RESEARCH ARTICLE Indian Journal of Animal Research, Volume 58 Issue 1: 13-20 (January 2024)

CRISPR/Cas Genome Editing Single Guide RNA (sgRNA) Design using Three Different Web Tool Platforms

B. Rajith Reddy[1], B. Ekambaram[2], P. Jaya Laxmi[3], CH. Harikrishna[4], T.K. Bhattacharya[5], G. Sushma[6]

10.18805/IJAR.B-5200

ABSTRACT

Background: Genome editing is a group of technologies that has the ability to change an organism's DNA. These technologies allow genetic material to be added, removed, or altered at particular locations in the genome. Several approaches to genome editing have been developed. A well-known one is called CRISPR/Cas9, which is short for clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9. Editing through CRISPR focuses mainly on the Protospacer-adjacent motif (PAM), a crucial region to identify the sgRNAs to target the desired gene or region of DNA. So designing precise single-guide RNAs (sgRNAs) to minimize off-target effects is critical for the success of gene editing without undesired results.

Methods: Hence, in this article, three different online tools are used to design sgRNA to target the prolactin (PRL) gene to knock out. The *Gallus gallus* prolactin (PRL) gene sequence is retrieved from NCBI and used for further downstream application. However, all three web tools may vary in design specifications and parameter choices, visualization, downstream analysis functionality, *etc*.

Result: While keeping a straightforward and interactive interface and running with default parameters, all web tools also accept a variety of advanced choices for more specialized searches. This maximizes user flexibility. Three tools produced several sgRNAs that satisfied various criteria for precise gene editing to boost the efficacy of the target prolactin gene. These online tools use a robust approach to identify off-target locations and the results are displayed in an interactive table and within the gene architecture.

Key words: CRISPR/Cas9, Genome editing, Protospacer-adjacent motif (PAM), Single guide RNA (sgRNA).

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and their CRISPR-associated (Cas) proteins system is an effective immune system among bacteria and archaea. It is an acquired immunity mechanism first discovered in the *E. coli* genome (Ishino *et al*., 1987). However based on its mechanism CRISPR/Cas system is classified into three major groups (I, II and III) (Makarova *et al*., 2011). CRISPR/Cas technology relies on DNA-RNA interaction as well as the simple design of RNA molecules for each specific sequence. However, Other emerging techniques for targeted genome editing including zinc-finger nucleases (ZFNs) and transcription-activator-like effector nucleases (TALENs), allow researchers to introduce permanent alterations by activating repair mechanisms by generating double-stranded breaks (Chandrasegaran and Carroll, 2016). Using Zinc finger proteins by changing amino acids in the α-helix, DNA binding specificity and affinity could be altered. Engineered zinc fingers were combined with the DNA cleavage domain of FokI, a type II restriction endonuclease, to form ZFNs, which allow for specific targeted double-strand breaks in DNA (Carroll, 2011). However, ZFNs are costly to design and not feasible for all laboratories (Gupta and Musunuru, 2014). Transcription activator-like effector (TALE) proteins when coupled with the nuclease domain of FokI, TALE nucleases (TALENs) emerged as a novel genome-editing tool (Bogdanove and Voytas, 2011; Nemudryi *et al*., 2014).

[1]Molecular Genetics Lab, ICAR-Directorate of Poultry Research, Rajendranagar, Hyderabad-500 030, Telangana, India.
[2]P.V. Narsimha Rao Telangana Veterinary University, Rajendranagar, Hyderabad-500 030, Telangana, India.
[3]Department of Animal Genetics and Breeding, College of Veterinary Science, Mamnoor, Warangal, Hyderabad-500 030, Telangana, India.
[4]Department of Livestock Farm Complex, College of Veterinary Science, Rajendranagar, Hyderabad-500 030, Telangana, India.
[5]National Research Centre on Equines, Hissar-125 001, Haryana, India.
[6]Department of Animal Genetics and Breeding, College of Veterinary Science, Tirupathi-517 502, Andhra Pradesh, India.

Corresponding Author: B. Rajith Reddy, Molecular Genetics Lab, ICAR-Directorate of Poultry Research, Rajendranagar, Hyderabad-500 030, Telangana, India. Email: rajith.bheem@gmail.com

How to cite this article: Reddy, B.R., Ekambaram, B., Laxmi, P.J., Harikrishna, C., Bhattacharya, T.K. and Sushma, G. (2024). CRISPR/Cas Genome Editing Single Guide RNA (sgRNA) Design using Three Different Web Tool Platforms. Indian Journal of Animal Research. 58(1): 13-20. doi: 10.18805/IJAR.B-5200.

