AsCRISPR: a web server for allele-specific sgRNA design in precision medicine

Guihu Zhao, Jinchen Li, Beisha Tang and Yu Tang*

Department of Geriatric Neurology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China; National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.

*Correspondence:
Dr. Yu Tang
tangyu-sky@163.com

Abstract

Allele-specific targeting by CRISPR provides a point of entry for personalized gene therapy of dominantly inherited diseases, by selectively disrupting the mutant alleles or disease-causing single nucleotide polymorphisms (SNPs), ideally while leaving normal alleles intact. Despite unprecedented specificity and tremendous therapeutic utility of allele-specific targeting by CRISPR, few bioinformatic tools have been implemented for the allele-specific purpose. We thus developed **AsCRISPR** (**A**llele-**s**pecific **CRISPR**), a web tool to aid the design of guide sequences that can discriminate between alleles. It can process with query sequences harboring single-base or short insertion-deletion (indel) mutations, as well as heterozygous SNPs deposited in the dbSNP database. Multiple CRISPR nucleases and their engineered variants including newly-developed Cas12b and CasX are included for users' choice. AsCRISPR provides the downloadable results of candidate guide sequences that may selectively target either allele. Meanwhile, AsCRISPR evaluates the on-target efficiencies, specificities and potential off-targets of those candidates, and also displays the restriction enzyme sites that might be disrupted upon successful genome edits. Other than designing allele-specific guide sequences for treating diseases, AsCRISPR could also be exploited to help studying the potential functions of genetic variants at certain gene loci by applying allele-specific editings. AsCRISPR is freely available at http://www.genemed.tech/ascrispr.

Keywords: Genomic Editing, CRISPR, Genetic Variants, SNP, Allele-specific, Precision Medicine

Introduction

Inherited diseases are caused by various types of mutations, insertions/deletions (indels), large genomic structural variations, as well as single nucleotide polymorphisms (SNPs) that are critical for personalized medicine. Of those, dominant inherited diseases present a special challenge for researchers to conduct gene therapies. Those patients inherited only one mutated allele and one normal allele on pairs of chromosomes. The treatment strategy typically involves an allele-specific manipulation by silencing or ablating the pathogenic alleles while exerting no aberrant effects on the wild-type ones. Previously, numerous studies used allele-specific short interference RNAs (siRNAs) to selectively suppress dominant mutant allele, and produced immense therapeutic benefits [1, 2]. Until recent years, allele-specific CRISPR genome editing has emerged to be a promising means to treat those dominant diseases. The CRISPR system provides a highly specific genome editing that is capable of discriminating disease-causing alleles from wild-type ones, whenever the genetic variants (1) generate unique protospacer adjacent motifs (PAMs); or (2) located within the spacer region, especially the seed region of the sgRNAs [3].

So far, allele-specific CRISPR has been increasingly employed in treating various diseases, including dominant diseases such as retinitis pigmentosa [4-7], corneal dystrophy [8] and dominant progressive hearing loss [9], as well as genome imprinting diseases [10]. And also it was used to alleviate haploinsufficiency by allele-specific CRISPR activation of wild-type alleles [11], and even was designed for manipulating HLA locus [12]. More excitingly, this strategy has been recently utilized to selectively inactivate mutant Huntington (HTT), taking advantage of novel PAMs created by SNPs flanking the HTT locus [13, 14]. Overall, allele-specific CRISPR is now believed to be a promising personalized strategy for treating genetic diseases.

However, it is always labor intensive and time consuming to figure out appropriate guide sequences that may discriminate between two alleles [8, 13]. Currently, most web servers only design sgRNAs from the reference genomes, without allele discriminations. Here we developed **AsCRISPR** (**A**llele-**s**pecific **CRISPR**), a web server to aid the design of sgRNAs, for allele-specific genome engineering. It has incorporated multiple CRISPR nucleases and can process flexibly with either input DNA sequences or heterozygous SNPs. To the best of our knowledge, AsCRISPR shows to be a valuable new resource for genome editing technologies that can design discriminating sgRNAs based on a short stretch of input sequences or SNP numbers, and involves newly-developed Cas nucleases such as Cas12b and CasX as well.

Implementation

AsCRISPR was developed using PHP and Perl on a Linux platform with an Apache web server. The front and back separation model was used; the front end is based on the Vue + Element, and the back end is based on the Laravel, a PHP web framework.

