Genes & Diseases

SHORT COMMUNICATION

In silico sgRNA tool design for CRISPR control of quorum sensing in *Acinetobacter* species

Abraham Peele Karlapudi [a],*, Venkateswarulu T.C [a], Jahnavi Tammineedi [a], Krupanidhi Srirama [a], Lohit Kanumuri [a], Vidya Prabhakar Kodali [b]

[a] *Department of Biotechnology, Vignan's Foundation for Science Technology and Research, Vadlamudi, Guntur 522213, India*
[b] *Department of Biotechnology, Vikrama Simhapuri University, Nellore, India*

Abstract   CRISPR genome editing utilizes Cas9 nuclease and single guide RNA (sgRNA), which directs the nuclease to a specific site in the genome and makes a double-stranded break (DSB). Design of sgRNA for CRISPR-Cas targeting, and to promote CRISPR adaptation, uses a regulatory mechanism that ensures maximum CRISPR-Cas9 system functions when a bacterial population is at highest risk of phage infection. *Acinetobacter baumannii* is the most regularly identified gram-negative bacterium infecting patients. Recent reports have demonstrated that the extent of diseases caused by *A. baumannii* is expanding and, in a few cases, now surpasses the quantity of infections caused by *P. aeruginosa*. Most *Acinetobacter* strains possess biofilm-forming ability, which plays a major role in virulence and drug resistance. Biofilm bacteria use quorum sensing, a cell-to-cell communication process, to activate gene expression. Many genes are involved in biofilm formation and the mechanism to disrupt the biofilm network is still not clearly understood. In this study, we performed in silico gene editing to exploit the *AbaI* gene, responsible for biofilm formation. The study explored different tools available for genome editing to create gene knockouts, selecting the *A. baumannii AbaI* gene as a target. Copyright © 2018, Chongqing Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

---

* Corresponding author. Vignan's Foundation for Science Technology and Research, Vadlamudi, India.
  E-mail address: youngscholar2013@gmail.com (A.P. Karlapudi).

Introduction

The relatively recent discovery of RNA-related resistance mechanisms among bacteria, which specifically targets foreign DNA, has been a significantly exciting development in molecular genetics. The aim of genome engineering is to change the genetic makeup of an organism in order to achieve desired outputs. Homologous recombination has been used widely as a traditional method for gene modification. However, this method has certain limitations, including the fact that it is not applicable to genetically modified organisms. Several restriction endonucleases are available for gene editing, such as Zinc-Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs), Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-Associated Protein-9 nuclease (Cas9). Chronic diseases associated with wound infections are related to biofilm development, and therefore with a corresponding relationship to drug-resistance mechanisms. Methicillin Resistance *Staphylococcus aureus* and *Pseudomonas aeruginosa* are the biofilm-producing microorganisms which most often infect patients. *Acinetobacter baumannii* infections have been increasing globally and have even occasionally comprised a higher rate of infection than those due to *P. aeruginosa*.[1] To study the gene involved in biofilm formation and quorum sensing regulation (*AbaI*), we explored the online available Insilco web tools; these tools can be used to investigate a relatively recent technology termed CRISPR Cas-9 to produce gene knockouts more effectively with increased accuracy and efficiency.[2] The CRISPR/Cas-9 framework permits specific target cleavage of genomic DNA guided by Cas9 nuclease when guide RNA (gRNA) forms a complex and binds to a 20-nucleotide complementary sequence adjacent to the protospacer adjacent motif (PAM) sequence at the 5′-NGG near the protospacer sequence.[3] CRISPR-based approaches have become an important tool in designing gRNA. The system generates gene knock-outs in various species by directly introducing sgRNA, leading to a DNA double-strand break (DSB) at a specified locus. In addition to its enhanced specificity, CRISPR technology allows complete functional knockout, which generates phenotypes more uniformly than other systems such as RNAi, which results in incomplete functional loss in some cases; in many experiments, CRISPR has replaced RNAi approaches for gene knock-out studies. However, TALENs also provide an efficient option, if the goal is to edit a specific gene, because they can create a cleavage site anywhere in the genome. CRISPR-Cas9 systems can be controlled to become capable tools for genome editing due to their simple design and cleavage activity. CRISPR is preferred over TALENs because it is less laborious and more affordable. Newly designed insilico targets for sgRNA expression in *A. baumannii* would enable a CRISPR—Cas9 genome editing system for bacteria. This system has been widely used to disrupt genes involved in biofilm biosynthesis, specifically the AbaI gene, which was found to be critical for the production of Acyl-homoserine lactone.[4] To determine the effects of AbaI on biofilm production, wild-type and mutant strains will be cultured: growth profile, gene expression and biofilm production will be monitored using wet laboratory analysis. AbaI

