ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 18(5): 471-477 (2022)
©Biomedical Informatics (2022)

BIOINFORMATION
Discovery at the interface of physical and biological sciences

OPEN ACCESS GOLD

BIOMEDICAL INFORMATICS

www.bioinformation.net
Volume 18(5)

Research Article

Received April 14, 2022; Revised May 31, 2022; Accepted May 31, 2022, Published May 31, 2022

DOI: 10.6026/97320630018471

Declaration on Publication Ethics:
The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:
The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

License statement:
This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Comments from readers:
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Edited by P Kangueane
Citation: Jameel *et al.* Bioinformation 18(5): 471-477 (2022)

From design to validation of CRISPR/gRNA primers towards genome editing

Mohd Rizwan Jameel[1,2,3], Zubaida Ansari[2] & Mohammad Irfan Qureshi[1]

[1]Department of Biotechnology, Jamia Millia Islamia, New Delhi – 110025, India; [2]Centre for Interdisciplinary Research in Basic Science, Jamia Milia Islamia, New Delhi – 110025, India; [3]International Centre for Genetic Engineering and Biotechnology, New Delhi - 110067, India

Author contacts:
Mohd Rizwan Jameel - E-mail: mohdrizwanjameel@gmail.com (+91-9718723180)
Zubaida Ansari - E-mail: zaansari@jmi.ac.in
M. Irfan Qureshi - E-mail: miqureshi@jmi.ac.in

Abstract:
CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 (CRISPR-associated system) is used to edit specific genomic sequences with precision and efficacy. There are many online platforms/software for the design of gRNAs and related primers. However, there are concerns in design regarding off-site deletions besides knocking out sequences in the target genes. Nonetheless, a well known robust platform for CRISPR/gRNA primers design is CRISPRdirect. We demonstrate the use of this tool in the design of CRISPR/gRNA

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 18(5): 471-477 (2022) ©Biomedical Informatics (2022)

primers for soluble starch synthases (SSS) II-1, 2, and 3 genes in the *Oryza sativa* genome followed by the PCR-mediated amplification of SSS genes with corresponding confirmation towards genome editing having improved phenotype features.

Keywords: CRISPR, CRISPdirect, primers, knock out, and soluble starch synthase (SSS).

Background:

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 endonuclease is a genomic editing tool. CRISPR/Cas9 is known for its simplicity, specificity, and versatility [**1-5**]. The functions of endonuclease Cas9 could be utilized to delete a specific site(s) of a gene using an engineered sequence single guide RNA (gRNA). The gRNA usually has 20 nucleotides against a specific target sequence in the genome [**5**]. Mutations in soluble starch synthases (SSS) II [**6-9**] have altered the starch biosynthesis pathway in one way or the other in rice. Variation in SSS isoform profiles could affect the starch biosynthesis pathway and control phenotypes in the natural varieties of rice [**10**]. Thus, design of specific CRISPR/gRNA primers is a challenge.

There are several online platforms/software offering services to design gRNAs and related primers (**Table 1**). We used the CRISPRdirect (http://crispr.dbcls.jp/) [**11**] platform to select target sites of gRNA and NGG as PAM sites in SSS II (1, 2, and 3) genes with possible off-target sites. The CRISPRdirect examined the whole genome template of *Oryza sativa* at length width of 20-mer, 12-mer and 8-mer adjacent to the PAM (protospacer adjacent motif) of the flanking region [**11**]. Therefore, it is of interest to describe the design of CRISPR/gRNA primers [**11**] for soluble starch synthases (SSS) II-1, 2, and 3 genes in the *Oryza sativa* genome using CRISPRdirect followed by the PCR-mediated amplification of SSS genes with corresponding confirmation.

Table 1: CRISPR/Cas tools for designing gRNAs and predicting off-site targets.

