Chapter 19

CRISPR Guide RNA Design Guidelines for Efficient Genome Editing

Patrick Schindele, Felix Wolter, and Holger Puchta

Abstract

The simple applicability and facile target programming of the CRISPR/Cas9-system abolish the major boundaries of previous genome editing tools, making it the tool of choice for generating site-specific genome alterations. Its versatility and efficacy have been demonstrated in various organisms; however, accurately predicting guide RNA efficiencies remains an organism-independent challenge. Thus, designing optimal guide RNAs is essential to maximize the experimental outcome. Here, we summarize the current knowledge for guide RNA design and highlight discrepancies between different experimental systems.

Key words Genome editing, CRISPR, Cas9, gRNA design, gRNA secondary structure, Mismatch tolerance, CRISPR prediction tool

1 Introduction

1.1 Genome Editing with CRISPR/Cas9

Targeted manipulation of DNA through site-specific double-strand breaks (DSBs) embodies the cornerstone of modern biotechnology. The challenge to target sites of interest continuously decreased over time with the discovery and development of novel tools, such as engineered nucleases [1]. The CRISPR/Cas9 system with its two-component setup accompanied by the simple target programming constitutes the current gold standard within the available toolbox. In this system, a complex of the Cas9 nuclease and a guide RNA (gRNA) mediates DSB induction at a selected target site. The gRNA is of major relevance, mediating target DNA recognition and binding on the one hand and activation of target DNA cleavage by Cas9 on the other hand [2]. The variable region of the gRNA (guide) determines the site of target DNA binding and can be adjusted to the sequence of interest. Employing this system, applications such as single and multiplex editing, epigenetic and transcriptional regulation, visualization of genomic loci, and base editing are feasible within a large number of organisms (for details *see* reviews [3–6]). The target selection solely requires an

Manfred Heinlein (ed.), *RNA Tagging: Methods and Protocols*, Methods in Molecular Biology, vol. 2166,
https://doi.org/10.1007/978-1-0716-0712-1_19,

abundant protospacer adjacent motif (PAM), apparently providing a wide variety of potential target sites, and gRNAs. However, indiscriminately selecting gRNAs can minimize and even prevent experimental success. Thus, optimization of gRNA design is required for improving target specificity and maximizing editing efficiency. Pre-experimental procedures therefore often involve the screening of different guides to determine the optimal target site. While animal cell lines allow high-throughput screenings, evaluating gRNA-editing efficiency in other organisms can be an elaborate process demanding longer periods for the generation of transgenic individuals. Especially for many crops like maize and wheat, generation of transgenic plants is very time and cost intensive, and being able to estimate gRNA efficiency before engaging in the laborious process of transgenic plant production would be most desirable. For this very reason, guidelines helping to define efficient gRNAs are of tremendous importance to ensure an optimal experimental progress. Both target site and gRNA features determine the on-target and off-target activity of the CRISPR nuclease. Length and sequence composition or structural features of the guide and target, respectively, have been reported as main contributors to overall efficiency.

1.2 Guide Length Variation for Reduced Off-Target Activity

The various Cas9 orthologs put different demands on PAM composition and guide length, yet, for each individual ortholog, an optimal PAM and guide length required for maximum on-target activity were identified [7]. Based on this observation, the most commonly used *Streptococcus pyogenes* Cas9 (SpCas9) is employed combining a 20-nt guide and the 5′-NGG-3′ PAM. However, due to the abundance of the PAM and a certain tolerance toward mismatches, off-target activity was frequently detected in human and animal cell lines [8, 9]. More recent studies on human 293T cells and *Drosophila* reported strongly reduced off-target activity by employing truncated guides of 17 to 18 nt in length while maintaining the editing efficiency of full-length guides [10–12]. The strong decrease in off-target activity can presumably be explained by a much stronger disruptive impact of mismatches on truncated guides. Unfortunately, contradictory results have been reported in stem cells and plants where truncated guides were less efficient than full-length guides [13, 14]. Apart from that, full-length guides are highly precise in plants and off-target effects only detectable for highly similar targets or targets with PAM distal mismatches [15–19].

