Guidelines for manipulating CRISPR/Cas system

Introduction

In 2013, CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeat/ CRISPR-associated) emerges as a gene engineering technology in many species, which has made genetic modification (adding, disrupting or changing the sequence of specific genes; formally called gene editing) more efficient and easier (1). By delivering appropriate sgRNAs (small guide RNAs) and Cas9 protein into cells, any desired site in the genome can be specifically cut, and this cleavage would trigger site-specifically genomic modifications such as non-homologous end joining (NHEJ) or homology directed repair (HDR) (Fig.1). It should be noted that NHEJ is a mechanism that repairs double-strand breaks (DSBs) around the targeted sites, whereas HDR repairs the broken DNA strand(s), either nicks or DSBs, by using of a provided DNA template.

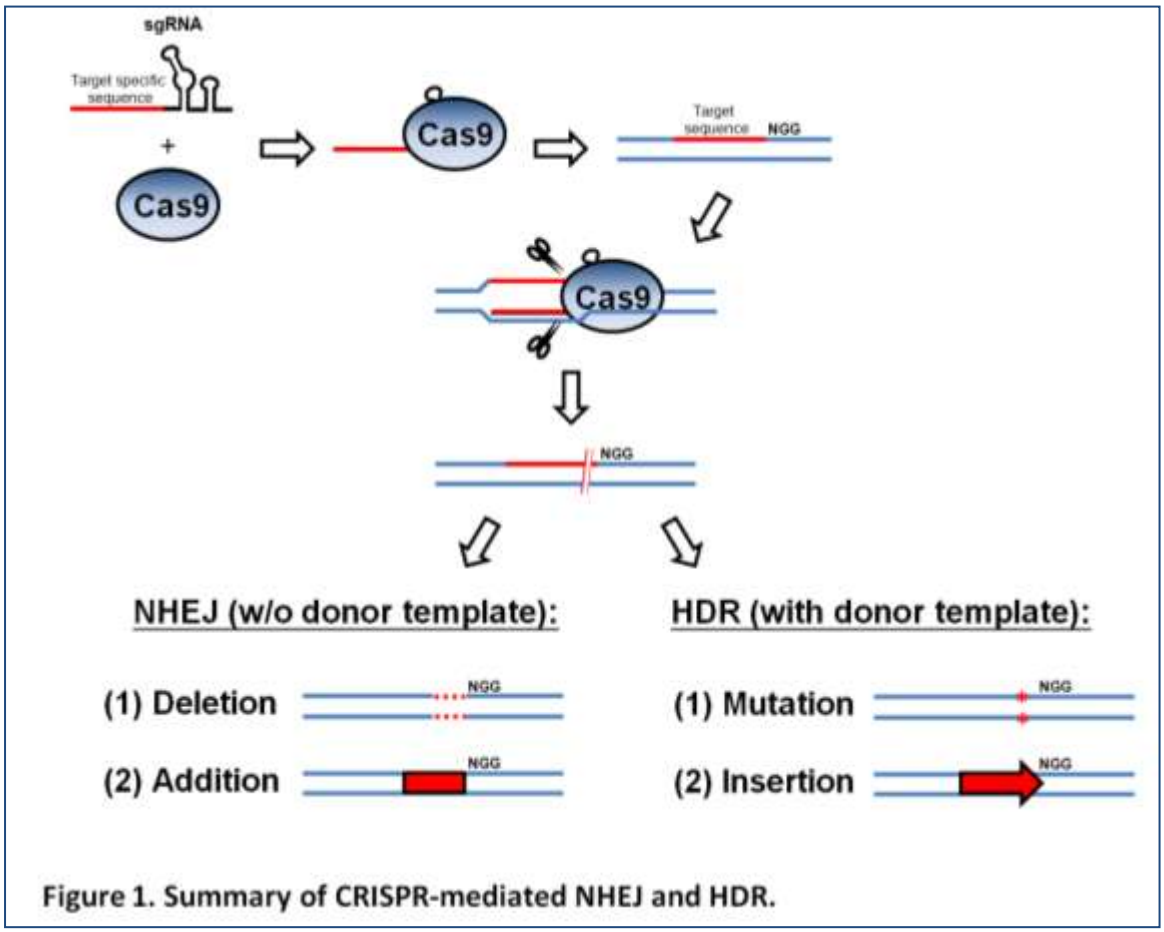

Figure 1. Summary of CRISPR-mediated NHEJ and HDR.

