Multiple Focused CRISPR Libraries from a Single Oligo Pool

ABSTRACT

Functional genetic screening using the CRISPR/Cas9 system is a powerful tool for interrogating gene function at scale, but access to and implementation of this technology is constrained by existing pre-defined CRISPR sgRNA libraries. Here, a single Twist Oligo Pool was used in combination with **C**ustom **L**ibrary M**ul**tipl**e**xed Cloning (CLUE; crispr-clue.de), an easy-to-use bioinformatics and wet-lab workflow, to generate numerous high-quality, custom sgRNA libraries. The exceptional fidelity of Twist Oligo Pools preserves the uniform representation of sgRNAs in the resulting CLUE libraries without cross-library contamination, meeting the stringent requirements for success in CRISPR screening applications.

INTRODUCTION

CRISPR/Cas9 genome engineering can be used to functionally interrogate gene function at virtually any scale. This is especially true for pooled screening applications in which sgRNAs are synthesized, cloned, delivered, and screened *en masse*. Although genome-scale sgRNA libraries enable unbiased hit identification, their vast size and sheer complexity can pose obstacles to certain experimental systems (i.e. *in vivo*) or less experienced CRISPR screeners.

Every aspect of the CRISPR screening process — from the synthesis of oligos used to construct sgRNA libraries to hit identification — requires careful consideration and optimization before initiating a screen (Doench, 2018). The complexity of CRISPR screening experiments increases with the complexity of the sgRNA library used. Highly complex sgRNA libraries can be skewed prior to a screen during library preparation by non-uniform oligo synthesis, excessive PCR amplification, and bacterial overgrowth leading to competition (Joung et al., 2017).

Focused, custom sgRNA libraries are desirable due to their increased experimental fidelity. Such libraries are also preferable for screens performed in finite or complex systems such as primary cells, organoids, and whole animals. CRISPR is increasingly being used on such systems as they better recapitulate the biological phenomena under study and, consequently, generate more biologically relevant and impactful insights (Doench, 2018). Yet, the high cell numbers required by genome-scale screens precludes the widespread use of these systems for CRISPR screening.

Improved access to custom, targeted sgRNA libraries would facilitate the application of CRISPR screening to more biologically relevant models. However, many laboratories lack the technical and bioinformatics expertise needed to generate such libraries. Thus, a workflow that couples an easy-to-use sgRNA library design tool with high-quality synthesis of oligonucleotide pools would simplify CRISPR screens for all users.

Researchers at the Helmholtz Zentrum München and the Ludwig-Maximillians University of Munich have developed CLUE (crispr-clue.de), a pipeline for efficient construction of several subgenome-scale sgRNA libraries, pooled into a single larger library (Becker et al., 2020). sgRNA oligos are flanked with layered barcodes to enable whole pool amplification and library-specific cloning for targeted CRISPR screening applications.

Starting from high-quality oligo pools, this application note shows how the CLUE workflow preserves sub-pool quality, yielding uniform sgRNA libraries virtually devoid of cross-library contamination. The high-fidelity of Twist Oligo Pools maintains sgRNA representation in PCR-amplified and cloned sub-libraries. Thus, an end-to-end pipeline for generating numerous high-quality, custom sgRNA libraries is provided.

METHODS

sgRNA library generation

Ten CLUE libraries were generated from a single Twist Oligo Pool using the CLUE pipeline, as described below.

To create a master sgRNA library for the CLUE pipeline, custom Python scripts were used to select sgRNAs from existing genome-wide sgRNA libraries (Wang et al., 2015; Doench et al., 2016; Horlbeck et al., 2016). These reference libraries include sgRNAs that span both mouse and human genomes as well as multiple CRISPR perturbation strategies (i.e., CRISPR knockout, CRISPRi, and CRISPRa). Selected sgRNA sequences were appended with promoter (U6 or H1), sgRNA scaffold, and multiple adapter sequences to enable oligo pool and library-specific amplification. The resulting oligo sequences were written to an output file that was submitted to Twist Bioscience® for oligo synthesis.