transcriptional patterns should suggest the way in which this gene is critical for activation of biofilm production. The *A. baumannii* strain isolated by Niu et al (2008) produces an Acyl-homoserine lactone molecule due to the action of *AbaI*. Mutagenesis of *AbaI* and its subsequent inactivation results in biofilm reduction.[4] An in silico tool based on the CRISPR approach provides flexibility, automation and is rapid enough to evaluate a large number of genes; however, in vivo cell- or animal-based models are used to observe phenotypic changes as a result of knocking out one or more genes.

SgRNA web tools and *AbaI* reference gene

The sequence of *Acinetobacter baumannii* strain M2 auto inducer synthase (AbaI) was retrieved from NCBI (Gen bank Accession ID: EU334497.1). The sequence was selected as a reference to study the in-silico approach for the design of sgRNA. Designing tools include integrated genomes of specific organisms with their Gene IDs retrieved from the ENSEMBL, UCSC genome browsers. The candidate search of genomic targets includes the 20-nucleotide gene sequence, which is unique compared to the rest of the genome and the target sequence, which should be immediately upstream of a Protospacer Adjacent Motif (PAM). This study will analyze the *AbaI* gene as a query target sequence by using a variety of tools available online in order to select a potential target sequence which meets the above two conditions. There are various in silico tools now available to assist with sgRNA selection based on on- and off-target activity considerations (Table 1).

We compared five on-target sgRNA design tools (CHOPCHOP, CCTop, E-CRISP, CRISPRdirect, CRISPR-ERA), and one off-target tool (Off-Spotter), as shown in Table 1. CRISPRdirect and CCTop are alignment based, whereas sgRNAs are aligned from the given genome purely by locating the PAM; E-CRISP, CHOPCHOP and CRISPR-ERA are hypothesis driven, whereas sgRNAs are aligned and scored with the contribution of specific factors (e.g., GC content and exon position) to sgRNA on-target efficacy. Table 1 provides a comparison of tools for the design of CRISPR guide RNA with genomic tools, featuring qualities ensuring the best genome editing method of choice. Most sgRNA design tools apply the most fundamental basis for high on-target action by distinguishing all PAM sites for the predetermined Cas9. These tools have different degrees of adaptability with respect to the genome and PAM site choices, and permit clients to enter any gene of interest. To suit elective Cas9 PAM prerequisites, a few design tools offer choices for selecting client-predefined PAMs. In total, more than one sgRNA is utilized for each target because not all designed sgRNAs are effective, even with the best efficacy-prediction tools.

CHOPCHOP

The CHOPCHOP tool (https://chopchop.rc.fas.harvard.edu) is widely used for the analysis of gRNA sequence and off-target numbers. The CHOPCHOP web server is the most broadly utilized for CRISPR-based genome modifications. CHOPCHOP initially requires a nucleotide sequence of a selected gene of interest. Users can specifically target

Table 1   Tools for the design of guide RNAs and prediction of off-target effects.

