



CrisPam: SNP-Derived PAM Analysis Tool for Allele-Specific Targeting of Genetic Variants Using CRISPR-Cas Systems

Roy Rabinowitz^{1,2*}, Shiri Almog^{1,3}, Roy Darnell¹ and Daniel Offen^{1,2,3*}

OPEN ACCESS

Edited by:

Marko Djordjevic,
University of Belgrade,
Serbia

Reviewed by:

Mikhail Gelfand,
Institute for Information Transmission
Problems (RAS), Russia
Olga Musharova,
Skolkovo Institute of Science and
Technology, Russia

*Correspondence:

Roy Rabinowitz
royr2@mail.tau.ac.il
orcid.org/0000-0002-9113-3233

Daniel Offen
doffen@post.tau.ac.il

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 04 May 2020

Accepted: 13 July 2020

Published: 18 August 2020

Citation:

Rabinowitz R, Almog S,
Darnell R and Offen D (2020)
CrisPam: SNP-Derived PAM
Analysis Tool for Allele-Specific
Targeting of Genetic Variants Using
CRISPR-Cas Systems.
Front. Genet. 11:851.
doi: 10.3389/fgene.2020.00851

¹Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel, ²Felsenstein Medical Research Center, Tel Aviv University, Tel Aviv, Israel, ³Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

Clustered regularly interspaced short palindromic repeats (CRISPR) is a promising novel technology that holds the potential of treating genetic diseases. Safety and specificity of the treatment are to be further studied and developed prior to implementation of the technology into the clinic. The guide-RNA (gRNA) allows precise position-specific DNA targeting, although it may tolerate small changes such as point mutations. The permissive nature of the CRISPR-Cas system makes allele-specific targeting a challenging goal. Hence, an allele-specific targeting approach is in need for future treatments of heterozygous patients suffering from diseases caused by dominant negative mutations. The single-nucleotide polymorphism (SNP)-derived protospacer adjacent motif (PAM) approach allows highly allele-specific DNA cleavage due to the existence of a novel PAM sequence only at the target allele. Here, we present CrisPam, a computational tool that detects PAMs within the variant allele for allele-specific targeting by CRISPR-Cas systems. The algorithm scans the sequences and attempts to identify the generation of multiple PAMs for a given reference sequence and its variations. A successful result is such that at least a single PAM is generated by the variation nucleotide. Since the PAM is present within the variant allele only, the Cas enzyme will bind the variant allele exclusively. Analyzing a dataset of human pathogenic point mutations revealed that 90% of the analyzed mutations generated at least a single PAM. Thus, the SNP-derived PAM approach is ideal for targeting most of the point mutations in an allele-specific manner. CrisPam simplifies the gRNAs design process to specifically target the allele of interest and scans a wide range of 26 unique PAMs derived from 23 Cas enzymes.

CrisPam is freely available at <https://www.danioffenlab.com/crispam>.

Keywords: CrisPam, allele-specific, clustered regularly interspaced short palindromic repeats, single-nucleotide polymorphism-derived protospacer adjacent motif, guide RNA design

INTRODUCTION

The clustered regularly interspaced short palindromic repeats (CRISPR)-Cas system enables precise genome editing mediated by the guide RNA (gRNA), an RNA molecule that directs the CRISPR associated (Cas) protein to the target DNA within the genome. The Cas enzyme, the catalytic unit of the CRISPR-Cas system, generates a DNA double-strand break in the presence of a DNA:gRNA complementary and a protospacer-adjacent motif (PAM) in immediate proximity to the target DNA (Jinek et al., 2012; Cong et al., 2013). The diversity of Cas proteins, derived from distinct bacterial strains, differ in several properties such as PAM sequence, cleavage pattern and position, sequence length, efficacy in different organisms, off-targets, and substrate (DNA or RNA). The standard Cas protein has been modified to broaden its applications to base-editing (Komor et al., 2016; Gaudelli et al., 2017), transcription repression and activation (Bikard et al., 2013; Gilbert et al., 2014; Konermann et al., 2015), epigenomic modifications (Hilton et al., 2015), visualization of genomic loci (Chen et al., 2013), and DNA nicking (Ran et al., 2013; single-strand cleavage). In an experiment design, the PAM sequence and size of the designated Cas should be taken under consideration; presence of a PAM is a limiting step in targeting unique loci, and the Cas size determines the optional delivery systems (e.g., viral vectors).

