CRISPy-web: An online resource to design sgRNAs for CRISPR applications

Kai Blin a, Lasse Ebdrup Pedersen a, Tilmann Weber a,*, Sang Yup Lee a,b,**

a The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Alle 6, 2970 Hørsholm, Denmark
b Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 Plus Program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 17 October 2015
Received in revised form 5 January 2016
Accepted 10 January 2016
Available online

Keywords:
CRISPR
Cas9
sgRNA
Guide RNA
Genome editing
Genome engineering

A B S T R A C T

CRISPR/Cas9-based genome editing has been one of the major achievements of molecular biology, allowing the targeted engineering of a wide range of genomes. The system originally evolved in prokaryotes as an adaptive immune system against bacteriophage infections. It now sees widespread application in genome engineering workflows, especially using the Streptococcus pyogenes endonuclease Cas9. To utilize Cas9, so-called single guide RNAs (sgRNAs) need to be designed for each target gene. While there are many tools available to design sgRNAs for the popular model organisms, only few tools that allow designing sgRNAs for non-model organisms exist. Here, we present CRISPy-web (http://crispy.secondarymetabolites.org), an easy to use web tool based on CRISPy to design sgRNAs for any user-provided microbial genome. CRISPy-web allows researchers to interactively select a region of their genome of interest to scan for possible sgRNAs. After checks for potential off-target matches, the resulting sgRNA sequences are displayed graphically and can be exported to text files. All steps and information are accessible from a web browser without the requirement to install and use command line scripts.

© 2016 The authors. Production and hosting by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

“Clustered regularly interspaced short palindromic repeats” (CRISPR) and their associated RNA-guided endonucleases are bacterial adaptive immune systems protecting the bacteria from infections with bacteriophages.1 The biotechnological application of this system is currently revolutionizing molecular biology and provides new opportunities for synthetic biology applications.2 The CRISPR system allows targeted genome engineering of bacteria,3–5 and also eukaryotes including yeast,6,7 plants,8 human cell lines9 and many more. Although several alternative CRISPR systems have recently been described (e.g., References 10–13), most CRISPR systems for genome engineering are based on the Streptococcus pyogenes type II CRISPR/Cas9 system. Cas9 is an RNA-guided endonuclease, which introduces double strand DNA breaks at positions that are complementary to a crRNA sequence that binds to Cas9 in a duplex with a tracrRNA. It has also been demonstrated that Cas9 accepts artificially fused crRNA:tracrRNA-hybrids termed “single guide RNAs” (sgRNAs) to direct it to the target cut sites.14 Thus CRISPR can be used as an “in vivo programmable restriction enzyme,” which cuts the target DNA within an exactly defined target sequence determined by the sgRNA.

One essential step in the design of CRISPR experiments is to identify suitable sgRNA sequences within the target gene(s), which have to fulfill certain conditions: (i) the 20 bp target sequence has to be directly upstream of a “protospacer adjacent motif” (PAM). For the S. pyogenes PAM, this motif is “NGG”, (ii) it is desirable that the sequence of the target motif is unique within the genome of the organisms to prevent off-target activity, i.e. the cleavage of the chromosome at wrong positions.

Many programs and web-servers exist to assist biologists in finding such CRISPR/Cas9 target sites. However, most of these tools (e.g. CCTop,15 CHOPCHOP16 CRISPR Design,17 WU-CRISPR,18 WGE CRISPR Finder,19 and CRISPy CHO20) are limited to a narrow set of – mostly eukaryotic – target genomes of model organisms. Only very few tools (e.g. sgRNAcas921) allow running Cas9 target searches on
A. Overview sgRNAs in *Streptomyces coelicolor* A3(2) actinorhodin biosynthetic gene cluster

![Diagram showing CRISPy web interface for identifying sgRNAs in the actinorhodin biosynthetic gene cluster](image)

B. Zoom in on gene SCO5087 (*actIORF1*, KSα)

![Diagram showing CRISPy web interface for zoomed view of SCO5087](image)

C. List of selected sgRNAs for export

![Diagram showing list of sgRNAs for export](image)

Fig. 1. Example output of a CRISPy-web run to identify sgRNAs in the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor* A3(2) (NCBI GenBank ID: NC_003888.3). (A) Overview of sgRNAs identified in the complete actinorhodin gene cluster. (B) Zoom view of SCO5087 (*actIORF1*). (C) List of sgRNAs selected for export.
user-specified genomes. Unfortunately, to the best of our knowledge, none of these generally applicable tools are provided as user-friendly web services, which can be used by non-computer scientists.

Here, we therefore present a web server implementation of the sgRNA prediction software CRISPy (http://crispy.secondarymetabolites.org/) that was originally designed for CRISPR-based engineering of Chinese hamster ovary (CHO) cells.26 It is implemented as a standalone web application for Cas9 target prediction that can design sgRNA for any gene/region of interest in user-submitted genome data. In addition, CRISPy-web tightly connects with the popular natural product prediction tool antiSMASH22–24 to assist researchers working on microbial secondary metabolites in leveraging the CRISPR/Cas9 system. Using antiSMASH or working on natural products of course is not a prerequisite for using CRISPy-web.