S.N.	Tool Name	Description	Reference	Website
1.	CRISPR.mit	Tool to facilitate the design of gRNAs.	[12]	http://crisp r.mit. edu/
2.	sgRNA designer	Online tool for effective sgRNAs designing.	[13]	https://portals.broad institute.org/gpp/public/analysis-tools /sgrna-design
3.	E-CRISP	Web application to design gRNA sequences.	[14]	http://www.ecrisp.org/ECRISP
4.	CRISPRseek	Part of R programming package for designing gRNAs.	[15]	http://www.bioconductor.org/packages/release/bioc/html/CRISPR seek.html
5.	Cas-OFFinder	Algorithm for identifying potential off-target sites in a genome.	[16]	http://www.rgenome.net/cas-offinder/
6.	CHOPCHOP	Online tool for predicting off-target binding of sgRNAs.	[17]	https://chopc hop.rc.fas.harvard.edu/
7.	CRISPRscan	sgRNA-scoring algorithm that effectively captures the activity of CRISPR/Cas9 *in vivo*.	[18]	http://www.crisprscan .org/
8.	CRISPRdirect	Web server for selecting rational CRISPR/Cas targets based on input sequence.	[11]	http://crispr.dbcls.jp/
9	PROTOSPACER	Web interface for finding, evaluating and sharing Cas9 guide-RNA designs.	[19]	http://www.protospace r.com/
10.	sgRNA Scorer 1.0 and 2.0	*In vivo* library methodology to assess sgRNA activity.	[20]	https ://crispr.med.harvard.edu/sgRNA ScorerV1/
11.	CRISPR Multi-Targeter	Online tool to find sgRNA targets.	[21]	http://www.multicrisp r.net/
12.	Off Spotter	An algorithm to assist in designing optimal gRNAs.	[22]	https://cm.jefferson.edu/Off-Spotter/
13.	WU-CRISPR	Web tool for the genome-wide design of sgRNAs.	[23]	http://crisp r.wustl.edu
14.	Breaking-Cas	Web tool to facilitate the design of guide RNA for CRISPR/Cas technique.	[24]	http://bioinfogp.cnb.csic.es/tools/breakingcas
15.	CHOPCHOP v2	An updated version of CHOPCHOP which improves the targeting power, usability, and efficiency of CHOPCHOP by offering new options for sgRNA design.	[25]	http://chopchop.cbu.uib.no/
16.	CRISPOR	Web tool to find guide RNAs from an input sequence.	[26]	http://crispor.tefor.net/
17.	CCTop	Online, intuitive user interface for designing of guide RNAs.	[27]	https ://crisp r.cos.uni-heide lberg .de/index .html
18.	CRISPR-P 2.0	Web-services for computer-aided sgRNA designing with minimal off-target activity.	[28]	http://crispr.hzau.edu.cn/CRISPR2/
19.	GuideScan	Software for the designing of gRNA libraries for various genomic regions.	[29]	http://www.guide scan.com/
20.	CRISPR-GE	Convenient, integrated toolkit to expedite all experimental designs and analyses of mutation for CRISPR/Cas/Cpf1- based genome editing in plants and other organisms	[30]	http://skl.scau.edu.cn
21.	CRISPR-Local	High throughput tool for designing single -guide RNAs in plants and other organisms.	[31]	http://crispr.hzau.edu.cn/CRISPRLocal/
22.	CRISPR-PLANT v2	Tool to predict off-target sites found in unbiased genome-wide studies.	[32]	http://www.genome.arizona.edu/crispr2/

Material and Methods:

Workflow:

The workflow in the CRISPR/gRNA primer design for engineering target sites in genomes using the CRISPRdirect software is shown in **Figure 1**.

Sequence search using BLAST:

Various DNA and protein sequence databases (NCBI, PDB, TAIR, Swissport, Rice Genome Annotation Project at MSU, etc.) are used for sequence search using BLASTN and BLASTX programs [**33**]. The rate likenesses among DNA and protein sequences were acquired utilizing MacVector programming (Acceleris, GmbH, Germany) and Bioedit programming (variant 7.25).

Alignments of DNA sequences:

Alignments of the DNA sequences were completed using the Bioedit software (version 7.25) and ClustalW (version 2.0) program [**34**] at EMBL.

Primer design:

The primers were designed using Mac-Vector (Acceleris, GmbH), Snap Gene (form 1.1.3), or using electronic programming, for example, Primer blast (http://www.ncbi.nlm.nih.gov/

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 18(5): 471-477 (2022)
©Biomedical Informatics (2022)

instruments/ groundwork impact/list) and Primer 3 plus (https://primer3plus.com/) with default values. Oligo-analyzer programming (http://www.idtdna.com) was used to check for the probability of self or hetero-dimer development in the planned set of Primers. These Primers were confirmed for their uniqueness using BLASTN at the NCBI database. The primers were further combined using IDT, Germany, or Sigma-Aldrich, India.

gRNA synthesis:
The target site to be knocked out was picked in the exonic region involved in gene expression. Thus, to pick the target sites in the cDNA sequence of the gene of interest was used. Both the forward and turn around single-stranded oligos of 20 ntds synthesized at IDT (USA) were used.