SNP-DERIVED PAM

The CRISPR-Cas system can tolerate mismatches between the gRNA and the target DNA. The bases at positions 8–13 at the PAM-proximal end of the spacer (regarding type II Cas proteins) are termed the seed sequence along with the first base at the 5' end (Jinek et al., 2012; Cong et al., 2013; Hsu et al., 2013; Anderson et al., 2015; Doench et al., 2016). Mismatches within the seed sequence are thought to be intolerable and abolish DNA cleavage (Cong et al., 2013; Anderson et al., 2015). However, previous studies have shown that a single mismatch between the gRNA and the target DNA results in a non-specific knockdown of both the mutant and the wild-type alleles in some proportion (Hsu et al., 2013; Doench et al., 2016; Capon et al., 2017; Christie et al., 2017). Thus, the nature of specificity of the Cas enzyme is seemingly insufficient to specifically target single-nucleotide variations (SNVs). A single-nucleotide polymorphism (SNP)-derived PAM approach overcomes this potential limitation of targeting the disease-causing allele, while leaving the wild-type allele intact. This method dramatically increases the specificity of targeting the mutant allele alone by identifying a PAM sequence that is present only at the mutant sequence (György et al., 2019). Meaning, the SNV causes the formation of the PAM sequence, thus allows cleavage of the target allele and prevents the unintended cleavage of the reference allele (Wu et al., 2013; Courtney et al., 2016; Christie et al., 2017). Moreover, studies have shown that allele-specific targeting in heterozygous cells results in high rates of gene conversion (14–47%), where the homologous chromosome serves as a template for homologous recombination to resolve the double-strand break (Wu et al., 2013; Filler Hayut et al., 2017).

In such cases, targeting the pathogenic allele results in its correction by the wild-type allele. These findings emphasize the potential of pathogenic allele-specific targeting in heterozygous patients, suffering from haploinsufficiency or gain-of-function associated diseases. Here, we discuss the utility of the SNP-derived PAM approach and use the terminology as termed by Courtney et al. (2016), although we refer to SNVs in general regardless of their frequency in the population. Most Cas proteins may be appropriate candidates for targeting a gene without a location preference of DNA cleavage, mainly used for gene-knockout experiments. However, when targeting a SNV in general, or if utilizing the SNP-derived PAM approach in particular, the selection of Cas proteins is limited. This is mostly due to the condition of PAM presence in proximity to the SNV or having a PAM generated by the SNV. Current gRNA design tools are based on reference genomes and, therefore, are not suitable for SNV targeting (Bae et al., 2014; Stemmer et al., 2015; Doench et al., 2016; Haeussler et al., 2016; Labun et al., 2016). In a recent study, AlleleAnalyzer, a bioinformatic tool was reported to target SNPs by obtaining sequences data from the 1,000 Genomes project. By utilizing disease-associated haplotypes, AlleleAnalyzer designs allele-specific dual gRNAs (Keough et al., 2019). Scott and Zhang (2017) and Lessard et al. (2017) also developed bioinformatic tools to design gRNAs that target conserved genomic loci to avoid gRNA incompatibility due to genetic variations. Furthermore, both studies emphasize the need of common gRNA designs due to regulations and its consequential costs. Another recently published web tool, SNP-CRISPR, designs gRNAs for non-reference genomes to support allelic targeting (Chen et al., 2020). SNP-CRISPR calculates the gRNA efficiency score for the variant and the reference sequences. However, it does not support SNP-derived PAM targeting, nor Cas enzymes options other than SpCas9 with varied PAM sequences to expand the targeting scope. Manual design of SNP-derived PAM-based gRNAs is challenging due to the wide repertoire of available Cas enzymes, their different PAMs and the complexity of some PAMs allowing variable bases in certain positions (e.g., NNGRRT). To that end, we established CrisPam, a web-tool that identifies Cas enzymes capable of allele-specific targeting *via* SNP-derived PAM. CrisPam does not refer to a reference genome and supports multiple input methods to facilitate simple user experience. Experimental researchers attempting to specifically target a SNV, either for development of CRISPR-based therapeutics or disease modeling, will benefit the utility of CrisPam for successful allele-specific gRNA designs.

MATERIALS AND METHODS

Parsing and SNP Data Analysis

CrisPam is a pythonic code that performs sequence analysis of SNVs. The input of the code is the variant position (reference and variation nucleotides) and its flanking sequences (upstream and downstream). The code is available at <https://github.com/RoyRabinowitz/CrisPam>.

