BIO Web of Conferences **127**, 01008 (2024) https://doi.org/10.1051/bioconf/202412701008
ICW Biotech 2024

Strategy to select the CRISPR/Cas9 target sequences of the *sucrose-phosphate synthase* gene for assembling an efficient gRNA spacer in shallot

Tri Joko Santoso[1*], *Atmitri* Sisharmini[2], *Aniversari* Apriana[2], *Kristianto* Nugroho[1], *and Alberta Dinar* Ambarwati[1]

[1]Research Center for Horticulture, Research Organization for Agriculture and Food, National Research and Innovation Agency. Jl. Raya Jakarta-Bogor Km. 46, Cibinong, Bogor 16915, West Java, Indonesia
[2]Research Center for Food Crops, Research Organization for Agriculture and Food, National Research and Innovation Agency. Jl. Raya Jakarta-Bogor Km. 46, Cibinong, Bogor 16911, West Java, Indonesia.

Abstract. Bacterial leaf blight (BLB), caused by *Xanthomonas axonopodis pv. Allii*, is a newly identified pathogen that infects shallots. This disease can lead to yield losses ranging from 20% to 100%. The genome editing system via CRISPR/Cas-9 is a technology that can be used to modify susceptibility genes accurately and precisely. The target gene for CRISPR/Cas-9 genome editing system to develop a BLB-resistant shallot is the *sucrose-phosphate synthase* (SPS). One of the critical factors in the CRSPR/Cas-9 genome editing includes the preparation of single guide RNA (sgRNA) design and construction of it into an expression plasmid vector. The study aimed to develop strategies for selecting the SPS gene's CRISPR/Cas9 target sequences to assemble an efficient gRNA spacer in shallot. The gRNA design of the SPS gene was carried out using the software at http://crispor.tefor.net/. The efficiency of sgRNAs is then filtered and predicted based on GC contents, the last four nucleotides, secondary RNA structure, and stem loops. The construction of the CRISPR/Cas9 module was carried out using the Golden Gate method. The results showed that based on the selection of target gRNA sequences in the SPS gene, one of the best gRNAs, gRNA-54/forw, has been produced for use in SPS gene editing research. The best-selected gRNA of the SPS gene was then successfully inserted into the CRISPR/Cas9 pRGEB32 cassette vector after verification using DNA sequencing analysis. Based on this result, the pRGEB32_gRNA-AcSPS construct is ready to be introduced into shallot to develop a BLB-resistant shallot variety.

[*] Corresponding author: trij002@brin.go.id

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

1 Introduction

Shallot (*Allium cepa* L. *aggregatum*) is an essential commodity in Indonesia. It is widely used as a spice to enhance food flavour and holds strategic economic value nationally. Shallots provide income for farmers, create job opportunities, and contribute to regional economic development. In addition to being a cooking spice, shallots are also used in the food and pharmaceutical industries. They possess medicinal properties that make them valuable for treating various diseases due to their antibacterial, antifungal, antimicrobial, and antiviral properties [1, 2].

Biotic stress is a major challenge in shallot cultivation. One of the biotic stress factors in shallot planting is bacterial leaf blight (HDB) caused by *Xanthomonas axonopodis* pv. *allii*. This disease is a new pathogen that attacks and infects shallots. It is highly dangerous and can result in yield losses ranging from 20% to 100% [3]. The disease has spread to several countries in the United States, Asia, Latin America, Africa, and Europe [4]. Efforts to control this disease have not yet been successful. The use of resistant shallot varieties offers an opportunity to minimize yield losses, and this technology is environmentally friendly.

Genome editing is a cutting-edge technology in genomics that can improve plant characteristics. It allows for the development of new superior plant varieties without altering their original superior traits, except for the desired target characteristics. This technology offers more efficient and precise plant improvement, as gene modification can be carried out accurately and in a targeted manner. This has the potential to speed up plant breeding [5]. The CRISPR/Cas9 genome editing technique is the most widely developed targeted genome modification system due to its efficiency, cost-effectiveness, ease of use, and simplicity [6, 7]. This system is the most powerful tool for modifying genes and genomes. It enables gene knockout, base replacement, multiplex gene editing, and regulation of gene transcription in plants.

The CRISPR/Cas9 system based on RNA-guided engineered nucleases (Cas9 protein) and a guide RNA complex (gRNA). The most widely used system is the type II CRISPR/Cas9 employing a Cas9 endonuclease from *Streptococcus pyogenes* [8]. The system allows the creation of double stranded breaks (DSBs), which can lead to gene mutations due to non-homologous end-joining (NHEJ) repair or gene replacement or correction because of homologous recombination-based repair (HR). In most cases, NHEJ causes random insertions or deletion (indels), which can result in frameshift mutations if they occur in the coding region of the gene, effectively creating a gene knockout.

The rapid rise of CRISPR as a technology for genome engineering has created a demand for algorithms and online tools that assist in designing targeted and effective guide RNA (gRNA). Successful use of the CRISPR-Cas9 system in research applications relies on high-quality genome sequences, gene annotations, and data on off-target metrics and effectiveness [9]. The CRISPR-Cas9 system targets nucleases to specific loci based on the guide RNA (gRNA) and the presence of the target DNA protospacer adjacent motif (PAM) sequence region [10].