Single-base mutations, short indels, and SNP IDs are the formats for input (**Figure 1**). The SNP information was downloaded from dbSNP v150 database (https://www.ncbi.nlm.nih.gov/SNP) and stored in MySQL database. To optimize the SNP query performance, an index on SNP table was added. Sequence can be extracted from the .2bit file (hg19/GRCh37, hg38/GRCh38, or mm10/GRCm38) with the twoBitToFa command base on the SNP information (chromosome, start genomic position and end genomic position, reference allele and alternate allele). AsCRISPR displayed the SNP sites located at both flanking nucleotides of a query SNP ID, which was implemented using D3.

In principle, AsCRISPR proceeds to figure out if (i) query variants give rise to novel PAMs, which confers stringent allele-specific targeting, or (ii) query variants locate within the seed region of guide sequences, which may abolish the Cas cleavage (**Figure 1**). AsCRISPR then outputs the candidate guide sequences, after performing the stringent search and filter. For example, those guide sequences with novel PAMs generated by variants that constitute an ambiguous genetic code (such as R and Y in the CjCas9:NNNNRYAC), will be excluded.
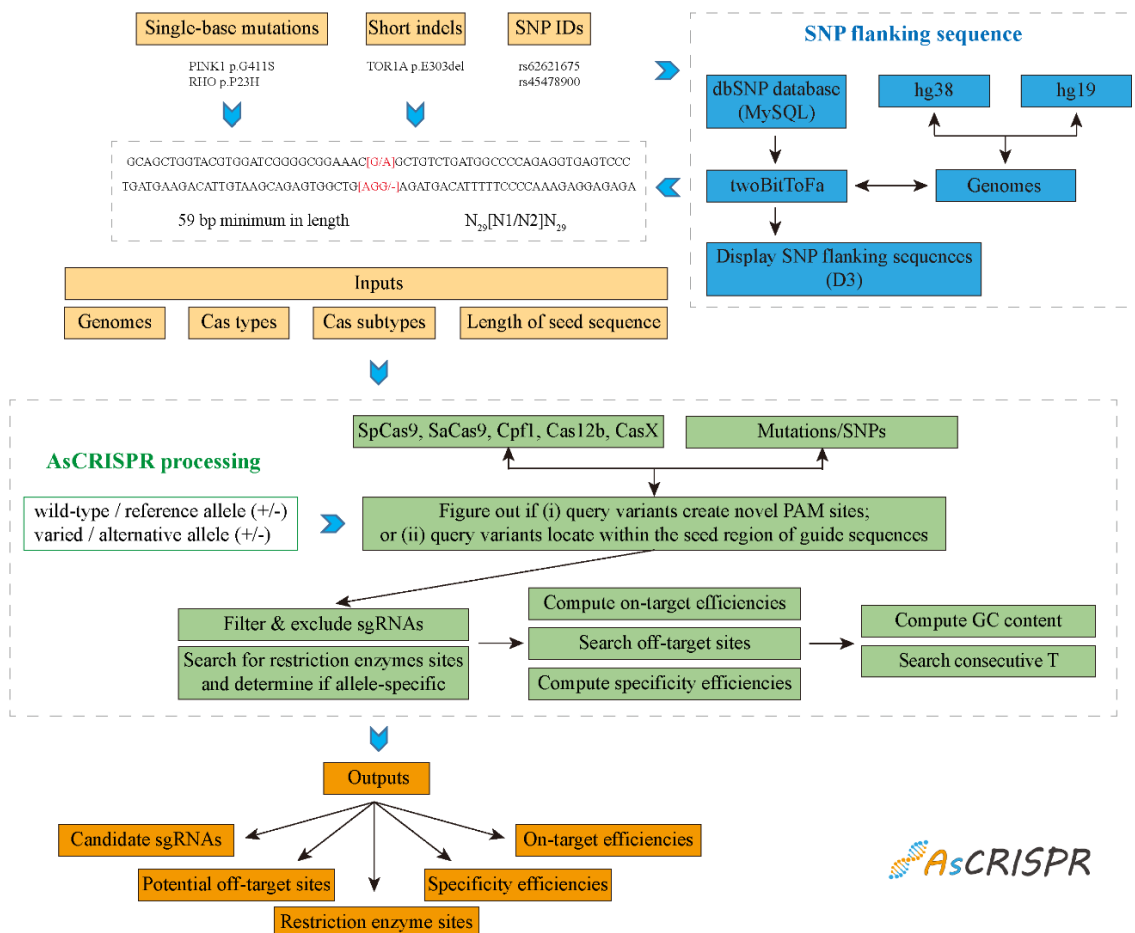

Figure 1. The workflow of AsCRISPR.

