

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


CRISPR Guide RNA Design for Research Applications

Citation

Mohr, Stephanie, Yanhui Hu, Benjamin Ewen-Campen, Benjamin Housden, Raghuvir Viswanatha, Norbert Perrimon. "CRISPR Guide RNA Design for Research Applications." FEBS Journal 283, no. 17 (2016): 3232-3238. DOI: 10.1111/febs.13777

Published version

https://doi.org/10.1111/febs.13777

Link

https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37372803

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

The FEBS Journal
FEBS PRESS science publishing by scientists
REVIEW ARTICLE

CRISPR guide RNA design for research applications

Stephanie E. Mohr[1,2], Yanhui Hu[1,2], Benjamin Ewen-Campen[2], Benjamin E. Housden[2], Raghuvir Viswanatha[2] and Norbert Perrimon[1,2,3]

1 Drosophila RNAi Screening Center, Harvard Medical School, Boston, MA, USA
2 Department of Genetics, Harvard Medical School, Boston, MA, USA
3 Howard Hughes Medical Institute, Boston, MA, USA

Keywords
bioinformatics; CRISPRa; CRISPR-Cas9; CRISPRi; gene editing; genome annotation; genome engineering

Correspondence
S. E. Mohr, *Drosophila* RNAi Screening Center, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
Fax: +1 617 432 7688
Tel: +1 617 432 5626
E-mail: stephanie_mohr@hms.harvard.edu

(Received 15 March 2016, revised 24 May 2016, accepted 7 June 2016)

doi:10.1111/febs.13777

The rapid rise of CRISPR as a technology for genome engineering and related research applications has created a need for algorithms and associated online tools that facilitate design of on-target and effective guide RNAs (gRNAs). Here, we review the state of the art in CRISPR gRNA design for research applications of the CRISPR-Cas9 system, including knockout, activation, and inhibition. Notably, achieving good gRNA design is not solely dependent on innovations in CRISPR technology. Good design and design tools also rely on availability of high-quality genome sequence and gene annotations, as well as on availability of accumulated data regarding off-targets and effectiveness metrics.

Introduction

The essentials of CRISPR gRNA design

With previous genome engineering applications, based on zinc finger nucleases and TALENs, a pair of proteins conferred DNA sequence specificity [1,2]. With the CRISPR-Cas9 system, however, targeting of the nuclease to a particular locus relies on the guide RNA (gRNA), as well as the presence at the target DNA region of a protospacer adjacent motif (PAM) sequence [3]. To design a gRNA, the following must be defined: (a) the target region or gene; (b) the version of Cas9 protein to be used, including what PAM sequence(s) is recognized; (c) what promoter will be used for *in vitro* or *in vivo* expression of the gRNA, i.e. so that the terminator sequence for the promoter can be excluded from the gRNA sequence; and (d) the cloning strategy, which might put additional constraints on gRNA sequence and will impact the design of the specific oligonucleotides to be synthesized and either used as a template for RNA production or cloned into an expression vector.

Importantly, design of gRNAs must also be coupled with the intended application, e.g. knockout via non-homologous end joining (NHEJ), knock-in, CRISPR activation (CRISPRa), or CRISPR interference (CRISPRi) [4,5]. For example, when generating a knockout in a genetic system, the importance of limiting off-targets on the same chromosome might outweigh the importance of limiting off-targets on other chromosomes that can be 'cleaned up' in subsequent backcrosses or outcrosses. Moreover, for applications such as CRISPRa or CRISPRi, the appropriate position of gRNA(s) relative to the transcription or translation

Abbreviations
bp, base pairs; CRISPRa, CRISPR activation; CRISPR, clustered regularly interspaced short palindromic repeats; CRISPRi, CRISPR interference; gRNA, guide RNA; NHEJ, non-homologous end joining; PAM, protospacer adjacent motif; TSS, transcription start site.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

start site of the target gene might be different than the appropriate position of gRNA(s) intended to be used in generation of a knockout or knock-in allele. In addition, some approaches require the design of multiple gRNAs for a single target, such as design of a pair of gRNAs for use with a nickase version of the Cas9 protein; design of nearby or distal gRNAs for use with a knock-in construct; or design of two or more gRNAs upstream of the transcription start site for CRISPRa. A summary of design considerations for CRISPR knockout, CRISPRa, and CRISPRi approaches for *Streptococcus pyogenes* Cas9 [6] is shown in Table 1. We note that guidelines and design rules for gRNAs used with alternative Cas9 proteins such as *Staphylococcus aureus* Cas9 or Cpf1 [7,8] might be different from those presented here.