Restriction map analysis:
Restriction maps of DNA fragments were generated using the Snap Gene software (version 1.1.3) to recognize the restriction sites in the required DNA sequence for either cloning into a range of vectors or for the construction of primers containing restriction sites.

Polymerase Chain Reaction:
The amplification of target DNA sequence was performed using standard protocol with all required components (PCR mixture) including Taq polymerase, and pair of primers at overhang regions at the gRNA target sites.

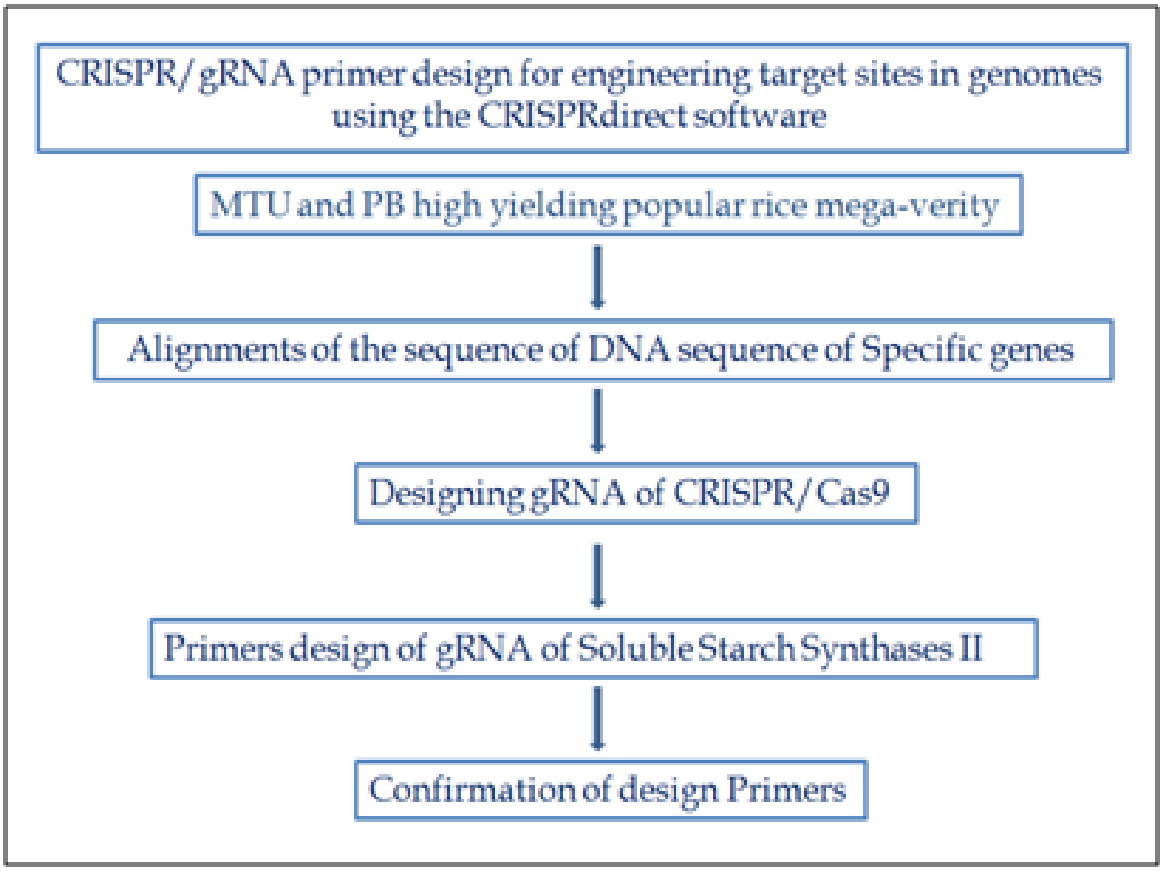

Figure 1: Flowchart of CRISPR/gRNA primer design for engineering target sites in genomes using the CRISPRdirect software.