The dataset of human pathogenic SNPs was generated from dbSNP (build 153), incorporating data from UCSC Genome Browser and ClinVar, as previously described Rabinowitz et al. (2020).

Implementation

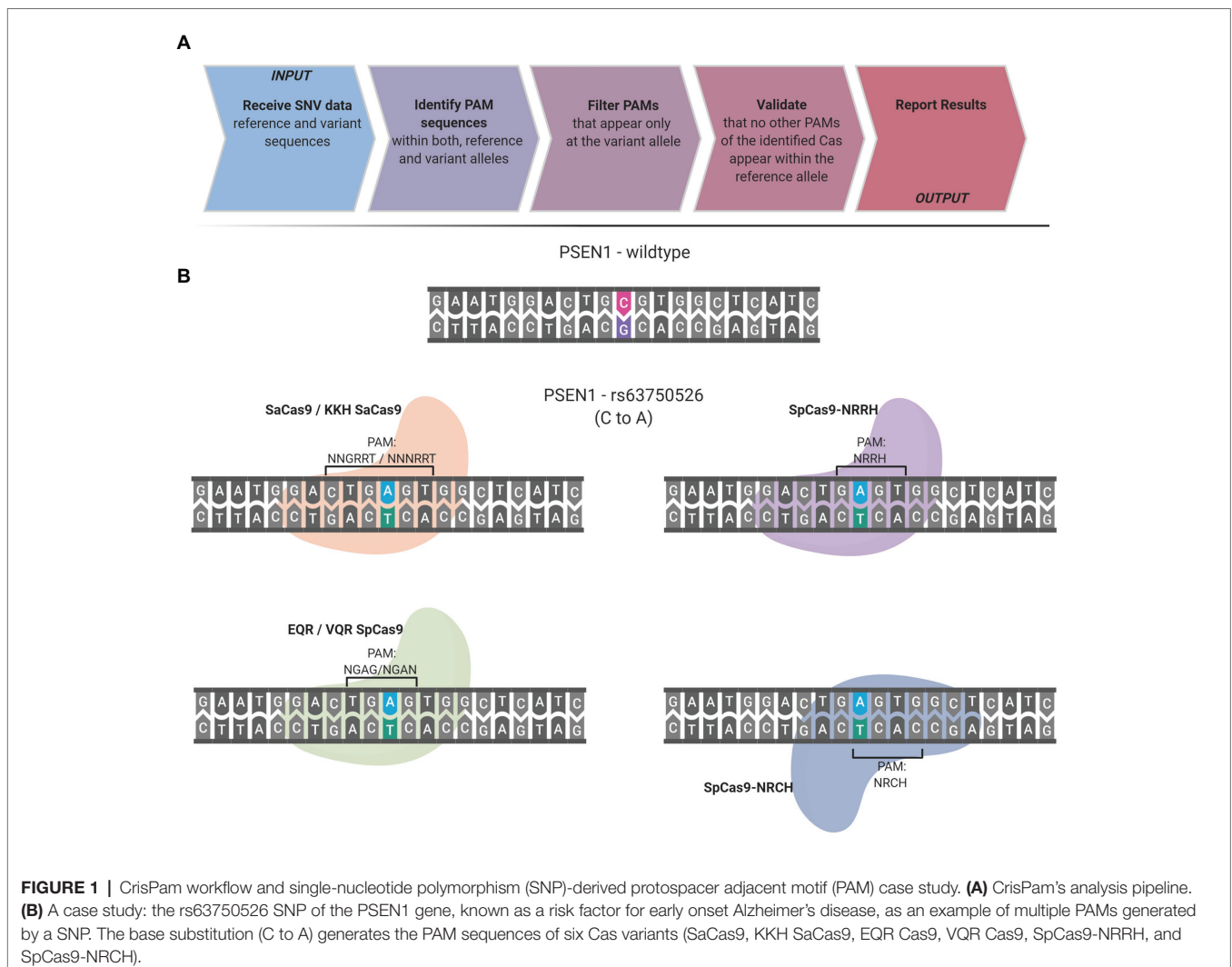
CrisPam is webserver-based tool. Thus, no software installation effort is required. The CrisPam DB is a.xlsx file and can be opened by Excel. The CrisPam script is written in Python 3.6 and uses standard libraries.