The lyophilized Twist Oligo Pool was resuspended and PCR-amplified using Kapa® Hifi DNA Polymerase (Roche) and initial amplification primers with the following parameters (PCR 1): 98°C

for 3 minutes; (98°C for 30 seconds, 62°C for 15 seconds, 72°C for 10 seconds) x 15 cycles; 72°C for 2 minutes. The PCR reaction was cloned directly using the Zero Blunt® TOPO® PCR Cloning Kit (Life Technologies Corp.). TOPO plasmids were purified, transformed in Endura® electrocompetent cells (Lucigen Corp.), and isolated. The resulting TOPO library was quality controlled using agarose gel electrophoresis and next-generation sequencing (NGS).

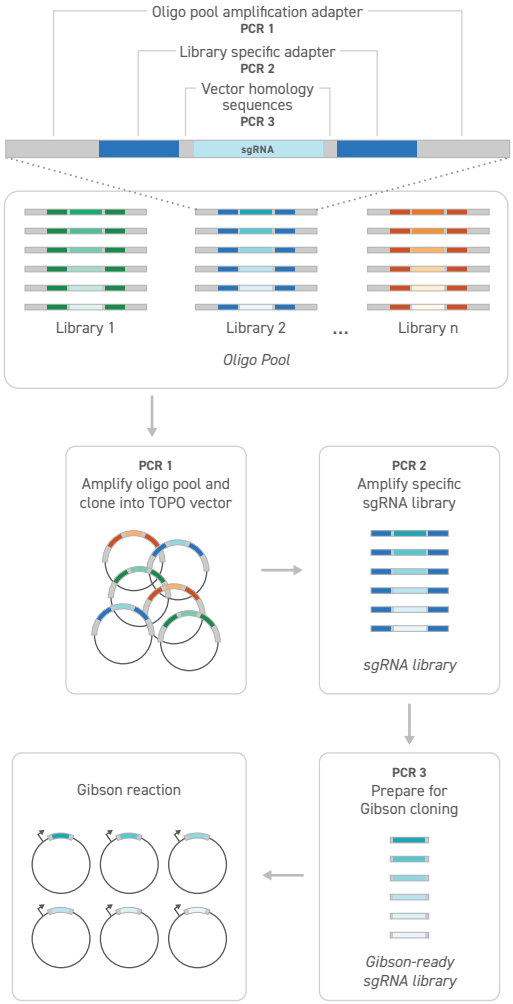
The TOPO library served as a template for amplifying each CLUE library for subsequent cloning. CLUE libraries were PCR-amplified using library-specific primers and the following thermocycling parameters (PCR 2): 98°C for 2 minutes; (98°C for 20 seconds, 57–62°C for 15 seconds, 72°C for 1 second) x 30 cycles; 72°C for 1 minute. Amplified CLUE libraries were column-purified and used to amplify cloning fragments for Gibson assembly® (PCR 3) (Synthetic Genomics, Inc.). Thermocycling conditions for PCR 3 were identical to PCR 2, always using a 62°C annealing temperature. Fragments generated by PCR 3 were cloned into linearized sgRNA expression vectors using NEBuilder HiFi DNA Assembly® Master Mix (New England Biolabs, Inc.). Cloning products were transformed in Endura electrocompetent cells and purified using NucleoBond® Xtra Maxi Kit (Macherey, Nagel, GmbH).

The resulting CLUE libraries were quality-controlled using NGS. For each library, a PCR reaction was performed for every million sgRNAs to append sequencing adapters and barcodes. Amplicons were purified and sequenced on a Illumina HiSeq® 2000 (Illumina, Inc.) (50 bp single-end reads) at a depth of >500 reads per sgRNA.

