Interdisciplinary Sciences: Computational Life Sciences (2018) 10:455–465
https://doi.org/10.1007/s12539-018-0298-z

REVIEW

Review of CRISPR/Cas9 sgRNA Design Tools

Yingbo Cui[1] · Jiaming Xu[1] · Minxia Cheng[2] · Xiangke Liao[1] · Shaoliang Peng[1,2,3]

Received: 6 March 2018 / Revised: 2 April 2018 / Accepted: 4 April 2018 / Published online: 11 April 2018

Abstract
The adaptive immunity system in bacteria and archaea, Clustered Regularly Interspaced Short Palindromic Repeats, CRISPR-associate (CRISPR/Cas), has been adapted as a powerful gene editing tool and got a broad application in genome research field due to its ease of use and cost-effectiveness. The performance of CRISPR/Cas relies on well-designed single-guide RNA (sgRNA), so a lot of bioinformatic tools have been developed to assist the design of highly active and specific sgRNA. These tools vary in design specifications, parameters, genomes and so on. To help researchers to choose their proper tools, we reviewed various sgRNA design tools, mainly focusing on their on-target efficiency prediction model and off-target detection algorithm.

Keywords CRISPR · CRISPR/Cas9 · SgRNA design · On-target efficiency · Off-target detection

1 Introduction

CRISPR/Cas is the adaptive immunity system in many bacteria and most archaea [1–3], which provides immunity against viruses and plasmids. In this system, the DNA of invaded virus or plasmid will be cleaved into novel spacer and stored in an array in DNA. When the same virus or plasmid invades again, the corresponding invading DNA will be recognized and interfered [2].

In 2012, researchers ported the CRISPR/Cas system to genome editing and explain the basic mechanism of CRISPR/Cas9 [4, 5]. CRISPR/Cas9 consists of two parts: sgRNA and Cas9 endonuclease. The two components form a complex to cleave target DNA sites, as shown in Fig. 1. sgRNA is derived from a fusion of the tracrRNA (trans-activating crRNA) and crRNA (CRISPR RNA) [2, 4]. From the engineering perspective, sgRNA consists of two parts: constant part (purple in Fig. 1), which forms a scaffold by several stem-loop for Cas9 binding, and a 5′-end 20-nt altered part (pink in Fig. 1) complementary to target DNA sequence, which is programmable to target different DNA sites [6]. Cas9 is nuclease derived from *Streptococcus pyogenes* (*S. pyogenes*) [7]. The target site in DNA mainly contains two parts: protospacer (black in Fig. 1) complementary to the 5′-end 20-nt sequence in sgRNA and protospacer adjacent motif (PAM, yellow in Fig. 1) bound by Cas9. PAM is usually short (for SpCas9 is 5′-NGG-3′) and directly adjacent to protospacer. The Cas9 protein will not cleave sequence without the presence of PAM.

If the protospacer pairs with the 5′-end 20-nt sequence, and the Cas9 binds with PAM, a double-strand break (DSB) will be made. After DSB, DNA repair machinery will start and catalyze non-homologous end joining (NHEJ) or homology-directed repair (HDR), as shown in Fig. 1. For NHEJ, the lost sequence may cannot be recovered and results in sequence insertions or deletions (indels), which produce gene loss of function, while for HDR, an introduced exogenous DNA template will fill the gap from DSB [8].

Except for gene loss of function (CRISPRko) by making a DSB, CRISPR/Cas9 system can also repress or boost the expression of specific gene [9–11]. Gene repression (CRISPRi) can be achieved by binding a catalytically dead Cas9 lacking cleaving activity to the transcription factor binding sites [9], and gene activation (CRISPRa) can be realized by

Yingbo Cui and Jiaming Xu are equal contributors.

✉ Shaoliang Peng
pengshaoliang@nudt.edu.cn

[1] College of Computer, National University of Defense Technology, Changsha 410073, China

[2] College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China

[3] National Supercomputing Center in Changsha, Changsha 410082, China

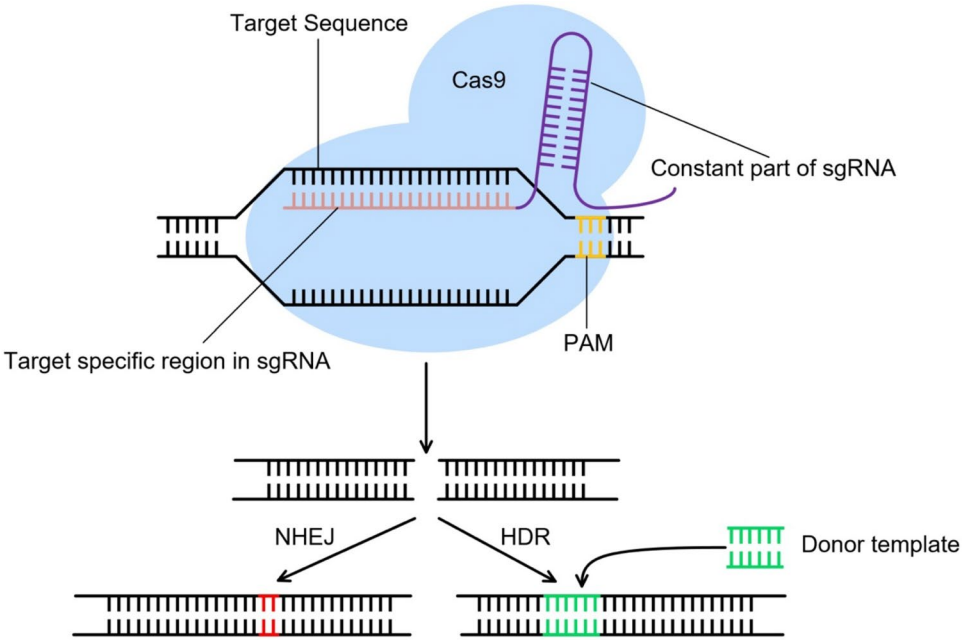

Fig. 1 CRISPR/Cas9 genome editing system

the fusion of an inactive Cas9 to the transcriptional activation domain [10, 11].