A strategy to create shallot plants resistant to HDB involves using the CRISPR/Cas9 system to disrupt or modify susceptibility genes (S gene) in the host. One potential S gene to target for editing is the sucrose-phosphate synthase (SPS) gene. Sucrose phosphate synthase (SPS) is a critical enzyme that controls plant sucrose biosynthesis. This biosynthesis will produce sugar product, which is required for the pathogen's nutritional needs. Plant pathogens use a set of endogenous effectors to divert host sugars form and use sucrose sugar for virulence. By modifying or suppressing the SPS gene, sugar products in cells can be reduced, preventing pathogen colonization and infection, meanwhile the truncated SPS protein or other family of SPS genes remain functionally active to support the plant growth and development..

The study aimed to develop strategies for selecting the SPS gene's CRISPR/Cas9 target sequences to assemble an efficient gRNA spacer in shallot.

2 Material and Method

2.1 Research material and plasmid

The coding sequences (CDS) of *Allium cepa*-SPS gene obtained from The NCBI GenBank was used as input for gRNA designing using the online software. The vector plasmid pRGEB32 (Addgene) was used to construct the CRISPR/Cas9-gRNA-SPS1 vector. This plasmid contains a binary vector for *Agrobacterium*-mediated rice transformation containing a maize *ubiquitin* 1 promoter driving the *Cas9* gene and a CaMV 35S promoter driving *hpt*II gene expression for hygromycin resistance.

The methodology to develop strategies for selecting the SPS gene's CRISPR/Cas9 target sequences to assemble an efficient gRNA spacer in shallot is illustrated in Figure 1.

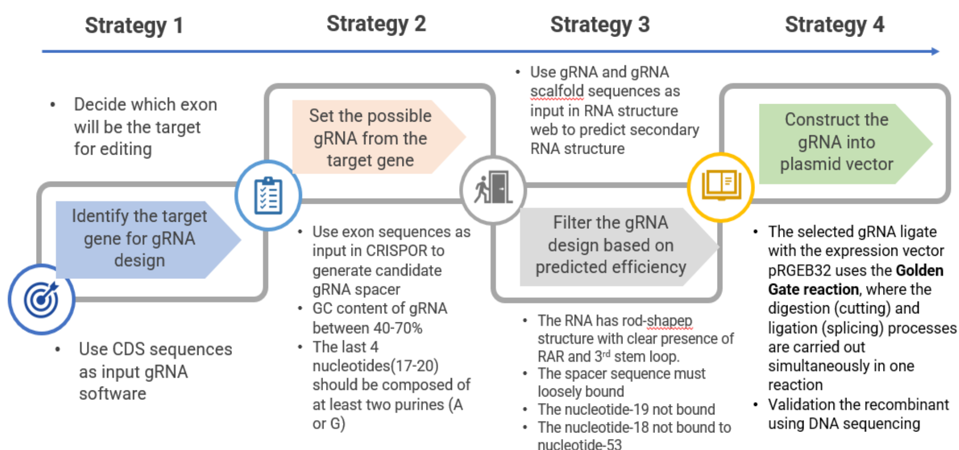


Fig. 1. Methodology has four strategies for designing an efficient gRNA spacer.

2.2 Identify the SPS target gene for gRNA design (strategy 1)

Identification of the SPS gene in shallots was conducted by searching the UniProt website (https://www.uniprot.org/) using the query "Allium cepa." After getting the SPS gene information from UniProt, we looked for the CDS sequence information on the NCBI GenBank website and downloaded the CDS-SPS gene sequence in FASTA format. I then used the CDS-SPS gene sequence as input to design the gRNA spacer.

2.3 Design and select the gRNA from the target gene (strategy 2)

The process of designing the gRNA primer starts with creating a guide RNA (gRNA) sequence based on the coding sequence or exons of the SPS gene. The gRNA sequence was designed using the online software CRISPOR (URL: http://crispor.tefor.net/). The set of gRNAs obtained from the online application was then analyzed for the GC content of the gRNA, which should fall within the range of 40-70%. Additionally, the last four nucleotides (17-20) of the gRNA should consist of at least two purines (A or G).

2.4 Filter the gRNA design based on predicted efficiency (strategy 3)

The best gRNA sequence is selected based on specificity, efficiency, and the number of off targets produced. For this purpose, each gRNA spacer (target sequence minus PAM sequence) was checked for gRNA secondary structure formation using the online tool RNAfold Webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold. cgi, [11, 12]. The selected gRNA sequence must meet the following three criteria, namely 1) nucleotides at positions 18 to 20 (seed region) of the unpaired spacer and can be accessed freely [16], 2) nucleotides at positions 51 to 53 from tracr-RNA unpaired with the seed region and freely accessible [17], and 3) the presence of all three stem-loop structures required to interact with Cas9 endonuclease can be identified.

The selected gRNA sequence is then used to design gRNA primers. The gRNA primer was designed by adding a BsaI enzyme restriction site to simplify inserting the gRNA into the expression vector. The expression vector used in this research is pRGEB32, which was obtained from the Addgene database (https://www.addgene.org/).