2. Design and implementation

CRISPy-web consists of multiple decoupled service components tied together by a message queue. The main services are the CRISPy worker service, the web application programming interface (API) server and the web user interface (UI). Sequences to be scanned for Cas9 target sites can be uploaded directly (in GenBank format) or transferred from antiSMASH22–24 by specifying the antiSMASH job ID. In the next version of antiSMASH, a direct link to select as target region. After a click on “Find targets”, the scan for sgRNAs starts.

Depending on the size of the selected region and the overall genome size, this sgRNA scan will take several minutes. After the scan is completed, a page graphically displays the scanned region (Fig. 1A). Genes are shown as gray arrows, sgRNAs on the forward strand are displayed as red boxes above the genes, sgRNAs on the reverse strand are displayed below the genes. A table below the graphical overview gives details on the identified sgRNAs, ranked by the least number of potential off-target matches. Hovering over the table row of an sgRNA hit highlights the corresponding sgRNA box in the graphics, so you can quickly locate the hit on the region.

As there are a lot of hits for this region, it is possible to zoom in on a particular gene of interest by clicking on the gene arrow and selecting “show results for this gene only”. Because the whole region already was scanned, this operation is instantaneous (Fig. 1B). On both the overview and on the zoomed view, sgRNAs you want to save for download can be added to the download basket by clicking the corresponding table row. The download basket button on the upper right will track the number of selected sgRNAs.

A click on the download basket icon will display the download overview page (Fig. 1C). There, the selected sgRNAs can be verified again. A click on the “Download CSV file” button will start the download of all selected sgRNAs in comma-separated value format. Files in this format can easily be used in a spreadsheet application, processed for downstream computational analysis or oligonucleotide orders.

Acknowledgment

This work was funded by a grant of the Novo Nordisk Foundation.

References

1. Horvath P, Barrangou R. CRISPR/Cas, the immune system of Bacteria and Archaea. Science 2010;327:167–70.
2. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 2014;346:1258096.
3. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 2013;31:233–9.
4. Tong Y, Charsantit P, Zhang L, Weber T, Lee SY. CRISPR-Cas9 based engineering of actinomycetal genomes. ACS Synth Biol 2015;4:1020–9.
5. Cobb RE, Wang Y, Zhao H. High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. ACS Synth Biol 2015;4:723–8.
6. Ronda C, Maury J, Jakociunas T, Jacobsen SA, Germann SM, Harrison SJ, et al. CeEdit: CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae. Microb Cell Fact 2013;14:97.
7. Jakociunas T, Bonde I, Herggard M, Harrison SJ, Kristensen M, Pedersen LE, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metall Eng 2015;28:213–22.
8. Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 2013;31:686–8.
9. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014;343:84–7.
10. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 2015;163(3):759–71.
11. Esvelt KM, Mali P, Balf JL, Moosburner M, young SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods 2013;10:1116–21.
12. Friedland AE, Baral R, Singhal P, Loveluck K, Shen S, Sanchez M, et al. Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adenovirus delivery and paired nickase applications. Genome Biol 2015;16:257.
13. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, et al. Discovery and functional characterization of diverse class 2 CRISPR-cas systems. Mol Cell 2015;60:385–97.
14. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012;337:813–18.
15. Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL. CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PLoS ONE 2015;10:e0124633.
16. 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:W401–7.
17. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 2013;31:827–32.
18. Wong N, Liu W, Wang X. Wu-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome Biol 2015;doi:10.1101/026971.
19. Hodgkins A, Farne A, Perera S, Grego T, Parry-Smith DJ, Skarnes WC, et al. WGE: a CRISPR database for genome engineering. Bioinformatics 2015;31:3078–80.
20. Ronda C, Pedersen LE, Hansen HG, Kalleshaug TB, Betenbaugh MJ, Nielsen AT, et al. Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISpy: a web-based target finding tool. Biotechnol Bioeng 2014;111:1604–16.
21. Xie S, Shen B, Zhang C, Huang X, Zhang Y. sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PLoS ONE 2014;9:e100448.
22. Weber T, Blin K, Duddela S, Krug D, Kim HJ, Bruccoleri R, et al. antiSMASH 3.0: a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 2015;43:W237–43.
23. Blin K, Medema MH, Kazempour D, Fischbach M, Breitling R, Takano E, et al. antiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res 2013;41:W204–12.
24. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, et al. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 2011;39:W339–46.
25. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 2009;25:1422–3.
26. Fielding RT, Taylor RN. Principled design of the modern Web architecture. ACM T Internet Techn 2002;2:115–50.