Compared with previous gene editing technologies, zinc-finger nucleases (ZFNs) [12–18] and transcription activator-like effector nucleases (TALENs) [19–22], which bind to specific DNA sequence by protein-DNA recognition, CRISPR/Cas9 system identifies specific site by the complementarity between the 20-nt sequence at the 5′-end of sgRNA and the protospacer in DNA, which is easier and cheaper to implemented in engineering. Besides, some studies show that CRISPR/Cas9 is more efficient than ZFNs and TALENs in gene editing [23–25]. Due to these advantages, it has been used in many species and mammalian cells [26–34]. In addition, some researchers have used the CRISPR/Cas9 system for genome-wide genetic screens in vitro and in vivo [35–43].

Currently, the large amount of sgRNA design tools vary in design specifications, parameters, genomes and so on [44, 45]. To help researchers choose the most suitable sgRNA design tools for their experiments, we review current sgRNA design tools, mainly focusing on the on-target efficiency prediction models and off-target prediction algorithms.

2 Overview of sgRNA Design Tools

The major work of sgRNA design is to find target sites in genome, which is pretty easy to implement by scanning PAM sequence (like 5′-NGG-3′ for SpCas9). However, several challenges exist for designing a good sgRNA: efficacy and specificity.

In theory, if the 5′-end 20-nt sequence in sgRNA is complementary to the target DNA sequence, the sgRNA-Cas9 complex should bind to the site and make a cleavage, but in practice, some studies suggest that the cutting efficiency varies significantly among sgRNAs [5, 26, 27, 33, 35–37, 46–60]. Predictive models to identify target sites with high efficiency are necessary.

Another challenge is off-target activity of sgRNA (specificity). The off-target is caused by both sgRNA and Cas9. A few mismatches between the 5′-end 20-nt sequence in sgRNA and target DNA sequence are tolerated [61]. Some studies have shown that CRISPR/Cas9 non-specially cleave DNA sites with several mismatches generating off-target mutations with considerable frequencies [10, 47, 52, 62–73]. The optimal PAM recognized by SpCas9 is 5′-NGG-3′. However, SpCas9 also binds to 5′-NAG-3′ or 5′-NGA-3′ with low frequencies [10, 57, 62, 74]. It is essential to find out potential off-target sites and improve sgRNA specificity.

To design sgRNA with high efficacy and specificity, many models and computational tools have been developed. Some representative ones are summarized in Table 1.

2.1 On-target Efficiency Prediction

The Root laboratory proposed a rule to predict on-target efficacy of sgRNAs [75]. They analyzed all possible target sites for six mouse genes and three human genes, including 1841 sgRNAs. In their model, they employed a support vector machine (SVM) model [90] to choose subsets with the best generalization error from 586 features [75]. The selected features by SVM were trained with a logistic regression classifier [91] to generate a sgRNA on-target

Table 1 Representative tools for sgRNA design

Tool	Genome specific	Nuclease	Input	On-target prediction	Off-target scoring	Websites	References
CRISPR.mit	Yes (17)	SpCas9	DNA sequence	No	Hsu et al., 2013 [62]	http://crispr.mit.edu/	[62]
sgRNA designer	No	SpCas9 SaCas9	DNA sequence Transcript ID Gene ID Gene Symbol	Rule Set 2 [50]	CFD [50]	https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design	[50, 75]
E-CRISP	Yes (55)	Cas9 Custom PAM	Gene ID Gene Symbol DNA sequence	Xu et al., 2015 [51] Doench et al., 2014 [75] E-score [76]	S-score [76]	http://www.e-crisp.org/E-CRISP/	[76]
CHOPCHOP	Yes (111)	7 PAMs Custom PAM	RefSeq Gene ID Genomic region	Rule Set 2 [50] Xu et al., 2015 [51] Chari et al., 2015 [56] Moreno-Mateos et al., 2015 [58] Doench et al., 2014 [75] G20	Cong et al., 2013 [26] Hsu et al., 2013 [62]	https://chopchop.rc.fas.harvard.edu/	[77, 78]
CRISPRseek	No	No	Software package	Doench et al., 2014 [75]	CFD [50] Hsu et al., 2013 [62]	http://www.bioconductor.org/packages/release/bioc/html/CRISPRseek.html (Local R Package)	[79]
Cas-OFFinder	Yes (36)	16 PAMs	crRNA sequences	No	Cas-OFFinder [80]	http://www.rgenome.net/cas-offinder/	[80]
CRISPRdirect	Yes (16)	Custom PAM	DNA sequence Genomic region Accession numbers	No	kmer + PAM [81]	http://crispr.dbcls.jp/	[81]
CCTop	Yes (65)	11 PAMs	DNA sequence	CRISPRater [55]	Stemmer et al., 2017 [80]	https://crispr.cos.uni-heidelberg.de/index.html	[68]
CRISPRscan	Yes (13)	SpCas9 LbCpf1 AsCpf1	DNA sequence Gene ID Gene symbol	Moreno-Mateos et al., 2015 [58]	Cong et al., 2013 [26] CFD [50] Hsu et al., 2013 [62]	http://www.crisprscan.org/	[58]
PROTOSPACER	No	SpCas9	Gene ID DNA Sequence Genomic Region	Doench et al., 2014 [75]	Hsu et al., 2013 [62]	http://www.protospacer.com/	[82]
sgRNA Scorer 1.0	No	SpCas9 StCas9	DNA Sequence	Chari et al., 2015 [56]	No	https://crispr.med.harvard.edu/sgRNAScorerV1/	[56]
WU-CRISPR	Yes (2)	SpCas9	DNA Sequence Gene ID Gene Symbol	Wong et al., 2015 [49]	Wong et al., 2015 [49]	http://crispr.wustl.edu	[49]
SSC	No	SpCas9	DNA Sequence	Xu et al., 2015 [51]	No	http://cistrome.org/SSC/	[51]

Table 1 (continued)