2.5 Construct gRNA into pRGEB32 using Golden Gate reaction (strategy 4)

The double-stranded gRNA ligation stage in the expression vector uses the Golden Gate reaction, where the digestion (cutting) and ligation (splicing) processes are carried out simultaneously in one reaction following the procedure previously used by [13]. The Golden Gate reaction was prepared in a total volume of 20 µl consisting of 1 µl pRGEB32 plasmid with a concentration of 50 ng/µl, 5 µl gRNA double-stranded oligonucleotide, 0.5 µL 1x BsaI buffer, 0.5 µl BsaI enzyme (2 U µl-1), 2 µl 1x T7 DNA ligase buffer, 1 µl T7 DNA ligase enzyme (2 U µl-1), and sterile ddH2O. The mixture was then incubated successively at 37 ºC for 5 minutes, 16 ºC for 10 minutes, 37 ºC for 5 minutes, 80 ºC for 5 minutes, and 25 ºC for 10 minutes using a thermocycler PCR machine.

The recombinant pRGEB32 plasmid carrying gRNA double-stranded oligonucleotides of gRNA-SPS was then transformed into E. coli DH5α competent cells. The transformation was carried out by reacting 5 µl of the Golden Gate reaction product with 50 µL of competent cells of E. coli strain DH5α. The mixture was then incubated on ice for 30 minutes, followed by heat shock at 42 ºC for 45 seconds and then incubated again on ice for 5 minutes. Next, 450 µl of liquid LB media was added to the bacterial suspension, followed by an incubation stage on an incubator shaker at 200 rpm at 37 oC for 1 hour. The bacterial suspension was then centrifuged at 6000 rpm for 2 minutes. A total of 350 µl of supernatant was then discarded, and the remainder was homogenized. Next, the bacterial suspension was spread in solid LB media containing 50 mg l-1 kanamycin and incubated at 37 ºC for 16 hours.

Recombinant plasmids were confirmed using DNA sequencing. A single colony of kanamycin-resistant *E. coli* bacteria was picked using a sterile tip, streaked thinly on replication media, then submerged in liquid LB media and cultured at 37°C and 200 rpm for 16 hours. The plasmid was then extracted from the colony using a plasmid isolation kit (Zymo Research, USA) and sent to the sequencing service lab.

3 Result and discussion

Target genes for genome editing with the CRISPR/Cas9 system in the NCBI GenBank database were identified, and coding sequence (CDS) information was obtained from the Allium cepa-SPS (AcSPS) gene. Important information regarding the AcSPS gene is as follows:

a. Source organism : Allium cepa
b. GenBank accession ID : EU164758.1

c. Molecule type : mRNA (CDS or exon)
d. Size of sequence : 3267 bp
e. Origin sequence

```
   1 ctgatctgca tagaacctgg attaaggttg ttgctactag aaatgcgaga gagaggagca
  61 gtcgtttgga gaatatgtgc tggagaattt ggcatcttgc tagaaaaaag aagcaggtcg
 121 agagtgaaga tgttcaaagg ttggcacatc gaaggtggga gcaagaacaa ggaaggagag
 181 aggcaacaga agacacgtca gaagatttat cagaaggaga gaaagagaag gctgatattg
 241 ttggtgaagt agatccaacc aagaggccgg taaaacaact ctctcaaacc ttctccgaaa
 301 ttccagcatg gcctgaggat aataaggaga agaaattgta tatagttctc ataagtttac
 361 atggactggt tcgtggagaa aacatggagc taggtcgtga ttcggacact ggtggtcaga
 421 ttaagtatgt agttgaactt gcacgtgctc tttcaatgat gcccggagtg tacagagtag
 481 acctcttcac tcgccaggtg tcgtctccag atgtagactg gagctacgga gagccaactg
 541 aaatgctgtc atcaggctcc tacgacacag aagggaatga tgcaggagaa agtgctggcg
 601 cttacatcgt gcgcatacca tgtgggcccc gtgataagta tctcagaaaa gagctactat
 661 ggccttacat tcaagaattt gtagatggag ctctagtgca tgtcttgaac atgtccaaag
 721 ttttaggaga acaagttgga ggtggtcagc cagtgtggcc ttatgtggta catggacact
 781 atgcggatct gggtgactgt gctgctcttc tttcaggtgc cttgaatgtt cctatggttt
 841 taactggtca ttctttagga agaaacaaac tagagcagct tttgaagcaa ggacgtcaat
 901 ctaaggaaga tataaattct acgtataaga taatgagacg tatagaagca gaagagttgt
 961 ctttagatgc ctcggagctt gttattacga gcactaggca ggagatcgat gagcagtggg
1021 gtttgtatga tgggtttgat gtgaagcttg agaaggtttt gagagctaga ataagaagag
1081 gggttagttc tcatggtcgg tacatgccta ggatggtggt tattcctcct ggtatggatt
1141 ttagtaacgt tgtagctcag gtagaagcaa gtgaaacgga tgatttggca tcaattttaa
1201 gtgctgatgg agctcctaag aagtcagtcc cacctatctg gtccgaggtg atgcgcttct
1261 ttacaaatcc tcacaaaccg atgattttgg ctttatcaag gccagatccc aaaaagaaca
1321 tcactactct cttaaaagca tttggagaat gtcgcccgtt aagggacctt gcaaatctta
1381 cattgattat gggaaataga gatgacatcg atgaaatggc atcagggagt gctggtgttc
1441 tgactacggt ccttaagttg atagataagt atgatcttta tggattggtg gctattccta
1501 aacatcacat ccagtctgat gtacccgaaa tttataggct tgctgcaaaa acaaagggtg
1561 tctttgttaa tcctgctttg gtggagcctt ttggcctcac tttgattgag gcggctgctc
1621 atggtcttcc aatggtggca acacataatg gcggtcctgt tgacattcat agggcactta
1681 acaatgggct acttgtagac ccacatgatc agaaagccat atcagatgca cttctaaagc
1741 ttgtagcaga caaaaaccta tggctagagt gcaagaaaaa cggtttaaaa aacatccatc
1801 tcttctcatg gcctgaacat tgtcgcacgt acctcttaag agtggcagca tgtcgaatga
1861 ggcatcctca atggcaaaac gatacttcat ccgatgatct tgcaaccgaa ggatctttag
1921 gagattctct aaaagatgtt catgaatcgt ctcttagact atctctagat ggagagaagg
1981 aatctatcaa cagctctgtc aattatgatg catctgaact caatcaagtg gcagaaggcg
2041 attcacaagt acaagaccaa gtcaagaaaa tccttgacag actcaagaag cagccatctg
2101 aagcagctaa taacactaaa aagtcggaaa attctagtaa tgttgtgcct aaatatcctc
2161 ttcttaggag gcgtagaagg ctttttgtca tagcacttga cagctacaac gacaagggcg
2221 aacctgaaaa aaagatgctt gatgtgataa aggaagtgtt taaggcgatc aaatcggaca
2281 ctcagatgtc gaggatttca gggtttgcgc tatcgactgc tatgcctgtt tcggagacca
2341 tagcactttt gaaattgggg aagattcagc ctgcagattt tgatgctttg atatgcagca
2401 gtggaggcga ggtgtactac ccaggcactt cccagtgcat ggatgagaat ggaaaattac
2461 gtgcagaccc ggattatgcg acgcacattg aatataggtg gggtataat ggggttaaaa
2521 ggacattgat gaagctgatg aactccgaag atggtcatgg agacaggaag tctttgaagg
2581 aggacgcatc tgctagcaac gctcactgtg tctcttttct cattaaagac ccatcaaagg
2641 caagaccagt agatgatctg cgccaaaagc taagaatgcg tgggatcagg tgccatctta
2701 tgtactgtag aaattctacg aggttgcaag ttatccctct tttggcttct cgagcacaag
2761 cactaaggta tctatttgtg cggtggggaa taagtgtagc gaacatatat gtgatactgg
2821 gagaaaaggg agatacagac catgaagaac tgatagctgg gacgcacaaa actctgatca
2881 tgagaggaat agtggagcgg ggatctgaag agttattgag aacagctgga agctatcaga
2941 gagaagacat tgttcctagt gacagtcctc ttgtagttta tacagaaggt tttgttaagt
3001 ctgaagagat catgaaggct ttcaaagaag tatcaaaagc tgcctctgca ctgtgaattc
3061 tgtgaataat aacagtacct tgtttatgca cttgtatacg attgagtcag atcgaccgcg
3121 attgcttatg caagatgatt aacgttgata tatgaaagag tctcacgaat cacgatcagt
3181 gacatgatga gatgtataat atattttctt tgatgtaatt atatatacag acactgtatt
3241 caaaagtata aaaaaaaaaa aaaaaaa
```

To design a gRNA from the CDS-SPS sequence (strategy 1), we used an online software application for the gRNA design tool. The CDS sequence was only around 2025 bp in size because the CRISPOR application requires an input sequence size of no more than 2300 bases to obtain a set of gRNA analysis results. Additionally, knocking out a gene using the CRISPR system will be more effective on base sequences from -50 to +300 bp from the transcription start site [14, 9].

The CRISPR tool was used to design guide RNA (gRNA) for a 2025 bp CDS sequence from the SPS gene, resulting in 90 gRNA sequences with corresponding PAMs in the form of NGG sequences (Table 1). Of these 90 gRNA sequences, 65 were in the forward position and 25 were in the reverse position.

In order to design gRNAs for CRISPR knockout, the process starts with identifying all possible gRNA designs based on a suitable adjacent PAM sequence. Next, the gRNAs are evaluated and ranked based on various criteria, and then filtered to emphasize or avoid specific features. The initial top ten gRNAs were filtered based on the criteria corresponding to strategy 2, which requires the GC content of the gRNA to be between 40-70% (Table 2). Additionally, the last 4 nucleotides (17-20) of the gRNA must contain at least two purines (A or G). Moreover, in order to ensure that the mutation effectively deactivates the protein encoded by the gene, target sequences above 50 base pairs from the start codon and below 65% of the length of the gene coding sequence are selected.