Tool	Input	Use case	Ref
CHOPCHOP	Click on input sequence, paste query sequence and click on find target sites	Analyze the target sequence, gRNA, PAM sequence, primer sequence and off-target numbers	5
CCTop	Paste the gene or exon sequence, and click on submit.	Analyze target sequence, direction, PAM sequence and high quality off target sites	6
E-CRISP	Paste fasta sequence of gene. User can choose the option and click on sgRNA design	Analyze the mutation sites and sgRNA for exon or gene sequence	7
CRISPR Direct	Paste query sequence and click on design	Analyze target sequence, direction, PAM sequence and number of off target and the mutation sequence	8
Off-Spotter	Paste the target or sgRNA sequence without PAM sequence and Click on ensemble id, protein id and on ccds	Analyze the off gets and gives the mutation analysis at transcription and translation level along with proteomics information	9
Crispr-era	Choose gene sequence and then paste the sequence and Click on Crispr-era search	Analyze the location of the target for knock-out analysis	10

exon sequences, splice sites and UTR regions. CHOPCHOP software automatically analyzes the query sequence and displays all possible 20bp sequences identified immediately after the PAM sequence (5′-NGG), providing gRNA scores from best to lowest. A significant part of sgRNA intended for a given quality will deliver a low or zero cutting rate, and numerous sgRNAs have the ability to tie indiscriminately in the genome, which can prompt off-target mutagenesis.

Retrieval of the biofilm-forming *AbaI* gene nucleotide fasta sequence from an NCBI search submitted to CHOPCHOP identified the 20 mer and PAM sequences in the query DNA sequence and determined the upstream starting point, followed by specifying the short guide RNA. Oligonucleotides required for the PCR assembly of the gRNA have also been identified by the tool. Output results in Fig. 1 illustrate sgRNA quality visualized in colours, while Table 2
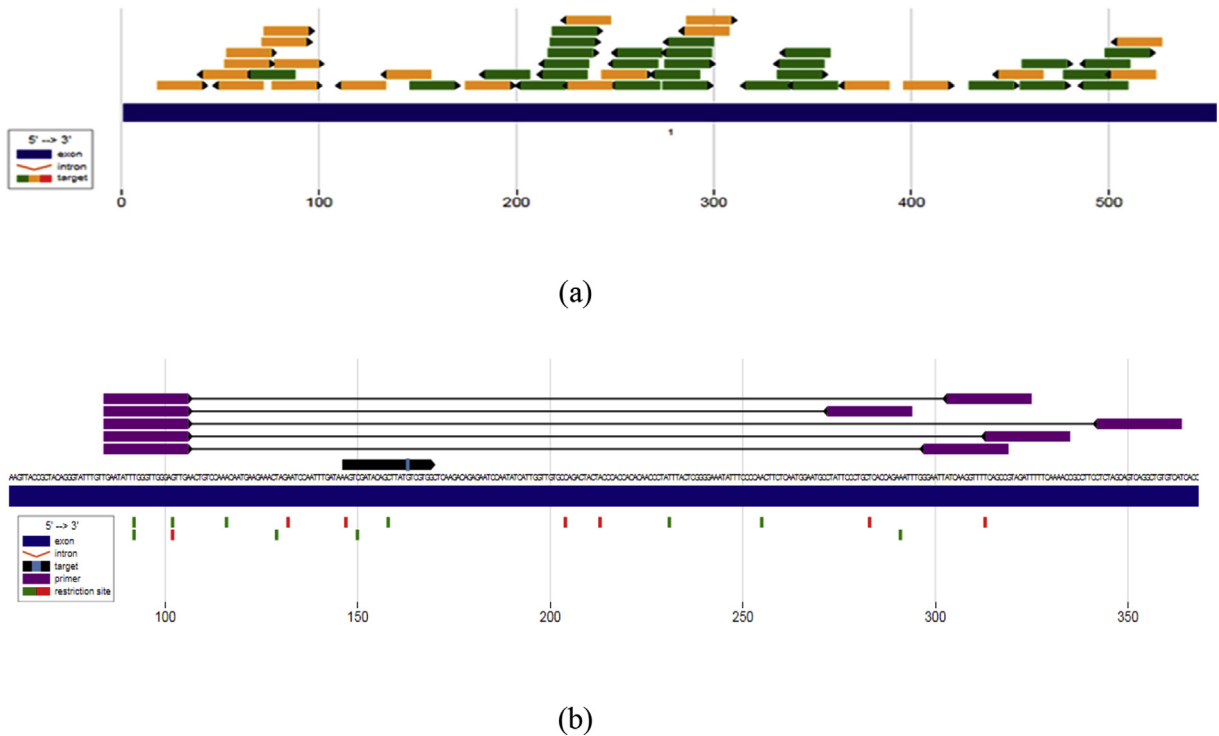