Different systems, different rules?

Alongside rapid development of CRISPR-based approaches, researchers have developed rules or algorithms for predicting specific properties, such as minimal off-targets or maximal effectiveness. Developers of off-target rules were able to borrow from approaches developed for limiting sequence-based off-targets of RNAi reagents or ensuring primer specificity [9]. The majority of off-target prediction algorithms for CRISPR are based on defining genomic sequences similar to the on-target sequence and then using a user-defined cutoff for the number of mismatched bases that are likely to prevent mutation. However, it is currently unclear how many mismatches are required to ensure mutation of an off-target site is prevented and this may vary depending on the system. In addition, it is possible that even off-target sequences that are not perfect matches, i.e. with a discontinuity in the sequence match creating a gap or a bulge, may be valid target sites [10] yet most current algorithms do not assess these. Several unbiased approaches have been developed to detect off-target cutting by CRISPR [11–13], although it is generally not feasible to apply these for each gRNA used. Nevertheless, the extent of off targets, as well as the different relevance of off-targets in approaches like CRISPRa and CRISPRi as compared with generation of NHEJ-mediated knockouts, suggest that additional and approach-specific off targets remain a concern; recent relevant reviews discussing off targets and their detection include Refs. [14–17].

Certain rules for gRNA effectiveness do not appear to always be generalizable for different Cas9 applications, or for the same application in different species. Even groups working in the same species have identified different criteria associated with effectiveness, for example, and reports of improvements relevant to one species (e.g. 3′ GG as reported for *Caenorhabditis elegans* [18]) await further validation in the same and other systems. In the case of CRISPRa and CRISPRi, researchers are working to define design rules not only relevant to the gRNA sequence but also regarding the number and position of gRNAs relative to specific annotated features of the target gene. For CRISPRa, two studies have found that activation is maximized when gRNAs fall within a window from − 400 to − 50 bp upstream of the TSS [19,20]. In contrast, CRISPRi knockdown is reportedly most effective when gRNAs target from − 50 to + 300 bp from the TSS [19], and when gRNAs do not target a portion of DNA that is bound by nucleosomes [21]. For CRISPRi, there appear to be different requirements in

Table 1. Guidelines for low-throughput gRNA design.

Approach	Common considerations (special considerations or exceptions)
CRISPR knockout via NHEJ	Target a gene-specific region (nested genes can complicate)
	Target a common exon (complex gene without a common exon or alternative translation start sites can complicate)
	Few or no off targets (for genetic systems, highest priority might be to avoid off targets on same chromosome)
	Consider frameshift likelihood [40]
	Filter with a predicted efficiency score
CRISPRa	Target a region 500–50 bp upstream of the TSS (alternative TSSs within the gene can complicate)
	Few or no off targets close to other genes (overlap with TSS-adjacent regions of another gene(s) can complicate)
CRISPRi	Target a region nearby the TSS (alternative TSSs can complicate)
	Few or no off targets close to TSS of other genes (overlap with TSS of another gene(s) can complicate)
All approaches	Avoid promoter terminator sequence
	Avoid restriction sites used for cloning the gRNA or homology arms (e.g. in the case of Gibson assembly)

The FEBS Journal (2016) © 2016 The Authors. The *FEBS* Journal published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

prokaryotes and eukaryotes: although CRISPRi in *Escherichia coli* requires that gRNAs target the non-template strand (a.k.a. the sense strand) [22], subsequent reports in eukaryotic cells have demonstrated that gRNAs binding to either strand can work [4,23,24]. Researchers are currently testing in parallel the impact of gRNA sequence, gRNA placement relative to transcription or translation start sites, and alternative Cas9 or Cas9 fusion proteins, with the goal of finding optimal parameters for on-target and robust activation or inhibition. Off-target effects are likely less of a concern than for CRISPR cutting approaches due to the limited range of effective binding sites relative to the target promoter. That is, even if binding occurs at off-target sites, it is unlikely to be within effective range of a promoter sequence and will therefore have little effect on gene expression.