RESULTS

CrisPam

The presence of a PAM solely within the desired target allele is the guiding principle of the SNP-derived PAM methodology. The following workflow occurs to detect unique PAMs generated by a SNP: the code obtains the reference and variation bases and the flanking DNA sequences upstream and downstream to the SNV. The complementary strand is generated and analyzed to detect unique PAMs on the complementary strand as well. Some SNVs have more than a single variation nucleotide; each variation is handled individually. CrisPam attempts to detect PAMs on both the reference and the variant alleles and suggests

suitable Cas variants, in case their PAM is present exclusively within the variant allele. For a given SNV, more than one PAM may be generated; therefore, CrisPam reports all the matches for the query SNV (**Figure 1A**). The suggested gRNA sequence for each matching Cas is the 20-23 nt upstream or downstream to the PAM, according to the Cas ortholog. The diverse PAM sequences were defined according to previous studies that characterized the unique properties and PAM compatibility of each Cas. Several Cas enzymes recognize multiple sequences as legitimate PAMs (e.g., enAsCas12a recognizes TGTV, VTTV, TTTT, and TTCN). For a hypothetical point mutation TT<C>TA to TT<T>TA (C to T), enCas12a might be erroneously suggested as a possible candidate Cas since the point mutation generates its TTTT PAM. However, the reference allele contains an alternative PAM – TTCN of the same Cas; therefore, an unintended targeting of the reference allele will presumably occur. For multiple-PAMs Cas enzymes, CrisPam validates that the reference sequence is not targetable by any PAM of the suggested Cas. The well-studied SpCas9 is known to be recognizing mainly NGG PAMs; however, it may exhibit some degree of binding to target DNA sequences



with NAG PAMs (Hsu et al., 2013). While CrisPam considers SpCas9 as a potential candidate only when the NGG PAM is generated by the variant allele, NAG is defined as a potential off-target PAM. Therefore, SpCas9 would not be identified as a possible match in case the NAG PAM is present within the reference sequence. Currently, CrisPam identifies 26 PAM sequences of 23 Cas enzymes (Table 1).

A Database of PAM-Generating SNPs

To assess the potential of the SNP-derived PAM approach to treat human pathogenic point mutations, we used CrisPam to analyze

TABLE 1 | Cas diversity and properties.

Type	Cas	PAM sequence	Size (bp)	PAM orientation
	SpCas9 (Jinek et al., 2012)	NGG (NAG)		
	VREER SpCas9 (Kleinstiver et al., 2015b)	NGCG		
	EQR SpCas9 (Kleinstiver et al., 2015b)	NGAG		
	VQR SpCas9 (Kleinstiver et al., 2015b)	NGAN		
	SpCas9-NG (Nishimasu et al., 2018)	NG, NANG	≈4,100	
	xCas9 (Hu et al., 2018)	NG		
	SpCas9-NRRH (Miller et al., 2020)	NRRH, NGGN		
	SpCas9-NRTH (Miller et al., 2020)	NRTH, NGGN		
II	SpCas9-NRCH (Miller et al., 2020)	NRCH, NGGN		Downstream of the target DNA
	SaCas9 (Ran et al., 2015)	NNGRRT		
	KKH SaCas9 (Kleinstiver et al., 2015a)	NNNRRT	≈3,150	
	NmCas9 (Esvelt et al., 2013)	NNNNGATT	≈3,240	
	St1Cas9 (Xu et al., 2015; Agudelo et al., 2020)	NNRGAAW		
	St3Cas9 (Magadán et al., 2012)	NGGNG	≈4,180	
	TdCas9 (Sun et al., 2018)	NAAAAC	≈4,200	
	CjCas9 (Kim et al., 2017)	NNNVRYM	≈2,950	
	SpCas9 (from <i>Streptococcus pasteurianus</i> ; Tan et al., 2018)	NNGTGA	≈3,400	
	AsCas12a and LbCas12a (Zetsche et al., 2015)	TTTN		
	AsCas12a RVR (Gao et al., 2017)	TATV		
	AsCas12a RR (Gao et al., 2017)	TYCV	≈3,685	
V	enAsCas12a (Kleinstiver et al., 2019)	TGTV, VTTV, TTTT, TTCN		Upstream of the target DNA
	Fncas12a (Fonfara et al., 2016)	TTV	≈3,900	
	Cas12e (CasX; Burstein et al., 2017)	TTCN	≈2,940	