Submitted: 18-07-2023 **Accepted:** 25-09-2023 **Online:** 11-10-2023

Both the ZFN and TALEN approaches have challenges in creating and verifying proteins that can recognize particular DNA sequences. The CRISPR system, in contrast, is RNA-mediated which consists of a small non-coding transactivating CRISPR RNA (tracrRNA), an operon that

codes for the Cas proteins and a repeat array encompassing crRNA units made up of a 5′-20-nucleotide targeting sequence and a 19-22 nucleotide repeat sequence make up the three distinct architectural components of the native CRISPR/Cas9 system (Sander and Joung, 2014). According to numerous studies, the highly conserved 3′ three nucleotide protospacer adjacent motif (PAM) that comes before the target sequence is necessary for Cas9 endonuclease activity (Shah *et al*., 2013; Wang *et al*., 2015). Depending on the CRISPR type or subtype, PAM sequence diversity is extremely wide, with NGG serving as the most potent trinucleotide for the Streptococcus pyogenes CRISPR/Cas9 system (Jiang and Doudna, 2017).

In the genome editing method using CRISPR/Cas the Cas9/sgRNA complex can produce double-stranded DNA breaks (DSBs) in predetermined genomic loci by binding a 20 bp target sequence followed by a 3 bp protospacer adjacent motif (PAM) *i.e*., NGG. A chimeric gRNA which is complementary to the target area and trans-activating CRISPR-RNA are usually included in this sgRNA. Most Cas systems require the predesigned sgRNA to anneal immediately upstream of a PAM, which in the case of SpCas9 is 5′NGG3′ (Doudna and Charpentier, 2014). The DNA repair mechanisms of the target organism then fix these double-stranded breaks (DSBs), which have the potential to change the target gene. Despite the fact that the fundamental mechanisms underlying all these genome editing techniques are similar, CRISPR/Cas9 technology has emerged as the favored method due to its simplicity, affordability, exceptional adaptability and capacity to simultaneously target many genes (Wang *et al*., 2016). The two primary intrinsic DNA repair mechanisms for DSBs are nonhomologous end joining (NHEJ) and homology-directed repair (HDR). NHEJ is error-prone, leading to minute insertions or deletions (indels) that disrupt the target region. It is the main DNA repair process in species and the most well-known and straight forward pathway in genome editing (Cubbon *et al*., 2018; Xue and Greene, 2021). In light of this, the current paper concentrates on designing sgRNA to target the desired gene (Prolactin) using three distinct web-based tools, as well as on how simple and quick it is to choose the best CRISPR/Cas target sequences. The optimal sgRNA will result in a gene deletion, a specific base edit, or a modification of gene expression. With a low off-target score, comparable DNA sequences elsewhere in the genome may be targeted, which would decrease efficiency. A high on-target score will suggest strong binding efficiency to the target. This article also covers how rigorous off-target search techniques are used to estimate the specificity of the target site in the genome.

MATERIALS AND METHODS

Target gene sequence

The present experiment on designing sgRNA using different web tools was done at ICAR-Directorate of Poultry Research, Rajendranagar, Hyderabad in the year 2021. To create sgRNAs for CRISPR/Cas9 genome editing, the chicken Prolactin (PRL) coding area was used in the current study. The coding sequence has been retrieved from NCBI (NCBI Reference sequence: NM_205466.3). However, depending on the requirements of the web tools under examination, any one of the three forms of input gene name, genomic coordinates, or DNA sequence can be utilized. All the web tools with web addresses are provided in Table 1. To start this online web tools a computer with an internet facility is needed with minimum knowledge of data retrieval from the database. However, the basic steps for most of the web tools involved are as follows: I) Choose the gene to alter. II) Access the DNA sequence using the genome database III) Download the format of the file and save it. IV) Select the type of online web tool V) Select the endonuclease and knock out/knock in based on the requirement VI) Before creating the gRNA/sgRNA, upload the target gene/select/Copy in the web tool provided area. VII) After adjusting the advance choices as necessary, click the finish button to see the results. VIII) Different sgRNA with a specific PAM will finally appear. IX) The sgRNA should be picked in accordance with the score or preferences as indicated by the web tools.