Tool	Genome specific	Nuclease	Input	On-target prediction	Off-target scoring	Websites	References
CROP-IT	Yes (2)	SpCas9	Genomic Region	No	Singh et al., 2015 [69]	http://cheetah.bioch.virginia.edu/AdliLab/CROP-IT/homepage.html	[69]
CRISPR Multi-Targeter	No	SpCas9 Custom PAM	DNA Sequence Gene Symbol Gene ID/ Transcripts ID RefSeq Sequence	Doench et al. 2014 [75]	GT-Scan [83] Cas-OFFinder [80]	http://www.multicrispr.net/	[84]
CLD	No	SpCas9	Gene ID Genomic Region	Xu et al., 2015 [51] Doench et al. 2014 [75] Custom	Heigwer et al., 2016 [68]	https://github.com/boutroslab/cld (Local Perl Package)	[85]
CRISPOR	Yes (258)	9 PAMs	DNA Sequence Genomic Region	Rule Set 2 [50] Xu et al., 2015 [51] Chari et al., 2015 [56] Doench et al., 2014 [75] Housden et al., 2015 [86] Prox GC [87, 88]	CFD [50] Hsu et al., 2013 [62]	http://crispor.tefor.net/	[89]
sgRNA Scorer 2.0	No	6 PAMs Custom PAM	DNA Sequence	Chari et al., 2015 [57]	No	https://crispr.med.harvard.edu/sgRNAScorerV2/	[57]
GuideScan	Yes (6)	SpCas9 AsCf1 LbCpf1	Genomic Region Gene Symbol	Rule Set 2 [50]	CFD [50]	http://www.guidescan.com/	[67]
CASPER	No	Custom PAM	DNA Sequence	CASPER [52]	CASPER [52]	https://github.com/TrinhLab/CASPER (Local Python Package)	[52]

efficacy prediction model. The efficiency score from this model is in the range of 0–1, where the higher values mean more effective. They implemented the prediction model in their online tool, sgRNA Designer [75]. Several tools use this model to predict sgRNA on-target efficiency, such as E-CRISP [76], CHOPCHOP [77, 78], PROTOSPACER [82], CLD [85], CRISPOR [89], and CRISPETa [92]. This tool supports medium-level throughput in query and one remarkable feature is that the model in this tool is based on the experimental data, so it is suitable for simultaneously generate several reliable guides. CRISPRpred, also resorts to learning algorithm to improve efficiency in predicting on-target activity and performs better than SVM model in Doench et al. [75, 93].

In 2016, the Root laboratory improved their model and proposed the Rule Set 2 [50]. Comparing with their previous model, the Rule Set 2 incorporated several new features, such as counts of position-independent nucleotide, location of target site in the corresponding gene, and melting temperature. Besides, the authors compared the performance of several classification-based models in machine learning and referred to the gradient-bossed regression trees in their reinforced model. GuideScan [67], CHOPCHOP [77, 78], CRISPOR [89], and Guide Picker [94] use the Rule Set 2 to evaluate activity of sgRNA.

The Liu laboratory assessed the features affecting sgRNA efficiency in CRISPR screen data [51]. They used Elastic-Net [95] to construct the predictive model and found some new features, mainly the preference for cytosine for sgRNA. They applied their model in SSC [51], CRISPR-FOCUS [53] and CRISPR-DO [60]. E-CRISP [76], CHOPCHOP [77, 78], CLD [85], CRISPOR [89],

and pgRNAFinder [96] also adopt this model. Besides, authors of SSC found the effect of sequence context on sgRNA efficiency differs significantly between CRISPRko and CRISPRa/i and further proposed a model to predict efficiency of sgRNA in CRISPRa/i. Due to the use of machine learning method, high efficiency of designed candidate guides is remarkable [97].

As the same demand for targeting in vivo, Moreno-Mateos et al. discovered that guanine enrichment and adenine depletion contribute to the stability, loading and activity of sgRNA by an analysis of target sites of 128 genes, where ten sgRNAs for each gene [58]. They used logistic regression model to select features associated with sgRNA activity, including features on mononucleotide and dinucleotide, and they used a linear regression model for sgRNA activity prediction. The model is integrated in their online tool CRISPRscan [58]. CHOPCHOP [77, 78] and CRISPOR [89] also use this model to predict sgRNA activity. Nevertheless, one factor should be noticed that this efficiency prediction model is based on the data of zebrafish and applying this model to other species will receive worse results compared with other tools.

The Church laboratory developed an in vivo methodology to evaluate sgRNA activity across thousands of genes [56]. The two libraries refer to target-site library and sgRNA library, respectively. They built a SVM classifier model to score the on-target activity of sgRNA sequences. They also observed the preference for guanine in position 20, next to the PAM in SpCas9 and St1Cas9 (*Streptococcus thermophilus*). The model was implemented in their online tool sgRNA Scorer. CHOPCHOP [77, 78] and CRISPOR [89] applied this model for on-target efficiency prediction. Its improved version sgRNA scorer 2.0 [57] established a more generalized model with a new combined SVM, which was proved to be useful for predicting sgRNA activity for multiple Cas9 orthologs and Cpf1 protein. The SVM model implemented in this tool highly improved on-target prediction power, whereas the analysis procedure is slow and the results have to be returned by email.

WU-CRISPR [49] identified some novel structural and sequence features from Doench's dataset [75], and built a sgRNA potency prediction model with SVM. They used the Chari dataset [56] to evaluate the model and got the best performance compared with three other tools [51, 56, 75]. This tool is recommended for its highly efficient sgRNA design using machine learning method and ease of use.

The Huang laboratory used Elastic-Net algorithm [95] in their model to predict sgRNA efficiency [59]. They borrowed knowledge from oligonucleotide design in microarray and found that including position-dependent features can help the design of effective sgRNA. The authors further found that the frequency of T and TT affects the efficiency of sgRNA significantly.

Although some studies found guanine enrichment can improve the stability and potency of sgRNA [35, 51, 58, 75], Malina et al. discovered that too many PAMs (5′-NGG-3′) within the target DNA sites inhibit the CRISPR/Cas9 activity *in vivo* [54]. Some other studies suggested the preference for guanines rather than adenines for the stability of sgRNA [35, 75]. Taking these into consideration, CASPER [52] integrate the CRISPRscan [58] score, PAM penalty score [54] and a score indicating guanines preference [35, 75] into its predictive model. Their tool also expanded the applications in multi-targeting analysis and multi-population analysis.