Table 2: A nucleotide BLAST data for Soluble Starch synthase (SSS) among *Oryza sativa indica* members

Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Oryza sativa indica	1770	4188	100%	0.0	99.90%	10603	GQ150964.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	25582588	CP018166.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	20709915	CP012618.1
Oryza sativa indica	1764	4182	100%	0.0	99.79%	9833	KF984391.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	25797731	CP056061.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	25690566	CP054685.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	10603	GQ150963.1
Oryza sativa indica	1764	4177	100%	0.0	99.79%	10603	GQ150968.1

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 18(5): 471-477 (2022)
©Biomedical Informatics (2022)

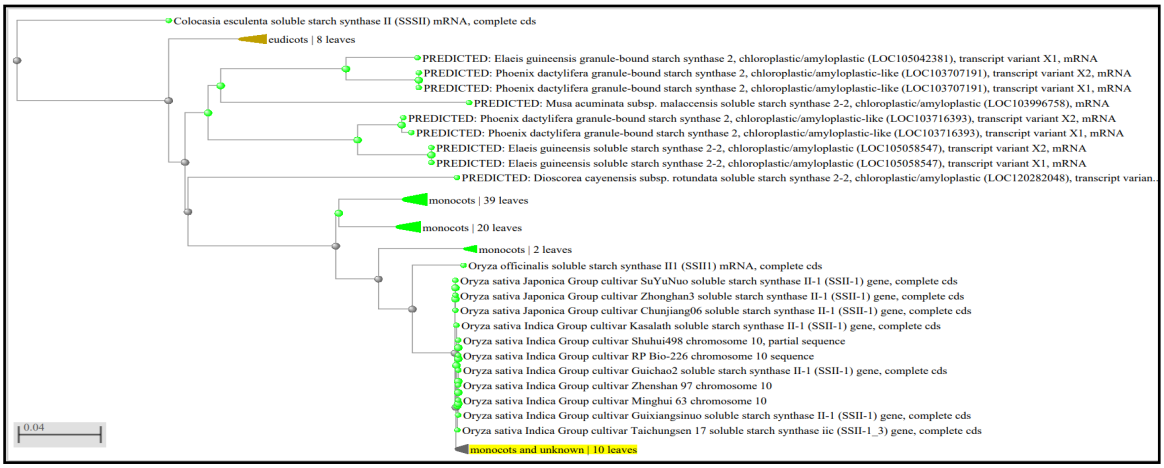

Figure 2: Phylogenetic tree for Soluble Starch synthase (SSS) among *Oryza sativa indica*