a total of 43,673 SNPs (pathogenic and likely-pathogenic SNPs) from dbSNP (Sherry et al., 1999). Successful matches of SNPs that generate at least a single PAM were found in 90% of the total SNPs (39,307 out of 43,673). **Figure 1B** represents a case study of a SNP (rs63750526 of the PSEN1 gene) that generates six PAMs. Such multiple-PAM generating SNPs confer the ability to opt the most suitable Cas depending on the application limitations such as vector size and on-target efficiency. We sought to assess the importance of synthetic Cas variants and their additive value to the variety of Cas enzymes. To that end, we compared the SNPs matching coverage of each Cas based on our analysis over the pathogenic SNPs dataset (**Figure 2**). Notably, synthetic variants account for a large proportion of the successful matches. In 6.65% of the analyzed SNPs (2,907 out of 43,673), a synthetic variant was reported to be solely Cas that is able to target the SNP in a SNP-derived PAM manner (**Figure 2B**). The full database (DB) of PAM-generating pathogenic and likely pathogenic SNPs is available as a Supplementary Material (**Supplementary Material S1**). Analyzing the data reveals that the 10% SNPs that were not targetable, share a certain type of base substitution: C→T or G→A (**Supplementary Material S2**). This can be rationalized by the abundance of G-rich and relaxed PAMs, compared to T-rich PAMs that tend to require more rigorous sequences.

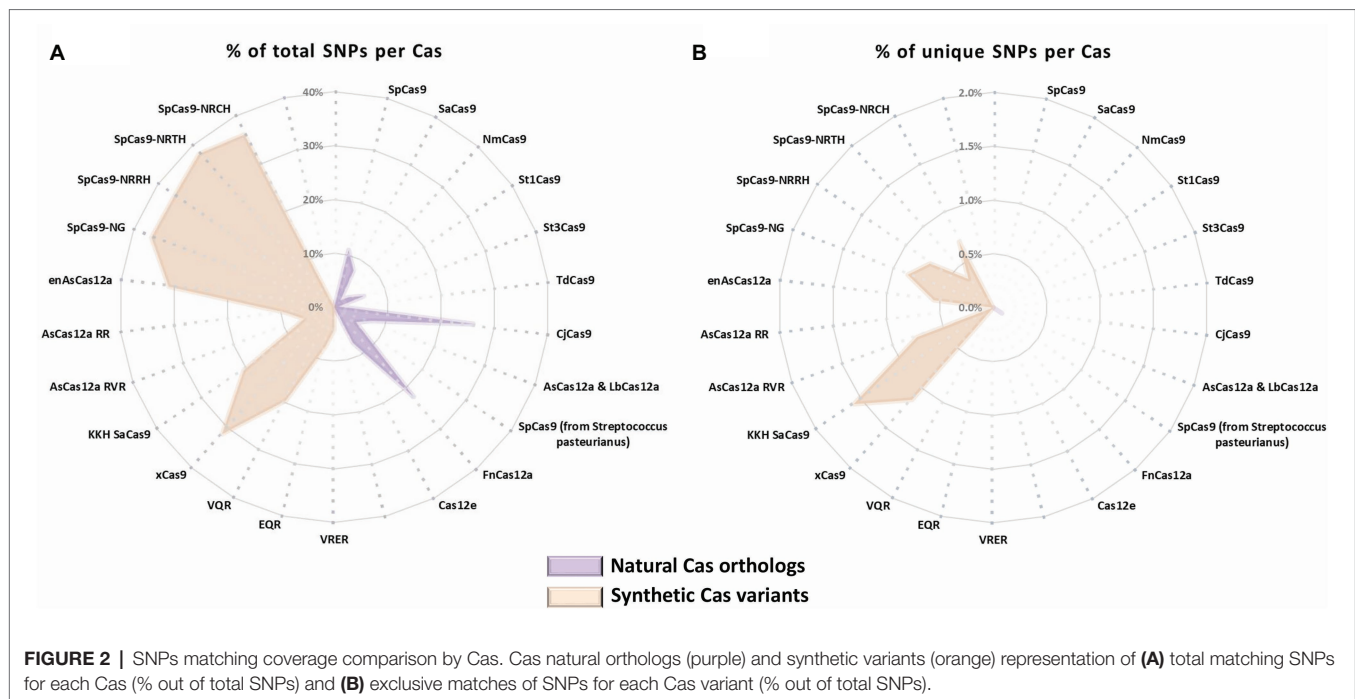
CrisPam – An Online SNP-Derived PAM Analysis Tool