1.3 Effects of Nucleotide Composition and Identity on Guide Efficiency

Concerning the effect of the nucleotide composition of the guide and target, rather inconsistent results have been published. It is agreed that guides having a very low or very high GC content are less effective [20–22]. However, analysis about the optimal GC content strongly vary between different organisms. In animal cell lines, a preferable GC content of 40–60% was reported [23]. In

plants, analysis of a significant amount of validated gRNAs revealed a spacious GC content ranging from 30% to 80% [22]. Though rather marginal, guides with a GC content of more than 50% showed slightly higher efficiencies than guides with a GC content under 50% [24]. The same was reported for *Drosophila*, where a GC content over 50% within the 6 nt proximal to the PAM was reported as beneficial, an effect that was also detected concerning germline transmission rates of heritable mutations [12]. A similar inconsistency is prevalent regarding nucleotide preferences. In animal and human cell lines, efficient gRNAs strongly prefer purines at the very 3′-end of the guide [21]. Whereas guanine is favored at positions 1 and 2 proximal to the PAM, thymine and cytosine are restrictive for efficient editing [20, 25]. Cytosine is also disfavored at position 18 distal to the PAM, however, strongly preferred at positions 3 and 5 proximal to the PAM and as variable nucleotide of the PAM (5′-CGG-3′ PAM), respectively. Adenine preference was detected in the midsection of the guide [20, 25]. In contrast, for plants no significant nucleotide preferences could be validated [22]. Merely an increased occurrence for guanine at the very 5′-end of the guides was detected, although this quite likely can be attributed to the common use of the U6 small nuclear RNA promoter constraining the first nucleotide to a guanine, at least if a one-nucleotide "G" overhang of the gRNA is not desired. Interestingly, irregular targets starting with H nucleobases still show comparable efficiencies in plants [24]. Analysis on target strand preference also provides contradictory results [20, 21].

1.4 Influence of Mismatch Nature and Position on Cleavage Activity

Off-target activity is largely determined by mismatch tolerance. Early studies showed that cleavage activity of Cas9 is preferentially abolished through mismatches in the PAM-proximal region [2, 26]. However, studies in human cells also revealed a strong impact of nucleotide identity on cleavage activity [9, 27]. Whereas a G:T mismatch between gRNA and target DNA in the PAM-proximal region only minimally affects cleavage, activity is barely detectable for a C:C mismatch between gRNA and target DNA [9]. The data from this study also indicate toward a significant relevance of nucleotides 5–7 proximal to the PAM with high levels of cleavage disruption independent of nucleotide identity. Consistent with this indication, a recent publication defined a new core region comprising nucleotides 4–7 proximal of the PAM with even single mismatches abolishing the majority of cleavage activity [27]. Crucial but rather neglected features concerning on-target and off-target activity are RNA and DNA bulges. These structures are formed when unpaired nucleotides reside in the otherwise consistent guide or target, respectively. In human cells, Cas9 can tolerate DNA bulges of 1 nt all along the target sequence, though exact positions are inconsistent between different gRNAs [11]. RNA bulges of 1 nt can also be tolerated; however, abolish

Cas9 cleavage when located in the PAM-proximal region. In contrast, DNA bulges of 2 or more nt completely abolish Cas9 cleavage independent of position, whereas larger RNA bulges at least outside the PAM-proximal region can be tolerated to a certain degree [11]. Guide-internal base pairs interfere with DNA target binding, as well. In plants, at least a certain degree of tolerance toward these pairings can be detected [22].

1.5 Preservation of gRNA Secondary Structure Is Essential for Proper Function

The structural characteristics of the gRNA are essential for interaction with the Cas protein. Structural analysis in human cell lines regarding SpCas9 revealed that an intact repeat:anti-repeat duplex and stem-loop 1 of the gRNA are of major relevance for Cas9 recognition, being less tolerant toward mismatches than stem-loop 2 and 3 [28]. Interestingly, structure analysis of validated gRNAs in plants revealed a minor relevance of stem-loop 1, indicating a disparity concerning the structural requirements in different species [22]. Mismatches or substitutions maintaining the stem-loop structures of the gRNA barely affect Cas9 function, rather emphasizing the significance of the global structure of the gRNA [28, 29]. Due to the variable nature of the guide sequence, base pairing and thus interference with gRNA secondary structure can occur (Fig. 1). Consequently, evaluating the guide-dependent gRNA secondary structure is inevitable. Structure analysis of various gRNAs in plants demonstrated a certain degree of tolerance toward base pairing between guide sequence and gRNA, with 12 base pairs or 7 consecutive base pairs, respectively, being sustainable for single gRNAs [22].