Scripts from CRISPOR (https://github.com/maximilianh/crisporPaper) were then integrated into AsCRISPR to assess sgRNA properties and scores. AsCRISPR also searched for possible sites recognized by restriction enzymes deposited in our database. In addition, guide sequences were further analyzed and reminded as "Not recommended" if (i) the GC contents are beyond 20%~80%; or (ii) containing four or more consecutive T, which might terminate the U6 or U3 promoter-drived transcription.
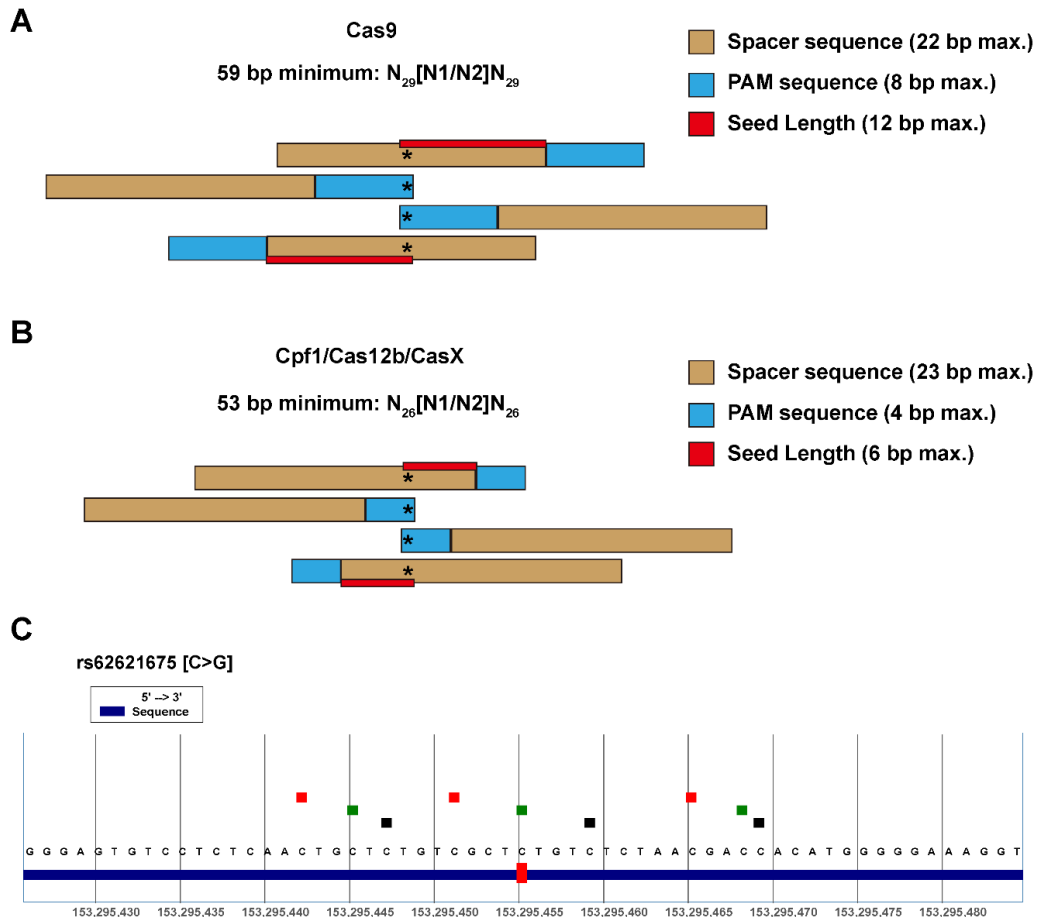

Figure 2. Input of query sequences. Demonstration of query DNA sequences for Cas9: 59 bp minimum **(A)** and Cpf1/Cas12b/CasX: 53 bp minimum **(B)**. To maximize users' demands, the length of input spacer sequence for Cas9 was set as 22 bp for query. Notably, AsCRISPR will only output 20 bp standard spacers for Cas9 instead. All in all, the input DNA sequence for AsCRISPR basically requires minimum 59 bp in length, with at least 29 bp flanking the mutation/variation site (*), to completely showcase the possible candidate sgRNAs. The format of input DNA sequences would be: $N_{29}[N1/N2]N_{29}$, in which N1 denotes the sequences in the wild-type allele and N2 denotes the sequences in the mutated or varied allele. **(C)** In query with an SNP ID, AsCRISPR will display other SNP sites located at both flanking 29 nucleotides.