Labuhn et al. utilized the surrogate fluorescent reporter knock-out assays to assess and analyze a dataset of 430 sgRNAs [55]. They found nine additional features affecting sgRNA activity and generated a linear model-based discrete system using R package lars [98]. Their model enables the exclusion of impotent sgRNA, implemented in their online tool CRISPRater.

Some other tools include more than one on-target efficiency prediction models and provide more choices for users, such as E-CRISP [76], CHOPCHOP [77, 78], CLD [85], and CRISPOR [89]. The details of the corresponding tools are shown in Table 1.

2.2 Off-target Detection

The Zhang laboratory investigated sgRNA target specificity by evaluating more than 700 sgRNA variants and predicated off-target loci which is induced by SpCas9 with indel mutation levels > 100 [62]. They found that mismatch tolerance between SpCas9 complex and DNA is influenced by the number, position and distribution of mismatches. Based on this observation, they proposed a penalty matrix (Table 2) to describe the effect of mismatch position. The penalty is between 0 and 1 where higher value means bigger effect on cleavage. Referring to this penalty matrix corresponding to each position, each sgRNA can be assigned a score according to its potential off-target sites, as an evidence to choose appropriate sgRNAs. The authors also implemented a web-based software tool to facilitate sgRNA design and validation (http://crispr.mit.edu). This method is widely used for sgRNA specificity score calculation, such as CRISPRscan [58], CRISPR-DO [60], CHOPCHOP [77, 78], and CRISPOR [89]. However, this tool is only suitable for short sequences (23–500nt) and the analysis is very slow.

Stemmer et al. noticed the limitation of previous sgRNA off-target detecting tools and developed an online sgRNA finding tool, CCTop [68]. This tool provides more custom options for users to choose. A graphical visualization interface makes it easier to identify the appropriate sgRNA among the candidates. Many researchers use third party sequence alignment tools such as BWA [99] and Bowtie

Table 2 Experimentally determined penalty weight matrix of mismatch position [62]

Mismatch relative position	Penalty weight
1	0
2	0
3	0.014
4	0
5	0
6	0.395
7	0.317
8	0
9	0.389
10	0.079
11	0.445
12	0.508
13	0.613
14	0.851
15	0.732
16	0.828
17	0.615
18	0.804
19	0.685
20	0.583

[100] for target site searching [66, 76, 77]. In CCTop, the candidate sgRNAs are all selected by Bowtie [100] and ranked by their off-target potential. The off-target score of each sgRNA is defined as follows:

$$\text{Score} = \sum_{off\_target} \left[\frac{\log_{10}(\text{dist}) + \text{score}_{off\_target}}{total\_off\_targets} \right] - total\_off\_target \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (1)$$

In this equation, dist means the distance between this off-target site and the closest exon, $\text{score}_{off\_target}$ is calculated by the relative position of mismatches in protospacer. *total_off_targets* means the total number of candidate off-target sites. The authors validated the predicted results with experiments and proved its effectiveness. This tool provides many custom options which are both favorable for primary users and advanced users. Additionally, the analysis speed in simple mode is fast, whereas advanced mode is slow.

The Root laboratory analyzed off-target effects on pooled-screening data [50], with all possible sgRNAs targeting the coding regions of human CD33 cell line, regardless of PAM. The author profiled many mismatch conditions between sgRNA and DNA, including alternative PAMs, mismatches, deletions and insertions. Using the experimental data, they proposed an algorithm to rank potential off-targets, which is called the Cutting Frequency Determination (CFD) score. They compared CFD score to off-target effect measurements of Hsu–Zhang [62]. and CCTop [68]. Throughout various different examinations, they concluded that CFD score performs better and could be able to avoid high-frequency off-target effects. Some other tools apply this scoring algorithm in their workflows, such as CRISPRScan [58], GuideScan [67], and CRISPOR [89]. Here a recent study indicates that CFD score performs better than CCTop [68], MIT-Zhang [62] and CROP-IT [69] score when analyzing less than four mismatches [89].

Mendoza et al. found that the previous off-target scoring method cannot quickly accommodate across organisms, so they proposed a novel algorithm to assess the off-target activity, named CASPER [52]. The output of this algorithm consists of three factors: S_H, S_T, and S_S. S_H is a score which is calculated by the experimental data of mismatch position and types derived from the Hsu–Zhang matrix [62] (Table 2). S_T is calculated by the distance of mismatch from PAM. S_S is a stepped score, associated with the position of mismatch in spacer. Overall, S_H is derived from the experimental data and S_T, S_S are derived from the properties associated with CRISPR ribonucleoprotein complexes. Furthermore, they adapted their scoring algorithm to as Cpf1 endonucleases and exhibited that CASPER can still effectively assess its off-target activity even without sufficient experimental data.

There are also many other tools devised their own algorithms to evaluate off-target potential of sgRNA, such as WU-CRISPR [49], CROP-IT [69], CT-Finder [73], E-CRISPR [76], CGAT [101], sgRNAcas9 [102], and CRISTA [103]. Overall, the off-target detection and evaluation algorithms in these tools also have instructive value in sgRNA design. Actually, after a lot of experiments, researchers found out that using single off-target model cannot always be effective and reliable in sgRNA design [89]. Based on this consideration, many tools provide advanced options for researchers to choose the off-target detection model or show several scores calculated by widely used models in their output. CHOPCHOP [77, 78], an online tool for sgRNA design, provides off-target scoring method including two models devised by Hsu et al. [62] and Cong et al. [26], separately. CRISPOR [89] offers scoring models by Hsu et al. [62] and Doench et al., 2016 [50] (i.e., CFD). CRISPR-RT [104] aims at crRNA design in CRISPR-C2c2/Cas13a system and provides a variety of custom options for users. It returns the candidate crRNA target sites by searching the reference transcriptome with user-input RNA/cDNA sequence. Users can rank all results by transcript-target numbers, gene-target numbers, GC content or more properties of target sites.

3 Other Considerations of sgRNA Design

Throughout various studies in CRISPR system, researchers have noticed that many other factors also influence the gene editing performance of CRISPR. Here we summarize some of these factors as follows:

Truncated sgRNAs with 17- to 19-nt spacer are more sensitive to mismatches than others with 20-nt length, which can effectively reduce off-target mutations [46]. In addition, the recent study has showed that shorter spacers such as 17 to 18 nt have higher specificity but perform less on-target effective than 19- or 20-nt spacers [51]. On the contrary, longer sgRNAs are less effective than sgRNAs with 20-nt spacers [58].

