CRISPR Frequently Asked Questions

Contents

Donor repair

Which repair pathway is most commonly used to repair CRISPR mediated double stranded breaks, NHEJ or HDR?

In eukaryotic cells, in the absence of a repair template, the non-homologous end joining (NHEJ) pathway generates insertions and deletions during double-stranded break (DSB) repair. However, in the presence of a DNA template with homology to the sequences flanking the DSB location, homology-directed repair (HDR) can seal the DSB in an error free manor [1]. In most cells both of these repair pathways are active, however the HDR pathway is generally less efficient that the NHEJ. The homologous template is naturally present as sister chromatin in late S phase and G2 phase [2], but it may also be added artificially as donor DNA [3, 4].

The efficiency of HDR is determined by the concentration of donor DNA present at the time of repair, the length of the homology arms of the donor DNA, the cell cycle, and the activity of the endogenous repair systems [4]. These factors contribute to the high variability of HDR efficiency observed across cell lines particularly in immortalized cells [5]. Therefore, it is important to determine the optimal HDR conditions for each cell line.

References:
1. Iliakis G, Wang H, et al. (2004) Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenetic and genome research. 104(1–4):14–20.
2. Heyer WD, Ehmsen KT, et al. (2010) Regulation of homologous recombination in eukaryotes. Annual review of genetics. 44:113–139.
3. Elliott B, Richardson C, et al. (1998) Gene conversion tracts from double-strand break repair in mammalian cells. Molecular and cellular biology. 18(1):93–101.
4. Lin S, Staahl BT, et al. (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife. 3:e04766.
5. Hasty P, Rivera-Perez J, et al. (1991) The length of homology required for gene targeting in embryonic stem cells. Molecular and cellular biology. 11(11):5586–5591.

When designing donor DNA for use in HDR, what are the optimal lengths of the left and right homology arms, and what is the maximum size of sequence that can be efficiently inserted in mammalian cells?

DNA with homology to the sequences flanking a double-stranded break (DSB) can serve as template for error-free homology directed-repair (HDR) of the DSB. The efficiency of HDR is determined by the concentration of donor DNA present at the time of repair, the length of the homology arms, the cell cycle, and the activity of the endogenous repair systems in the particular cell [1]. These factors contribute to the high variability of HDR efficiency observed across different cell lines, especially in immortalized cells [2]. Typically, in replicating mammalian cells, donor arms are at least 500 bp in length [3]. However, it is important to determine the optimal HDR conditions for your cell line.

Inserts between the homology arms are frequently in the 1–2 kb range [4]. longer inserts are possible, but the efficiency of recombination decreases as the insert size increases [5].

Single-stranded oligo DNA (ssODN) has also recently been identified as a substrate that is preferred by the HDR mechanism, and often achieves good efficiency with homology arms as short as 40 bp [6, 7]. The drawback to using ssODNs is that they are generally limited to a couple hundred bases in length, so the insert size is limited.

References:
1. Elliott B, Richardson C, et al. (1998) Gene conversion tracts from double-strand break repair in mammalian cells. Molecular and cellular biology. 18(1):93–101.
2. Lin S, Staahl BT, et al. (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife. 3:e04766.
3. Thomas KR, Folger KR, et al. (1986) High frequency targeting of genes to specific sites in the mammalian genome. Cell. 44(3):419–428.
4. Dickinson DJ, Ward JD, et al. (2013) Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nature methods. 10(10):1028–1034.
5. Li K, Wang G, et al. (2014) Optimization of genome engineering approaches with the CRISPR/Cas9 system. PloS one. 9(8):e105779.
6. Chen F, Pruett-Miller SM, et al. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nature methods. 8(9):753–755.
7. Davis L Maizels N (2014) Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. Proceedings of the National Academy of Sciences of the United States of America. 111(10):E924–932.

Can gBlocks® Gene Fragments be used for microinjection as sgRNAs?