RESULTS AND DISCUSSION

CHOPCHOP can be run with as few as three basic input options, or with additional advanced parameters. The basic input comprises a gene name, genomic coordinates, or a pasted sequence and type of organism. The advanced options allow the user to target a sub-region of the gene, such as the 5′UTR, 3′UTR, splice sites, full exons, or a specified subset of exons. The CHOPCHOP web tool interface is given in Fig 1A, 1B, which shows the website design and the outcome of the result. First, the Prolactin gene fasta sequence retrieved from NCBI is pasted in the given area and the organism is chosen by keeping all parameters as default. The majority of CHOPCHOP queries are executed within a matter of seconds and the results are displayed in an interactive table and interactive gene model. CHOPCHOP ranks the search results according to a number of off-targets, mismatches, perfect hits and GC content. Table 2 shows the top five predicted sgRNAs to target the prolactin gene. The CRISPR Design web server accepts an accession number, a genome coordinate, or an arbitrary nucleotide

Table 1: Different web tools used for designing SgRNA of prolactin gene of Gallus gallus.

Tool name	Web address	References
CHOP CHOP	https://chopchop.cbu.uib.no/	Montague *et al*. (2014)
CRISPR direct	http://crispr.dbcls.jp/	Naito *et al*. (2015)
CC TOP	https://cctop.cos.uni-heidelberg.de/	Stemmer *et al*. (2015)

sequence. Target sequences of 20 nucleotides (nt) adjacent to the PAM sequence are searched from both strands of the input sequence which is given in Fig 2A, 2B and 2C. The prolactin coding region is pasted as fasta sequence and the organism is chosen and all other parameters are placed as default. The output data is given as a table consisting of the start and end of the sequence along with GC content, melting temperature (Tm) and all other parameters. Table 3 shows 5 different sgRNAs generated using this tool. CRISPR/Cas9 target online predictor (CC Top) tool in which Prolactin coding region has given as input sequence all sgRNA target sites are identified according to adjustable parameters like the type of PAM, the identity of the two most 5′ nucleotides ('NN', 'GN' or 'GG') as well as the two most 3′ nucleotides ('NN' or 'GG'). All the parameters are kept as default except the PAM which is selected as NGG. The interface of this tool is given in Fig 3A, 3B and 3C. After processing, a results page is displayed containing the input parameters as a graphical representation of the query sequence with the identified sgRNA target sites as well as a full list of all candidates ranked by taking into account the number of total off-target sites, the distribution of mismatches and the proximity to exons. The top 5 sgRNAs are listed in Table 4. However, a fasta file containing all sgRNA target sites as well as a tab-separated file containing the full results can be downloaded.

The success of CRISPR/Cas editing depends on many practical steps, however, the primary prerequisite is to design a sgRNA and clone it into a suitable vector to deliver into the *in vitro* / *in vivo* system with decreased off-target effect in the genome and increase the specificity of the cleavage. So keeping this in view we analyzed the Chicken Prolactin (PRL) gene as an example to design the sgRNAs. All three tools are user-friendly and speedy in displaying the results

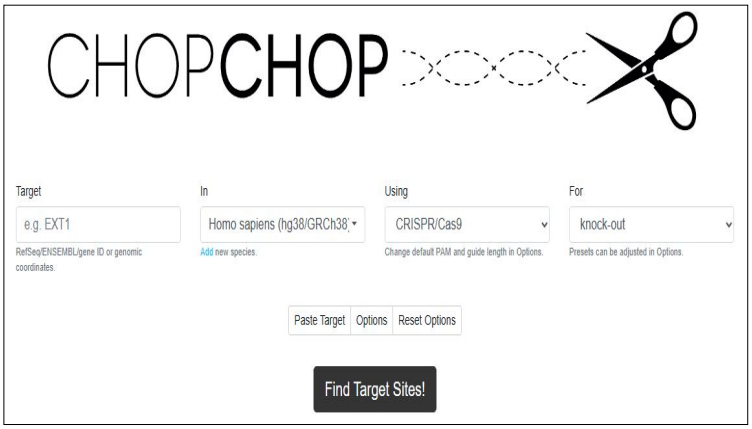


Fig 1A: The home page of CHOPCHOP allows users to enter a gene name, genomic coordinates or a DNA sequence and select an organism and CRISPR/Cas9 mode.



