Scan for PDF

Poster # 1694

ONE-seq for Variant-Aware Therapeutic Guide Selection

Douglas R. Smith, Alyssa Ferreira, Ashley Lester, Shankar Shankaracharya, Divya Ravindran, Vijetha Vemulapalli, Andrew Hollinger, Thomas E.Mullen

SeQure Dx, 1440 Main Street, Waltham, Massachusetts 02451

Summary

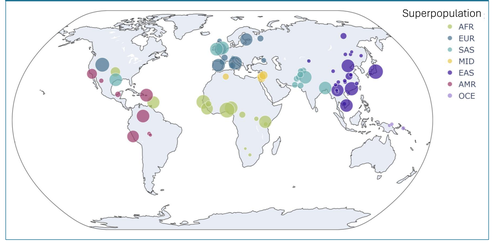
ONE-seq leverages computational tools and biochemical assays to nominate candidate off-target sites across thousands of genomes with high sensitivity. Here, we present an application of ONE-seq to identify guides with the lowest potential off-target editing risk. The biochemical *in vitro* cleavage data are augmented with biological annotation to enable prioritization of high scoring sites by their potential impact. We applied ONE-seq in a variant-aware manner to nominate off-target sites across global human populations for three therapeutically relevant guide RNAs targeting the PCSK9 gene. Comprehensive ONE-seq libraries were generated including sites with up to 6 differences relative to the on-target site by screening the HG38 reference sequence and over 4000 genomes from the 1000 Genomes and Human Genome Diversity Project data sets. Candidate off-target sites were identified by ONE-seq analysis and classified into multiple tiers based on their ONE-seq editing score and combined annotation concern score. An approach for screening multiple guides in a single run using ONE-seq Screen streamlines the guide selection process and reduces costs for analysis of larger numbers of candidate guides.

Candidate Guides

We evaluated three guide RNAs selected from 20 candidates reported by Musunuru et al., 2021. The three guides selected had the highest reported editing efficiencies and a range of MIT scores.

Name	Guide Sequence	MIT score	Editing %
PCSK9-1	CCCGCACCTTGGCGCAGCGG/TGG	90	26
PCSK9-3	GCTTACCTGTCTGTGGAAGC/GGG	66	25.4
PCSK9-4	TGCTTACCTGTCTGTGGAAG/CGG	63	23.8

ONE-seq incorporates global human genetic diversity



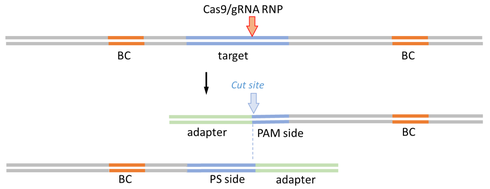
Nomination of candidate off-target sites incorporates genetic variation across diverse populations. Shown are ample collection sites for 1kG + HGDP samples (4150 genomes). Color: superpopulation (SP), dot size: number of samples.

Libraries are synthesized and cleaved *in vitro*

Comprehensive oligonucleotide libraries are synthesized, representing candidate off-target sites from over 4000 human genomes with up to 6 differences compared to the on-target site. Some sites occur multiple times in the genome.

Library >>	PCSK9-1	PCSK9-3	PCSK9-4
Reference targets	60,136	121,650	156,334
Variant targets	38,195	64,324	80,715
Total targets	98,331	185,974	237,049
Total genomic sites	505,371	220,263	279,816

Amplified libraries are subjected to *in vitro* cleavage in triplicate using complexed Cas9/gRNA at three RNP to DNA ratios (10:1, 1:1 and 0.1:1).



Schematic diagram of ONE-seq library and *in vitro* Cleavage. A pair of barcodes are uniquely associated with each candidate off-target site.

Candidate sites are scored and annotated

In vitro cleavage products are sequenced and the resulting reads (~4 million per sample) are processed using a custom analysis pipeline.

A ONE-seq score is calculated for each library member (ratio of off-target to on-target read counts). Functional annotation is added, including:

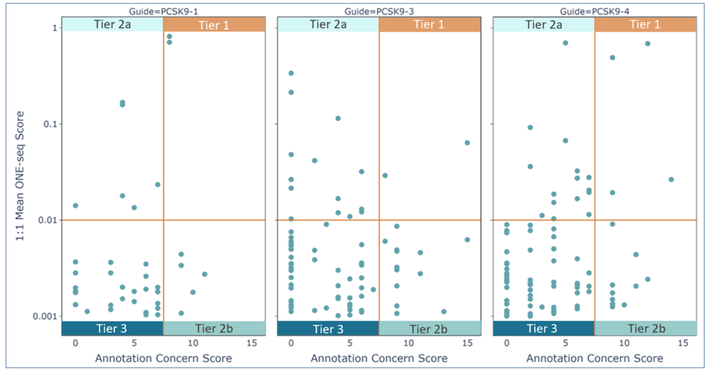
- Proximity of high-scoring sites to the on-target locus
- Variant effect analysis with gene feature, frameshift and splice site prediction
- Genetic constraint (tolerance to loss-of-function variation)
- Potential gene expression impact (ENCODE promoter and enhancer elements, LncRNAs, and GTEx tissue-specific expression profiles)
- Clinical cancer and disease gene information with classification data, CGC, ClinGen and Mondo links

A combined annotation concern score is assigned, based on component scores for each of these categories.