A recent report compared the efficiency of microinjection of DNA versus RNA in mouse embryos [1]. While it was shown that DNA is effective for producing sgRNA, in vitro transcribed RNA was observed to be more efficient. Typically, microinjections are performed using in vitro transcribed Cas9 and sgRNA rather than native dsDNA. gBlocks® Gene Fragments are ideal for use as template for in vitro transcription and will work well in these applications [2, 3].

References:
1. Horii T, Arai Y, et al. (2014) Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. Scientific reports. 4:4513.
2. Niu Y, Shen B, et al. (2014) Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell. 156(4):836–843.
3. Wang H, Yang H, et al. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 153(4):910–918.

Is the hairpin region of the gRNA always the same, and will the repetitiveness of the hairpin be a problem when constructing a gBlocks® Gene Fragment with multiple gRNAs?

The hairpin region of the gRNA is a conserved structure that binds to the Cas9 protein [1]. Creating multiple gRNAs on single gBlocks® Gene Fragment will likely create repeats that prevent accurate synthesis. It will probably be necessary to split the sequences into individual gRNA constructs and order those constructs as gene fragments. If there are any issues with sgRNA design, please contact the IDT synthetic biology group at genes@idtdna.com.

References:
1.   Sternberg SH, Redding S, et al. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 507(7490):62–67.

PAM

What is a PAM sequence and where is it located?

CRISPR/Cas9 mechanisms recognize DNA targets that are complementary to a short CRISPR sgRNA sequence. The part of the sgRNA sequence that is complementary to the target sequence is known as a protospacer. In order for Cas9 to function it also requires a specific protospacer adjacent motif (PAM) that varies depending on the bacterial species of the Cas9 gene. The most commonly used Cas9 nuclease, derived from *S. pyogenes*, recognizes a PAM sequence of NGG that is found directly downstream of the target sequence in the genomic DNA, on the non-target strand.
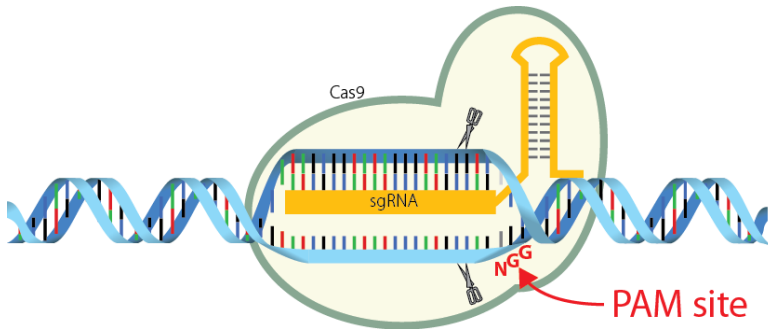
Recognition of the PAM by the Cas9 nuclease is thought to destabilize the adjacent sequence, allowing interrogation of the sequence by the sgRNA, and resulting in RNA-DNA pairing when a matching sequence is present [1, 2]. Cas9 nucleases with alternative PAMs have also been characterized and successfully used for genome editing [3]. It is important to note that the PAM is not present in the sgRNA sequence but needs to be immediately downstream of the target site in the genomic DNA.

References:
1. Anders C, Niewoehner O, et al. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 513(7519):569–573.
2. Sternberg SH, Redding S, et al. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 507(7490):62–67.
3. Esvelt KM, Mali P, et al. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods. 10(11):1116–1121.

If the target area does not contain PAM sequence, will the CRISPR/Cas9 complex bind?

No, the CRISPR/Cas9 complexes that have been characterized recognize target sequences lacking PAM sequences extremely poorly or not at all [1]. However, if the NGG sequence found in the

S. pyogenes CRISPR system is not present in your sequence, CRISPR systems from other bacterial species that recognized alternative PAM sequences work. For example, the PAM sequence for *S. thermophiles* is NNAGAA and NGGNG, and for N. meningiditis, it the PAM is NNNNGATT [2, 3].