Rank	Target sequence	Genomic location	Strand	GC content (%)	Self-complementarity	MM0	MM1	MM2	MM3	Efficiency
1	GGCAGTTGACTGATCCAATGGGG	chr2:58656754	-	50	0	0	0	0	0	69.72
2	TGGCAGTTGACTGATCCAATGGG	chr2:58656755	-	45	1	0	0	0	0	52.54
3	GAACGTTATGCTCAGGGTCGGGG	chr2:58657272	+	55	0	0	0	0	0	43.52
4	AAGATAGTTGTCGATTTTGTGGG	chr2:58660950	-	30	0	0	0	0	0	37.84
5	TCATCAGCGAGTTGCAGGGATGG	chr2:58660874	-	55	0	0	0	0	1	55.09
6	GACCTACTGAATTTAGTAGTGGG	chr2:58658738	+	35	1	0	0	0	1	51.73
7	AACAGAGGGGCTTCATTGAAAGG	chr2:58655140	+	45	0	0	0	0	1	43.51
8	TAAGATAGTTGTCGATTTTGTGG	chr2:58660951	-	30	0	0	0	0	1	41.61
9	CAGGATGAACGTTATGCTCAGGG	chr2:58657266	+	45	0	0	0	1	0	59.08
10	AATTGCTATCATGGATTAGGCGG	chr2:58660987	-	35	2	0	0	0	2	65.73
11	CGATTTTGTGGGAATCTCTGCGG	chr2:58660939	-	45	0	0	0	0	2	62.82

Fig 1B: The main results page presents the sgRNA or TALEN target sites within the gene architecture (exon, blue; intron, red), with each option color-coded according to ranking.

Table 2: Top 5 predicted sgRNAs targeting prolactin (PRL) gene by CHOPCHOP tool.

Rank	Target sequence	Genomic location	Strand	GC content (%)	Self-complementarity	MM0	MM1	MM2	MM3	Efficiency
1	GGCAGTTGACTGATCCAATGGGG	chr2:58656754	-	50	0	0	0	0	0	69.72
2	TGGCAGTTGACTGATCCAATGGG	chr2:58656755	-	45	1	0	0	0	0	52.54
3	GAACGTTATGCTCAGGGTCGGGG	chr2:58657272	+	55	0	0	0	0	0	43.52
4	AAGATAGTTGTCGATTTTGTGGG	chr2:58660950	-	30	0	0	0	0	0	37.84
5	TCATCAGCGAGTTGCAGGGATGG	chr2:58660874	-	55	0	0	0	0	1	55.09

Table 3: List of 5 predicted sgRNAs targeting prolactin (PRL) gene by CRISPR direct tool.

Position		Target sequence	Sequence information				Number of target sites		
Start-end	+ -	20 mer + PAM (Total 23 mer)	GC% of 20 mer	Tm of 20 mer	TTTT in 20 mer	Restriction sites	20 mer + PAM	12 mer + PAM	8 mer + PAM
7-29	+	AACAGAGGGGCTTCATTGAAAGG [gRNA]	45.00%	71.38°C	-		1 [detail]	10 [detail]	2120 [detail]
18-40	+	TTCATTGAAAGGTTTGTTTCTGG [gRNA]	30.00%	60.43°C	-		0 [detail]	18 [detail]	5407 [detail]
21-43	+	ATTGAAAGGTTTGTTTCTGGCGG [gRNA]	35.00%	63.86°C	-		0 [detail]	36 [detail]	2702 [detail]
30-52	+	TTTGTTTCTGGCGGTTCTTCTGG [gRNA]	45.00%	69.29°C	-		1 [detail]	2 [detail]	1667 [detail]
54-76	+	GTCCAACACACTTCTGACCAAGG [gRNA]	50.00%	71.61°C	-		1 [detail]	13 [detail]	1856 [detail]

with a wide range of selections for the type of organism with off-target analysis. However, there are differences in output and based on our experiment advanced options can be modified. In all the tools one or other sgRNAs has shown GC content of up to 50% because the reports suggest that high GC content in sgRNAs is most effective and increases its efficiency (Ren *et al*., 2014: Jamal *et al*., 2016). The insilico tools study located 23 bp target sites including the PAM motif. Guanine at position 20 in the target site, which is associated with improved activity has been reported (Wang *et al*., 2014: Labuhn *et al*., 2018), all the tools exhibited Guanine (G) in the list of resulted output.