Table 1. 90 gRNA sequences with corresponding PAMs resulted from CRISPOR tool

No.	#guideId	targetSeq	No.	#guideId	targetSeq
1	16rev	AGTAGCAACAACCTTAATCCAGG	46	519forw	TGTCGTCTCCAGATGTAGACTGG
2	25forw	TCTGCATAGAACCTGGATTAAGG	47	527forw	CCAGATGTAGACTGGAGCTACGG
3	54forw	CTACTAGAAATGCGAGAGAGAGG	48	534rev	CCTGATGACAGCATTTCAGTTGG
4	67forw	GAGAGAGAGGAGCAGTCGTTTGG	49	554forw	CCAACTGAAATGCTGTCATCAGG
5	81forw	GTCGTTTGGAGAATATGTGCTGG	50	559rev	ATCATTCCCTTCTGTGTCGTAGG
6	90forw	AGAATATGTGCTGGAGAATTTGG	51	572forw	TCAGGCTCCTACGACACAGAAGG
7	115forw	TCTTGCTAGAAAAAAGAAGCAGG	52	573forw	CAGGCTCCTACGACACAGAAGGG
8	138forw	TCGAGAGTGAAGATGTTCAAAGG	53	584forw	GACACAGAAGGGAATGATGCAGG
9	142forw	GAGTGAAGATGTTCAAAGGTTGG	54	596forw	AATGATGCAGGAGAAAGTGCTGG
10	153forw	TTCAAAGGTTGGCACATCGAAGG	55	618rev	TACTTATCACGGGGCCCACATGG
11	156forw	AAAGGTTGGCACATCGAAGGTGG	56	623forw	TACATCGTGCGCATACCATGTGG
12	157forw	AAGGTTGGCACATCGAAGGTGGG	57	624forw	ACATCGTGCGCATACCATGTGGG
13	170forw	CGAAGGTGGGAGCAAGAACAAGG	58	627rev	TTTCTGAGATACTTATCACGGGG
14	174forw	GGTGGGAGCAAGAACAAGGAAGG	59	628rev	TTTTCTGAGATACTTATCACGGG
15	181forw	GCAAGAACAAGGAAGGAGAGAGG	60	629rev	CTTTTCTGAGATACTTATCACGG
16	215forw	ACGTCAGAAGATTTATCAGAAGG	61	660forw	ATCTCAGAAAAGAGCTACTATGG
17	229forw	ATCAGAAGGAGAGAAAGAGAAGG	62	663rev	TCTACAAATTCTTGAATGTAAGG
18	242forw	AAAGAGAAGGCTGATATTGTTGG	63	686forw	TACATTCAAGAATTTGTAGATGG
19	255rev	TGTTTTACCGGCCTCTTGGTTGG	64	715rev	AACTTGTTCTCCTAAAACTTTGG
20	259rev	GAGTTGTTTTACCGGCCTCTTGG	65	725forw	TTGAACATGTCCAAAGTTTTAGG
21	264forw	GTGAAGTAGATCCAACCAAGAGG	66	737forw	AAAGTTTTAGGAGAACAAGTTGG
22	267rev	GTTTGAGAGAGTTGTTTTACCGG	67	740forw	GTTTTAGGAGAACAAGTTGGAGG
23	268forw	AGTAGATCCAACCAAGAGGCCGG	68	743forw	TTAGGAGAACAAGTTGGAGGTGG
24	289rev	CCATGCTGGAATTTCGGAGAAGG	69	750rev	TGTACCACATAAGGCCACACTGG
25	295rev	CTCAGGCCATGCTGGAATTTCGG	70	756forw	TTGGAGGTGGTCAGCCAGTGTGG
26	303rev	TTATTATCCTCAGGCCATGCTGG	71	759rev	TAGTGTCCATGTACCACATAAGG
27	309forw	CCTTCTCCGAAATTCCAGCATGG	72	766forw	TCAGCCAGTGTGGCCTTATGTGG
28	312rev	TTCTTCTCCTTATTATCCTCAGG	73	773forw	GTGTGGCCTTATGTGGTACATGG
29	316forw	CGAAATTCCAGCATGGCCTGAGG	74	784forw	TGTGGTACATGGACACTATGCGG
30	325forw	AGCATGGCCTGAGGATAATAAGG	75	790forw	ACATGGACACTATGCGGATCTGG
31	362forw	ATAGTTCTCATAAGTTTACATGG	76	791forw	CATGGACACTATGCGGATCTGGG
32	367forw	TCTCATAAGTTTACATGGACTGG	77	815forw	GACTGTGCTGCTCTTCTTTCAGG
33	374forw	AGTTTACATGGACTGGTTCGTGG	78	820rev	TAAAACCATAGGAACATTCAAGG
34	385forw	ACTGGTTCGTGGAGAAAACATGG	79	831rev	GAATGACCAGTTAAAACCATAGG
35	392forw	CGTGGAGAAAACATGGAGCTAGG	80	835forw	AGGTGCCTTGAATGTTCCTATGG
36	403forw	CATGGAGCTAGGTCGTGATTCGG	81	845forw	AATGTTCCTATGGTTTTAACTGG
37	410forw	CTAGGTCGTGATTCGGACACTGG	82	857forw	GTTTTAACTGGTCATTCTTTAGG
38	413forw	GGTCGTGATTCGGACACTGGTGG	83	890forw	CTAGAGCAGCTTTTGAAGCAAGG
39	462rev	AGGTCTACTCTGTACACTCCGGG	84	904forw	GAAGCAAGGACGTCAATCTAAGG
40	463rev	GAGGTCTACTCTGTACACTCCGG	85	970rev	GCTCGTAATAACAAGCTCCGAGG
41	464forw	CGTGCTCTTTCAATGATGCCCGG	86	973forw	AGAGTTGTCTTTAGATGCCTCGG
42	482rev	ACGACACCTGGCGAGTGAAGAGG	87	996forw	AGCTTGTTATTACGAGCACTAGG
43	494rev	CTACATCTGGAGACGACACCTGG	88	1000forw	TGTTATTACGAGCACTAGGCAGG
44	496forw	AGTAGACCTCTTCACTCGCCAGG	89	1017forw	GGCAGGAGATCGATGAGCAGTGG
45	507rev	CCGTAGCTCCAGTCTACATCTGG	90	1018forw	GCAGGAGATCGATGAGCAGTGGG