References:
1.   Hsu PD, Scott DA, et al. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nature biotechnology. 31(9):827–832.
2.   Esvelt KM, Mali P, et al. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nature methods. 10(11):1116–1121.
3.   Ran FA, Hsu PD, et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nature protocols. 8(11):2281–2308.

How often are the PAM sequences presented in the mammalian genome in average?

Using the frequency of 'GG' = 5.21% in the reference human genome [1] there would be an expected 161,284,793 NGG PAM sites in the human genome, or roughly one 'GG' dinucleotide every 42 bases.

References:
1.   Scherer S Cold Spring Harbor Laboratory. Press., A short guide to the human genome2008, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xiv, 173 p.

Is a PAM sequence present in every gene?

This is hard to say. Because the frequency of the dinucleotide 'GG' is 5.21% in the human genome [1], it is reasonable to expect that most genes will contain the NGG sequence of the *S. pyogenes* PAM.

References:
1.   Scherer S Cold Spring Harbor Laboratory. Press., A short guide to the human genome2008, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xiv, 173 p.

Is the PAM sequence part of the sgRNA sequence construct?

The PAM sequence is located on the non-complementary strand. In other words, it is on the strand of DNA that contains the same DNA sequence as the target sgRNA [1]. The PAM sequence should not be included in the design of the sgRNA.
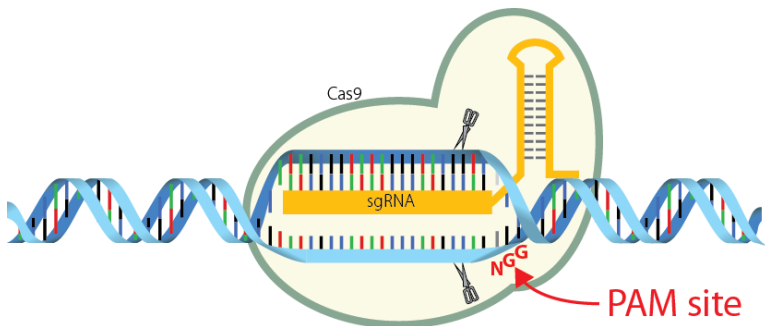
References:
1.   Anders C, Niewoehner O, et al. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 513(7519):569–573.

Cas9

If Cas9 is stably integrated into the genome, is it continually acting on the genome and causing mutations, even without the gRNA?

In cells stably expressing Cas9, it is likely that there is some nonspecific activity of the Cas9 nuclease based on observed toxicity [1, 2]. Therefore, controlling the expression of Cas9 through transient mechanisms is an important consideration [3, 4]. However, Platt et. al. (2014) also managed to produce mice that constitutively express Cas9 with no obvious abnormalities, suggesting that in at least some systems there may be little to no off target activity [5]. Cas9 activity requires a guide RNA that includes both the CRISPR gRNA and trans-activating crRNA components, to mediate its cleavage specificity. Without formation of this ternary complex, Cas9 is not thought to mediate non-specific DSB [6]. Therefore, the extent of off target activity is likely dependent on the expression level of Cas9 and the presence of RNA sequences that can act as a gRNA substrate.

References:
1. Jinek M, Chylinski K, et al. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 337(6096):816-821.
2. Gasiunas G, Barrangou R, et al. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America. 109(39):E2579–2586.
3. DiCarlo JE, Norville JE, et al. (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic acids research. 41(7):4336-4343.
4. Jiang W, Brueggeman AJ, et al. (2014) Successful transient expression of Cas9 and single guide RNA genes in *Chlamydomonas reinhardtii*. Eukaryotic cell. 13(11):1465–1469.
5. Platt Randall J, Chen S, et al. (2014) CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. Cell. 159(2):440–455.
6. Anders C, Niewoehner O, et al. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 513(7519):569–573.

How do I transfect the Cas9/gRNA?

There are a large number of methods currently used to transfer the Cas9 gene and the gRNAs into cells. Transfection methods vary depending on cell type and materials used. A recent protocol, from the lab of Dr George Church at Harvard University, includes methods for delivering CRISPR constructs into human-induced pluripotent stem cells and HEK293 cells, and is a good place to start [1]. The protocol also contains methods for assessing the success of the genome modification and isolating cells with the targeted changes.