Interestingly, a recently published set of design rules that greatly improve gRNA effectiveness for Cas9 cutting activity and for CRISPRi was unable to predict gRNA effectiveness for CRISPRa [25]. This suggests that separate design principals may be necessary for these two approaches. In addition, given that the effectiveness of CRISPRa is inversely correlated with the background expression level of a given gene (i.e. CRISPRa is more effective for genes that are expressed at low levels) [20,26], researchers should keep in mind that even when using an otherwise effective gRNA to activate a gene of interest, a high basal expression level will likely limit the impact CRISPRa can have. Altogether, it remains to be seen to what extent general rules will emerge for various CRISPR approaches; to what extent optimal parameters will be specific to a particular species, tissue, or cell type, in which endogenous factors might be relevant to specificity and/or effectiveness; and to what extent optimal parameters will be specific to a given delivery method (transfection, transduction, etc.), e.g. as rate-limiting steps might be different for different delivery methods.

Navigating the options

There are many tools available for gRNA design. Recent reviews listing tools include [14,27]. To help researchers identify relevant tools, lists of tools have also been compiled on web pages, including at AddGene (www.addgene.org), and a Google docs-based 'navigator' has been developed to help users find a match between their gRNA design needs and available software (http://tinyurl.com/matchcrispr). A subset of gRNA design tools developed by academic labs and supporting gRNA design for several species include E-CRISP (http://www.e-crisp.org/E-CRISP/) [28], CHOP-CHOP (https://chopchop.rc.fas.harvard.edu/) [29], CRISPR Direct (http://crispr.dbcls.jp/) [30], and CRISPR-ERA (http://crispr-era.stanford.edu/) [31]. Developers of one of these tools also recently published a protocol for gRNA design and cloning [32]. In addition to tools supporting design for multiple species, several species-specific gRNA design tools have been developed by various groups (e.g. for design of gRNAs targeting mouse or human genes, http://tinyurl.com/broadgRNA). Moreover, at least two companies, Benchling and Desktop Genomics, provide CRISPR gRNA design tools. Most tools filter out or flag gRNAs with potential sequence-specific off-targets and check for an appropriate PAM adjacent to the gRNA. Some, but not all, tools also apply filters or show 'scores' related to predicted effectiveness. In practice, researchers can use multiple tools, perhaps using one or more to build an initial list of gRNA designs, then selecting those in common, or using one tool for design and another to evaluate potential effectiveness. In addition, when designing genome-scale libraries, compromises are made in order to obtain near-full coverage. A workflow for design of gRNAs in batch mode (i.e. for several genes at a time) is shown in Fig. 1.

Importance of good gene annotations

Several online tools make it possible to view gRNA designs in the context of a genome browser, as in many cases, choosing an appropriate gRNA design is highly dependent upon the position of the gRNA relative to specific features of the gene, such as within 500–50 bp of the transcription start site (e.g. for CRISPRa) [19], nearby the transcription start site (e.g. for CRISPRi), in a common coding exon (e.g. for NHEJ-mediated knockout) [33,34], or within a specific exon, intron, protein domain-encoding sequence, or other (e.g. for knock-in). Moreover, for NHEJ-mediated knockout alleles, targeting of protein domains was recently shown to result in a higher proportion of loss-of-function mutations, presumably due to lower tolerance of amino acid deletions or substitutions within protein domains [35].