Pooled CRISPRi Screen

U-87 MG cells (ATCC® HTB-14™) were transduced with a lentivirus encoding nuclease-dead Cas9 (dCas9) tagged with enhanced green fluorescent protein (EGFP). The resulting population was enriched by fluorescence activated cell sorting (FACS) and subsequently transduced with a CLUE library at a low multiplicity of infection to ensure a single sgRNA integrant in each cell. U-87 MG dCas9-EGFP cells transduced with the CLUE library were enriched using FACS and briefly expanded in culture before initiation of the screen. Genomic DNA was harvested at t = 0 and t = 16 days using the Wizard® Genomic DNA Purification Kit (Promega). During passaging, cells were maintained at a coverage of 250 cells per sgRNA. Integrated sgRNA sequences were PCR-amplified from purified genomic DNA using Kapa Hifi DNA Polymerase (Roche) and the P5-H1 f and P7-EF1a primers. Following purification, the resulting amplicons were sequenced on a HiSeq 2000 (50 bp single-end reads). Custom Python scripts were used to map and count NGS reads, which were then analyzed using the MAGeCK algorithm.

THE CLUE WORKFLOW



STEP 1: Design a screen
- Choose organisms (e.g., homo sapiens)
- Choose gene perturbation method (CRISPRi, CRISPRa, CRISPRko)
- Choose groups of genes to be targeted or sgRNA sequences to be included
- Choose sgRNA vector
- Choose number of sgRNAs / gene

STEP 2: Upload information to CLUE (crispr-clue.de)

STEP 3: Download multiplexed oligo pool sequence information and directly submit to Twist's online ordering platform

STEP 4: Receive Oligo Pool

STEP 5: Preserve Oligo Pool and assess oligo pool quality (optional):
- Amplify total oligo pool employing the outermost primer binding sites ("PCR 1")
- Batch-Clone Amplicon into TOPO®-Vector (Life Technologies Corp.)
- NGS-Sequencing of TOPO-Vector insert to assess 90% oligo pool represented within 1 log

STEP 6: Clone individual sub-libraries:
- Perform library-specific PCR ("PCR 2")
- Prepare amplicon for cloning into sgRNA vector by performing PCR 3 with primers homologous to sgRNA vector ("PCR 3")
- Batch-clone amplicon into linearized sgRNA vector
- NGS-Sequence cloned library for quality assessment

LIBRARY NUMBER	INTENDED LIBRARY SIZE (sgRNAs)	NUMBER OF sgRNAs IDENTIFIED BY NGS	sgRNA CLONING EFFICIENCY (%)	PERFECTLY MATCHING NGS-READS (%)	READS MAPPING TO OTHER SUB-LIBRARY (%)	EMPTY VECTOR READS (%)
1	265	265	100.00	92.53	0.100	0.02
2	705	705	100.00	92.64	0.150	0.01
3	445	444	99.78	92.71	0.020	0.01
4	170	170	100.00	93.19	0.010	0.00
5	520	520	100.00	94.40	0.005	0.00
6	530	530	100.00	89.68	0.010	0.00
7	504	504	99.80	78.90	0.040	8.64
8	449	449	99.78	77.32	0.370	10.39
9	1024	1024	99.90	63.93	0.002	14.95
10	970	970	100.00	62.89	0.030	15.95

Table 1. NGS-based quality assessment of 10 pooled sgRNA libraries produced with the CLUE-pipeline.

RESULTS

Pool Uniformity During CLUE Workflow

The initial TOPO-cloned library was normally distributed, with >90% of the library represented within 1 log (Figure 1). Table 1 summarizes quality control metrics for the CLUE libraries generated from the TOPO library in this study. All CLUE libraries mapped to the original oligonucleotide pool at high rates (63–94%). sgRNAs were normally distributed in all 10 CLUE libraries. For each library, >90% of sgRNAs fell within one log of the mean of the distribution, indicating that sgRNAs were uniformly represented in every library (Figure 2). Six CLUE libraries were fully represented, and the remaining four lacked only a single sgRNA each. In every library, fewer than 0.5% of reads mapped to other libraries. Thus, the CLUE workflow can produce sgRNA libraries with uniform representation and without cross-library contamination.