References:
1. Yang L, Yang JL, et al. (2014) CRISPR/Cas9-Directed Genome Editing of Cultured Cells. Current protocols in molecular biology / edited by Frederick M. Ausubel, et al. 107:31 31 31-31 31 17.

How do I target multiple sites within a genome at the same time using the CRIPSR/Cas system?

There are various ways to multiplex genome editing using CRISPR/Cas9. An application in *S. cerevisiae* used a single plasmid containing the natural polycistronic CRISPR arrays containing multiple target sequences [1]. The transcribed pre-crRNA is then processed into crRNAs by mechanisms endogenous to the host cell, resulting in creation of gRNAs for multiple targets from a single construct. Multiplex genome editing in mouse ES cells was also performed by transfecting cells with 3 separate plasmids each expressing one target sgRNA or 5 PCR products that codes for one sgRNA [2].

References:
2.   Bao Z, Xiao H, et al. (2014) Homology-Integrated CRISPR-Cas (HI-CRISPR) System for One-Step Multigene Disruption in *Saccharomyces cerevisiae*. ACS synthetic biology.
3.   Wang H, Yang H, et al. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 153(4):910–918.

How can I check the off-target effects while using CRISPRs?

Three methods are commonly used to detect off target effects when using CRISPR: nuclease detection, such as the Surveyor® PLUS Mutation Detection Kit; Sanger sequencing of target and off target regions; and whole-genome or enriched-genome next generation sequencing.

Off-target activity occurs because the CRISPR/Cas9 complex lacks perfect specificity [1-3]. Because most off target effects are thought to occur at sites that share similarity to the target sequence, the genome can be screened for such sites. However, predicting and screening off target sites is complicated because the degree of similarity that will produce an off target cleavage event is not precisely known. CRISPR design tools such as the one developed by the Zhang lab and the CRISPR Design Tool from the Broad Institute both identify likely off target sites during the design process, giving researchers useful predication information.

Following CRISPR genome editing, both the target site and predicted off-target sites can be screened for sequence changes by Sanger sequencing or with the Surveyor PLUS Mutation Detection Kit. The latter detects mutations at within a PCR amplified fragment of the desired site by recognizing and cleaving mismatches to the wild type DNA. The results of the assay can be assessed by simply running the processed fragments on an agarose gel and looking for a cleavage product. However, a final verification by whole genome sequencing should be done prior to publishing or using the modified organism.

References:
1.   Hendel A, Fine EJ, et al. (2015) Quantifying on- and off-target genome editing. Trends in biotechnology. 33(2):132–140.
2.   Kim D, Bae S, et al. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nature methods. 12(3):237–243.
3.   Cho SW, Kim S, et al. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome research. 24(1):132–141.

Are there any design strategies you can use to reduce off-target effects?

Utilizing the online design tool may be the best resource to date, however we are just beginning to understand the complexities of specificity in the CRISPR system. The Zhang lab published a paper studying the off target effects of the CRISPR target sequences in 293T and 293FT cells [1]. These data are used for the algorithm that assesses off target effects and on target efficiencies for the design tool developed by the Zhang lab. The CRISPR Design Tool from the Broad Institute uses the same data as well as additional data from their in house experiments. There are several other online tools online as well [2-4]. IDTs in house research has tested the tool from the Broad Institute and found it to be good at identifying targets that have very poor on target efficiencies but it has mixed results with the prediction of targets with moderate to high efficiencies.

References:
1. Hsu PD, Scott DA, et al. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nature biotechnology. 31(9):827–832.
2. Bae S, Park J, et al. (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.
3. Naito Y, Hino K, et al. (2014) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. Bioinformatics.
4. Xiao A, Cheng Z, et al. (2014) CasOT: a genome-wide Cas9/gRNA off-target searching tool. Bioinformatics.