(a)

(b)

Figure 1   CHOPCHOP output results shown in Cas9 nickase mode with default parameters. The results are displayed across the gene; sgRNA target sites within the gene (exon, blue; intron, red), green (good), yellow (better) colors indicates according to ranking (b) Clicking on a specific target sequence result displays the predicted cut site in red, primer options. Different colors indicate the quality of each sgRNA, green indicates good quality.

Table 2 Shows the result interpretation of CHOPCHOP on providing the list of possible target sites, GC% and number of offsite targets with the best target site as rank one. columns '12 mer+PAM' and '8 mer+PAM' in order to show the number of hits with perfect matches for their seed sequence (12 or 8 mer, respectively) adjacent to the PAM. Note that the numbers of hits displayed here include both on-target and off-target sites.

Ranking	Target sequence	Genomic location (EU334497.1)	Exon	Strand	GC (%)	Off targets		
						0	1	2
1	AGTCGATACAGCTTATGTCGTGG	147	1	+	48	0	0	0
2	TCATTGGTGGACCTGCACGATGG	489	1	−	57	0	0	0
3	GAGTAAATAGGGTTGTGTGGTGG	215	1	−	48	0	0	0
4	ATTCAACAAATACCCTGTAGCGG	66	1	−	43	0	0	0
5	ACAGCCTGACTGCTAGAGGAAGG	337	1	−	57	0	0	0
6	AGGCGGTTTTGAAAAATCTACGG	317	1	−	43	0	0	0
7	GCCTGACTGCTAGAGGAAGGCGG	334	1	−	65	0	0	0
8	AGAACGTTTATTACGTGCTGCGG	456	1	+	48	0	0	0
9	CATTGGTGGACCTGCACGATGGG	488	1	−	61	0	0	0
10	CACACAACCCTATTTACTCGGGG	219	1	+	52	0	0	0

provides information regarding target sequence GC content (%), so that the best sequences are selected between 40 and 80%. Primer details to the target site are provided in Table 3.

CRISPRdirect

CRISPRdirect (http://crispr.dbcls.jp/) is an online tool which accepts nucleotide sequences up to 10 Kbp, and which analyzes the target sequence, direction, PAM sequence, number of off targets and the mutation sequence. It is a simple and functional web server for selecting rational CRISPR/Cas targets. The tool searches the entire genome for the best matches with each target sequence (20 mer) and PAM sequences (12 or 8 mer). CRISPRdirect provides efficient selection of CRISPR/Cas target sites with reduced numbers of potential off-target candidates. CRISPRdirect restores the outcomes more rapidly than other tools and gives a helpful interface for gRNA configuration, making it a capable apparatus for utilizing the CRISPR/Cas framework on a wide level. A typical output of CRISPRdirect is a display of a list of target candidate sequences, with '0' in '20mer+PAM' being shown in gray. Target sequences with TTTTs are also shown in gray. The avoidance of TTTTs in gRNA vectors is via use of the pol III promoter (Fig. 2).

E-CRISP

The E-CRISP (http://www.e-crisp.org/E-CRISP/) web server evaluates target sites of query sequences, allowing users to design CRISPR gRNA to knock out genes, while also including advanced options of De novo. E-CRISP designs gRNA sequences via multiple libraries and identifies target sequences complementary to the gRNA ending in a protospacer-adjacent motif (PAM); this tool typically searches the mutation sites and sgRNA for exon or gene sequence by rapidly finding binding sites. E-CRISP provides putative designs of targets and assesses their exons, transcripts and CpG islands (Fig. 3).