A CLUE Library Identifies Transcriptional Regulators of Glioma

A targeted CRISPR interference (CRISPRi) screen was designed to test the functionality of CLUE libraries. To this end, the U87-MG glioblastoma cell line was transduced with nuclease-dead Cas9 (dCas9) to mediate CRISPRi. The dCas9-expressing U87-MG glioblastoma cell line was subsequently transduced with a CLUE library containing 505 sgRNAs targeting 88 transcriptional regulators. The representation of sgRNAs in genomic DNA isolated at days 0 and 16 of the screen were compared to identify transcriptional regulators of glioblastoma proliferation. sgRNAs targeting known regulators of oncogenesis, including Zinc finger protein 217 (ZNF217) and E2F1, were depleted at the end of the screen. These hits were confirmed using the sgRNA ranking algorithm MAGeCK (Figure 3, next page). These results highlight the ability of CLUE libraries to identify biological meaningful hits in pooled CRISPR screens.

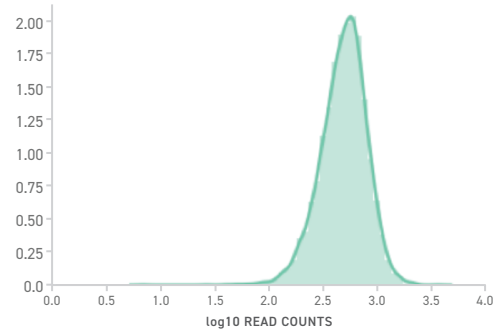
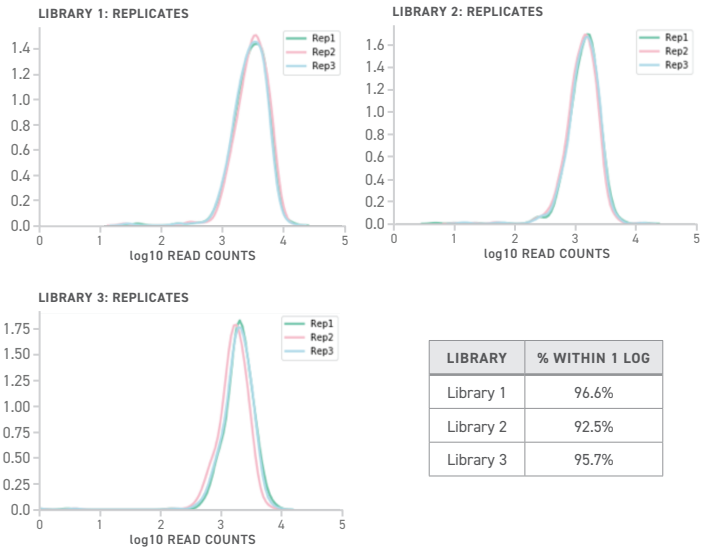


Figure 1: NGS analysis showing the distribution of oligos represented in the TOPO pool.



LIBRARY	% WITHIN 1 LOG
Library 1	96.6%
Library 2	92.5%
Library 3	95.7%

Figure 2. The distribution of libraries 1–3 from Table 1 showing that tight uniformity is maintained from the TOPO library.