The user may restrict the search to only target sites suited for synthesis using a particular polymerase, *e.g.* GG-

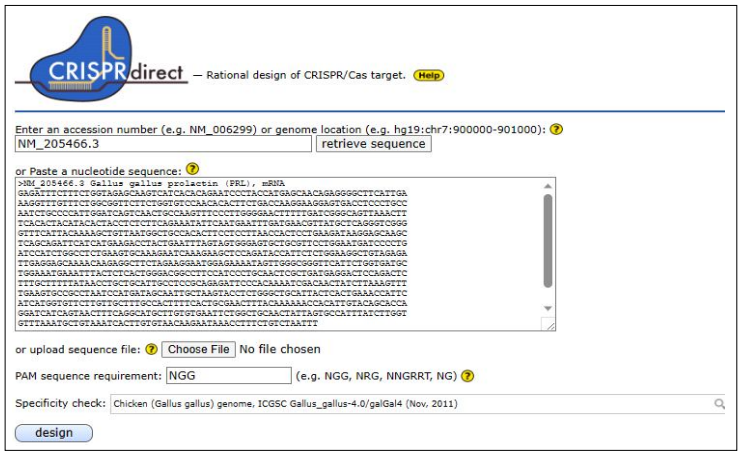


Fig 2A: Top page. The server accepts either an accession number or a nucleotide sequence as input.

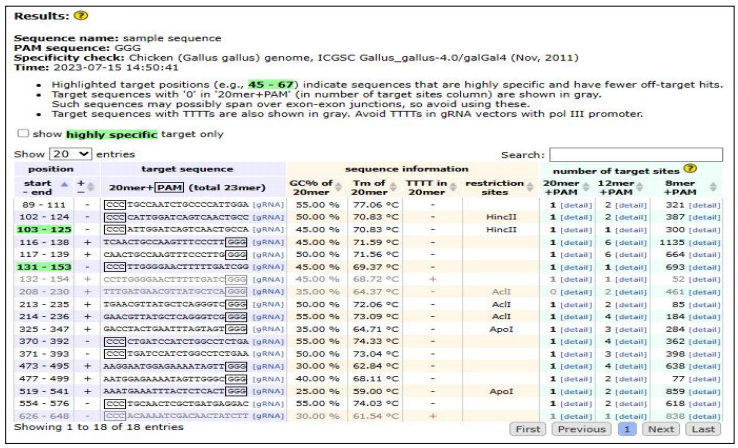


Fig 2B: Typical output of CRISPRdirect. A list of CRISPR/Cas target candidates is displayed.

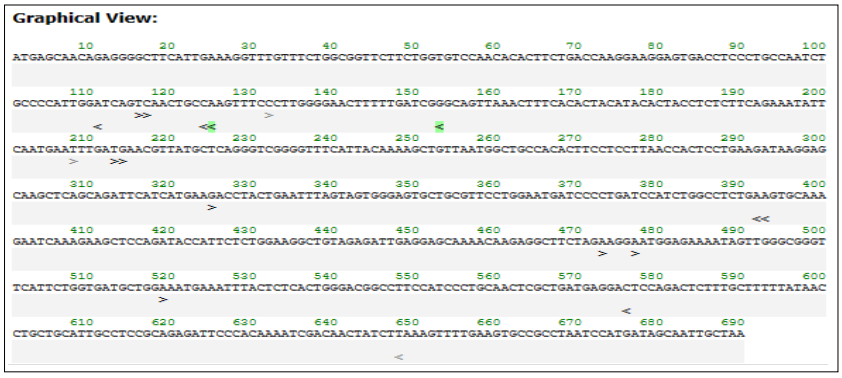


Fig 2C: A graphical view of target sites demonstrates the position and orientation of each site.

Table 4: List of 5 pedicted sgRNAs targeting prolactin (PRL) gene by CRISPR/Cas9 target online predictor (CC Top) tool.