CCTop

CRISPR/Cas9 target online predictor (CCTop, http://crispr.cos.uni-heidelberg.de) targets and identifies candidate sgRNA target sites according to the high quality of off-targets on selecting the required PAM sequence. CCTop has a sensible number of adaptable alternatives to provide a list of the best candidates, with users being able to select target sites in a simple way. CRISPR/Cas9 has off-target activity with indels in the DNA; exploring off-targets would be a more proficient approach for selecting down targets. CCTop gives the output results with a maximum number of mismatches that can be found near the PAM; mismatches do not allow double strand breaks to take place (Fig. 4).

Table 3 PAM sequence to that target site along with oligonucleotides or primer details.

Pair	Left primer coordinates (cds)	Left primer	Left primer Tm	Right primer coordinates (cds)	Right primer	Right primer Tm	Product size
1	85—107	GAATATTTGGGTTGGGAGTTGA	60.1	298—320	CGGCTGAAAACCTTGATAATTC	60	235
2	85—107	GAATATTTGGGTTGGGAGTTGA	60.1	314—336	GGTTTTGAAAAATCTACGGCTG	60	251
3	85—107	GAATATTTGGGTTGGGAGTTGA	60.1	343—365	ATGACACAGCCTGACTGCTAGA	60.1	280
4	85—107	GAATATTTGGGTTGGGAGTTGA	60.1	273—295	AATTTCTGGTGAGCAGGGAATA	60	210
5	85—107	GAATATTTGGGTTGGGAGTTGA	60.1	304—326	AATCTACGGCTGAAAACCTTGA	60.1	241

position			target sequence		sequence information			
start - end	+	-	20mer+PAM (total 23mer)	GC% of 20mer	Tm of 20mer	TTTT in 20mer	restriction sites	
333 - 355	+		ACCGCCTTCCTCTAGCAGTC AGG [gRNA]	60.00 %	78.33 ºC	-		
334 - 356	-		CCG CCTTCCTCTAGCAGTCAGGC [gRNA]	60.00 %	77.18 ºC	-		
478 - 500	+		GGTTTTCGTGCCCATCGTGC AGG [gRNA]	60.00 %	76.30 ºC	+	Bme1580I	
337 - 359	-		CCT TCCTCTAGCAGTCAGGCTGT [gRNA]	55.00 %	76.69 ºC	-		
341 - 363	-		CCT CTAGCAGTCAGGCTGTGTCA [gRNA]	55.00 %	74.93 ºC	-		
488 - 510	-		CCC ATCGTGCAGGTCCACCAATG [gRNA]	55.00 %	74.79 ºC	-		
489 - 511	-		CCA TCGTGCAGGTCCACCAATGA [gRNA]	55.00 %	75.35 ºC	-		

Figure 2 Screenshot of CRISPRdirect output showing AbaI gene candidates targeted by CRISPR/Cas9 system. The column '20 mer+PAM displayed the perfect match hits for each 20 mer target sequence adjacent to the PAM.

Name	Nucleotide sequence	SAE-Score	Target	Matchstring	Number of Hits
EU334497.1_2_0	GGCATTCCATTGAGA AGTTG NGG		*	NO-TARGET	1
EU334497.1_5_0	GGTTTTCGTGCCCAT CGTGC NGG		*	NO-TARGET	2
EU334497.1_5_0	GGTTTTCGTGCCCAT CGTGC NGG		*	NO-TARGET	2
EU334497.1_3_0	GCCTGACTGCTAGAG GAAGG NGG		*	NO-TARGET	3
EU334497.1_3_0	GCCTGACTGCTAGAG GAAGG NGG		*	NO-TARGET	3
EU334497.1_3_0	GCCTGACTGCTAGAG GAAGG NGG		*	NO-TARGET	3