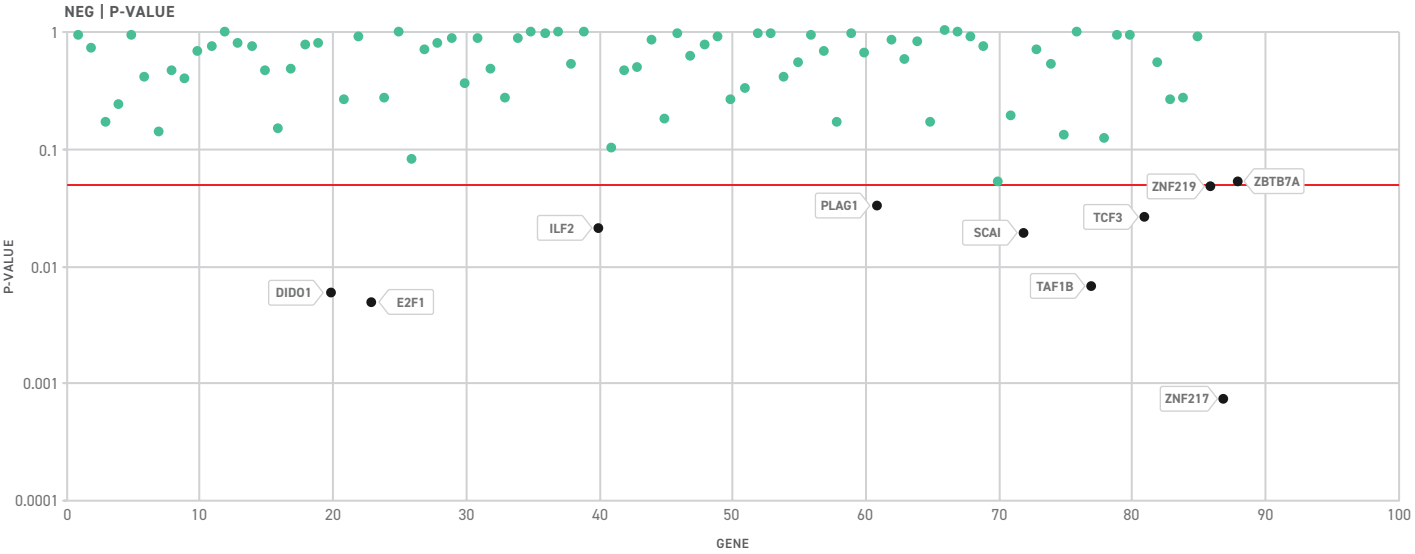


Figure 3. A targeted CRISPRi screen highlights transcription factors required for the proliferative fitness of malignant glioma cells. Significance of depletion after 16 days is plotted for targeted genes. Red line: p=0.05. Genes depleting with p<0.05 are indicated with black data points.

CONCLUSIONS

CRISPR screening is a highly scalable and effective tool for interrogating biological pathways and processes. In practice, however, genome-scale CRISPR screens can be difficult to implement. Although more focused sgRNA libraries are easier to use, few tools exist to generate such libraries, posing another hurdle for users interested in implementing this technology.

Access to multiplexed, focused libraries is beneficial for both cost and efficiency by allowing several experiments to be performed from a single pool. The pipeline described here provides an easy-to-use solution that enables the multiplexed generation of multiple targeted CRISPR libraries. CLUE allows users to customize the library size, target species, and perturbation strategy (i.e., CRISPR knockout, CRISPRi, or CRISPRa) to meet the exact requirements of their experiments.

Identifying hits in CRISPR screens occurs by comparing changes in sgRNA representation before and after a selection process. Thus, sequence fidelity and sgRNA uniformity are key parameters that must be assessed during library preparation. Poor oligonucleotide synthesis can skew sgRNA representation across a library and introduce sgRNA sequences that elicit unintended or no gene perturbation effects. These effects are exacerbated in cell-limited screens (e.g., *in vivo* screens) in which cell numbers and sgRNA representation are more difficult to monitor. The CLUE libraries generated in this study exhibited uniform sgRNA distributions and high mapping rates to the initial oligonucleotide pool from which they were derived. When CLUE is coupled with the extraordinary quality of Twist oligo synthesis, highly uniform sub-libraries can be extracted from a larger pool, allowing researchers to easily design and realize libraries of sgRNAs ready to use in high fidelity targeted CRISPR screens.

REFERENCES

Becker M, Noll-Puchta H, Amend D, Nolte F, Fuchs C, Jeremias I, Braun CJ (2020) CLUE: a bioinformatic and wet-lab pipeline for multiplexed cloning of custom sgRNA libraries. Nucleic Acids Res 48:e78.

Doench JG (2018) Am I ready for CRISPR? A user's guide to genetic screens. Nat Rev Genet 19:67–80.

Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 34:184–191.

Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, Fields AP, Park CY, Corn JE, Kampmann M, Weissman JS (2016) Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. Elife 5.

Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, Sanjana NE, Zhang F (2017) Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc 12:828–863.

Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM (2015) Identification and characterization of essential genes in the human genome. Science 350:1096–1101.