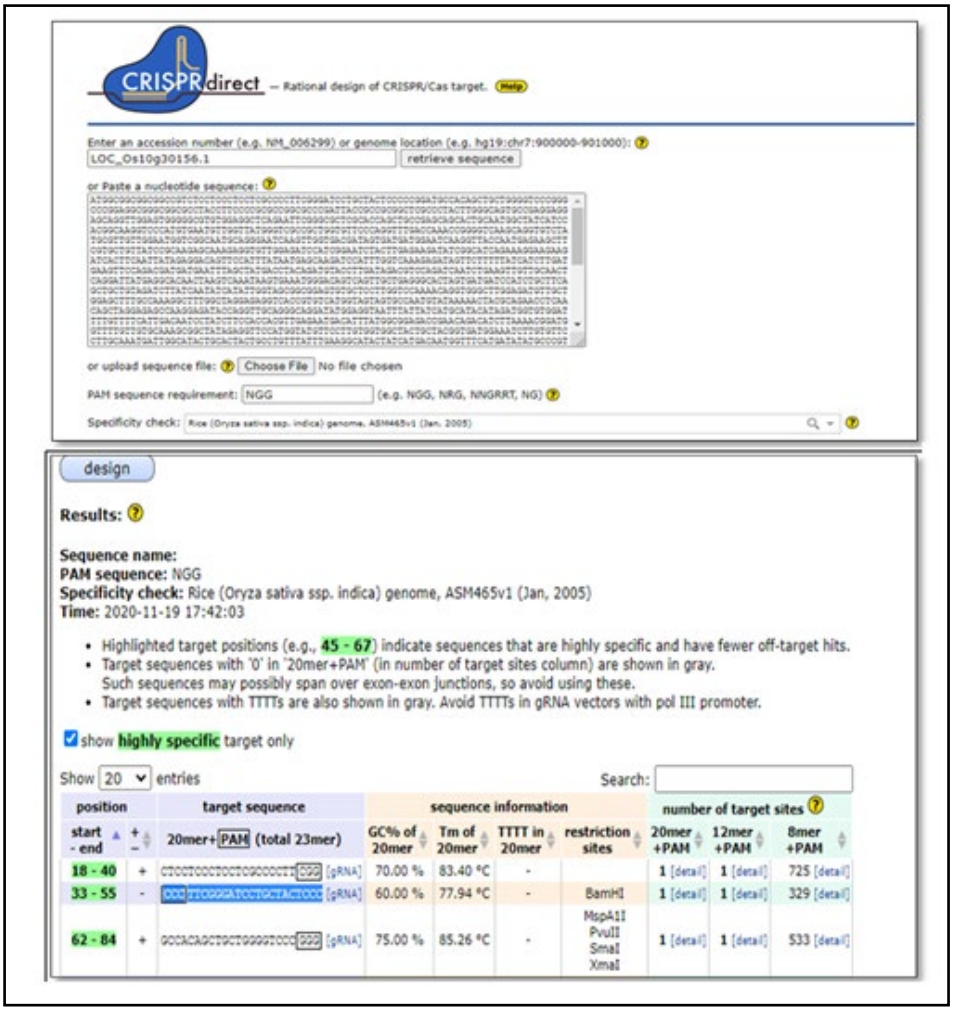

Figure 3: Data retrieval using the 'CRISPRdirect' software is shown. Sequence data of a gene with gene ID is entered. This is followed by the running of algorithm and display of results by CRISPR direct in the form of 20mer + PAM sequence. Percent (%) of GC content, melting temperature (Tm) of 20 mer, and available restriction sites are also shown.

Results:

Sequence search using BLAST:

A BLAST data for the target gene Soluble Starch Synthase (SSS) from *Oryza sativa indica* showed 8 closes matches with 100% query coverage, and 99.90% to 99.79% identity (**Table 2**). Distance tree for the data is shown in **Figure 2**.

Table 3: Description of Soluble Starch Synthase II (1, 2 and 3 isomers) data

S. No.	Gene (CDS)	Location (*Oryza sativa*)	Protein ID/Gene
1	Soluble Starch Synthase II-1	OsSSII-1 (LOC_Os10g30156)	Q7XE48/ Os10g0437600
2	Soluble Starch Synthase II-2	OsSSII-2 (LOC_Os02g51070)	Q6Z2T8/ Os02g0744700
3	Soluble Starch Synthase II-3	OsSSII-3 (LOC_Os06g12450)	Q0DDE3/ Os06g0229800

Primers of CRISPR/Cas9 gRNA:

CRISPR primers and gRNA were designed for SSS II 1, 2 and 3 (**Tables 3, 4, and 5**) using CRISPRdirect (**Figure 3**). Primers for other required genes such as hygromycin resistant gene (hygromycin phos-phor-transferase, *hpt*) and Cas9 were also designed (**Table 6**). Next-generation sequences (NGS) were used for generating targeted sequences. Blast search likeness of SSS II-1, 2 and 3 (**Table 3**) gene sequences were completed at several DNA databases (PDB, NCBI, TAIR, Swissport, Rice Genome Explanation at MSU). Alignment of the SSS II sequences were completed using the bio edit software (version 7.25) and ClustalW (version 2.0) program [**34**] of EMBL. The majority of the primers of CRISPR gRNA for target specific sequences of SSS II were done either using the Mac-Vector (Acceleris, GmbH), CRISPRdirect, Snap Gene (form 1.1.3), or electronic programming, for example, Primer blast (http://www.ncbi.nlm.nih.gov/instruments/groundwork impact/list) and Primer 3 plus (https://primer3plus.com/) with default limitations. Oligo-analyzer programming (http://www.idtdna.com) was used to check for the probability of self or heterodimer development in the designed set of Primers of CRISPR Cas9. These Primers were confirmed for their uniqueness using BLASTN at the NCBI database.