Results

AsCRISPR helps to design sgRNAs based on four major types of Cas nucleases including the commonly used Cas9, Cpf1 and also recently reported Cas12b [15, 16] and CasX [17], each type of which contains its variant subtypes with distinct PAM sites and seed lengths (**Table 1**). This allows the users to freely choose the optimal combination of Cas protein and sgRNA to meet their own needs.

Input Format

The inputs for AsCRISPR could be DNA sequences harboring single-base mutations or short indels, and also simply SNP IDs deposited in the dbSNP database. All inputs will be finally processed as the format of $N_{29}[N1/N2]N_{29}$, in which N1/N2 denotes the sequence in the wild-type/reference and mutated/varied allele, respectively. Thus, it means that the input sequence requires 59 bp minimum in length, with at least 29 bp flanking the mutation/variation site, to be processed for the output of a complete list of candidate discriminating sgRNAs (**Figure 2A, B**). Notably, when users are in query with an SNP ID, AsCRISPR will also display other SNP sites located at both flanking 29 nucleotides (**Figure 2C**), which provides extra variation information and would be of great value for designing personalized gene targeting.

Candidate Guides

AsCRISPR provides downloadable results with candidate sgRNAs that target only one allele (**Figure 3**). For better visualization, AsCRISPR ranks all guide sequences first by listing pairs with the same PAM sequence back-to-back. Furthermore, AsCRISPR evaluates their on-target efficiencies, specificity efficiencies and potential off-targets throughout the genome, taking advantage of the CRISPOR's scoring system [18]. Specifically, the on-target efficiencies were calculated with multiple reported algorithms and were normalized to 0-1. For SpCas9, efficiency scores were predicted according to Xu et al., 2015 [19]; Doench et al., 2016 [20]; Moreno-Mateos et al., 2015 [21]; and Listgarden et al., 2018 [22], respectively. For SaCas9, efficiency scores were predicted according to Najm et al., 2018 [23]. And for Cpf1, Cas12b and CasX, efficiency scores were predicted according to Kim et al., 2018 [24].

Off-targets

The potential off-target sequences throughout the genome are searched by 3-base mismatches maximum (**Figure 3**). AsCRISPR lists the number of off-targets for each guide sequence with 0, 1, 2 or 3 mismatches (0-1-2-3). Clicking on the (0-1-2-3) will reveal more information about the off-target information in the downstream data sheet, including the locations (exon, intron or intergenic region), sequence mismatches and so forth. Users can freely re-rank the off-targets by locations. The specificity score measures the uniqueness of a guide sequence in the genome. The higher the specificity score, the lower are off-target effects. Specificity scores were calculated based on Hsu scores [25] and CFD scores [20].

For Cpf1, Cas12b and CasX, no off-target ranking algorithms were available in the literatures so far, instead we just applied Hsu and CFD scores to their off-targets.

Restriction Sites

AsCRISPR also searches for possible sites recognized by restriction enzymes along the spacer sequences (**Figure 3**), which might be disrupted after gene targeting, and further determines whether those candidate enzymes are also allele-specific. This provides an important tool for the characterization and screening of targeted single colonies by restriction fragment length polymorphism (RFLP).

Exemplary Running

We have listed several typical sequences on the website for exemplary running. For example, heterozygous *PINK1 p.G411S* is one of the ideal mutations for allele-specific targeting, which was previously demonstrated to increase the risk of Parkinson's disease via a dominant-negative mechanism [26]. In the Cas9 mode, AsCRISPR outputs 11 discriminating sgRNAs in combination with 3 subtypes of Cas9 including SpCas9, SpCas9-V(R)QR and SaCas9-KKH (**Table 2**). One of those sgRNA exploits a novel PAM (5'-CgG-3') created by the mutation, and another 5 pairs of sgRNAs containing the mutation point within the seed region selectively target either wild-type allele or mutated allele (**Table 2**). Therefore, by using Cas9, totally 5 candidate sgRNAs might be specific to the mutated *PINK1 p.G411S* allele, which are ready for the users' experimental evaluations (**Table 2**). Besides, we also listed other exemplary mutations including the single mutations (*TGFBI p.L527R; RHO p.P23H; LMNA p.G608G*), 3-base delete mutations (*TOR1A p.E303del*) and short indel mutations (*COL7A1 c.8068_8084delinsGA*).