To mediate site-specific genome editing by CRISPR/Cas9 system, a sgRNA (combining crRNA and tracrRNA by a linker) guides the associated Cas9 to the complementary

sequence of target-DNA, which is then cleaved by Cas9-congenitial endonuclease (RuvC-like and HNH nuclease domains) (2). The only sequence constraint in the DNA target site is the presence of a protospacer adjacent motif (PAM) immediately proximal to the DNA target sequence; the sequence is NGG for *S. pyogenes* Cas9. Figure 2 shows the domain structure of Cas9 and the schematic interaction between sgRNA/Cas9 and its DNA target.

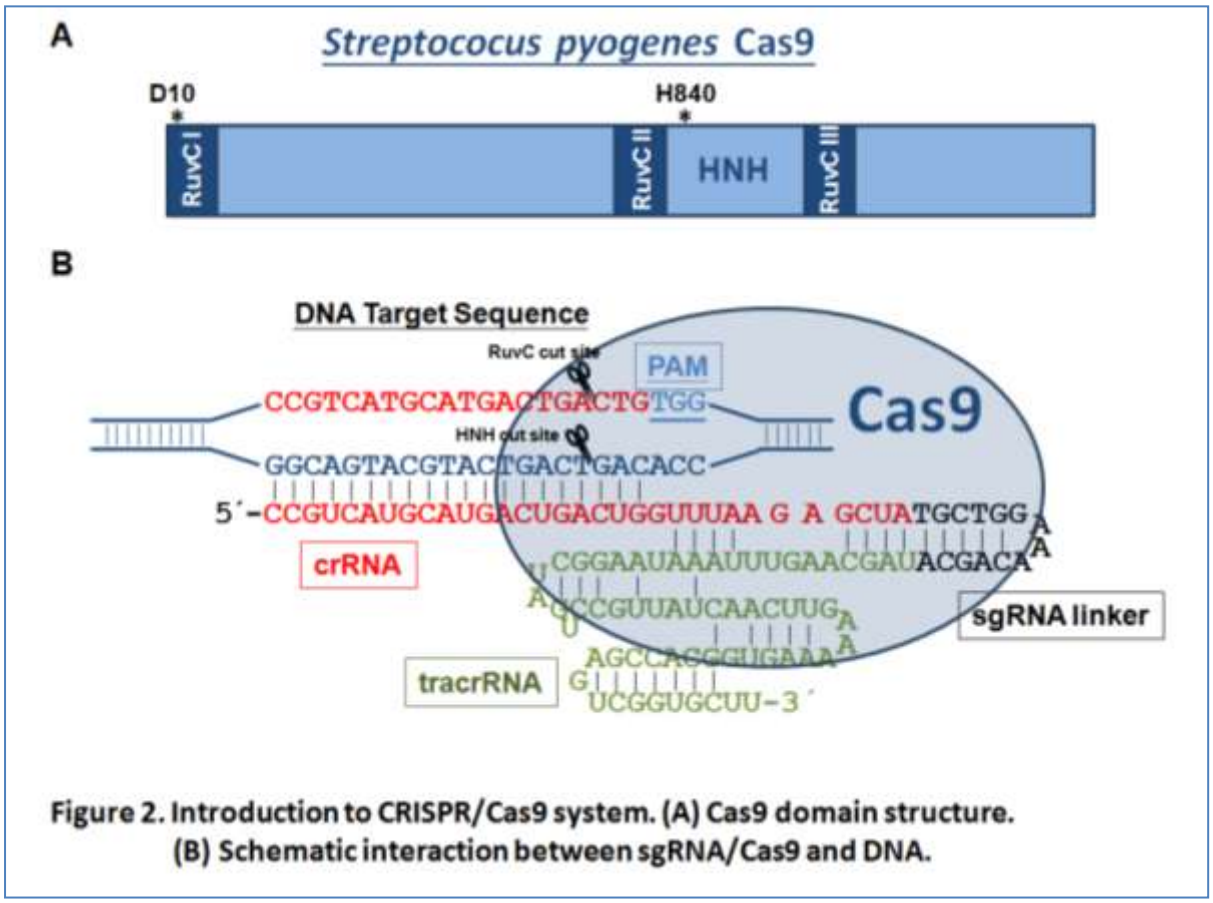

Figure 2. Introduction to CRISPR/Cas9 system. (A) Cas9 domain structure. (B) Schematic interaction between sgRNA/Cas9 and DNA.