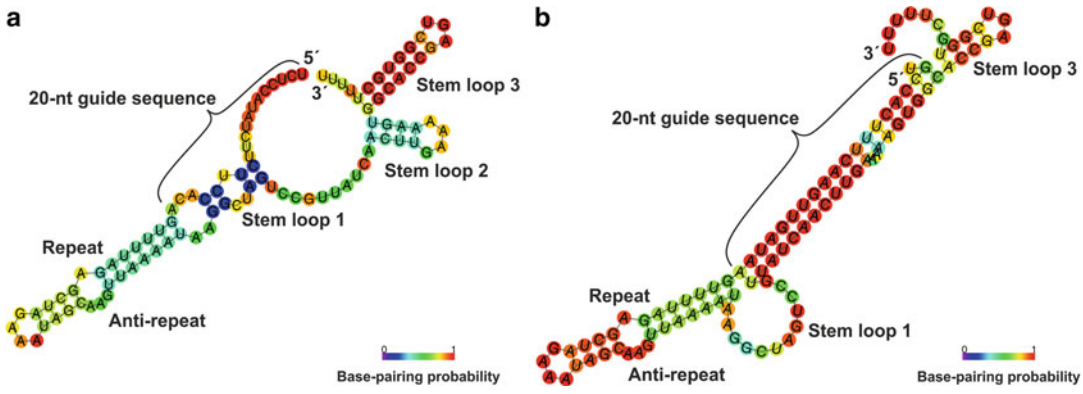

Fig. 1 Impact of guide sequence on gRNA secondary structure. (**a**) Illustration of a gRNA with intact secondary structure features. Guide sequence and gRNA exhibit marginal interactions only, not affecting the essential stem-loop structures of the gRNA. (**b**) Illustration of a gRNA with impaired secondary structure features. Guide sequence and gRNA exhibit a significant level of base pair interactions, compromising the formation of essential stem-loop structures (in this example stem-loop 2). The color scale displays the base-pairing probability. The RNA secondary structure was predicted using the RNAfold web server (http:/rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi)

1.6 Chromatin Accessibility Affects Genome Editing Efficiency

Epigenetic modifications influence chromatin state and thus accessibility of DNA; hence, they are a major factor affecting genome targeting ability. Restrictions due to epigenetic repression were already hypothesized for previous site-specific nucleases such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) [30, 31]. For SpCas9, in vitro and in vivo analysis demonstrated no restriction in cleavage activity when targeting methylated DNA [9, 32]. However, in vitro analysis revealed an impact of nucleosome occupancy on Cas9-mediated cleavage. A correlation between Cas9 binding and low nucleosome occupancy also indicates toward a contribution in vivo [33, 34]. Further experiments in human cells also demonstrated an impairment of genome editing at an epigenetically repressed reporter locus and data from zebrafish also suggests a negative correlation between chromatin accessibility and genome editing efficiency [35–37]. Additionally, the open chromatin state associated with transcriptionally active regions can have its own positive effect on Cas9 editing by displacing Cas9, thereby increasing the rate at which cleaved ends are exposed and accessible for DNA repair. On the other hand, this might have a negative effect for dCas9-based applications where extended binding is beneficial [32].