Table 4: Sequences for sgRNAs of SSS II -1, 2, and 3

Gene	gRNA	Orientation	Oligo Sequence (5`→3`)
SSS II-1	*gRNA1*	F	CAGGGGGAGTAGCAGGATCCCGAA
		R	AAACTTCGGGATCCTGCTACTCCC
	gRNA2	F	CAGGGCGTGTGGAGGCTCAGAATT
		R	AAACAATTCTGAGCCTCCACACGC
SSS II-2	*gRNA1*	F	CAGGGTTTGGTGGCACTTGTATGG
		R	AAACCCATACAAGTGCCACCAAAC
	gRNA2	F	CAGGTAGGTGTCCGCAAACGTTAC
		R	AAACGTAACGTTTGCGGACACCTA
		F	CAGGGTCACGGGGAGGGCCCAGGA
SSS II-3	*gRNA1*	R	AAACTCCTGGGCCCTCCCCGTGAC
		F	CAGGTAGATCTTCCCCTATTCCTG
	gRNA2	R	AAACCAGGAATAGGGGAAGATCTA

Table 5: Primers for sgRNAs targeted sequences of SSS II -1, 2, and 3.

Gene		Sequence (5`→3`)
OsSSSII-1		Forward primer - GACTCCACCACCTTCTCCT Reverse primer - ATAGACACCTGCTTGACCCC
OsSSSII-2	***Target 1***	Forward primer -GCGGCGGCGGATCCTATGTC
	Target 2	Reverse primer - GCACTAACTAATTGCACGTTCTGGC
		Forward primer –CCACAACGTAAGAGACTAAGAGGTGG Reverse primer – ATATCATTGTGCCGGTGTCGAAAGG
OsSSSII-3		
Target 1		Forward primer - CCATTTCATCCCCCGCACGGCAC Reverse primer - GCACGCGCAAGCACACGCGTC
Target 2		Forward primer - CGGAACGGATCGGTCGTTACC Reverse primer – AGCCAAAGGGCCCGAATCGTC

Table 6: Primers of hygromycin resistant gene (*hpt*) and Cas9 genes

S. No.	Gene	Primer sequence (5`--3`)
1	*Hpt*	**F**- GCCTGAACTCACCGCGACG **R**-CAGCCATCGGTCCAGACG
2	*Cas9*	**F**-TCGGCGAACTCCATGCTAT **R**-TCACCTTGTCGTCGAACAGG

Discussion:

The CRISPR/Cas9 application is used for gene editing. Gene editing in the intronic region will have no impact. Targeting exon regions is essential for effective approach to delete functional segment of gene in CRISPR/Cas9-mediated genome editing. The gRNA can be changed by digesting the scaffold vector with an appropriate restriction enzyme to clone sgRNA. cDNA sequences of three *OsSSSII-1, OsSSSII-2, and OsSSSII-3* genes were used in the CRISPRdirect program predict potential gRNAs. The CRISPRdirect identified few target sites which were promptly adjoining PAM sequences while checking the sense and antisense strands of both the cDNAs. The target sequence resembled 5′-N (20)- NGG-3′ or 5′-

CCN-N (20)- 3′. The target sequence which is 20 ntds, was available only near the PAM sequence which is 5'NGG3'. Both the forward and turn around single-stranded oligos of 20 ntds were obtained from IDT (USA). These data shows that CRISPRdirect [**11**] CRISPR-p [**35**] Cas-OF finder [**16**] and Cas-OT [**36**] are important tools for designing gRNAs and for predicting off targets. Design of primers of CRISPR/Cas9 gRNA using CRISPRdirect is simple and effective. The program generates a list of 20-22 bp spacers for targeted sequence for a gene with the search of PAMs (5`-NGG-3`) on both gene strands. Thus, we describe the enhanced amylose rice (data not shown) using CRISPR/Cas9 tool targeting isoforms 1, 2 and 3 of *SSSII*. Hence, designed gRNA respective targeted *SSSII-1, SSSII-2*, and *SSSII-3* in the Indica rice with transgenic-free homozygous *SSSII* mutant were generated. Therefore, food product made of flour with higher amylose content with resistant starch (RS) is possible from effective design of primers.

Conclusion:

We describe the use of the CRISPRdirect tool in the design of CRISPR/gRNA primers for soluble starch synthases (SSS) II-1, 2, and 3 genes in the *Oryza sativa* genome followed by the PCR-mediated amplification of SSS with corresponding confirmation towards genome editing. Thus, the use of the CRISPRdirect tool in the design of CRISPR/gRNA primers towards genome editing having improved phenotype features is illustrated.