Zetsche et al. found that Cpf1 (CRISPR from Prevotella and Francisella 1) protein might be more convenient in gene editing than Cas9 system, because it only need mature crRNAs for targeting, while CRISPR/Cas9 system requires both tracrRNA and crRNA. T-enriched PAM and staggered DSB at distal end of spacer are notable features in Cpf1, bringing new options in experimentation [105].

Fonfara et al. have studied the Cas9 homologs and identified specific PAM sequences and dual-RNA (i.e., tracrRNA-crRNA) in some Cas9 orthologous proteins (for example, *S. thermophilus, S. mutans, P. multocida* and *N. meningitidis*). They also verified the coevolution between dual-RNA and Cas9 protein through experimentation. Their study indicates that the Cas9 orthologs and associated PAM and dual-RNA can become candidate components in CRISPR/Cas9 system [106].

Ran et al. devised a double nicking approach by a pair of sgRNA-guided Cas9n (Cas9 D10A) nickases, which can effectively reduce the off-target effects. In this gene editing approach, two modified effectors will bind to two opposite strands and surround 10 to 31 nt target sequence and then make a DSB [107]. Some recent studies have applied this technique in their research [108, 109].

Tsai et al. described a dimeric architecture in CRISPR system called dimeric RNA-guided FokI nucleases (RFNs). This kind of dimeric nickases implement cleavage by two guide RNAs (gRNAs) binding to the opposite strands. This dimeric architecture can promise any nucleotide at 5′ end of gRNA and relative to paired nickases, which makes steps forward in improving gRNA specificity [110].

CRISPR gene editing system can be divided into two categories, CRISPRko and CRISPRa/i. The DNA target sites vary for these two categories. In CRISPRko, target sites are usually protein-coding region and Cas9 nucleases produce loss-of-function gene knockouts by inducing DSB. In CRISPRa/i, sgRNA guides catalytically inactivate Cas9 (dCas9) to effector domains to activate or repress gene transcription without DNA cleavage [50, 51].

Canonical sgRNA including 20-nt sequence and PAM has limited target sites *in vivo*, because T7 or SP6 promoters in CRISPR/Cas9 system restrict GG or GA in the 5′-end of sgRNA sequence [58]. Experimental results showed that there are two kinds of alternative truncated sequences which can become efficient substitutes to replace the 20-nt sequence, one is truncated sgRNA at 5′ end by less than 2-nt and the other is the 20-nt sgRNA with one mismatch in 5′ end [58].

Some researchers investigated the epigenetics features, and found that chromatin accessibility affects the binding of dCas9 to DNA sites [111, 112]. Chari et al. also observed this phenomenon [56]. Whereas Moreno-Mateos et al. did not notice a strong effect of chromatin accessibility on CRISPR/Cas9 activity [58].

Four consecutive T (i.e., TTTT) is a signal of termination for pol III promoter thus the presence of TTTT in target DNA sequence sites cannot be adoptable. It is necessary to avoid TTTT in target sites when using pol III promoter [81].

4 Conclusion

CRISPR/Cas9 technology has been used more and more widely for genome editing. Its efficacy, specificity, easy to use, cost-effectiveness and versatility, will boost this technique on more fronts. In this review, we examined various aspects of sgRNA design tools, including activity prediction models, and off-target detection algorithms. Almost all of these models or algorithms depend on large-scale experimental datasets and systematic analysis. Therefore, with more CRISPR/Cas9 datasets available, more novel sgRNA design tools will be developed to facilitate the research community.

Acknowledgements We would like thank to Jingyu Amy Peng and Chen-Hao Chen from Harvard T.H. Chan School of Public Health, Shenglin Mei and Jian Ma from Tongji University for their discussion about this work. This work was supported by National Key R&D Program of China 2017YFB0202602, 2017YFC1311003, 2016YFC1302500, 2016YFB0200400, 2017YFB0202104; NSFC Grants 61772543, U1435222, 61625202, 61272056; the Funds of State Key Laboratory of Chemo/Biosensing and Chemometrics; the Fundamental Research Funds for the Central Universities; and Guangdong Provincial Department of Science and Technology under grant No. 2016B090918122.

References

1. Koonin EV, Makarova KS (2009) CRISPR-Cas: an adaptive immunity system in prokaryotes. F1000 Biology Reports 1:95. https://doi.org/10.3410/B1-95
2. Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. Science 327(5962):167–170. https://doi.org/10.1126/science.1179555