Species such as *Drosophila*, *C. elegans*, zebrafish, mouse, and rat have the support of model organism databases (MODs) that include high-quality and regularly updated gene annotations based on experimental data. The human genome is also supported by high-quality and regularly updated gene annotations, such as those presented at NCBI, Ensembl, and UCSC Genome Bioinformatics [36–38]. For any of these species, those designing gRNAs should be aware that

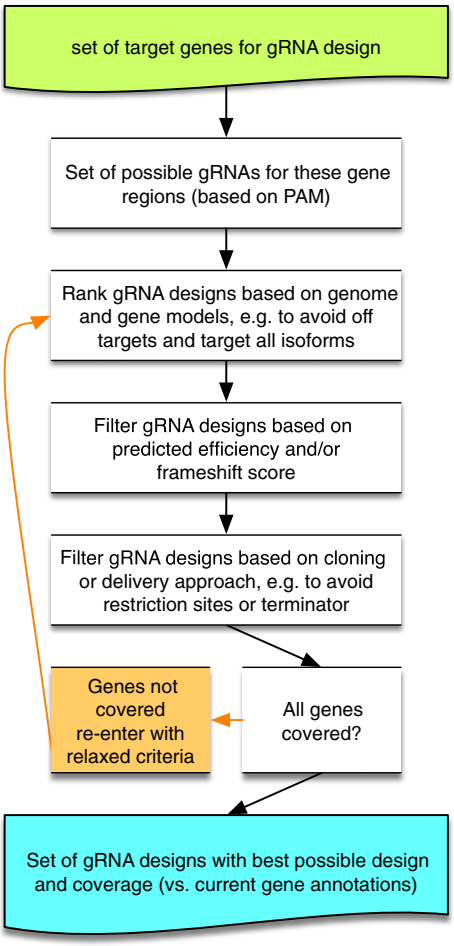


Fig. 1. Workflow for batch-mode design of gRNAs for CRISPR knockout. Starting with a set of all possible gRNA designs based on an appropriate adjacent PAM sequence, gRNAs are first ranked based on various criteria, then filtered to enrich for or avoid specific features. When a gene is not covered with the most stringent set of criteria, it re-enters the pipeline for design with relaxed criteria applied. Notably, if gene annotations are updated between the time of library design and experimental data analysis, interpretation of off targets, isoform coverage, etc. should be updated based on the current annotation in order to facilitate the most up-to-date interpretation of data.

when gene annotations are updated, the interpretation of positions of gene features such as the TSS or isoforms can change. Moreover, making updates of gRNA design tool databases can be a challenge, and users of a given gRNA design tool should be aware of how recently a given tool was updated relative to updates to MOD or human genome annotations. A specific example of the potential impact of annotation updates on gRNA design or interpretation of results is the human gene FAM83A. Information at NCBI Gene for FAM83A as viewed in May 2016 includes coordinates on a recent genome assembly (chr 8, 123,179,047–123,210,079) and a previous one (chr8, 124,194,752–124,222,318). At Ensembl, the annotated starting position is different (Chr8: 123,178,960–123,210,079 forward strand). Moreover, based on NCBI genome browser, the isoform 'FAM83A isoform a' and 'FAM83A isoform b' have different TSSs than 'FAM83A isoform c,' such that a gRNA designed for to target isoform a and b, e.g. for CRISPRa, would not be predicted to efficiently activate isoform c. In addition, the annotation of isoforms of FAM83A as presented at NCBI has changed over time, with different annotations for 'FAM83A isoform a' presented in a March 29, 2016 release (http://www.ncbi.nlm.nih.gov/nuccore/NM_032899.6) as compared with a release on November 20, 2015 (http://www.ncbi.nlm.nih.gov/nuccore/NM_032899.5). Notably, the interpretation of the transcription start site for these two annotations of 'FAM83 isoform a' are different.