Figure 3 E-CRISP targets different regions of the AbaI gene sequence and displayed the filtered results. Off-targets are analyzed on the basis of sequence alignment of each design to the reference genome; SAE refers to Specificity, Annotation and Efficiency scores.
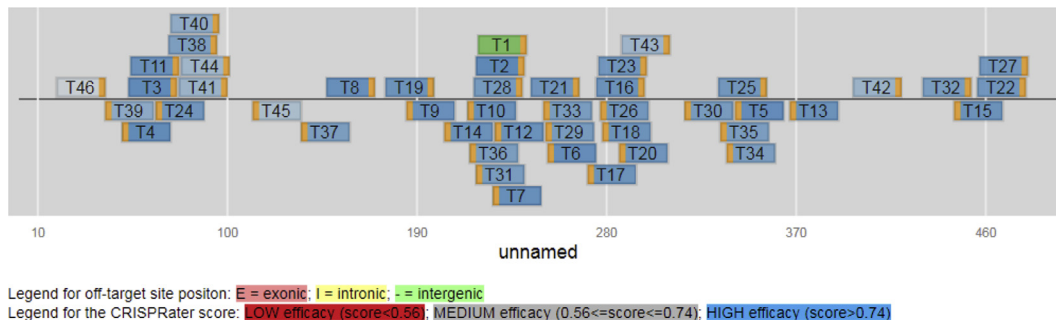
Figure 4 Screenshot providing the information on all sgRNA target sites using CCTop server. The output consists of a forwarded result that can be downloaded as it is or in a FASTA file and also the query sequence can be visualized with color coded sgRNA.

# of occurences	Strand	20mer	Reverse complement of	Starting position on input	# of results per mismatch	
23	+	AGTCGATACAGCTTATGTCG		147	0 / 1 / 2 / 3 / 4 / 5	0 / 0 / 0 / 0 / 2 / 21
37	-	TCATTGGTGGACCTGCACGA	TCGTGCAGGTCCACCAATGA	492	0 / 1 / 2 / 3 / 4 / 5	0 / 0 / 0 / 0 / 5 / 32
42	+	CACACAACCCTATTTACTCG		219	0 / 1 / 2 / 3 / 4 / 5	0 / 0 / 0 / 0 / 8 / 34

Figure 5 OFF SPOTTER displayed the number of off-targets and mismatches for the corresponding gRNA and its genomic location. Output includes a histogram showing the number of potential off-targets as a function of the number of mismatches at the site's genomic location.

Table 4   Top ranking sgRNA design of *Acinetobacter baumannii*, AbaI gene for genome editing.

Length	Start	End	Strand	Nucleotide sequence	S-Score	E-Score	Genomic start	Genomic End
23	249	272	minus	GGCATTCCATTGAGAAGTTG NGG	100	56.637	−251	−228
23	477	500	plus	GGTTTTCGTGCCCATCGTGC NGG	80	61.9142	−23	0
23	477	500	plus	GGTTTTCGTGCCCATCGTGC NGG	80	61.9142	−23	0
23	333	356	minus	GCCTGACTGCTAGAGGAAGG NGG	60	73.0403	−167	−144
23	333	356	minus	GCCTGACTGCTAGAGGAAGG NGG	60	73.0403	−167	−144
23	333	356	minus	GCCTGACTGCTAGAGGAAGG NGG	60	73.0403	−167	−144
23	456	479	plus	GAACGTTTATTACGTGCTGC NGG	60	58.4173	−44	−21

Off-Spotter

Off-Spotter (https://cm.jefferson.edu/Off-Spotter/) is a rapid and thorough algorithmic answer for identifying off-targets and all genomic sites that satisfy PAM conditions to provide gRNA information. Users can easily access the site and can utilize checkmark catches to characterize the number of nucleotides the query contains. The Off-Spotter algorithm helps users with ideal gRNA design at the point of applying gRNA succession and PAM constraint. Off-Spotter rapidly and thoroughly recognizes all genomic destinations that fulfil satisfactory PAM conditions. Off-Spotter is quick, adaptable and can help in the design of ideal gRNAs by providing several PAM decisions, a run-time meaning of the essential seed sequence and permitted number of mismatches, and an adaptable yield interface that permits arrangement of the outcomes. Each 20-mer target would be processed and the off-target results reported in separate tables (Fig. 5).