Activation and repression of gene expression by CRISPR technique

In addition to the applications of site-specific gene editing triggered by DSB or nick (Fig. 3A-3C), several CRISPR-related applications have been developed using a catalytically inactive dCas9 (Fig. 3D-3F). Since the catalytically inactive dCas9 acts as a sgRNA-dependent DNA-binding protein, CRISPR/Cas system offers a flexible platform for site-specific gene targeting (3). Fusion of dCas9 protein to the domain of transcriptional effector, e.g., VP16/VP64 or KRAB, enables efficient transcriptional activation or repression of target gene in the cells (4). Coupling of dCas9 to fluorescent protein, e.g., GFP or mCherry, can visualize specific genomic loci in the living cells by sgRNA-targeting (5). With proper modification, CRISPR technique can even bind to a RNA target and stimulate its cleavage (6). These studies set up a framework for new applications of CRISPR/Cas system for precise gene regulation and DNA-binding in the target cells. Here are some important

tips for the design of activator-like and repressor-like effectors by using CRISPR/Cas9 system:

(i) CRISPR-mediated transcriptional activation (CRISPRa). In a previous study, active sgRNAs for CRISPRa are found to be located at -400 bp ~ -50 bp region upstream from the transcription start site (TSS) (7). Taken advantage of an optimized CRISPRa system, Konemann et al. found that the highest levels of gene activation could be achieved by targeting within the -200 bp ~ +1 bp region of the TSS (8).

(ii) CRISPR interference (CRISPRi). CRISPRi provides alternative method for gene repression at the level of transcription (by blocking transcription initiation or transcription elongation) within 1~2 weeks (9). Strong CRISPRi activity is obtained by targeting -50 bp ~ +300 bp region (optimal +50 bp ~ +100 bp downstream) of the TSS (7). Neither the DNA strand (template or non-template strand) nor gRNA G/C content correlated with the CRISPRi repression activity (7). Up to 50% of designed sgRNAs can efficiently repress gene expression by more than 70% (please see CRISPR technique supports).

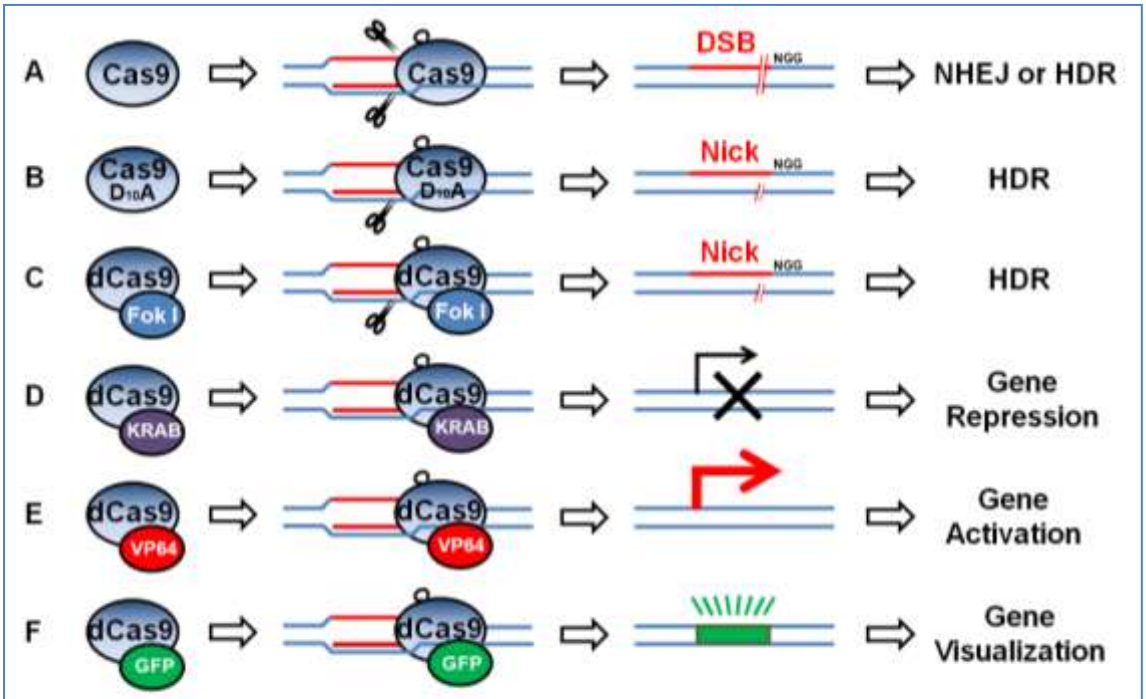

Figure 3. Applications of CRISPR technique.