Similarly, for heterozygous SNPs, AsCRISPR processes with the input SNP numbers and translates them into DNA sequences (59 bp) after retrieving the genomic database. As an example, we used AsCRISPR to analyze one of the SNPs, rs62621675:[C>G], only with Cas9, and successfully obtained 13 discriminating sgRNAs in combination with 4 subtypes of Cas9 including SpCas9-V(R)QR, SpCas9-EQR, SpCas9-VRER and CjCas9 (**Table 2**). Three of those sgRNA exploits novel PAMs (5'-AgA-3'; 5'-AgAG-3'; 5'-AGAGACAc-3') created by the SNP, and another 5 pairs of sgRNAs containing the variant point within the seed sequence (**Table 2**).

Users can freely select the candidate sgRNAs by Cas types, on-target efficiencies, specificity scores, off-targets properties, and others. We recommend to select the sgRNAs with novel PAMs, since they contribute to the most stringent discrimination. For a more detailed demonstration, users can also find a AsCRISPR tutorial on the website, which can be read online or downloaded as a pdf document.

Figure 3. Snapshot of the graphic user interface of AsCRISPR. (A) Search page. The AsCRISPR homepage allows two types of input (DNA sequence or SNP IDs) with default parameters optimized for novice users. **(B)** Result pages including the output of guide sequences, off-targets information and restriction enzyme information.

Future Developments

So far, AsCRISPR has integrated the genomes of Homo sapiens (hg19/GRCh37), Homo sapiens (hg38/GRCh38) and Mus musculus (mm10/GRCm38). We are planning to upload more genomes for analysis in the near future, to expand its allele-specific utilities. Notably, our understanding of on- and off-target sgRNA efficiencies is evolving rapidly. Although the on-target efficiencies in AsCRISPR were calculated with multiple reported algorithms, the scoring algorithms have been continuously improved. Cas12b and CasX may have their special efficiency scoring algorithms that are different from those of Cas9 and Cpf1, however, to our best knowledge, there are still lack of published studies working on it. We will thus incorporate the convincing scoring algorithms, which predict either on- or off-target efficiencies, into AsCRISPR as they became available. We also welcome any constructive feedback from users for improving our web server.

Conclusion and Discussion

We have thus developed AsCRISPR, which is an easy-to-use and streamlined web tool for designing potential discriminating sgRNAs between alleles to facilitate the CRISPR-based personalized therapy. Particularly, we incorporated two recently reported types of Cas nucleases, Cas12b and CasX, which show to be promising for genetic engineering due to their smaller size and higher specificity.

As we just finished the AsCRISPR implementation, another software termed AlleleAnalyzer was published, aiming to identify optimized pairs of personalized and allele-specific sgRNAs [27]. AlleleAnalyzer also leverages patterns of shared genetic variation across thousands of publicly available genomes to design sgRNA pairs that will have the greatest utility in large populations [27].

However, the difference is that AsCRISPR as a web tool can process either query sequences or SNP numbers, which is more likely demand-driven for research studies and clinical therapeutics. Moreover, AsCRISPR only outputs single sgRNAs instead of pairs of sgRNAs, although the users may also freely use AsCRISPR to manually design another sgRNA to make an sgRNA pair. Note that numerous non-coding RNAs or regulatory elements are widespread, dual-sgRNA excision of a large DNA fragment might bring about extra risks for disease treatment. Thus, we believe that AsCRISPR possesses extra allele-specific utilities and add to the bioinformatic repositories for allele-specific genomic editings.

Interestingly, people have been avoiding genetic variants when designing sgRNAs for therapeutic genome editing in large populations. Previous studies have performed comprehensive analysis on the Exome Aggregation Consortium (ExAC) and 1000 Genomes Project (1000GP) data sets, and determined that genetic variants could negatively impact sgRNA efficiency, as well as both on- and off-target specificity at therapeutically

implicated loci [28, 29]. Thus, for the CRISPR-based therapy in large patient populations, genetic variations should be considered in the design and evaluation of sgRNAs to minimize the risk of treatment failure and/or adverse outcomes. To address that, people thus endeavor to identify universal/platinum sgRNAs located in the low-variation regions, with the help of, for example, the ExAC browser, to maximize their population efficacy [28].