3. Koonin EV, Makarova KS (2013) CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. RNA Biol 10(5):679–686. https://doi.org/10.4161/rna.24022
4. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471(7340):602–607. https://doi.org/10.1038/nature09886
5. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821. https://doi.org/10.1126/science.1225829
6. Cong L, Zhang F (2015) Genome engineering using CRISPR-Cas9 system. Methods Mol Biol 1239:197–217. https://doi.org/10.1007/978-1-4939-1862-1_10
7. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, Marraffini LA (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature 519(7542):199–202. https://doi.org/10.1038/nature14245
8. Wyman C, Kanaar R (2006) DNA double-strand break repair: all's well that ends well. Annu Rev Genet 40(1):363–383. https://doi.org/10.1146/annurev.genet.40.110405.090451
9. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152(5):1173–1183. https://doi.org/10.1016/j.cell.2013.02.022
10. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31(9):833–838. https://doi.org/10.1038/nbt.2675
11. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS (2014) Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159(3):647–661. https://doi.org/10.1016/j.cell.2014.09.029
12. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435(7042):646–651. https://doi.org/10.1038/nature03556
13. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11:636–646. https://doi.org/10.1038/nrg2842
14. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA (2008) Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol 26(6):695–701. https://doi.org/10.1038/nbt1398
15. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA (2011) Zinc finger protein-dependent and-independent contributions to the in vivo off- target activity of zinc finger nucleases. Nucleic Acids Res 39(1):381–392. https://doi.org/10.1093/nar/gkq787
16. Chu SW, Noyes MB, Christensen RG, Pierce BG, Zhu LJ, Weng Z, Stormo GD, Wolfe SA (2012) Exploring the DNA-recognition potential of homeodomains. Genome Res 22(10):1889–1898. https://doi.org/10.1101/gr.139014.112
17. Enuameh MS, Asriyan Y, Richards A, Christensen RG, Hall VL, Kazemian M, Zhu C, Pham H, Cheng Q, Blatti C, Brasefield JA, Basciotta MD, Ou J, McNulty JC, Zhu LJ, Celniker SE, Sinha S, Stormo GD, Brodsky MH, Wolfe SA (2013) Global analysis of Drosophila Cys2-His2 zinc finger proteins reveals a multitude of novel recognition motifs and binding determinants. Genome Res 23(6):928–940. https://doi.org/10.1101/gr.151472.112
18. Shekar M, Venugopal MN (2017) Insight into a transcriptional adaptor zinc finger encoded by a putative protein in the white spot syndrome virus genome. Interdiscip Sci Comput Life Sci. https://doi.org/10.1007/s12539-017-0268-x
19. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512. https://doi.org/10.1126/science.1178811
20. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501. https://doi.org/10.1126/science.1178817
21. Bogdanove AJ, Voytas DF (2011) TAL effectors: customizable proteins for DNA targeting. Science 333:1843–1846. https://doi.org/10.1126/science.1204094
22. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148. https://doi.org/10.1038/nbt.1755
23. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K (2013) Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell 12(4):393–394. https://doi.org/10.1016/j.stem.2013.03.006
24. Ikmi A, McKinney SA, Delventhal KM, Gibson MC (2014) TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. Nat Commun 5:5486. https://doi.org/10.1038/ncomms6486
25. Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, Wang Y, Brodsky RA, Zhang K, Cheng L, Ye Z (2014) Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell 15(1):12–13. https://doi.org/10.1016/j.stem.2014.06.011
26. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823. https://doi.org/10.1126/science.1231143
27. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826. https://doi.org/10.1126/science.1232033
28. Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA (2013) Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat Methods 10(8):741–743. https://doi.org/10.1038/nmeth.2532
29. Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM (2013) Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194(4):1029–1035. https://doi.org/10.1534/genetics.113.152710
30. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, Thomson JA (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci USA 110(39):15644–15649. https://doi.org/10.1073/pnas.1313587110
31. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31(3):227–229. https://doi.org/10.1038/nbt.2501
32. Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, Li Y, Gao N, Wang L, Lu X, Zhao Y, Liu M (2013) Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol 31(8):681–683. https://doi.org/10.1038/nbt.2661
33. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome

engineering. Cell 154(6):1370–1379. https://doi.org/10.1016/j.cell.2013.08.022
34. Chakrapani V, Rasal KD, Kumar S, Mohapatra SD, Sundaray JK, Jayasankar P et al (2017) In silico analysis of nsSNPs of carp TLR22 gene affecting its binding ability with Poly I:C. Interdiscip Sci Comput Life Sci. https://doi.org/10.1007/s12539-017-0247-2
35. Wang T, Wei JJ, Sabatini DM, Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343(6166):80–84. https://doi.org/10.1126/science.1246981
36. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343(6166):84–87. https://doi.org/10.1126/science.1247005
37. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera MC, Yusa K (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32(3):267–273. https://doi.org/10.1038/nbt.2800
38. Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM (2015) Identification and characterization of essential genes in the human genome. Science 350(6264):1096–1101. https://doi.org/10.1126/science.aac7041
39. Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, Mupo A, Grinkevich V, Li M, Mazan M, Gozdecka M, Ohnishi S, Cooper J, Patel M, McKerrell T, Chen B, Domingues AF, Gallipoli P, Teichmann S, Ponstingl H, McDermott U, Saez-Rodriguez J, Huntly BJP, Iorio F, Pina C, Vassiliou GS, Yusa K (2016) A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. Cell reports 17(4):1193–1205. https://doi.org/10.1016/j.celrep.2016.09.079
40. Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, Fields AP, Park CY, Corn JE, Kampmann M, Weissman JS (2016) Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. ELife 5:e19760. https://doi.org/10.7554/eLife.19760
41. Aguirre AJ, Meyers RM, Weir BA, Vazquez F, Zhang CZ, Ben-David U, Cook A, Ha G, Harrington WF, Doshi MB, Kost-Alimova M, Gill S, Xu H, Ali LD, Jiang G, Pantel S, Lee Y, Goodale A, Cherniack AD, Oh C, Kryukov G, Cowley GS, Garraway LA, Stegmaier K, Roberts CW, Golub TR, Meyerson M, Root DE, Tsherniak A, Hahn WC (2016) Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. Cancer Discov 6(8):914–929. https://doi.org/10.1158/2159-8290.CD-16-0154
42. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell 160(6):1246–1260. https://doi.org/10.1016/j.cell.2015.02.038
43. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, Goodale A, Lee Y, Ali LD, Jiang G, Lubonja R, Harrington WF, Strickland M, Wu T, Hawes DC, Zhivich VA, Wyatt MR, Kalani Z, Chang JJ, Okamoto M, Stegmaier K, Golub TR, Boehm JS, Vazquez F, Root DE, Hahn WC, Tsherniak A (2017) Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. Nat Genet 49(12):1779. https://doi.org/10.1038/ng.3984
44. Yennmalli RM, Kalra S, Srivastava PA, Garlapati VK (2017) Computational tools and resources for crispr/cas 9 genome editing method. MOJ Proteom Bioinform 5(4):00164. https://doi.org/10.15406/mojpb.2017.05.00164
45. Zhu LJ (2015) Overview of guide RNA design tools for CRISPR-Cas9 genome editing technology. Front Biol 10(4):289–296. https://doi.org/10.1007/s11515-015-1366-y
46. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 32(3):279–284. https://doi.org/10.1038/nbt.2808
47. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31(9):822–826. https://doi.org/10.1038/nbt.2623
48. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J (2013) RNA-programmed genome editing in human cells. ELife 2:e00471. https://doi.org/10.7554/eLife.00471
49. Wong N, Liu W, Wang X (2015) WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. Genome Biol 16:218. https://doi.org/10.1186/s13059-015-0784-0
50. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 34(2):184. https://doi.org/10.1038/nbt.3437
51. Xu H, Xiao T, Chen CH, Li W, Meyer CA, Wu Q, Wu D, Cong L, Zhang F, Liu JS, Brown M, Liu XS (2015) Sequence determinants of improved CRISPR sgRNA design. Genome Res 25(8):1147–1157. https://doi.org/10.1101/gr.191452.115
52. Mendoza BJ, Trinh CT (2018) Enhanced guide-RNA design and targeting analysis for precise CRISPR genome editing of single and consortia of industrially relevant and non-model organisms. Bioinformatics 34(1):16–23. https://doi.org/10.1093/bioinformatics/btx564
53. Cao Q, Ma J, Chen CH, Xu H, Chen Z, Li W, Liu XS (2017) CRISPR-FOCUS: a web server for designing focused CRISPR screening experiments. PLoS One 12(9):e0184281. https://doi.org/10.1371/journal.pone.0184281
54. Malina A, Cameron CJ, Robert F, Blanchette M, Dostie J, Pelletier J (2015) PAM multiplicity marks genomic target sites as inhibitory to CRISPR-Cas9 editing. Nat Commun 6:10124. https://doi.org/10.1038/ncomms10124
55. Labuhn M, Adams FF, Ng M, Knoess S, Schambach A, Charpentier EM, Schwarzer A, Mateo JL, Klusmann JH, Heckl D (2018) Refined sgRNA efficacy prediction improves large- and small-scale CRISPR-Cas9 applications. Nucleic Acids Res 46(3):1375–1385. https://doi.org/10.1093/nar/gkx1268
56. Chari R, Mali P, Moosburner M, Church GM (2015) Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods 12(9):823. https://doi.org/10.1038/nmeth.3473
57. Chari R, Yeo NC, Chavez A, Church GM (2017) sgRNA Scorer 2.0: a species-independent model to predict CRISPR/Cas9 activity. ACS Synth Biol 6(5):902–904. https://doi.org/10.1021/acssynbio.6b00343
58. Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK, Giraldez AJ (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods 12(10):982. https://doi.org/10.1038/nmeth.3543
59. Kuan PF, Powers S, He S, Li K, Zhao X, Huang B (2017) A systematic evaluation of nucleotide properties for CRISPR sgRNA design. BMC Bioinform 18(1):297. https://doi.org/10.1186/s12859-017-1697-6
60. Ma J, Köster J, Qin Q, Hu S, Li W, Chen C, Cao Q, Wang J, Mei S, Liu Q, Xu H, Liu XS (2016) CRISPR-DO for genome-wide CRISPR design and optimization. Bioinformatics 32(21):3336–3338. https://doi.org/10.1093/bioinformatics/btw476