Table 2. The top ten gRNAs filtered by based on criteria corresponding to strategy 2

No.	gRNA	Sequence	PAM	% GC content	4 last nucleotide
1	25forw	TCTGCATAGAACCTGGATTA	AGG	40	A, A
2	54forw	CTACTAGAAATGCGAGAGAG	AGG	45	A, A, G, G
3	81forw	GTCGTTTGGAGAATATGTGC	TGG	45	G, G
4	138forw	TCGAGAGTGAAGATGTTCAA	AGG	40	A, A
5	142forw	GAGTGAAGATGTTCAAAGGT	TGG	40	A, G, G
6	153forw	TTCAAAGGTTGGCACATCGA	AGG	45	G, A
7	156forw	AAAGGTTGGCACATCGAAGG	TGG	50	A, A, G, G
8	157forw	AAGGTTGGCACATCGAAGGT	GGG	50	A, G, G
9	170forw	CGAAGGTGGGAGCAAGAACA	AGG	50	A, A, A
10	174forw	GGTGGGAGCAAGAACAAGGA	AGG	55	A, G, G, A

Based on the analysis of secondary structure formation, one of the best gRNAs out of the 10 analysed is the 54/for (Figure 2) in the SPS gene. This best-selected gRNA (54forw) meets the criteria outlined by [15, 16] (strategy 3). Specifically, the nucleotides at positions 18–20 (seed region) of the spacer are unpaired and freely accessible, and nucleotides 51–53 of the tracr-RNA are unpaired with the seed region and can be accessed freely. Additionally, the gRNA's secondary structure contains the three stem-loop structures required to interact with the Cas9 endonuclease. These selected gRNAs are suitable for use in SPS gene editing research. Furthermore, the GC content of this gRNA sequence falls within the 40-70% range, and it contains at least 2 purine bases (A or G) in the last 4 nucleotides, meeting the criteria for a good gRNA.

The oligonucleotide of gRNA-54forw, which will be inserted into the pRGEB32 vector, was illustrated in Figure 3. The gRNA-SPS oligonucleotides (F and R) were added with the BsaI restriction enzyme site and then annealed into double strands through an oligonucleotide annealing process. The ds-gRNA-SPS-54/forw fragment was subsequently combined with the pRGEB32 expression vector plasmid using the golden gate reaction (strategy 4) to generate a recombinant plasmid carrying the cas9 and gRNA-SPS-54/forw genes. The golden cloning approach has several advantages including simpler, more effective, and efficient cloning process. Previous studies, this approach has also be applied successfully to construct the CRISPR/Cas9 vector [13, 17]. The recombinant plasmid resulted in golden gate reaction was then introduced into *E. coli* competent cells, resulting in the production of several single bacterial colonies. Recombinant plasmid DNA was extracted from the bacterial colony, and four samples of the pRGEB32-gRNA-SPS recombinant plasmid were sent to PT. Genetic Science to analyse the DNA sequence for validating the success of gRNA-SPS insertion.

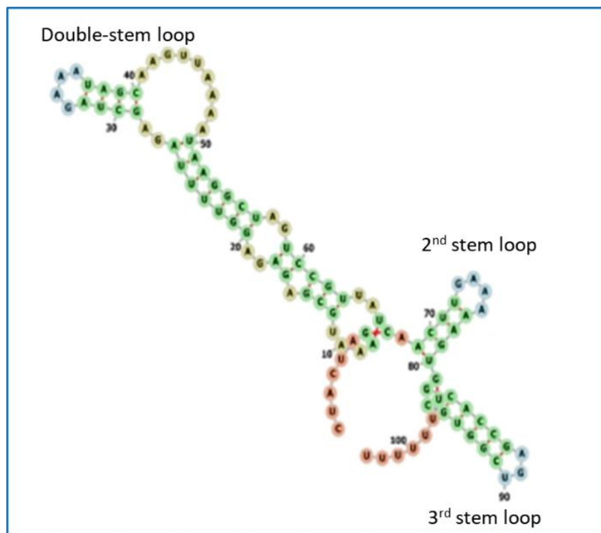
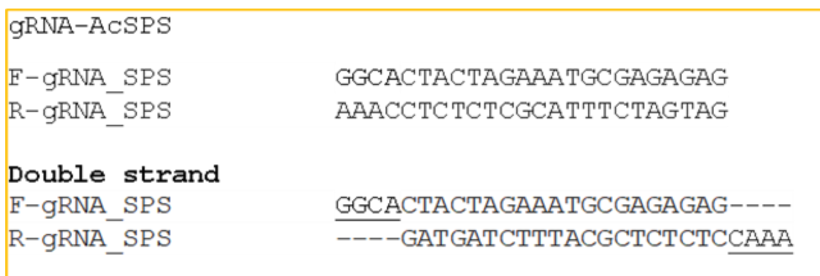