Nevertheless, although gene models are far from perfect even in the best-annotated species, for species supported by MODs, as well as for the human genome, the pieces are in place to design gRNAs in an informed manner relative to the positions of specific gene features. By contrast, the lack of availability of good gene annotations for non-model species will be a roadblock to selecting 'good' gRNA designs. However, one of the exciting things about CRISPR technology is that it enables gene editing in additional species, allowing for application of genetic approaches to new systems and topics, perhaps opening the door to a significant expansion in the number and variety of 'model' species in use in the near future. What might be possible in emerging and new research species? For approaches like CRISPRa and CRISPRi, transcript data mapped to the genome would appear to be a minimum requirement for design. For knockout approaches, annotation of open reading frames (ORFs) might suffice, such as for the design of gRNAs appropriate for production of NHEJ-mediated mutant alleles of the target gene. In either case, off targets will be a concern, even when full-genome sequence is available, as the relationship between predicted off-targets and gene features will not be clear.

Importance of experimental validation

Whether working *in vivo* or in cell culture, and regardless of the species, the application of CRISPR approaches should be considered exploratory, requiring additional follow-up to confirm initial results. Best practices in a genetic system include comparing the phenotypes of independent CRISPR-generated

The FEBS Journal (2016) © 2016 The Authors. The *FEBS* Journal published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

knockouts either independently or in combination to confirm an association between a phenotype and genotype. For example, phenotypes associated with homozygotes can be compared to the phenotypes of hemizygotes or transheterozygotes (e.g. generated via multiple independent gRNA designs). In addition, testing for 'rescue' of phenotypes associated with a knockout with a wild-type transgene can be a valuable approach to validation. Best practices in a cultured cell line would include use of several unique designs per gene in high-throughput studies (e.g. pooled screens) and for low-throughput studies, production of multiple unique mutant cell lines based on different gRNA designs, and perhaps most critically, testing of transgene rescue.

Better design tools depend on availability of more data

Researchers undertaking gRNA design using an online tool developed by others should be aware of what if any rules for limiting off targets and maximizing effectiveness are being applied by a given gRNA design tool or algorithm, as well as in particular from what study, species, application, delivery method, the rules were derived. For example, for NHEJ-mediated knockout pooled screens, a recent study searching to map new essential genes used data from the first genome-scale CRISPR fitness screen in cell lines to improve gRNA selection in their subsequent screen. The authors showed that the new gRNA pool, although similar in size to the original pool, was better able to identify essential genes than the original [39]. This bootstrapping technique is likely to add greater precision to gRNA selection for knockout screening. Related to this, sharing of data regarding good designs and, importantly, sharing of information regarding poorly performing designs as well, again in association with specific species, application, etc. will help the community improve design criteria in the future. Given enough data, and provided the data are associated in particular with meta-data regarding species, approach, and delivery system, bioinformatics experts and others should be able to better understand what are the key universal or system-specific factors relevant to gRNA design.

Summary and conclusions

The CRISPR-Cas9 system offers exciting new opportunities for research using established model systems and cultured cells, as well as in non-model species. Design constraints are different for different applications, including for true genetic approaches such as knockout or knock-in, and perturbation approaches such as CRISPRa or CRISPRi. Anyone designing gRNAs using available online tools should be aware of the underlying rules for off-targets and effectiveness applied by the developers of those tools, and aware of updates to gene annotations and their quality. In the future, sharing additional data on successes and failures regarding gRNA off targets and effectiveness should improve our ability to define universal and species-specific rules for effective CRISPR gRNA design.

Acknowledgements

This work was supported by NIH NIGMS R01 GM067761, NIGMS R01 GM084947, and ORIP R24 RR032668 (N. Perrimon, PI). SEM is additionally supported in part by NCI Cancer Center Support Grant NIH 5 P30 CA06516 (E. Benz, PI). NP is an investigator of the Howard Hughes Medical Institute.

Author contributions

SEM and YH defined the concept of the review and all authors contributed to writing the manuscript.