61. Lin Y, Cradick TJ, Brown MT, Deshmukh H, Ranjan P, Sarode N, Wile BM, Vertino PM, Stewart FJ, Bao G (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res 42(11):7473–7485. https://doi.org/10.1093/nar/gku402
62. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31(9):827–832. https://doi.org/10.1038/nbt.2647
63. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 31:839–843. https://doi.org/10.1038/nbt.2673
64. Cradick TJ, Fine EJ, Antico CJ, Bao G (2013) CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res 41(20):9584–9592. https://doi.org/10.1093/nar/gkt714
65. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24(1):132–141. https://doi.org/10.1101/gr.162339.113
66. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP, Aryee MJ, Joung JK (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33(2):187. https://doi.org/10.1038/nbt.3117
67. Perez AR, Pritykin Y, Vidigal JA, Chhangawala S, Zamparo L, Leslie CS, Ventura A (2017) GuideScan software for improved single and paired CRISPR guide RNA design. Nat Biotechnol 35(4):347–349. https://doi.org/10.1038/nbt.3804
68. Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL (2017) CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. Plos One 12(4):e0176619. https://doi.org/10.1371/journal.pone.0176619
69. Singh R, Kuscu C, Quinlan A, Qi Y, Adli M (2015) Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. Nucleic Acids Res 43(18):e118. https://doi.org/10.1093/nar/gkv575
70. Graham DB, Root DE (2015) Resources for the design of CRISPR gene editing experiments. Genome Biol 16:260. https://doi.org/10.1186/s13059-015-0823-x
71. Jakočiūnas T, Jensen MK, Keasling JD (2015) CRISPR/Cas9 advances engineering of microbial cell factories. Metab Eng 34:44–59. https://doi.org/10.1016/j.ymben.2015.12.003
72. Oliveros JC, Mònica F, Daniel TM, David SL, Lluis M, Pilar C et al (2016) Breaking-Cas—interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic Acids Res 44(W1):W267. https://doi.org/10.1093/nar/gkw407
73. Zhu H, Misel L, Graham M, Robinson ML, Liang C (2016) CT-Finder: a web service for CRISPR optimal target prediction and visualization. Sci Rep 6:25516. https://doi.org/10.1038/srep25516
74. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31(3):233–239. https://doi.org/10.1038/nbt.2508
75. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32(12):1262. https://doi.org/10.1038/nbt.3026
76. Heigwer F, Kerr G, Boutros M (2014) E-CRISP: fast CRISPR target site identification. Nat Methods 11(2):122–123. https://doi.org/10.1038/nmeth.2812
77. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res 42(Web Server issue):W401–W407. https://doi.org/10.1093/nar/gku410
78. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res 44(W1):W272–W276. https://doi.org/10.1093/nar/gkw398
79. Zhu LJ, Holmes BR, Aronin N, Brodsky MH (2014) CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. Plos One 9(9):e108424. https://doi.org/10.1371/journal.pone.0108424
80. Bae S, Park J, Kim JS (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30(10):1473. https://doi.org/10.1093/bioinformatics/btu048
81. Naito Y, Hino K, Bono H, Ui-Tei K (2015) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics 31(7):1120–1123. https://doi.org/10.1093/bioinformatics/btu743
82. MacPherson CR, Scherf A (2015) Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. Nat Biotechnol 33(8):805. https://doi.org/10.1038/nbt.3291
83. O'Brien A, Bailey TL (2014) GT-Scan: identifying unique genomic targets. Bioinformatics 30(18):2673–2675. https://doi.org/10.1093/bioinformatics/btu354
84. Prykhozhij SV, Rajan V, Gaston D, Berman JN (2015) CRISPR multiTargeter: a web tool to find common and unique CRISPR single guide RNA targets in a set of similar sequences. PLoS One 10(3):e0119372. https://doi.org/10.1371/journal.pone.0119372
85. Heigwer F, Zhan T, Breinig M, Winter J, Brügemann D, Leible S, Boutros M (2016) CRISPR library designer (CLD): software for multispecies design of single guide RNA libraries. Genome Biol 17(1):55. https://doi.org/10.1186/s13059-016-0915-2
86. Housden BE, Valvezan AJ, Kelley C, Sopko R, Hu Y, Roesel C, Lin S, Buckner M, Tao R, Yilmazel B, Mohr SE, Manning BD, Perrimon N (2015) Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. Sci Signal 8(393):rs9. https://doi.org/10.1126/scisignal.aab3729
87. Ren X, Yang Z, Xu J, Sun J, Mao D, Hu Y, Yang SJ, Qiao HH, Wang X, Hu Q, Deng P, Liu LP, Ji JY, Li JB, Ni JQ (2014) Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. Cell Rep 9(3):1151–1162. https://doi.org/10.1016/j.celrep.2014.09.044
88. Farboud B, Meyer BJ (2015) Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics 199(4):959–971. https://doi.org/10.1534/genetics.115.175166
89. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly JS, Concordet JP (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol 17(1):148. https://doi.org/10.1186/s13059-016-1012-2
90. Labaj W, Papiez A, Polanski A, Polanska J (2017) Comprehensive analysis of MILE gene expression data set advances discovery of leukaemia type and subtype biomarkers. Interdiscip Sci Comput Life Sci 9(1):24–35. https://doi.org/10.1007/s12539-017-0216-9
91. Pei Z, Liu J, Liu M, Zhou W, Yan P, Wen S et al (2018) Risk-predicting model for incident of essential hypertension based on environmental and genetic factors with support vector machine. Interdiscip Sci Comput Life Sci. https://doi.org/10.1007/s12539-017-0271-22016
92. Pulido-Quetglas C, Aparicio-Prat E, Arnan C, Polidori T, Hermoso T, Palumbo E, Ponomarenko J, Guigo R, Johnson AK (2017) Scalable design of paired CRISPR guide RNAs for