CRISPR-ERA

CRISPR-ERA (http://crispr-era.stanford.edu) is an automated sgRNA design tool for CRISPR-mediated editing, repression and activation. This tool was designed for construction of sgRNA for regulation of genes and is used to predict nucleotide sequences. The tool uses a rapid calculation algorithm to scan for sgRNA restricting destinations present genome wide for evaluating efficiency and specificity using a set of rules from a database. CRISPR-ERA provides the option of defined target sequences as input to the user in the FASTA format. sgRNA are ranked based on E (efficacy) and S (specificity) scores and on providing details based on off-target binding sites (Table 4).

Conclusion

Tools such as E-CRISPR, CRISPR-ERA provide distinct and attractive features and are able to adapt to a wide range of genomes. Among the six tools used, E-CRISPR provides a rapid and highly straightforward approach for identification of sgRNA targets. Basic functions are the first choice of determining a tool's suitability; tools such as E-CRISP, CCTop, CHOPCHOP, CRIPSR Direct and Off-spotter range from 31 to 41% in terms of their basic functions, such as single and multi-target design, off-target awareness, mismatch limit, single and multi-PAM design. Among these tools, E-CRISP provides the highest number of functions (41%), with shared advanced functions such as SNP, secondary structure and micro homology awareness of 7—31%. The CHOPCHOP tool has the most advanced functions, with 10—20% of user utility functions that help increase the speed of sgRNA primer design. All six software tools have input flexibility in terms of being user-friendly and show good output diversity in different formats. Most of the tools are exclusive to a certain subset of organisms.[11] Considering the present research scenario, it is evident that knowledge on the use of sgRNA on-target design tools in various genomes other than humans and mice is unclear. The final goal of using the different set of web tools discussed is to achieve more efficient and predictable gene editing. This gene knockout approach using CRISPR technology allows molecular biologists to easily engineer large numbers of cloned cell lines and re-engineer Cas9 for producing desirable characteristics. CRISPR-Cas9 creates advantages for functional genomic research applications due to its simplicity and specificity. In silico sgRNA design provides a key step for the CRISPR system and will allow CRISPR studies to take advantage of bioinformatics and computational techniques. Continuing efforts will be necessary to improve in silico sgRNA design with high on-target efficacy and reduced off-target effects. All the existing tools are designed for general sgRNA design without considering cell-type heterogeneity; personalized sgRNA design is an important future direction for in silico sgRNA design. Continuing efforts to evaluate this technology with respect to prokaryotic and eukaryotic model systems will be beneficial for creating error-prone disease-free organisms and new drugs and therapies at the gene level in the coming years.

Conflict of the interest

The authors declare no conflicts of interest.

Acknowledgement

References

1. Gaj T, Gersbach CA, Barbas III CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 2013;31:397—405.

2. Tsai SQ, Joung JK. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet*. 2016; 17:300—312.

3. Bialk P, Rivera-Torres N, Strouse B, Kmiec EB. Regulation of gene editing activity directed by single-stranded oligonucleotides and CRISPR/Cas9 systems. *PLoS One*. 2015;8:10.

4. Niu C, Clemmer KM, Bonomo RA, Rather PN. Isolation and characterization of an autoinducer synthase from *Acinetobacter baumannii*. *J Bacteriol*. 2008;190:3386—3392.

5. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res*. 2014;42:401—407.

6. Stemmer M, Thumberger T, Del Sol KM, Wittbrodt J, Mateo JL. CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One*. 2015;10.

7. Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014;11:122—123.

8. Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*. 2015;31:1120—1123.

9. Pliatsika V, Rigoutsos I. "Off-Spotter": very fast and exhaustive enumeration of genomic lookalikes for designing CRISPR/Cas guide RNAs. *Biol Direct*. 2015;29:10.

10. Liu H, Wei Z, Dominguez A, Li Y, Wang X, Qi LS. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics*. 2015;31: 3676—3678.

11. MacPherson CR, Scherf A. Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. *Nat Biotechnol*. 2015,June 29;33:805.