Fig. 2. Secondary RNA structure of gRNA 54forw



gRNA-AcSPS

F-gRNA_SPS	GGCACTACTAGAAATGCGAGAGAG
R-gRNA_SPS	AAACCTCTCTCGCATTTCTAGTAG

Double strand

F-gRNA_SPS	GGCACTACTAGAAATGCGAGAGAG----
R-gRNA_SPS	----GATGATCTTTACGCTCTCTCCAAA

Fig. 3. The illustration of the double-stranded oligonucleotide of gRNA-SPS

The DNA sequencing analysis of four CRISPR/Cas9 cassette recombinant plasmid samples revealed that only one sample, sample no. 2, carried the desired target gRNA-SPS. The gRNA sequence in this sample was highlighted in yellow (Figure 4). Plant SPS consists of three domains: glycosyltransferase domain, C-terminal domain, and N-terminal domain [18]. Based on gRNA design, the CRISPR/Cas9 cassette vector constructed in this study will be used for the introduction of frameshift insertion-deletions (InDels) adjacent to the N-terminal domain of the SPS gene, due to the gRNA-SPS is located at 54 bp from start codon site (part of N-terminal domain)This study successfully produced a red onion gRNA-SPS construct, which was inserted into the plasmid vector pRGEB32 containing the Cas9 gene and the hptII gene. The plasmid map of the resulting pRGEB32-gRNA-SPS construct can be seen in Figure 5. This plasmid construct obtained from the research will be transformed into the *Agrobacterium tumefaciens* vector for introduction it into the shallot genome through genetic transformation. In addition, these strategies developed in this research can be used to a new user to quickly set up an efficient gRNA spacer to be used for CRISPR/Cas-based editing in plant improvement purposes.

```
        1st_BASE_4949716_SPS-2

CGNCCTTAGGATCTTTAAACATACGAACAGATCACTTAAAGTTCTTCTGAAGCAACTTAAAGTTATCAGGCATGCATGGATCTTGGAGGAA
TCAGATGTGCAGTCAGGGACCATAGCACAAGACAGGCGTCTTCTACTGGTGCTACCAGCAAATGCTGGAAGCCGGGAACACTGGGTACGTT
GGAAACCACGTGATGTGAAGAAGTAAGATAAACTGTAGGAGAAAAGCATTTCGTAGTGGGCCATGAAGCCTTTCAGGACATGTATTGCAGT
ATGGGCCGGCCCATTACGCAATTGGACGACAACAAAGACTAGTATTAGTACCACCTCGGCTATCCACATAGATCAAAGCTGATTTAAAAGA
GTTGTGCAGATGATCCGTGGCACTACTAGAAATGCGAGAGAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA
CTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTTTTAGCGCGTGCAT
GCCTGCAGGTCCACAAATTCGGGTCAAGGCGGAAGCCAGCGCGCCACCCCACGTCAGCAAATACGGAGGCGCGGGGTTGACGGCGTCACCC
GGTCCTAACGGCGACCAACAAACCAGCCAGAAGAAATTACAGTAAAAAAAAAGTAAATTGCACTTTGATCCACCTTTTATTACCTAAGTCT
CAATTTGGATCACCCTTAAACCTATCTTTTCAATTTGGGCCGGGTTGTGGTTTGGACTACCATGAACAACTTTTCGTCATGTCTAACTTCC
CTTTCAGCAAACATATGAACCATATATAGAGGAGATCGGCCGTATACTAGAGCTGATGTGTTTAAGGTCGTTGATTGCACGAGAAAAAAAA
ATCCAAATCGCAACAATAGCAAATTTATCTGGTTCAAAGTGAAAAGATATGTTTAAAGGTAGTCCAAAGTAAAACTTATAGATAATAAAAT
GTGGTCCAAAGCGTAATTCACTCAAAAAANTCAACGAGACGTGTACCAAACGGAGACAAACGGCATCTTCTCGAAATTTCCCAACCGCTCG
CTCGCCGCCTCGTCTTCCCGGAAACCGCGGTGGTTTCACCTGGCGGATTCNCCAGCAGACGGAAACTTCCGGCACGGGAATCCTCCCCCAC
CCAACGCCATAATACCAGCCCCCCTTCCCCTCCTCCATANCTCCACCCNAGAAATTTNCCCCATNTCCCAGGGTCTCTGTTCGAATAAATC
CTCGGTCCCAAGGAACTGGTTTCTCC
```

Fig. 4. The recombinant plasmid confirmed carrying the gRNA-SPS target.

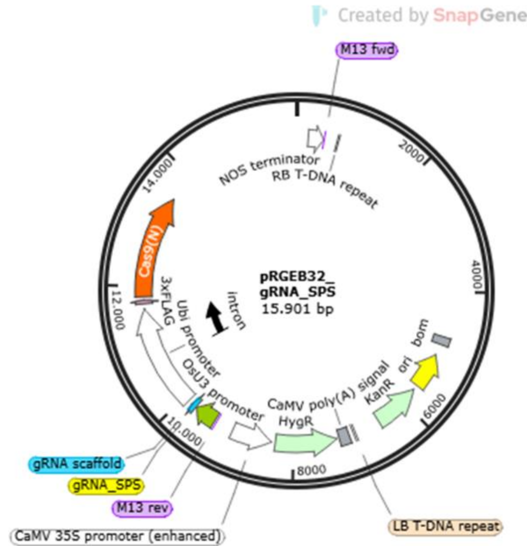


Fig. 5. Map of the plasmid vector CRISPR/Cas9-gRNA-SPS (pRGEB32-gRNA-SPS)

4 Conclusion

1. Based on the strategies applied, it has obtained the best gRNA spacer to be used for gene editing AcSPS gene in shallot, namely gRNA-54forw
2. The selected gRNA-SPS has successfully constructed into vector expression pRGEB32 and combined with Cas9 and hptII gene.