On-target/ off-target effects of sgRNA and tools for design sgRNA

Many efforts have been devoted to investigate general rules for rational design of highly active sgRNAs. However, the picture of the features of active sgRNAs is not fully understood yet. Doench et al. has investigated the sequence features of active sgRNAs-targeting sites

including PAM and provided a free sgRNA on-target web tool, namely sgRNA Designer, for scoring any sequence of interest (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design) (10). For designing of a common and unique sgRNA for highly similar or identical target sequences in multiple genes, a free web-based tool called CRISPR MultiTargeter (http://www.multicrispr.net/) was developed (11). However, the efficacy of a given sgRNA needs to be validated by experimentation.

Off-target effects of CRISPR/Cas system

Accumulated data showed that sgRNA would still be able to pair with target DNA that is not perfect match; thus, it raises the concern of off-target effect by using CRISPR/Cas9 system. It has been estimated that each sgRNA might bind to 10 ~ 1,000 off-target sites in the whole genome (12). However, the binding specificity/capability of a sgRNA is highly dependent on the PAM-proximal sequence, also known as a “seed-like” region (see below) (12). Genome wide analysis showed that more than 50% active sgRNA binding sites are found in open chromatin regions, of which the highest DNA-binding frequency is located within 5’UTR, followed by exon, intron, 3’UTR and intergenic regions (12). Wild type Cas9 modifies some but not all bound target sites indicating that binding of sgRNA to its target DNA would not necessarily induce genome editing (12). By comparison of the cleavage efficiency of an active sgRNA and its binding sequence, researchers noticed that up to 5 mismatches between a sgRNA and target DNA would significantly reduce its efficiency of CRISPR-mediated cleavage (13, 14). A full complementarity between the “seed-like” region (7-12 PAM-proximal bases) of sgRNA and its target DNA was found to be critical for the binding specificity/capability of a sgRNA (15, 16). However, one mismatch within the seed region could be tolerated (17). Moreover, one bulge (1 base-skipping) formed between sgRNA and its target sequence is also well tolerated (17). Several sgRNA prediction algorisms/web tools have been developed to minimize the possible off-target effects. A list of these algorisms/web tools are showing below:

(1) Optimized CRISPR Design (http://crispr.mit.edu/) (by the Zhang Lab);

(2) CRISPRdirect (http://crispr.dbcls.jp/) (18);

(3) Cas9 online designer (http://cas9.wicp.net/);

(4) CHOPCHOP (https://chopchop.rc.fas.harvard.edu/) (19);

(5) E-CRISP (http://www.e-crisp.org/E-CRISP/designcrispr.html) (20);

(6) sgRNAcas9 (www.biootools.com) (21);

(7) CRISPRseek (22)

(http://www.bioconductor.org/packages/release/bioc/html/CRISPRseek.html);

(8) CCTop (http://crispr.cos.uni-heidelberg.de) (23);

(9) CasOT (http://eendb.zfgenetics.org/casot/) (24);

(10) Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (25);

(11) GT-Scan (http://gt-scan.braembl.org.au/gt-scan/) (26);

(12) WGE (http://www.sanger.ac.uk/htgt/wge/) (27).

Minimization of off-target effects of sgRNA

As these off-target effects might give rise to unexpected and misleading results, several tips have been suggested to minimize the off-target effects in the use of CRISPR/Cas system. They include:

(i) Select sgRNA with high target specificity:

Use sgRNA design algorisms/web tools (mentioned above) to evaluate an input DNA sequence and select sgRNA(s) with high target specificity (avoid possible off-targets at the level of genome-wide).