1.7 CRISPR Prediction Tools for Approving Target Selection

Applying in silico tools may assist in predicting on-target and minimizing off-target activity. However, some tools do not necessarily cover all contributing factors by the current state of knowledge. Furthermore, depending on the experimental system the data are based on, discrepancies between prediction and outcome can occur. To obtain an optimal consensus, the use of multiple prediction tools is recommended. For RNA secondary structure prediction, free-available online tools, such as Mfold [38] and RNAfold [39], are reliable to exclude potential issues from RNA structure. Computational prediction tools for the identification of optimal guide sequences are available on a large scale; however, they might differ concerning their parameters. While SSC [25] only allows for variation between different guide lengths, CRISPR-P 2.0 [40] and CCTop [41] allow the choice of a variety of CRISPR orthologs and target organisms. CRISPR RGEN Tools [42, 43] additionally offers crucial off-target prediction criteria such as RNA and DNA bulges. While these tools are highly useful to assist in target site selection, their limitations should always be kept in mind. Predictive power is often limited and efficiency prediction is based solely on the target sequence, whereas local chromatin context cannot be taken into account [44].

1.8 Nontrivial Considerations for Designing CRISPR Knockout Experiments

As extensively described above, designing efficient gRNAs is one of the major concerns when conducting CRISPR experiments. However, depending on the experimental goal, further criteria have to be taken into account. As the perhaps most frequent used CRISPR application, the following paragraph concentrates on the

prerequisites for designing knockout experiments. The majority of knockout mutants are generated through nonhomologous end joining (NHEJ)-mediated DSB repair and therefore based on the introduction of Indel mutations. Thus, selecting pertinent target sites within exons is of major concern. In particular, mutagenesis within exons aims for either mutations in essential protein domains or generation of frameshift mutations, the latter one being preferred when aiming for complete knockouts. Targeting regions too close to the C- or N-terminus of the encoded protein is not recommended, either increasing the probability of maintaining the majority of essential domains or, if the start codon is compromised, resulting solely in a minor displacement of transcription initiation. Nevertheless, mutations in the encoded N-terminal region are preferred for frameshift mutations, affecting the majority of the coding sequences and impeding a distortion through potential splice variants.

In the following, the design of a gRNA will be explained in detail comprising helpful bioinformatic design tools and guide sequence requirements based on current knowledge of gRNA design.

2 Materials

2.1 Bioinformatic Online Tools

1. CCTop.
2. CRISPR RGEN tools.
3. RNAfold.

2.2 Sequence Information

1. Query sequence.

3 Methods

3.1 Guide Sequence Selection for CRISPR-Mediated Mutagenesis Using Bioinformatic Tools

The first and essential step of designing the CRISPR experiment is the identification of the optimal guide sequence. Due to the variety of criteria to be considered for this purpose, the use of online tools is recommended. These tools cover the majority of the design criteria and thereby definitely help to choose the optimal guide sequence. To obtain an optimal consensus, the use of multiple prediction tools is recommended. In the following, the online tools CCTop and Cas-Designer (CRISPR RGEN tools) are employed, both comprising a solid extent of selectable options to prevent off-target activity and promote on-target activity.

3.1.1 CCTop

1. Go to https://crispr.cos.uni-heidelberg.de/ to open the CCTop tool.

2. Enter your gene/sequence of interest into the query sequence field. For knockout experiments, concentrate on the exons of the gene (*see* Subheading 1.8).
3. Select the PAM type. The PAM depends on the employed CRISPR/Cas-system and respective Cas ortholog (e.g., 5′-NGG-3′ for *Streptococcus pyogenes* Cas9).
4. Select the guide length of the respective CRISPR ortholog. This sequence corresponds to the nucleotides upstream of the PAM for Cas9 (*see* **Note 1**).
5. Optional: Enter 5′ and 3′ target site limitations. Dependent on the promoter and/or experimental system, the occurrence of specific nucleotides at these positions increases editing efficiency (*see* Subheading 1.3).
6. Optional: Specify the 5′ guide sequence overhangs. Cloning of the guide is commonly realized via oligonucleotide annealing and sticky-end ligation (*see* **Note 2**).
7. For off-target prediction, select the number of total mismatches between guide and target site to be considered. Recommendation: Select four (*see* Subheading 1.4) (*see* **Note 3**).
8. Optional: Define the core length of the respective CRISPR ortholog and the number of total core mismatches between guide and target site to be considered. Recommendation: seven for core length and two for core mismatches (*see* Subheading 1.4) (*see* **Note 3**).
9. Select the target genome and submit the request.
10. Choose the guide(s) with the best efficacy score/off-target ratio (*see* **Note 4**). Compare the results with Cas-Designer to identify the optimal guide(s).

