An algorithm for selecting highly functional and specific guide RNAs for CRISPR-Cas9 gene knockout

Louise C. Baskin, Emily M. Anderson, Shawn McClelland, Elena Maksimova, Tyler Reed, Steve Lenger, Žaklina Strezoska and Anja van Brabant Smith
Dharmacon, part of GE Healthcare, 2650 Crescent Drive, Suite #100, Lafayette, CO 80026, US

Abstract

While gene disruption using crRNAs to target one or a few genes can often be chosen in an ad hoc manner, performing high throughput loss-of-function screens requires gRNAs that have consistently high functional knockout efficiencies. In order to understand the parameters affecting CRISPR-Cas9 gene editing efficiency, we systematically transfected synthetic tracrRNA and crRNAs targeting components of the proteasome into a reporter cell line in which knockout of proteasome function results in fluorescence of a ubiquitin-EGFP fusion protein that is normally degraded by the proteasome pathway. We evaluated the functionality of >1100 crRNA sequences in this system to identify parameters that are important for DNA cleavage and subsequent functional gene disruption. Using this data, we developed and trained an algorithm to score crRNAs based on how likely they are to produce functional knockout of targeted genes. We further tested our algorithm by designing synthetic crRNAs to genes unrelated to the proteasome and examined their ability to knock out gene function using additional phenotypic assays, as well as their cleavage efficiency using next-generation sequencing analysis. Our results demonstrate that high-scoring crRNAs have increased functionality, thereby validating the CRISPR algorithm.

To augment our functionality algorithm, we developed an optimized alignment program to perform rapid, flexible, and complete specificity analysis of crRNAs, including detection of gapped alignments. Recent work has demonstrated gene editing by crRNAs containing bulges of up to four nucleotides, but most design tools are unable to detect putative off-targets based on gapped alignments. We have combined this comprehensive specificity check with our functionality algorithm to select and score highly specific and functional gRNAs for any given gene target.

High-throughput assay for functional protein knockout

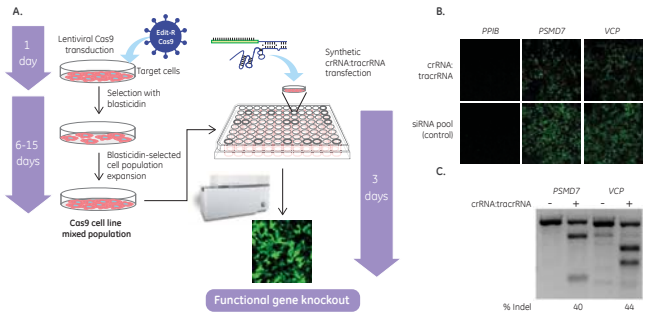


Figure 1. A. A stable Cas9-expressing recombinant U2OS cell line (Ubi(G76V)-EGFP U2OS) was generated and then transfected with synthetic crRNA:tracrRNAs (25 nM) targeting components of the proteasome. The recombinant U2OS cells constitutively express EGFP fused to a mutant ubiquitin (Gly76Val). When the proteasome is functioning, ubiquitin-EGFP is degraded (no signal), and when proteasome function is disrupted, ubiquitin-EGFP accumulates (GFP signal). **B.** Synthetic crRNAs targeting *PSMD7* and *VCP*, known components of the proteasome, show EGFP-positive cells indicating functional protein knockout. **C.** crRNAs targeting *PSMD7* and *VCP* resulted in high gene editing efficiencies estimated by a DNA mismatch detection assay using T7EI endonuclease.

crRNA functionality is position and sequence dependent

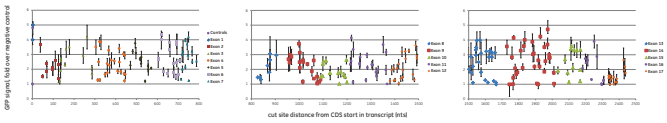


Figure 2. Ubiquitin(G76V)-EGFP U2OS cells stably expressing Cas9 were transfected with 266 synthetic crRNA:tracrRNA complexes targeting the coding region of the VCP gene. EGFP fluorescence was measured 72 hours post-transfection; an increase in EGFP fluorescence indicates functional knockout of the *VCP* gene resulting in disruption of proteasome function. crRNAs in different exons are indicated by the different colors. The data indicate that crRNAs vary in their ability to cause functional gene disruption.

crRNAs that have high functionality scores show high editing efficiency

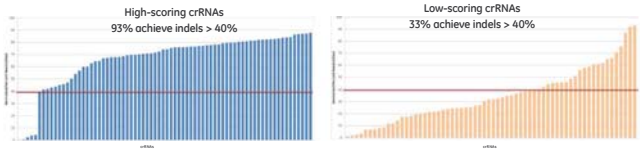


Figure 3. HEK293T-CAG-Cas9 cells were transfected with either high-scoring or low-scoring crRNAs (50 nM crRNA:tracrRNA) using DharmaFECT 1 transfection reagent (0.25 µL/well) in 96-well format. Gene editing efficiencies were determined using next-generation sequencing. 93% of the top 10 high-scoring crRNAs targeting ten different genes have > 40% indel formation and only 33% of the 10 lowest scoring designs have > 40% indel formation.