Acknowledgments:

We would like to thank Dr. MK Reddy, International Centre for Genetic engineering and Biotechnology, New Delhi, Co-supervisor of M. Rizwan Jameel, for his generous help in this study.

References:

[1] Zhou H *et al. Nucleic Acids Res* 2014 **42:**10903 [PMID: 25200087]
[2] Liu H *et al. Bioinformatics* 2015 **31:**3676 [PMID: 26209430]
[3] Paul J & Qi Y. *Plant Cell Rep* 2016 **35:**1417 [PMID: 27114166]
[4] Sun Y *et al. Front Plant Sci* 2016 **7:**1928 [PMID: 28066481]
[5] Mehta S *et al. Stress Management* 2000 **427** [DOI: https://doi.org/10.1007/978-981-15-1322-0_23]
[6] Edward A *et al. Plant J* 1999 **17:**251 [PMID: 32202019]
[7] Shomura A *et al. Theor Appl Genet* 2002 **104:**1 [PMID: 12579422]
[8] Cmiel M *et al. Plant J.* 2003 **34:**173185 [PMID: 12694593]
[9] Myers A *et al. Plant Mol Biol* 2004 **54:**865 [PMID: 15604657]
[10] Crofts N *et al. J Exp Bot*. 2015 **66:**4469 [PMID: 25979995]
[11] Naito Y *et al. Bioinformatics* 2015 **31:**1120 [PMID: 25414360]
[12] Hsu PD *et al. Nature Biotechnol* 2013 **31:**827 [PMID: 23873081]
[13] Doench JG *et al. Nature Biotechnol* 2014 **32:**1262 [PMID: 25184501]
[14] Heigwer F *et al. Nature Methods* 2014 **11:**122 [PMID: 24481216]
[15] Zhao H *et al. Nucl Acids Res* 2014 **43:**D1018 [DOI: 10.1093/nar/gku894]
[16] Bae S *et al. Bioinformatics* 2014 **30:**1473 [PMID: 24463181]
[17] Montague TG *et al. Nucl Acids Res* 2014 **42**:W401 [PMID: 24861617]
[18] Moreno-Mateos MA *et al. Nature Methods* 2015 **12:**982 [PMID: 26322839]
[19] MacPherson CR *et al. Nature Biotechnol* 2015 **33:**805 [PMID: 26121414]
[20] Chari R *et al. Nature Methods* 2015 **12:**823 [PMID: 26167643]
[21] Prykhozhij SV *et al. PloS one* 2015 **10:**e0119372 [PMID: 25742428]
[22] Pliatsika V & Rigoutsos I. *Biology Direct* 2015 **10:**4[PMID: 25630343]
[23] Wong N *et al. Genome Biol* 2015 **16:**218 [PMID: 26521937]
[24] Oliveros JC *et al. Nucleic Acids Res* 2016 **44**:W267 [PMID: 27166368]
[25] Labun K *et al. Nucleic Acids Res* 2016 **44:**W272 [PMID: 27185894]
[26] Haeussler M *et al. Genome Biol* 2016 **17:**148 [PMID: 27380939]
[27] Stemmer M *et al. PloS one* 2015 **10:**e0124633. [PMID: 25909470]
[28] Liu H *et al. Molecular Plant* 2017 **10:**530 [PMID: 28089950]
[29] Perez AR *et al. Nature Biotechnol* 2017 **35:**347[PMID: 28263296]
[30] Xie X *et al. Mol Plant* 2017 **10:**1246[PMID: 28624544]
[31] Sun J *et al. Bioinformatics* 2018 **34:**1953 [PMID: 29365045]
[32] Minkenberg B *et al. Plant Biotechnol* J 2019 **17:**5 [PMID: 30325102]
[33] Altschul SF *et al. J Mol Biol*. 1990 **215:**403 [PMID: 2231712]
[34] Larkin MA *et al. Bioinformatics* 2007 **23:**2947 [PMID: 17846036]
[35] Lei Y *et al. Mol Plant* 2014 **7:**1494 [PMID: 24719468]
[36] Xiao *et al. Bioinformatics* 2014 **30:**1180 [PMID: 24389662]

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 18(5): 471-477 (2022)
©Biomedical Informatics (2022)