Although the genetic variations would be a challenge for platinum sgRNA design, it provides a promising entry for designing allele-specific or personalized sgRNAs in treating individual patients. Deciphering genetic variations helps to seek common platinum sgRNAs for the treatment in large populations, whereas AsCRISPR, as well as AlleleAnalyzer, go towards the opposite direction that exploits the discrimination abilities of heterozygous genetic variants to facilitate the design of allele-specific targets for individuals in the era of precision medicine.

Authors' contributions

YT conceived of, designed, and directed the study. GZ wrote the scripts and implemented the website. YT and GZ wrote the paper. All authors read and approved the final manuscript.

Competing interests

The authors have declared no competing interests.

Acknowledgments

References

1. Monga, I., et al., *ASPsiRNA: A Resource of ASP-siRNAs Having Therapeutic Potential for Human Genetic Disorders and Algorithm for Prediction of Their Inhibitory Efficacy.* G3 (Bethesda), 2017. **7**(9): p. 2931-2943.

2. Hohjoh, H., *Disease-causing allele-specific silencing by RNA interference.* Pharmaceuticals (Basel), 2013. **6**(4): p. 522-35.

3. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.

4. Burnight, E.R., et al., *Using CRISPR-Cas9 to Generate Gene-Corrected Autologous iPSCs for the Treatment of Inherited Retinal Degeneration.* Mol Ther, 2017. **25**(9): p. 1999-2013.

5. Giannelli, S.G., et al., *Cas9/sgRNA selective targeting of the P23H Rhodopsin mutant allele for treating retinitis pigmentosa by intravitreal AAV9.PHP.B-based*

delivery. Hum Mol Genet, 2018. **27**(5): p. 761-779.

6. Li, P., et al., *Allele-Specific CRISPR-Cas9 Genome Editing of the Single-Base P23H Mutation for Rhodopsin-Associated Dominant Retinitis Pigmentosa.* CRISPR J, 2018. **1**: p. 55-64.

7. Bakondi, B., et al., *In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa.* Mol Ther, 2016. **24**(3): p. 556-63.

8. Christie, K.A., et al., *Towards personalised allele-specific CRISPR gene editing to treat autosomal dominant disorders.* Sci Rep, 2017. **7**(1): p. 16174.

9. Gyorgy, B., et al., *Allele-specific gene editing prevents deafness in a model of dominant progressive hearing loss.* Nat Med, 2019. **25**(7): p. 1123-1130.

10. Bashtrykov, P., G. Kungulovski, and A. Jeltsch, *Correction of aberrant imprinting by allele-specific epigenome editing.* Clin Pharmacol Ther, 2016. **99**(5): p. 482-4.

11. Matharu, N., et al., *CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency.* Science, 2019. **363**(6424).

12. Xu, H., et al., *Targeted Disruption of HLA Genes via CRISPR-Cas9 Generates iPSCs with Enhanced Immune Compatibility.* Cell Stem Cell, 2019. **24**(4): p. 566-578 e7.

13. Monteys, A.M., et al., *CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo.* Mol Ther, 2017. **25**(1): p. 12-23.

14. Shin, J.W., et al., *Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9.* Hum Mol Genet, 2016. **25**(20): p. 4566-4576.

15. Strecker, J., et al., *Engineering of CRISPR-Cas12b for human genome editing.* Nat Commun, 2019. **10**(1): p. 212.

16. Teng, F., et al., *Repurposing CRISPR-Cas12b for mammalian genome engineering.* Cell Discov, 2018. **4**: p. 63.

17. Liu, J.J., et al., *CasX enzymes comprise a distinct family of RNA-guided genome editors.* Nature, 2019. **566**(7743): p. 218-223.

18. Haeussler, M., et al., *Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR.* Genome Biol, 2016. **17**(1): p. 148.

19. Xu, H., et al., *Sequence determinants of improved CRISPR sgRNA design.* Genome Res, 2015. **25**(8): p. 1147-57.

20. Doench, J.G., et al., *Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9.* Nat Biotechnol, 2016. **34**(2): p. 184-191.

21. Moreno-Mateos, M.A., et al., *CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo.* Nat Methods, 2015. **12**(10): p. 982-8.

22. Listgarten, J., et al., *Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs.* Nat Biomed Eng, 2018. **2**(1): p. 38-47.

