Bioinformatics, 31(24), 2015, 4014–4016
doi: 10.1093/bioinformatics/btv537
Advance Access Publication Date: 10 September 2015
Applications Note

OXFORD

Sequence analysis

Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites

Jeongbin Park[1], Sangsu Bae[1,2,*] and Jin-Soo Kim[3,4,*]

[1]Department of Chemistry, [2]Institute for Materials Design, Hanyang University, Seoul, South Korea, [3]Center for Genome Engineering, Institute for Basic Science, Seoul, South Korea and [4]Department of Chemistry, Seoul National University, Seoul, South Korea

*To whom correspondence should be addressed.

Associate Editor: John Hancock

Received on July 3, 2015; revised on August 27, 2015; accepted on September 6, 2015

Abstract

Summary: We present Cas-Designer, a user-friendly program to aid researchers in choosing appropriate target sites in a gene of interest for type II CRISPR/Cas-derived RNA-guided endonucleases, which are now widely used for biomedical research and biotechnology. Cas-Designer rapidly provides the list of all possible guide RNA sequences in a given input DNA sequence and their potential off-target sites including bulge-type sites in a genome of choice. In addition, the program assigns an out-of-frame score to each target site to help users choose appropriate sites for gene knockout. Cas-Designer shows the results in an interactive table and provides user-friendly filter functions.

Availability and implementation: Free access at http://rgenome.net/cas-designer/.

Contact: sangsubae@hanyang.ac.kr or jskim01@snu.ac.kr

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Targeted genome editing based on the type II clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) protein system from the prokaryotic adaptive immune response is now widely used for biomedical research and biotechnology (Kim and Kim, 2014). CRISPR RNA-guided endonucleases (RGENs) consist of the Cas9 protein and guide RNAs (gRNAs), which are provided as two small RNAs [sequence-variable crRNA and invariable trans-activating CRISPR RNA (tracrRNA)] or as single-chain gRNAs in which essential portions of crRNA and tracrRNA are linked. RGENs cleave target DNA in a sequence-dependent manner; the repair of the resulting double-strand breaks (DSBs) by endogenous mechanisms gives rise to targeted genome modifications in cells and whole organisms. In addition to the base pairing between gRNAs and target DNA, Cas9 also contributes to RGEN specificity by recognizing a protospacer-adjacent motif (PAM) downstream of the gRNA target sequence. Cas9 proteins derived from different organisms recognize distinct PAM sequences. However, RGENs can also cleave off-target DNA sites with several mismatches (Cho *et al.*, 2014; Fu *et al.*, 2013) or with missing or additional nucleotides (Lin *et al.*, 2014), compared to gRNA sequences.

We and others have presented computer programs that help researchers choose appropriate target sites for Cas9 nucleases (Aach *et al.*, 2014; Cradick *et al.*, 2014; Doench *et al.*, 2014; Bae *et al.*, 2014a; Heigwer *et al.*, 2014; Hsu *et al.*, 2013; Lei *et al.*, 2014; Montague *et al.*, 2014; O'Brien and Bailey, 2014; Naito *et al.*, 2014; Sander *et al.*, 2010; Upadhyay and Sharma, 2014; Xie *et al.*, 2014; Zhu *et al.*, 2014). To the best of our knowledge, however, there is no web-based tool for providing the list of all possible gRNAs in a given input DNA sequence with their potential off-target sites with a DNA or RNA bulge in a genome of choice and with customizable PAM sequences appropriate for Cas9 derived from species other than S. pyogenes.