3.1.2 CRISPR RGEN Tools: Cas-Designer

1. Go to http://www.rgenome.net/cas-designer/ to open the Cas-Designer tool.
2. Select the PAM type. The PAM depends on the employed CRISPR/Cas-system and respective Cas ortholog (e.g., 5′-NGG-3′ for *Streptococcus pyogenes* Cas9).
3. Select the target genome.
4. Enter your target gene/sequence of interest into the query sequence field. For knockout experiments, concentrate on the exons of the gene (*see* Subheading 1.8).
5. Select the guide length (here: crRNA length) of the respective CRISPR ortholog (*see* **Note 1**).
6. Optional: Allow integration of 1 nt bulge for off-target analysis (*see* Subheading 1.4). However, this is not recommended in this step (*see* **Note 5**).
7. Submit the request.

8. Select a GC content of 25–75% and submit the filter (*see* Subheading 1.3).
9. Choose the guide(s) with a "out-of-frame score" above 66 (*see* **Note 6**) and with no potential mismatch targets (*see* **Note 7**). Compare with CCTop results to identify the optimal guide(s).

3.2 Extended Off-Target Analysis

Online tools for guide sequence selection often include mismatches only for off-target prediction. However, DNA and RNA bulges also contribute to off-target activity. Therefore, an extended off-target analysis using tools including these criteria might be useful to improve off-target prediction even further.

3.2.1 CRISPR RGEN Tools: Cas-OFFinder

1. Go to http://www.rgenome.net/cas-offinder/ to open the Cas-OFFinder tool.
2. Select the PAM type. The PAM depends on the employed CRISPR/Cas-system and respective Cas ortholog (e.g., 5′-NGG-3′ for *Streptococcus pyogenes* Cas9).
3. Select the target genome.
4. Enter the guide sequence(s) into the query sequence field. The guide sequence(s) equal the guide(s) selected with CCTop/Cas-Designer.
5. Select the number of total mismatches and the DNA/RNA bulge size between guide and target site to be considered. Recommendation: Select three for number of total mismatches, two for DNA bulge, and one for RNA bulge (*see* Subheading 1.4) (*see* **Note 8**).
6. Submit the request.
7. Choose the guide(s) with no off-target sites or at least high discrepancy to the predicted off-target sites (*see* **Notes 3** and **8**).

3.3 Analysis of gRNA Secondary Structure

The secondary structure of the gRNA considerably contributes to overall activity. Being strongly affected by the variable guide sequence, verification of the secondary structure is inevitable.

3.3.1 RNAfold

1. Go to http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi to open the RNAfold tool.
2. Enter the complete gRNA sequence including the guide (s) determined by the previous analysis into the sequence query field. The guide is upstream of the gRNA backbone.
3. Keep the default settings.
4. Submit the request.

5. Analyze the predicted gRNA secondary structure by comparing it to its optimal structure (Fig. 1) (*see* Subheading 1.5) (*see* **Note 9**).
6. Guide(s) of suitable gRNAs can subsequently be used for cloning of the CRISPR constructs and experimental procedure.

4 Notes

1. The optimal guide length depends on the employed CRISPR/Cas-system, respective Cas ortholog, as well as experimental system. The most commonly used SpCas9 and SaCas9 show solid efficiency among the majority of experimental systems with a 20-nt guide.
2. Cloning of the guide is commonly realized by its synthesis as oligonucleotides with subsequent oligonucleotide annealing and ligation into the linearized gRNA expression vector. The addition of 5′-overhangs to the oligonucleotides enables sticky-end cloning which guarantees integration of the guide in the correct orientation.
3. In general, a total amount of $\geq$4 mismatches between guide and target site or $\geq$2 mismatches between guide and target site within the first seven nucleotides proximal to the PAM is sufficient to prevent the majority of cleavage activity. Thus, if the amount of total mismatches is $\leq$4, at least two mismatches should be inside this so-called core region to prevent off-target activity. The "core" region is defined as the region being the most sensitive toward mismatches.
4. A high efficacy score is desirable. If off-target sites exist, $\geq$4 mismatches in total or $\geq$2 mismatches in the "core" region prevent the majority of cleavage. Select the target(s) that have the highest efficacy score while showing the lowest likelihood for off-target activity.
5. This option is only available for the Cas-Designer tool and consequently complicates the comparison between the CCTop and Cas-Designer results in this step.
6. The "out-of-frame score" describes the likelihood for the emergence of out-of-frame mutations caused by the microhomology-mediated end-joining pathway. Out-of-frame mutations are desired for knockout experiments.
7. Targets with two mismatches within the "core" region can still be selected.
8. Prevent selecting targets that show potential off-target sites by only having one 1 nt DNA bulge in total or one 1 nt RNA bulge outside the "core" region. One 2 nt DNA bulge in total