(ii) Use Cas9 nickase instead of wild type Cas9:

In combination of a Cas9-related nickase (Cas9 $D_{10}A$ mutant) with appropriate offset sgRNAs can efficiently generate NHEJ by paired double nicking, which is found to reduce off-target activity by 50 ~ 1,500 folds and thus enhance the specificity of genome editing (28-30). Offset (the distance between paired sgRNAs) from -4 bp to +20 bp (or -10 bp to +30 bp) between sgRNA pairs was suggested to induce robust NHEJ (28, 29), whereas offset by up to 100 bp apart showed only modest indels formation (28). Of note, sgRNA pairs creating 5' overhangs (PAM-out orientation) were able to mediate efficient genome editing (28-30). On the other hand, sgRNA pairs targeting to adjacent sites (PAM head-to-tail orientation) on the same DNA strand did not show indels by NHEJ (29).

(iii) Lower concentrations of sgRNA and Cas9:

To minimize nonspecific cleavage, the amounts of sgRNA and Cas9 can be titrated to avoid off-target effects (14).

(iv) Use optimized sgRNA:

Although most sgRNAs were designed with a 20-nt protospacer target sequence, truncated sgRNAs (with 17-nt of complementarity in length) have been shown to lower

off-target effects without affecting the on-target genome editing efficiencies (31). On the other hand, some studies showed that 20-nt sgRNAs with two extra guanines at the 5′ terminus can efficiently discriminate on-target sites from off-target sites (32, 33).

(v) Modulation of NHEJ and HDR:

For precise genome editing by promoting HDR but not NHEJ, several strategies have been examined, including (1) inhibition of NHEJ by targeting the key enzyme in the NHEJ pathway (34); (2) control the timing of cell cycle (HDR occurs only during S and G2 phase, whereas NHEJ occurs throughout the cell cycle) (35); and use of small molecular compounds to enhance precise genome engineering by HDR (36).

Perspectives of CRISPR/Cas

In sum, the programmability of CRISPR/Cas system has made gene targeting and editing much more flexible. With rapid progress in this new technology, there is no doubt that other exciting CRISPR/Cas-related applications would appear in the near future. Moreover, the CRISPR/Cas researches in mammalian cells and animals are providing new insight into clinical applications and promising a new tool for gene therapy.

Reference

1. **Doudna JA, Charpentier E.** 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science **346:**1258096.
2. **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.
3. **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.
4. **Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS.** 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell **154:**442-451.
5. **Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B.** 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell **155:**1479-1491.
6. **O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA.** 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature **516:**263-266.
7. **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:**647-661.
8. **Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib**

N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F. 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature **517:**583-588.

9. **Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS.** 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc **8:**2180-2196.
10. **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:**1262-1267.
11. **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:**e0119372.
12. **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:**677-683.
13. **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:**822-826.
14. **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:**827-832.
15. **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.
16. **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:**819-823.
17. **Wang X, Wang Y, Wu X, Wang J, 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.
18. **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.
19. **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:**W401-407.
20. **Heigwer F, Kerr G, Boutros M.** 2014. E-CRISP: fast CRISPR target site identification. Nat Methods **11:**122-123.
21. **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:**e100448.
22. **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:**e108424.

23. **Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL.** 2015. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. PLoS One **10:**e0124633.
24. **Xiao A, Cheng Z, Kong L, Zhu Z, Lin S, Gao G, Zhang B.** 2014. CasOT: a genome-wide Cas9/gRNA off-target searching tool. Bioinformatics doi:10.1093/bioinformatics/btt764.
25. **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:**1473-1475.
26. **O'Brien A, Bailey TL.** 2014. GT-Scan: identifying unique genomic targets. Bioinformatics **30:**2673-2675.
27. **Hodgkins A, Farne A, Perera S, Grego T, Parry-Smith DJ, Skarnes WC, Iyer V.** 2015. WGE: A CRISPR database for genome engineering. Bioinformatics doi:10.1093/bioinformatics/btv308.
28. **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:**1380-1389.
29. **Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, Skarnes WC.** 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods **11:**399-402.
30. **Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ, Joung JK.** 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol **32:**569-576.
31. **Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK.** 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol **32:**279-284.
32. **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:**132-141.
33. **Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI, Kim JS.** 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods **12:**237-243, 231 p following 243.
34. **Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL.** 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol **33:**538-542.
35. **Lin S, Staahl BT, Alla RK, Doudna JA.** 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife **3:**e04766.
36. **Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y, Liu H, La Russa M, Xie M, Ding S, Qi LS.** 2015. Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell **16:**142-147.