnology* **25**(8): 930.
- Sparkes I A, Runions J, Kearns A and Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* **1**(4): 2019.