Coordinates	Strand	MM	Target_seq	PAM	Distance		Gene name	Gene id
2:58573411-58573433	+	0	GAACGTTA[TGCTCAGGGTCG]	GGG	0	E	prl	ENSGALG00000012671
2:57931650-57931672	+	4	GGACGTTG[TGTTGAGGGTCG]	TGG	4658	-	DCDC2	ENSGALG00000042454
10:20557838-20557860	-	4	GGAAGTGA[TGCTCAGCGTCG]	AGG	125	-	MFAP1	ENSGALG00000028295
6:8282476-8282498	-	4	GTATGTTA[TGGTCAGGGACG]	TGG	33250	-		ENSGALG00000048540
5:57252885-57252907	-	4	CAAATTTA[TGCTCAGGGTAG]	GGG	4058	-	MDGA2	ENSGALG00000012228

or GN-/NG- at the 5end of the sgRNA with CHOPCHOP tool. In the CRISPR Direct tool, the presence or absence of TTTT in the target sequence is also indicated because four consecutive T's cause pol III terminations in gRNA vectors with pol III promoters. Further in the CRISPR Direct tool the column 20 mer+PAM shows the number of hits with perfect matches for each target sequence (20 mer) adjacent to the PAM. In the CC TOP tool adjacent to the PAM a 12 bp default length core sequence is displayed. Since the seed sequence alterations at positions 8 to 12 nt next to the PAM are known to affect cleavage, it is likely that this region plays the most important role in determining target specificity. The selecting sites predicted to have the most specific seed regions with the fewest possible off-target mismatches may be crucial to improving on-target efficiency (DiCarlo *et al*., 2013; Sander *et al*., 2014;

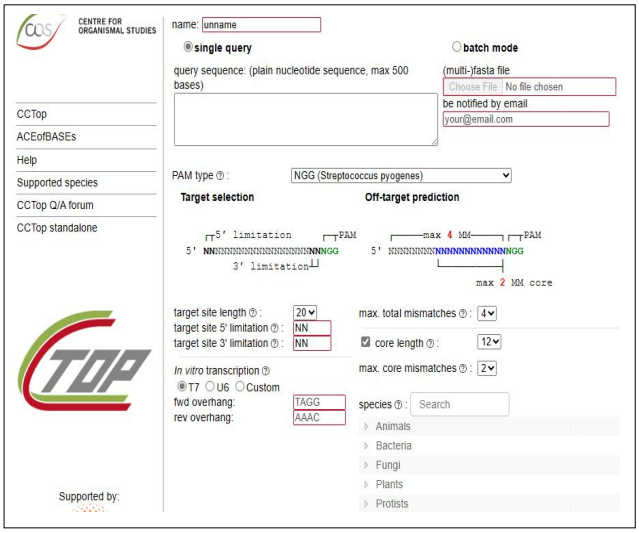


Fig 3A: CCTop web interface CCTop web interface.

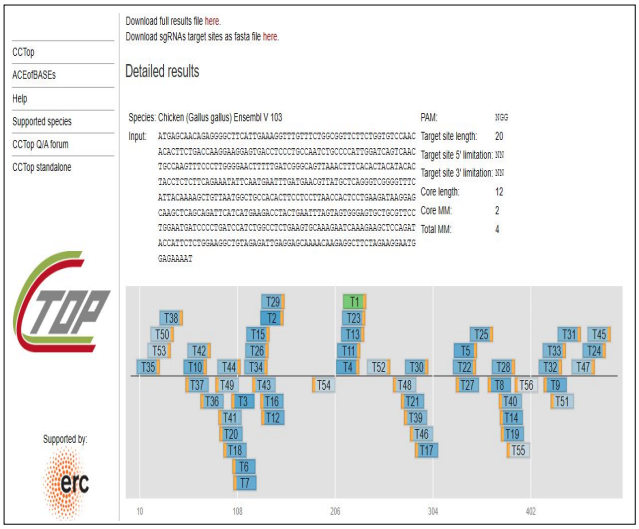
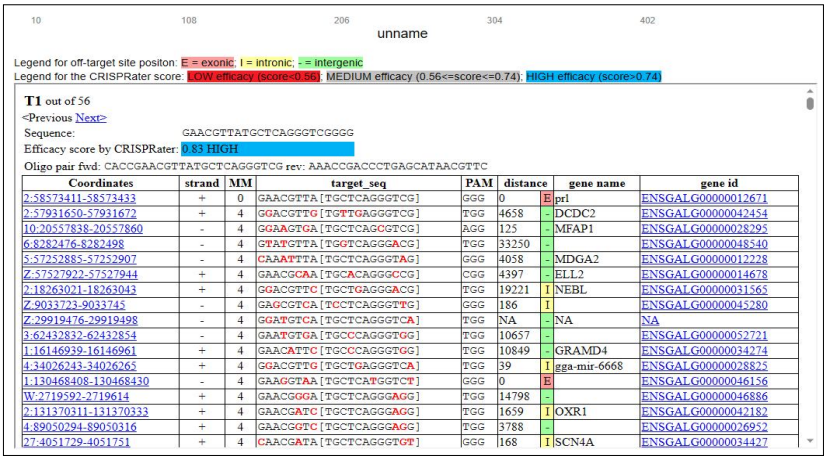