We established a web tool that performs CrisPam's SNP-derived PAM targeting abilities on user data. Multiple input methods are available to allow analysis of SNVs that are yet to be reported and included in NCBI's dbSNP or non-pathogenic SNVs for research purposes. SNV data are accepted in one of the following methods: user manual input, fetch SNV data by rsID, fetch SNV by genomic coordinates (from 56 genomes), and a batch file that allows analysis of multiple SNVs. User-defined PAM is supported to enable researchers to identify suitable Cas proteins that are yet to be reported. The CrisPam web tool also supports single-nucleotide insertion and deletion variations. For NGG-based Cas variants, off-targets assessment is available via CRISTA (Abadi et al., 2017).

CrisPam is available at <https://www.danioffenlab.com/crispam>.

DISCUSSION

The SNP-derived PAM targeting approach for enhancing allele specificity is a promising method for future CRISPR-based novel genome-editing therapies. As many patients suffering from genetic conditions are heterozygous, carrying one copy of a pathogenic allele, it is essential to develop SNP customized treatments for increasing treatment's safety. Currently, CrisPam supports off-targets assessment only for NGG-based Cas enzymes. For non-NGG enzymes, it is suggested to use complementary tools to assess potential off-targets (Stemmer et al., 2015; Abadi et al., 2017; Hwang et al., 2018). We recommend using CrisPam as the first step in the pipeline of gRNA design. For multiple-PAM generating



SNPs, considerations such as delivery vector capacity [e.g., adeno-associated virus (AAV) or lentivirus] and efficiency may also define the most suitable Cas enzyme for the experiment. The webtool is ideal for researchers that intend to specifically target certain variations, while it may be less effective as a starting point for identifying target SNPs. Though, the CrisPam DB presents targetable SNPs and the relevant Cas enzymes to perform allele-specific targeting. Not all SNPs' clinical significance is commonly agreed or yet to be known and such SNPs may not be included in the dataset we used to generate the CrisPam DB. For that reason, it is suggested to use the CrisPam webtool if the SNP of interest does not appear in the CrisPam DB. Previous studies reported that genetic variations, in rare cases, may change the genomic target DNA and form mismatches between the target DNA and the gRNA or even disrupt the PAM site. Not only that genetic variations influence on-target activity, they may also cause altered potential off-targets sites. Therefore, for clinical purposes, it is essential to perform whole-genome sequencing for patients and validate on-target sequence integrity and detect potential personalized off-targets (Lessard et al., 2017; Scott and Zhang, 2017).

By utilizing CrisPam, we demonstrate the high compatibility of the SNP-derived PAM approach to preform allele-specific targeting using varied CRISPR-Cas systems. As demonstrated on human pathogenic SNPs data, 90% of the SNPs were found to have a SNP-derived PAM. Our findings underline the relevance of this allele-specific targeting approach in developing genome-editing therapeutic strategies. While CRISPR applications have been widely expanded, the SNP-derived PAM approach may be utilized for gene silencing (using inactive Cas), genetic screening, and more applications other

than allele-specific DNA cleavage. This study emphasizes the emerging importance of broadening PAM compatibility of Cas proteins to enable allele-specific targeting and overcome the PAM limitation. Furthermore, CrisPam offers a simple interface to design an allele-specific targeting experiment using the CRISPR-Cas system.

DATA AVAILABILITY STATEMENT

Project home page: <https://www.danioffenlab.com/crispam>. Programming language: Python 3.6. The script is available at <https://github.com/RoyRabinowitz/CrisPam>. The SNPs dataset that was used to generate the CrisPam DB for PAM-generating pathogenic \ likely-pathogenic SNPs was previously described in "Prediction of Synonymous Corrections by the BE-FF Computational Tool Expands the Targeting Scope of Base Editing" (Rabinowitz et al. 2020). The script is available at https://github.com/shiranab/parse_dbSNP. The obtained dataset of SNP-derived PAM targetable SNPs is available at the supplementary files (**sup. file S1**). License: Free for academic end-users solely for non-commercial research purposes. Any restrictions to use by non-academics: Contact Prof. Dani Offen or Ramot at Tel Aviv University LTD. The datasets presented in this study can be found in online repositories.