Here we present Cas-Designer, a quick gRNA designing tool for various type II Cas9 nucleases, which lists all possible gRNA sequences from a given input DNA sequence along with their potential off-target sites in a genome of interest. Notably, Cas-Designer rapidly searches for potential off-target sites with a DNA or RNA bulge (Fig. 1A). Furthermore, Cas-Designer provides an out-of-frame score (Bae *et al.*, 2014b) for each gRNA to help choose appropriate

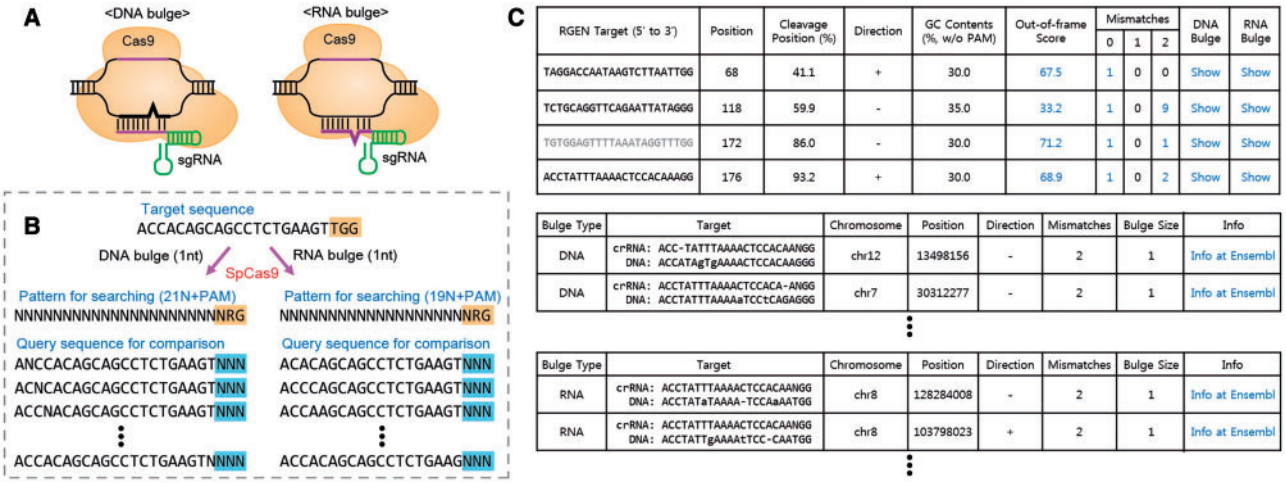

Fig. 1. (**A**) Schematic showing off-target sites with a DNA or RNA bulge. (**B**) The strategy for allowing 1-nt DNA or RNA bulge based on Cas-OFFinder. (**C**) An example of an output table. Cas-Designer shows all possible gRNAs from input sequences along with useful information (top). If a user clicks on a blue colored number, word, or phrase, more detailed information, such as DNA (middle) or RNA (bottom) bulge targets, is provided. In addition, a user can obtain the relevant genomic information via the Ensembl genome browser (Flicek *et al.*, 2011), by clicking on the 'Info at Ensembl' button

sites for gene knockout. Cas-Designer is available as a command-line program or at our website.

2 Implementation

2.1 gRNA selection

First, Cas-Designer finds possible target sites with a user-defined PAM sequence [5′-NGG-3′ or 5′-NRG-3′ for SpCas9, 5′-NNAGAAW-3′ for StCas9 (Cong *et al.*, 2013), 5′-NNNNGMTT-3′ for NmCas9 (Hou *et al.*, 2013) and 5′-NNGRRT-3′ for SaCas9 (Ran *et al.*, 2015)] in a given DNA sequence. Second, Cas-Designer rapidly calculates the microhomology-associated out-of-frame score that is positively correlated with the frequency of frame-shift mutations (Bae *et al.*, 2014b). Cas-Designer shows cleavage positions, GC content and out-of-frame scores in this step. The tool also shows target sequences that contain more than four thymidine nucleotides in tandem in gray to indicate that these sequences are often recognized by RNA polymerase III as transcription termination signals (Braglia *et al.*, 2005).

2.2 Searching for potential off-target sites

Next, Cas-Designer invokes Cas-OFFinder (Bae *et al.*, 2014a) to rapidly identify potential off-target sites in a user-defined genome. To find potential off-target sites with a DNA or RNA bulge, we used a simple brute-force approach as shown in Figure 1B. Briefly, the program inserts up to three additional 'N' nucleotides or deletes up to 3 nucleotides in the target sequence to identify potential off-target sites with a DNA bulge or an RNA bulge, respectively. To this end, we developed a lightweight wrapper script of Cas-OFFinder written in Python, which can be separately downloaded at our website.