High-scoring crRNAs correlate with stronger phenotypes

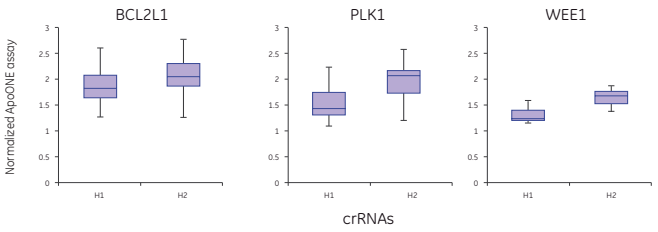
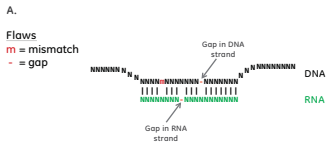


Figure 4. Box plot representation of the functionality of crRNAs targeting *BCL2L1*, *PLK1* or *WEE1* as determined by the ApoONE homogeneous assay (Promega) 48 hours after transfection. For the box plots, crRNAs were divided into bottom half (H1) and top half (H2) based on their functionality score. The medians as well as the distribution of data between the lower and upper quartile demonstrate that high-scoring crRNAs have increased functionality.

Comprehensive identification of mismatches & gaps is important for crRNA specificity



B.

Target/Off-target	Sequence	PAM
Intended target	GGTCATCTGGGAGAAAAGCG	TGG
OT1	GGTCCTCTGGGAGAAAAGACG	CAG
OT2	GGT-ATCTGGGAGAAAAGCA	TGG
OT3	GGTC-TCTGGGAGAAAAG-G	AAG

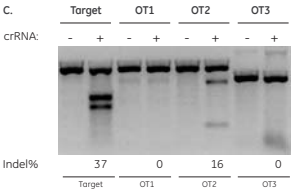
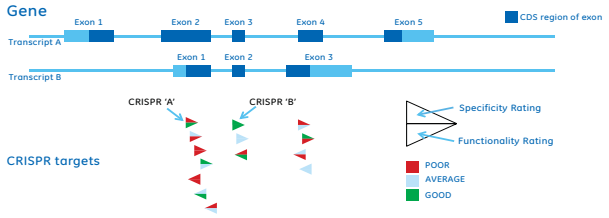


Figure 5. A. Potential off-target sites in the genome for any given crRNA include not just mismatches but gaps as well. Gaps can exist in the crRNA or the DNA target strand. The novel Dharmacon specificity tool rigorously detects both mismatches AND gaps between DNA target and guide RNA to identify all possible off-targets. **B.** Target sites with mismatches can lead to off-target editing. Ubi(G76V)-EGFP U2OS cells stably expressing Cas9 were transfected with synthetic crRNA:tracrRNA (25 nM) and genomic DNA was isolated 72 hours post-transfection. The Dharmacon specificity tool identified potential off-target (OT) sites containing mismatches and gaps in the RNA or the DNA strand that were not identified by other specificity tools. The sites were analyzed for off-target cleavage with a mismatch detection assay (T7EI). Red nucleotides indicate mismatches, red dashes indicate gaps and underlined red nucleotides indicate insertions relative to the target DNA sequence.

The optimal crRNAs balance functionality and specificity



CRISPR target 'A' – May function, but poor specificity
CRISPR target 'B' – OPTIMAL PICK, scores well in both functionality & specificity

Figure 6. Schematic of how any given crRNA may differ with regard to its specificity and functionality. Our algorithm balances these two attributes to pick the crRNAs predicted to have the highest specificity and functionality.

Conclusions

Edit-R predesigned crRNA reagents offer improved function and specificity

- The Edit-R CRISPR algorithm was trained on functional gene knockout data and can be applied across genome-wide guide RNA designs
- The Edit-R specificity tool detects gaps and mismatches to avoid potential off-targets for increased specificity

gelifesciences.com/dharmacon