AUTHOR CONTRIBUTIONS

Conceived and designed the study, analyzed the data, and wrote the manuscript: RR. Programing: SA, RD, and RR. Webserver and user interface integration: SA. Principle

investigator: DO. All authors contributed to the article and approved the submitted version.

FUNDING

Funding for open access charge: Sackler School of Medicine, Tel Aviv University.

ACKNOWLEDGMENTS

RR is supported by PhD fellowships provided by the “Prais-Drimmer Institute for the Development of Anti-Degenerative

REFERENCES

- Abadi, S., Yan, W. X., Amar, D., and Mayrose, I. (2017). A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. *PLoS Comput. Biol.* 13:e1005807. doi: 10.1371/journal.pcbi.1005807
- Agudelo, D., Carter, S., Velimirovic, M., Durringer, A., Rivest, J. F., Levesque, S., et al. (2020). Versatile and robust genome editing with *Streptococcus thermophilus* CRISPR1-Cas9. *Genome Res.* 30, 107–117. doi: 10.1101/gr.255414.119
- Anderson, E. M., Haupt, A., Schiel, J. A., Chou, E., Machado, H. B., Strezoska, Ž., et al. (2015). Systematic analysis of CRISPR-Cas9 mismatch tolerance reveals low levels of off-target activity. *J. Biotechnol.* 211, 56–65. doi: 10.1016/j.jbiotec.2015.06.427
- Bae, S., Park, J., and Kim, J. -S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048
- Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41, 7429–7437. doi: 10.1093/nar/gkt520
- Burstein, D., Harrington, L. B., Strutt, S. C., Probst, A. J., Anantharaman, K., Thomas, B. C., et al. (2017). New CRISPR-Cas systems from uncultivated microbes. *Nature* 542, 237–241. doi: 10.1038/nature21059
- Capon, S. J., Baillie, G. J., Bower, N. L., da Silva, J. A., Paterson, S., Hogan, B. M., et al. (2017). Utilising polymorphisms to achieve allele-specific genome editing in zebrafish. *Biol. Open* 6, 125–131. doi: 10.1242/bio.020974
- Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G. -W., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479–1491. doi: 10.1016/j.cell.2013.12.001
- Chen, C. L., Rodiger, J., Chung, V., Viswanatha, R., Mohr, S. E., Hu, Y., et al. (2020). SNP-CRISPR: a web tool for SNP-specific genome editing. *G3 Genes, Genomes, Genet.* 10, 489–494. doi: 10.1534/g3.119.400904
- Christie, K. A., Courtney, D. G., DeDionisio, L. A., Shern, C. C., De Majumdar, S., Mairs, L. C., et al. (2017). Towards personalised allele-specific CRISPR gene editing to treat autosomal dominant disorders. *Sci. Rep.* 7:16174. doi: 10.1038/s41598-017-16279-4
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143
- Courtney, D. G., Moore, J. E., Atkinson, S. D., Maurizi, E., Allen, E. H. A., Pedrioli, D. M. L., et al. (2016). CRISPR/Cas9 DNA cleavage at SNP-derived PAM enables both in vitro and in vivo KRT12 mutation-specific targeting. *Gene Ther.* 23, 108–112. doi: 10.1038/gt.2015.82
- Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191. doi: 10.1038/nbt.3437
- Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J., and Church, G. M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–1121. doi: 10.1038/nmeth.2681
- Drugs” and the “Aufzien Family Center for the Prevention and Treatment of Parkinson’s Disease”. The authors wish to thank Shiran Abadi for her help and support in the generation of the SNPs dataset. This manuscript has been released as a pre-print at Research Square (Rabinowitz et al., 2019). **Figure 1** was created with BioRender.com.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00851/full#supplementary-material>