or one 1 nt RNA bulge within the "core" region is sufficient to prevent cleavage. If at least three mismatches in total are present, additional bulges should abolish cleavage.

9. Efficient gRNAs show only few interactions between the guide sequence and gRNA backbone. Furthermore, intact stem-loop structures are crucial for high activity.

5 Conclusion

Many considerations and useful tools are available to aid for the selection of suitable CRISPR targets. However, a considerable inconsistency is reported between experimental systems and currently available prediction tools are far from predicting gRNA efficiency with high fidelity. While transient protoplast assays can also give relatively high levels of confidence, experimental validation inducing heritable changes in individuals remains the only way to achieve certainty regarding gRNA efficiency.

References

1. Urnov FD (2018) Genome editing B.C. (Before CRISPR): lasting lessons from the "old testament". CRISPR J 1(1):34–46. https://doi.org/10.1089/crispr.2018.29007.fyu
2. Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821. https://doi.org/10.1126/science.1225829
3. Doudna JA, Charpentier E (2014) Genome editing the new frontier of genome engineering with CRISPR-Cas9. Science 346 (6213):1258096. https://doi.org/10.1126/science.1258096
4. Knott GJ, Doudna JA (2018) CRISPR-Cas guides the future of genetic engineering. Science 361(6405):866–869. https://doi.org/10.1126/science.aat5011
5. Schindele P, Wolter F, Puchta H (2018) Transforming plant biology and breeding with CRISPR/Cas9, Cas12 and Cas13. FEBS Lett 592(12):1954–1967. https://doi.org/10.1002/1873-3468.13073
6. Kumlehn J, Pietralla J, Hensel G et al (2018) The CRISPR/Cas revolution continues: from efficient gene editing for crop breeding to plant synthetic biology. J Integr Plant Biol 60 (12):1127–1153. https://doi.org/10.1111/jipb.12734
7. Cebrian-Serrano A, Davies B (2017) CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. Mamm Genome 28 (7–8):247–261. https://doi.org/10.1007/s00335-017-9697-4
8. Fu Y, Foden JA, Khayter C et al (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
9. Hsu PD, Scott DA, Weinstein JA et al (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31(9):827–832. https://doi.org/10.1038/nbt.2647
10. Fu Y, Sander JD, Reyon D et al (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 32(3):279–284. https://doi.org/10.1038/nbt.2808
11. Lin Y, Cradick TJ, Brown MT et al (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res 42(11):7473–7485. https://doi.org/10.1093/nar/gku402
12. Ren X, Yang Z, Xu J et al (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
13. Zhang J-P, Li X-L, Neises A et al (2016) Different effects of sgRNA length on CRISPR-

mediated gene knockout efficiency. Sci Rep 6:28566. https://doi.org/10.1038/srep28566