References

1 Cathomen T & Joung JK (2008) Zinc-finger nucleases: the next generation emerges. *Mol Ther* **16**, 1200–1207.
2 Joung JK & Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49–55.
3 Wright AV, Nunez JK & Doudna JA (2016) Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44.
4 Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451.
5 Sander JD & Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347–355.
6 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.
7 Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS *et al.* (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191.
8 Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J,

van der Oost J, Regev A *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771.
9 Barrangou R, Birmingham A, Wiemann S, Beijersbergen RL, Hornung V & Smith A (2015) Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic Acids Res* **43**, 3407–3419.
10 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**, 7473–7485.
11 Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, Chang T, Huang H, Lin RJ & Yee JK (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* **33**, 175–178.
12 Kim D, Kim S, Kim S, Park J & Kim JS (2016) Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res* **26**, 406–415.
13 Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP *et al.* (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187–197.
14 Peng R, Lin G & Li J (2015) Potential pitfalls of CRISPR/Cas9-mediated genome editing. *FEBS J* **283**, 1218–1231.
15 Hartenian E & Doench JG (2015) Genetic screens and functional genomics using CRISPR/Cas9 technology. *FEBS J* **282**, 1383–1393.
16 Wu X, Kriz AJ & Sharp PA (2014) Target specificity of the CRISPR-Cas9 system. *Quant Biol* **2**, 59–70.
17 Hendel A, Fine EJ, Bao G & Porteus MH (2015) Quantifying on- and off-target genome editing. *Trends Biotechnol* **33**, 132–140.
18 Farboud B & Meyer BJ (2015) Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics* **199**, 959–971.
19 Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC *et al.* (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661.
20 Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H *et al.* (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588.
21 Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R & Weissman JS (2016) Nucleosomes impede Cas9 access to DNA *in vivo* and *in vitro*. *Elife* **5**, e12677.
22 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**, 1173–1183.
23 Ghosh S, Tibbit C & Liu JL (2016) Effective knockdown of *Drosophila* long non-coding RNAs by CRISPR interference. *Nucleic Acids Res* **44**, e84.
24 Mandegar MA, Huebsch N, Frolov EB, Shin E, Truong A, Olvera MP, Chan AH, Miyaoka Y, Holmes K, Spencer CI *et al.* (2016) CRISPR interference efficiently induces specific and reversible gene silencing in human iPSCs. *Cell Stem Cell* **18**, 541–553.
25 Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184–191.
26 Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ *et al.* (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326–328.
27 Jamal M, Khan FA, Da L, Habib Z, Dai J & Cao G (2015) Keeping CRISPR/Cas on-target. *Curr Issues Mol Biol* **20**, 1–20.
28 Heigwer F, Kerr G & Boutros M (2014) E-CRISP: fast CRISPR target site identification. *Nat Methods* **11**, 122–123.
29 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**, W401–W407.
30 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**, 1120–1123.
31 Liu H, Wei Z, Dominguez A, Li Y, Wang X & Qi LS (2015) CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics* **31**, 3676–3678.
32 Du D & Qi LS (2016) CRISPR technology for genome activation and repression in mammalian cells. *Cold Spring Harb Protoc* **2016**, pdb.prot090175.
33 Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87.
34 Wang T, Wei JJ, Sabatini DM & Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84.
35 Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB & Vakoc CR (2015) Discovery of cancer drug targets by

The FEBS Journal (2016) © 2016 The Authors. The *FEBS* Journal published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* **33**, 661–667.

36 Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM *et al.* (2014) RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res* **42**, D756–D763.

37 Herrero J, Muffato M, Beal K, Fitzgerald S, Gordon L, Pignatelli M, Vilella AJ, Searle SM, Amode R, Brent S *et al.* (2016) Ensembl comparative genomics resources. *Database (Oxford)* **2016**, doi: 10.1093/database/bav096.

38 Speir ML, Zweig AS, Rosenbloom KR, Raney BJ, Paten B, Nejad P, Lee BT, Learned K, Karolchik D, Hinrichs AS *et al.* (2016) The UCSC Genome Browser database: 2016 update. *Nucleic Acids Res* **44**, D717–D725.

39 Hart T, Chandrashekhar M, Aregger M, Steinhart Z, Brown KR, MacLeod G, Mis M, Zimmermann M, Fradet-Turcotte A, Sun S *et al.* (2015) High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526.

40 Bae S, Kweon J, Kim HS & Kim JS (2014) Microhomology-based choice of Cas9 nuclease target sites. *Nat Methods* **11**, 705–706.