23. Najm, F.J., et al., *Orthologous CRISPR-Cas9 enzymes for combinatorial genetic*

screens. Nat Biotechnol, 2018. **36**(2): p. 179-189.

24.   Kim, H.K., et al., *Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity.* Nat Biotechnol, 2018. **36**(3): p. 239-241.

25.   Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nat Biotechnol, 2013. **31**(9): p. 827-32.

26.   Puschmann, A., et al., *Heterozygous PINK1 p.G411S increases risk of Parkinson's disease via a dominant-negative mechanism.* Brain, 2017. **140**(1): p. 98-117.

27.   Keough, K.C., et al., *AlleleAnalyzer: a tool for personalized and allele-specific sgRNA design.* Genome Biol, 2019. **20**(1): p. 167.

28.   Scott, D.A. and F. Zhang, *Implications of human genetic variation in CRISPR-based therapeutic genome editing.* Nat Med, 2017. **23**(9): p. 1095-1101.

29.   Lessard, S., et al., *Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci.* Proc Natl Acad Sci U S A, 2017. **114**(52): p. E11257-E11266.

Table 1. Cas nucleases and their variants used in AsCRISPR

Cas & Variants	Species	PAM Sequence	PAM Location	Max. Seed Length	Output Spacer Length
Cas9 & Variants					
SpCas9	*Streptococcus pyogenes*	**NGG**	3' end	12	20
SpCas9-VRER	*Streptococcus pyogenes*	**NGCG**	3' end	12	20
SpCas9-EQR	*Streptococcus pyogenes*	**NGAG**	3' end	12	20
SpCas9-VQR	*Streptococcus pyogenes*	**NGA**	3' end	12	20
SpCas9-VRQR	*Streptococcus pyogenes*	**NGA**	3' end	12	20
SaCas9	*Staphylococcus aureus*	**NNGRRT**	3' end	12	20
SaCas9-KKH	*Staphylococcus aureus*	**NNNRRT**	3' end	12	20
NmCas9	*Neisseria meningitidis*	**NNNNGATT**	3' end	12	20
StCas9	*Streptococcus thermophilus*	**NNAGAAW**	3' end	12	20
TdCas9	*Treponema denticola*	**NAAAAC**	3' end	12	20
CjCas9	*Campylobacter jejuni*	**NNNNRYAC**	3' end	12	20
Cpf1 & Variants					
AsCpf1	*Acidaminococcus sp. BV3L6*	**TTTV**	5' end	6	23
LbCpf1	*Lachnospiraceae bacterium ND2006*	**TTTV**	5' end	6	23
FnCpf1	*Francisella novicida U112*	**TTV**	5' end	5	21
FnCpf1	*Francisella novicida U112*	**KYTV**	5' end	5	21
AsCpf1-RR	*Acidaminococcus sp. BV3L6*	**TYCV**	5' end	6	23
AsCpf1-RVR	*Acidaminococcus sp. BV3L6*	**TATV**	5' end	6	23
Cas12b & Variants					
AaCas12b	*Alicyclobacillus acidiphilus*	**TTN**	5' end	6	23
AkCas12b	*Alicyclobacillus kakegawensis*	**TTTN**	5' end	6	23
BhCas12b v4	*Bacillus hisashii*	**ATTN**	5' end	6	23
CasX & Variants					

DpbCasX	*Deltaproteobacteria*	**TTCN**	5' end	6	20
PlmCasX	*Planctomycetes*	**TTCN**	5' end	6	20

Table 2. A partial list of potential discriminating sgRNAs identified by AsCRISPR in Cas9 mode