14. Sugano SS, Nishihama R, Shirakawa M et al (2018) Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. PLoS One 13(10):e0205117
15. Peterson BA, Haak DC, Nishimura MT et al (2016) Genome-wide assessment of efficiency and specificity in CRISPR/Cas9 mediated multiple site targeting in Arabidopsis. PLoS One 11(9):e0162169. https://doi.org/10.1371/journal.pone.0162169
16. Feng C, Su H, Bai H et al (2018) High-efficiency genome editing using a dmc1 promoter-controlled CRISPR/Cas9 system in maize. Plant Biotechnol J 16(11):1848–1857. https://doi.org/10.1111/pbi.12920
17. Lee K, Zhang Y, Kleinstiver BP et al (2019) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. Plant Biotechnol J 17(2):362–372. https://doi.org/10.1111/pbi.12982
18. Tang X, Liu G, Zhou J et al (2018) A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. Genome Biol 19(1):84. https://doi.org/10.1186/s13059-018-1458-5
19. Zhang Q, Xing H-L, Wang Z-P et al (2018) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. Plant Mol Biol 96 (4-5):445–456. https://doi.org/10.1007/s11103-018-0709-x
20. Doench JG, Hartenian E, Graham DB et al (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32 (12):1262–1267. https://doi.org/10.1038/nbt.3026
21. Wang T, Wei JJ, Sabatini DM et al (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343 (6166):80–84. https://doi.org/10.1126/science.1246981
22. Liang G, Zhang H, Lou D et al (2016) Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Sci Rep 6:21451. https://doi.org/10.1038/srep21451
23. Liu X, Homma A, Sayadi J et al (2016) Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. Sci Rep 6:19675. https://doi.org/10.1038/srep19675
24. Ma X, Zhang Q, Zhu Q et al (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol Plant 8 (8):1274–1284. https://doi.org/10.1016/j.molp.2015.04.007
25. Xu H, Xiao T, Chen C-H et al (2015) Sequence determinants of improved CRISPR sgRNA design. Genome Res 25 (8):1147–1157. https://doi.org/10.1101/gr.191452.115
26. Le C, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823. https://doi.org/10.1126/science.1231143
27. Zheng T, Hou Y, Zhang P et al (2017) Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. Sci Rep 7:40638. https://doi.org/10.1038/srep40638
28. Nishimasu H, Ran FA, Hsu PD et al (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156 (5):935–949. https://doi.org/10.1016/j.cell.2014.02.001
29. Konermann S, Brigham MD, Trevino AE et al (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517(7536):583–588. https://doi.org/10.1038/nature14136
30. Sander JD, Dahlborg EJ, Goodwin MJ et al (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8(1):67–69. https://doi.org/10.1038/nmeth.1542
31. Sanjana NE, Le C, Zhou Y et al (2012) A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7 (1):171–192. https://doi.org/10.1038/nprot.2011.431
32. Verkuijl SA, Rots MG (2019) The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. Curr Opin Biotechnol 55:68–73. https://doi.org/10.1016/j.copbio.2018.07.005
33. Horlbeck MA, Witkowsky LB, Guglielmi B et al (2016) Nucleosomes impede Cas9 access to DNA in vivo and in vitro. elife 5:e12677. https://doi.org/10.7554/eLife.12677
34. Wu X, Scott DA, Kriz AJ et al (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
35. Daer RM, Cutts JP, Brafman DA et al (2017) The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. ACS

Synth Biol 6(3):428–438. https://doi.org/10.1021/acssynbio.5b00299

36. Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA et al (2018) Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). PLoS One 13(4): e0196238. https://doi.org/10.1371/journal.pone.0196238
37. Jensen KT, Fløe L, Petersen TS et al (2017) Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. FEBS Lett 591 (13):1892–1901. https://doi.org/10.1002/1873-3468.12707
38. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–3415. https://doi.org/10.1093/nar/gkg595
39. Hofacker IL (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31 (13):3429–3431. https://doi.org/10.1093/nar/gkg599
40. Liu H, Ding Y, Zhou Y et al (2017) CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. Mol Plant 10 (3):530–532. https://doi.org/10.1016/j.molp.2017.01.003
41. Stemmer M, Thumberger T, Del Sol KM et al (2015) CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. PLoS One 10(4):e0124633. https://doi.org/10.1371/journal.pone.0124633
42. Bae S, Park J, 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 (10):1473–1475. https://doi.org/10.1093/bioinformatics/btu048
43. Park J, Bae S, Kim J-S (2015) Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites. Bioinformatics 31 (24):4014–4016. https://doi.org/10.1093/bioinformatics/btv537
44. Lee CM, Davis TH, Bao G (2018) Examination of CRISPR/Cas9 design tools and the effect of target site accessibility on Cas9 activity. Exp Physiol 103(4):456–460. https://doi.org/10.1113/EP086043