genomic deletion. PLOS Comput Biol 13(3):e1005341. https://doi.org/10.1371/journal.pcbi.1005341
93. Rahman MK, Rahman MS (2017) CRISPRpred: a flexible and efficient tool for sgRNAs on-target activity prediction in CRISPR/cas9 systems. Plos One 12(8):e0181943. https://doi.org/10.1371/journal.pone.0181943
94. Hough SH, Kancleris K, Brody L, Humphryes-Kirilov N, Wolanski J, Dunaway K, Ajetunmobi A, Dillard V (2017) Guide Picker is a comprehensive design tool for visualizing and selecting guides for CRISPR experiments. BMC Bioinform 18:167. https://doi.org/10.1186/s12859-017-1581-4
95. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. J Roy Stat Soc 67(2):301–320
96. Xiong Y, Xie X, Wang Y, Ma W, Liang P, Songyang Z, Dai Z (2017) pgRNAFinder: a web-based tool to design distance independent paired-gRNA. Bioinformatics 33(22):3642–3644. https://doi.org/10.1093/bioinformatics/btx472
97. Chuai GH, Wang QL, Qi L (2016) In silico meets in vivo: towards computational CRISPR-based sgRNA design. Trends Biotechnol 35(1):12. https://doi.org/10.1016/j.tibtech.2016.06.008
98. Efron B, Hastie T, Johnstone I, Tibshirani R (2004) Least angle regression. Ann Stat 32:407–451
99. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324
100. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25. https://doi.org/10.1186/gb-2009-10-3-r25
101. Brazelton VA Jr, Zarecor S, Wright DA, Wang Y, Liu J, Chen K, Yang B, Lawrence-Dill CJ (2015) A quick guide to CRISPR sgRNA design tools. Gm Crops Food 6(4):266–276. https://doi.org/10.1080/21645698.2015.1137690
102. Xie S, Shen B, Zhang C, Huang X, Zhang Y (2014) sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. Plos One 9(6):e100448. https://doi.org/10.1371/journal.pone.0100448
103. Abadi S, Yan WX, Amar D, Mayrose I (2017) A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. PLoS Comput Biol 13(10):e1005807. https://doi.org/10.1371/journal.pcbi.1005807
104. Zhu H, Richmond E, Liang C (2018) CRISPR-RT: a web application for designing CRISPR-C2c2 crRNA with improved target specificity. Bioinformatics 34(1):117–119. https://doi.org/10.1093/bioinformatics/btx580
105. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163(3):759. https://doi.org/10.1016/j.cell.2015.09.038
106. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécrivain AL, Bzdrenga J, Koonin EV, Charpentier E (2014) Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Res 42(4):2577–2590. https://doi.org/10.1093/nar/gkt1074
107. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154(6):1380–1389. https://doi.org/10.1016/j.cell.2013.08.021
108. Mccaffrey J, Sibert J, Zhang B, Zhang Y, Hu W, Riethman H et al (2016) CRISPR-Cas9 D10A nickase target-specific fluorescent labeling of double strand DNA for whole genome mapping and structural variation analysis. Nucleic Acids Res 44(2):e11-e11. https://doi.org/10.1093/nar/gkv878
109. Chang L, Sun C, Chen X, Yang W, Zhang J, Zhang Y et al. (2017). Knocking out of human DNAH2 gene in U2OS cells by CRISPR/Cas9n double nick system. Sheng Wu Gong Cheng Xue Bao. https://doi.org/10.13345/j.cjb.160273
110. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK (2014) Dimeric CRISPR RNA-guided FokI-dCas9 nucleases directed by truncated gRNAs for highly specific genome editing. Nat Biotechnol 32(6):569–576. https://doi.org/10.1038/nbt.2908
111. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, Jaenisch R, Zhang F, Sharp PA (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32(7):670–676. https://doi.org/10.1038/nbt.2889
112. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat Biotechnol 32(7):677–683. https://doi.org/10.1038/nbt.2916