Spacer Sequence	PAM	Cas9 Type	Targeting Strand	Direction	GC (%)	Specificity Score	Doench Score	Najm Score	Self-complementarity	Off-targets 0-1-2-3
Inquiry sequence: Human PINK1 p.G411S Mutation GCAGCTGGTACGTGGATCGGGGCGGAAAC[G/A]GCTGTCTGATGGCCCCAGAGGTGAGTCCC										
CATCAGACAGCcGTTTCCGC[#]	CCCGAT	SaCas9-KKH	Wild-type	-	60	1	--	0.64	1	0-0-0-4
CATCAGACAGCtGTTTCCGC[#]	CCCGAT	SaCas9-KKH	Varied	-	55	0.96	--	0.64	0	0-1-1-10[$]
TCAGACAGCcGTTTCCGCCC[#]	CGA	SpCas9-V(R)QR	Wild-type	-	65	0.97	0.51	--	1	0-0-1-2
TCAGACAGCtGTTTCCGCCC[#]	CGA	SpCas9-V(R)QR	Varied	-	60	0.5	0.51	--	0	0-1-0-7[$]
TACGTGGATCGGGGCGGAAA	CgG[#]	SpCas9	Wild-type	+	60	1	0.48	--	1	0-0-0-1
GATCGGGGCGGAAACgGCTG[#]	TCTGAT	SaCas9-KKH	Wild-type	+	70	1	--	0.58	2	0-0-0-2
GATCGGGGCGGAAACaGCTG[#]	TCTGAT	SaCas9-KKH	Varied	+	65	0.94	--	0.58	1	0-1-0-2[$]
TCGGGGCGGAAACgGCTGTC[#]	TGA	SpCas9-V(R)QR	Wild-type	+	70	0.99	0.4	--	2	0-0-0-1
TCGGGGCGGAAACaGCTGTC[#]	TGA	SpCas9-V(R)QR	Varied	+	65	0.84	0.4	--	1	0-1-0-3[$]
GGGCGGAAACgGCTGTCTGA[#]	TGG	SpCas9	Wild-type	+	65	0.94	0.5	--	2	0-0-0-10
GGGCGGAAACaGCTGTCTGA[#]	TGG	SpCas9	Varied	+	60	0.61	0.5	--	1	0-1-0-18[$]
Inquiry sequence: rs62621675:[C>G] GGGAGTGTCCTCTCAACTGCTCTGTCGCT[C/G]TGTCTCTAACGACCACATGGGGGAAAGGT										
Spacer Sequence	**PAM**	**Cas9 Type**	**Targeting Strand**	**Direction**	**GC (%)**	**Specificity Score**	**Doench Score**	**Najm Score**	**Self-complementarity**	**Off-targets 0-1-2-3**
CCCATGTGGTCGTTAGAGAC	AgA[#]	SpCas9-V(R)QR	Reference	-	55	0.99	0.52	--	0	0-0-0-3
GTCGTTAGAGACAgAGCGAC[#]	AGA	SpCas9-V(R)QR	Reference	-	55	0.96	0.46	--	1	0-0-0-6
GTCGTTAGAGACAcAGCGAC[#]	AGA	SpCas9-V(R)QR	Alternative	-	55	0.86	0.46	--	1	0-1-0-1[$]
CCCATGTGGTCGTTAGAGAC	AgAG[#]	SpCas9-EQR	Reference	-	55	0.99	0.52	--	0	0-0-0-3

GTCGTTAGAGACAgAGCGAC[#]	AGAG	SpCas9-EQR	Reference	-	55	0.97	0.46	--	1	0-0-0-5
GTCGTTAGAGACAcAGCGAC[#]	AGAG	SpCas9-EQR	Alternative	-	55	0.87	0.46	--	1	0-1-0-0[$]
CTTTCCCCCATGTGGTCGTT	AGAGACAc[#]	CjCas9	Alternative	-	55	0	0.43	--	0	0-0-0-0
CATGTGGTCGTTAGAGACAg[#]	AGCG	SpCas9-VRER	Reference	-	50	1	0.69	--	0	0-0-0-0
CATGTGGTCGTTAGAGACAc[#]	AGCG	SpCas9-VRER	Alternative	-	50	0.71	0.69	--	0	0-1-0-0[$]
TGTGGTCGTTAGAGACAgAG[#]	CGA	SpCas9-V(R)QR	Reference	-	50	0.8	0.74	--	0	0-0-2-20
TGTGGTCGTTAGAGACAcAG[#]	CGA	SpCas9-V(R)QR	Alternative	-	50	0.75	0.74	--	1	0-1-0-14[$]
CTCTGTCGCTcTGTCTCTAA[#]	CGA	SpCas9-V(R)QR	Reference	+	50	0.68	0.47	--	0	0-0-4-42
CTCTGTCGCTgTGTCTCTAA[#]	CGA	SpCas9-V(R)QR	Alternative	+	50	0.61	0.47	--	0	0-1-0-17[$]

[#] Mutations/SNPs are labeled in red and turned into lowercase.

[$] Note: When processing with the guide sequences from varied/alternative alleles, it is highly possible that one of those off-targets (containing 1-base mismatch labeled with [$]) would be exactly the sequence of their corresponding wild-type/reference alleles. This will compromise the specificity scores of guide sequences from varied/alternative alleles to some extent.