Fig 3B: Main page containing the input fields to customize the identification of sgRNA target sites and the off-target prediction.



Coordinates	strand	MM	target_seq	PAM	distance		gene name	gene id
2:58573411-58573433	+	0	GAACGTTA[TGCTCAGGGTCG]	GGG	0	E	prl	ENSGALG00000012671
2:57931650-57931672	+	4	GGACGTTG[TGTTGAGGGTCG]	TGG	4658	-	DCDC2	ENSGALG00000042454
10:20557838-20557860	-	4	GGAAGTGA[TGCTCAGCGTCG]	AGG	125	-	MFAP1	ENSGALG00000028295
6:8282476-8282498	-	4	GTATGTTA[TGGTCAGGGACG]	TGG	33250	-		ENSGALG00000048540
5:57252885-57252907	-	4	CAAATTTA[TGCTCAGGGTAG]	GGG	4058	-	MDGA2	ENSGALG00000012228
Z:57527922-57527944	+	4	GAACGGAA[TGCACAGGGCCG]	CGG	4397	-	ELL2	ENSGALG00000014678
2:18263021-18263043	+	4	GGACGTTC[TGCTGAGGGACG]	TGG	19221	I	NEBL	ENSGALG00000031565
Z:9033723-9033745	-	4	GAGCGTCA[TCCTCAGGGTTG]	GGG	186	I		ENSGALG00000045280
Z:29919476-29919498	-	4	GGATGTCA[TGCTCAGGGTCA]	TGG	NA	-	NA	NA
3:62432832-62432854	-	4	GAATGTGA[TGCCCAGGGTGG]	TGG	10657	-		ENSGALG00000052721
1:16146939-16146961	+	4	GAACATTC[TGCCCAGGGTGG]	TGG	10849	-	GRAMD4	ENSGALG00000034274
4:34026243-34026265	+	4	GGACGTTG[TGCTGAGGGTCA]	TGG	39	I	gga-mir-6668	ENSGALG00000028825
1:130468408-130468430	-	4	GAAGGTAA[TGCTCATGGTCT]	GGG	0	E		ENSGALG00000046156
W:2719592-2719614	+	4	GAACGGGA[TGCTCAGGGAGG]	TGG	14798	-		ENSGALG00000046886
2:131370311-131370333	+	4	GAACGATC[TGCTCAGGGAGG]	TGG	1659	I	OXR1	ENSGALG00000042182
4:89050294-89050316	+	4	GAACGGTC[TGCTCAGGGAGG]	TGG	3788	-		ENSGALG00000026952
27:4051729-4051751	+	4	CAACGATA[TGCTCAGGGTGT]	GGG	168	I	SCN4A	ENSGALG00000034427

Fig 3C: Results page providing detailed information of all identified sgRNA target sites.

Liu *et al*., 2016). CRISPR/Cas9 system has off-target activity with insertions or deletions between target DNA and gRNA sequences (Lin *et al*., 2014), mismatches were tolerated at any position except within the PAM motif (Hsu *et al*., 2013). Another study demonstrated that the number of mismatches tolerated is dependent upon the specific sgRNA (Fu *et al*., 2013). All the studied tools take into consideration of mismatch position along the sgRNA/DNA heteroduplex as default. So the study showed all the tools are effective in generating sgRNAs considering all the criteria to target the gene of interest for increased efficiency of editing by CRISPR/Cas 9 mechanism. However, all the tools are updated as per the scientific findings to make the target even more precise.