- Filler Hayut, S., Melamed Bessudo, C., and Levy, A. A. (2017). Targeted recombination between homologous chromosomes for precise breeding in tomato. *Nat. Commun.* 8:15605. doi: 10.1038/ncomms15605
- Fonfara, I., Richter, H., Bratović, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517–521. doi: 10.1038/nature17945
- Gao, L., Cox, D. B. T., Yan, W. X., Manteiga, J. C., Schneider, M. W., Yamano, T., et al. (2017). Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* 35, 789–792. doi: 10.1038/nbt.3900
- Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644
- Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159, 647–661. doi: 10.1016/j.cell.2014.09.029
- György, B., Nist-Lund, C., Pan, B., Asai, Y., Karavitaki, K. D., Kleinstiver, B. P., et al. (2019). Allele-specific gene editing prevents deafness in a model of dominant progressive hearing loss. *Nat. Med.* 25, 1123–1130. doi: 10.1038/s41591-019-0500-9
- Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J. -B., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 17:148. doi: 10.1186/s13059-016-1012-2
- Hilton, I. B., D’Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., et al. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510–517. doi: 10.1038/nbt.3199
- Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647
- Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155
- Hwang, G. H., Park, J., Lim, K., Kim, S., Yu, J., Yu, E., et al. (2018). Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics* 19:542. doi: 10.1186/s12859-018-2585-4
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829
- Keough, K. C., Lyalina, S., Olvera, M. P., Whalen, S., Conklin, B. R., and Pollard, K. S. (2019). AlleleAnalyzer: a tool for personalized and allele-specific sgRNA design. *Genome Biol.* 20:167. doi: 10.1186/s13059-019-1783-3
- Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H. -Y., et al. (2017). In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat. Commun.* 8:14500. doi: 10.1038/ncomms14500
- Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Topkar, V. V., Zheng, Z., et al. (2015a). Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* 33, 1293–1298. doi: 10.1038/nbt.3404

- Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., et al. (2015b). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485. doi: 10.1038/nature14592
- Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, J. E., Hsu, J. Y., Clement, K., et al. (2019). Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* 37, 276–282. doi: 10.1038/s41587-018-0011-0
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946
- Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588. doi: 10.1038/nature14136
- Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* 44, W272–W276. doi: 10.1093/nar/gkw398
- Lessard, S., Francioli, L., Alfoldi, J., Tardif, J. -C., Ellinor, P. T., MacArthur, D. G., et al. (2017). Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc. Natl. Acad. Sci. U. S. A.* 114, E11257–E11266. doi: 10.1073/pnas.1714640114
- Magadán, A. H., Dupuis, M. È., Villion, M., and Moineau, S. (2012). Cleavage of phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas system. *PLoS One* 7:e40913. doi: 10.1371/journal.pone.0040913
- Miller, S. M., Wang, T., Randolph, P. B., Arbab, M., Shen, M. W., Huang, T. P., et al. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* 38, 471–481. doi: 10.1038/s41587-020-0412-8
- Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262. doi: 10.1126/science.aas9129
- Rabinowitz, R., Abadi, S., Almog, S., and Offen, D. (2020). Prediction of synonymous corrections by the BE-FF computational tool expands the targeting scope of base editing. *Nucleic Acids Res.* 48, W340–W347. doi: 10.1093/nar/gkaa215
- Rabinowitz, R., Darnell, R., and Offen, D. (2019). CrisPam: SNP-derived PAM analysis web tool and human pathogenic SNPs database for CRISPR allele-specific targeting. *Res. Sq.* [Preprint]. doi: 10.21203/RS.2.9413/V1
- Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., et al. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191. doi: 10.1038/nature14299
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi: 10.1038/nprot.2013.143
- Scott, D. A., and Zhang, F. (2017). Implications of human genetic variation in CRISPR-based therapeutic genome editing. *Nat. Med.* 23, 1095–1101. doi: 10.1038/nm.4377
- Sherry, S. T., Ward, M., and Sirotkin, K. (1999). dbSNP—database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* 9, 677–679.
- Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J., and Mateo, J. L. (2015). CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* 10:e0124633. doi: 10.1371/journal.pone.0124633
- Sun, B., Yang, J., Ye, R. D., Jiang, Y., Chen, D., and Yang, S. (2018). A CRISPR-Cpf1-assisted non-homologous end joining genome editing system of *Mycobacterium smegmatis*. *Biotechnol. J.* 13:e1700588. doi: 10.1002/biot.201700588
- Tan, S. Z., Reisch, C. R., and Prather, K. L. J. (2018). A robust CRISPR interference gene repression system in *Pseudomonas*. *J. Bacteriol.* 200, e00575–e00517. doi: 10.1128/JB.00575-17
- Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., et al. (2013). Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* 13, 659–662. doi: 10.1016/j.stem.2013.10.016
- Xu, K., Ren, C., Liu, Z., Zhang, T., Zhang, T., Li, D., et al. (2015). Efficient genome engineering in eukaryotes using Cas9 from *Streptococcus thermophilus*. *Cell. Mol. Life Sci.* 72, 383–399. doi: 10.1007/s00018-014-1679-z
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Rabinowitz, Almog, Darnell and Offen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.