Because this step is the most time-consuming, we focused on improving Cas-OFFinder to achieve faster analysis. First, a new version now supports a 2 bit sequence format, which is smaller in size than is a FASTA format. Second, the internal OpenCL (Open Computing Language) kernels now use the local memory of the graphic card rather than the global memory. Third, the new version now uses atomic operation in OpenCL kernels and OpenMP (Open Multi-Processing) in the CPU operating part. These changes reduced the average analysis time per target by a factor of about 30 (Supplementary Fig. S1).

2.3 Output tables and filter options

Cas-Designer shows the results in an interactive table as shown in Figure 1C. A user can click on a blue-colored number or a 'Show' button to access the genomic loci of potential off-target sites.

In addition, we implemented a user-friendly filter function on the results page for convenience. Output data can be filtered according to GC content, out-of-frame scores, or mismatch numbers. The filtered results are shown immediately, without whole-page refreshing, by using the AJAX (Asynchronous JavaScript and Extensible Markup Language) web development technique.

Funding

The Plant Molecular Breeding Center of Next Generation BioGreen 21 Program (PJ01119201) and the research fund of Hanyang University (HY-2015) to S.B. and Institute for Basic Science (IBS-R021-D1) to J.-S.K.

Conflict of Interest: none declared.

References

Aach,J. *et al.* (2014) CasFinder: Flexible algorithm for identifying specific Cas9 targets in genomes. *bioRxiv*, doi:10.1101/005074.

Bae,S. *et al.* (2014a) Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*, **30**, 1473–1475.

Bae,S. *et al.* (2014b) Microhomology-based choice of Cas9 nuclease target sites. *Nat. Methods*, **11**, 705–706.

Braglia,P. *et al.* (2005) Sequence context effects on oligo(dT) termination signal recognition by *Saccharomyces cerevisiae* RNA polymerase III. *J. Biol. Chem.*, **280**, 19551–19562.

Cho,S.W. *et al.* (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.*, **24**, 132–141.

Cong,L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.

Cradick,T.J. *et al.* (2014) COSMID: A web-based tool for identifying and validating CRISPR/Cas Off-target Sites. *Mol. Ther. Nucleic Acids*, **3**, e214.

Doench,J.G. *et al*. (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.*, **32**, 1262–1267.

Flicek,P. *et al*. (2011) Ensembl 2011. *Nucleic Acids Res.*, **39**, D800–D806.

Fu,Y. *et al*. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, **31**, 822–826.

Heigwer,F. *et al*. (2014) E-CRISP: fast CRISPR target site identification. *Nat. Methods*, **11**, 122–123.

Hou,Z. *et al*. (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA.*, **110**, 15644–15649.

Hsu,P.D. *et al*. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.

Kim,H. and Kim,J.-S. (2014) A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.*, **15**, 321–334.

Lei,Y. *et al*. (2014) CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol. Plant*, **7**, 1494–1496.

Lin,Y. *et al*. (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.*, **42**, 7473–7485.

Montague,T.G. *et al*. (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.*, **42**, W401–W407.

Naito,Y. *et al*. (2014) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, **31**, 1120–1123.

O'Brien,A. and Bailey,T.L. (2014) GT-Scan: identifying unique genomic targets. *Bioinformatics*, **30**, 1–3.

Ran,F.A. *et al*. (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*, **520**, 186–191.

Sander,J.D. *et al*. (2010) ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res.*, **38**, W462–W468.

Upadhyay,S.K. and Sharma,S. (2014) SSFinder: high throughput CRISPR-Cas target sites prediction tool. *Biomed Res. Int.*, **2014**, 742482.

Xie,S. *et al*. (2014) sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS One*, **9**, e100448.

Zhu,L.J. *et al*. (2014) CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. *PLoS One*, **9**, e108424.