CONCLUSION

A thorough and accurate off-target prediction improves the implementation of the CRISPR/Cas technology. All the tools have different design processes with fast run times, powerful off-target prediction and integrated primer design. These insilico tools accept a wide range of inputs - gene identifiers, genomic regions, or pasted sequences making them suitable for a broad range of uses, providing a dynamic graphical result presentation with an interactive gene visualization. The visualization of all possible target sites makes the selection of the optimal candidate easy. Users can copy and paste all the sgRNAs from the tools under study. In order to analyze whether Cas9 has successfully cleaved the target locus, users may need to amplify the region of interest for further analysis by methods such as deep sequencing or a T7E1 assay. All the tools match the needs of both beginners and experts.

ACKNOWLEDGEMENT

All the authors would like to thank the Director, ICAR-DPR, Hyderabad and P.V.Narsimha Rao Telangana Veterinary University, Rajendranagar, Hyderabad.

Conflict of interest

The authors declare that there are no conflicts of interest.

REFERENCES

Bogdanove, A.J. and Voytas, D.F. (2011). TAL effectors: Customizable proteins for DNA targeting. Science. 333: 1843-1846.

Carroll, D. (2011). Genome engineering with zinc-finger nucleases. Genetics. 188: 773-782.

Chandrasegaran, S. and Carroll, D. (2016). Origins of programmable nucleases for genome engineering. Journal of Molecular Biology. 428: 963-989.

Cubbon, A., Ivancic-Bace, I., Bolt, E.L. (2018). CRISPR-Cas immunity, DNA repair and genome stability. Bioscience Reports. 38: BSR20180457. doi: 10.1042/BSR20180457.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Research. 41: 4336-4343.

Doudna, J.A. and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science. 346: 1258096. DOI: 10.1126/science.1258096.

Fu,Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nature Biotechnology. 31: 822-826.

Gupta, R.M. and Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs and CRISPR-Cas9. The Journal of Clinical Investigation. 124: 4154-4161.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology. 31: 827-832.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli* and identification of the gene product. Journal of Bacteriology. 169: 5429-5433.

Jamal, M., Khan, F.A., Da, L., Habib, Z., Dai, J., Cao, G. (2016). Keeping CRISPR/Cas on-target. Current Issues in Molecular Biology. 20: 1-12.

Jiang, F. and Doudna, J.A. (2017). CRISPR-Cas9 structures and mechanisms. Annual Review of Biophysics. 46: 505-529.

Labuhn, M., Adams, F.F., Ng, M., Knoess, S., Schambach, A., Charpentier, E.M., Heckl, D. (2018). Refined sgRNA efficacy prediction improves large-and small-scale CRISPR-Cas9 applications. Nucleic Acids Research. 46(3): 1375-1385.

Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Bao, G. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Research. 42: 7473-7485.

Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J. and Takumi, T. (2016). Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. Scientific Reports. 6(1): 19675. https://doi.org/10.1038/srep19675.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Koonin, E.V. (2011). Evolution and classification of the CRISPR-Cas systems. Nature Reviews Microbiology. 9: 467-477.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., Valen, E. (2014). CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Research. 42(W1): W401-W407.

Naito, Y., Hino, K., Bono, H., Ui-Tei, K. (2015). CRISPR direct: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics. 31(7): 1120-1123.

Nemudryi, A.A., Valetdinova, K.R., Medvedev, S.P., Zakian, S.M. (2014). TALEN and CRISPR/Cas genome editing systems: Tools of discovery. Acta Naturae. 6: 19-40.

Ren, X., Yang, Z., Xu, J., Sun, J., Mao, D., Hu, Y., Ni, J.Q. (2014). Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. Cell Reports. 9: 1151-1162.

Sander, J.D. and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology. 32: 347-355.

Shah, S.A., Erdmann, S., Mojica, F.J., Garrett, R.A. (2013). Protospacer recognition motifs: Mixed identities and functional diversity. RNA Biology. 10: 891-899.

Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J., Mateo, J.L. (2015). CCTop: An intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PloS One. 12(4): e0176619. https://doi.org/10.1371/Journal.Pone.0176619.

Wang, H., La Russa, M., Qi, L.S. (2016). CRISPR/Cas9 in genome editing and beyond. Annual Review of Biochemistry. 85: 227-264.

Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., Wang, Y. (2015). Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. Cell. 163: 840-853.

Wang, T., Wei, J.J., Sabatini, D.M., Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science. 343: 80-84.

Xue, C. and Greene, E.C. (2021). DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. Trends in Genetics. 37: 639-656.