Candidate off-target sites are classified into multiple tiers based on their ONE-seq and combined annotation concern scores.

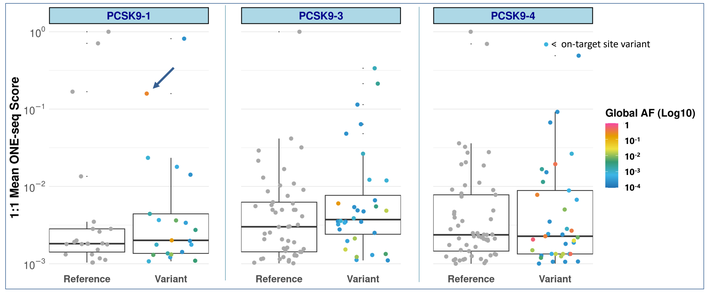
Tiered off-target editing analysis

Classifying ONE-seq results in the context of annotation concern scores enables assessment of off-target editing risk and biological impact potential. The Figure shows results including reference and variant sites from each of the three targets. Sites with higher in vitro cleavage efficiency are at the top, and sites with higher potential biological impact are towards the right.



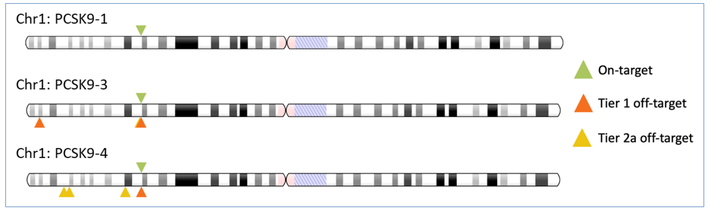
Variant off-target sites with significant AF and editing impact

A graphical summary of variant and reference sites is shown below with variants of higher global allele frequency being displayed in warmer colors. This provides a visual indication of the likelihood of encountering high scoring variant off target sites for each guide. Accompanying information on the allele frequencies in different populations permits a more detailed analysis in relation to the ethnogeographic prevalence of the disease being treated. The dot indicated by the arrow represents a variant off-target site with a global allele frequency of 26% and elevated editing of the variant allele by 874-fold. This variant occurs in a LncRNA gene, ENSG00000232325, that is expressed at significant levels in reproductive tissues and whole blood - based on GTEx data.



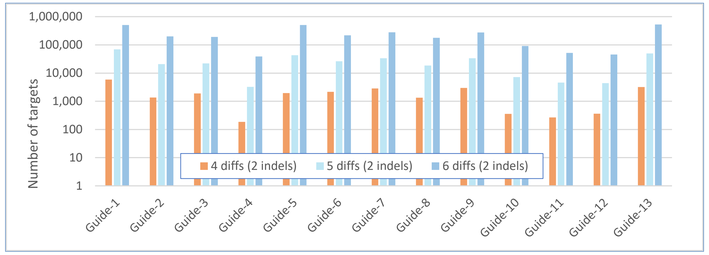
Off-target proximity to target site

High-frequency off-target editing at sites close to the on-target locus increase the probability of intrachromosomal rearrangements. The Figure below highlights candidate off-target sites with high ONE-seq scores that are present on the same chromosome as the on-target locus. PCSK9-1 has the lowest risk.



Reduced complexity libraries power ONE-seq Screen

Variant-aware ONE-seq libraries vary significantly in size depending on the guide sequence. The chart below shows 13 Cas9 libraries with various complexity cutoffs including some highly specific and highly promiscuous guides. The average size of libraries with up to 4 differences and 2 indels is approximately 1900 sites. Multiple reduced-complexity libraries such as these can be combined and processed in a single ONE-seq run for cost-effective selection of guides based on in vitro editing and annotation, as described in this poster.



Highlights

- ONE-seq enables comprehensive, variant-aware analysis for guide selection and off-target site nomination for therapeutic genome editing targets.
- Variant off-target sites with elevated *in vitro* editing efficiency and prevalence in specific global populations are common.
- Tiered analysis of ONE-seq scores combined with potential biological and clinical impact assessment enables prioritization of nominated sites for follow-up testing in cells.
- ONE-seq Screen enables cost-effective comparison of multiple guides at sites most likely to be edited in cells at high frequency

Contact: doug.smith@sequre-dx.com, thomas.mullen@sequre-dx.com | Booth 2150