References

1. L. Mobin, M.A. Haq, R. Ali, S. Naz, S.G. Saeed, Antibacterial and antioxidant potential of the phenolic extract and its fractions isolated from Allium ascalonicum (onion) peel. Nat. Pro. Res. **0**, 1–5 (2021). https://doi.org/10.1080/14786419.2021.1948040
2. E.M. Lengbiye, C.M. Mbadiko, C.M. Falanga et al., Antiviral Activity, Phytochemistry and Toxicology of Some Medically Interesting Allium Species: A Mini Review. Int. J. pathog. **5,** 64-77 (2020). https://doi.org/10.9734/ijpr/2020/v5i430145
3. J.J Serfontein, Xanthomonas blight of onion in South Africa. Plant Dis. **85**, 442-442 (2001). https://doi.org/10.1094/PDIS.2001.85.4.442A
4. P. Roumagnac, O. Pruvost, F. Chiroleu, G. Hughes, Spatial and Temporal analyses of bacterial of onion caused by Xanthomonas axonopodis pv. Allii. Phytopathology, **94**, 138-146 (2004). https://doi.org/10.1094/PHYTO.2004.94.2.138
5. L. Bortesi, R. Fisher, The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol Adv. **33**, 41-52 (2015). https://doi.org/10.1016/j.biotechadv.2014.12.006
6. N.A. Abdallah, C.S. Prakash, A.G. McHughen, Genome editing for crop improvement: Challenges and opportunities. GM Crops Food. **6,** 183-205 (2015). https://doi.org/10.1080/21645698.2015.1129937
7. R. Mishra, K. Zhao, Genome editing technologies and their applications in crop. Plant Biotech. Rep. **12**, 57–68 (2018). https://doi.org/10.1007/s11816-018-0472-0
8. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E.A. Charpentier, Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. **337,** 816–821. (2012). https://doi.org/10.1126/science.1225829
9. S.E. Mohr, Y. Hu, B. Ewen-Campen, B.E. Housden, R. Viswanatha, N. Perrimon, CRISPR guide RNA design for research applications. FEBS J. **283,** 3232–3238 (2016). https://doi.org/10.1111/febs.13777
10. A.V. Wright, J.K. Nunez, J.A. Doudna, Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. Cell. **164**, 29–44 (2016). https://doi.org/10.1016/j.cell.2015.12.035
11. A.R. Gruber, R. Lorenz, S.H. Bernhart, R. Neubock, I.L. Hofacker, The Vienna RNA websuite. Nucleic Acids Res. **36**, W70–W74. (2008). https://doi.org/10.1093/nar/gkn188
12. R. Lorenz, S.H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, et al., Vienna RNA package 2.0. Algorith. Mol. Biol. **6,** 26 (2011). https://doi.org/10.1093/nar/gkn188
13. A.H. Mubarok, A. Sisharmini, A Apriana, T.J. Santoso, Suharsono. CRISPR/Cas9 Cassette Construction for *OsARF2* Gene Editing and Development of Transgenic Rice Nipponbare Containing CRISPR/Cas9-OsARF2. J. AgroBiogen. **18,** 45–56 (2022).
14. L.A. Gilbert, M.A. Horlbeck, B. Adamson, J.E. Villalta, Y. Chen, E.H. Whitehead, C. Guimaraes, B. Panning, H.L. Ploegh, M.C. Bassik, et al., Genome-scale CRISPR-mediated control of gene repression and activation. Cell. **159**, 647–66 (2014). https://doi.org/10.1016/j.cell.2014.09.029
15. D.W.S. Wong, The ABCs of Gene Cloning, 2nd edition. Nucleic Acids Res. **42,** Web Server issue W401–W407 (2014). https://doi.org/10.1093/nar/gku410
16. J.G. Doench, E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B.L. Ebert, R.J. Xavier, D.E. Root, Rational design of highly active sgRNAs for

CRISPR-Cas9-mediated gene inactivation. Nat. Biotechnol. **32**, 1262–1267 (2004). https://doi.org/10.1038/nbt.3026.

17. A.A.Md Yusof, A.A. Tamizi, N.A. Nurul Asyikin Mohd-Zim, S.S.A.Sattar, M.S. Salleh, N.S.A. Nur Azmi, Z. Zainal, Z. Zainuddin, N.N. Samsulrizal, Development of CRISPR/Cas9 Construct in Rice (*Oryza sativa* subsp. *indica*) Using Golden Gate Cloning Method Towards Drought Tolerance. JTLS. **13**, 257-276. (2022). http://dx.doi.org/10.11594/jtls.13.02.04
18. K.C. Teck, J.M. Bujnicki, T.C. Tan, F. Huynh, B.K. Patel, J. Sivaraman. 2008. The structure of sucrose phosphate synthase from Halothermothrix orenii reveals its mechanism of action and binding mode. Plant Cell. 20 (4):1059–1072. http://doi:10.1105/tpc.107.051193.