Next-generation sequencing of the amplified sgRNA library to determine sgRNA distribution

Timing 3–5 d

32

*Library PCR for NGS*. We have provided NGS primers that amplify the sgRNA target region with Illumina adaptor sequences (Table 3). To prepare the sgRNA library for NGS, set up a reaction for each of the 10 NGS-Lib-Fwd primers and 1 NGS-Lib-KO-Rev or NGS-Lib-SAM-Rev barcode primer as follows:

**Table 3 Primer sequences for amplifying sgRNA library and NGS.**

From: Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening

The 10 different NGS-Lib-Fwd primers each contain 1–10 additional nucleotides after the Illumina Read 1 Sequencing primer designed to increase the diversity of the NGS library. All 10 NGS-Lib-Fwd primers are required when amplifying the sgRNA library for NGS. The eight NGS-Lib-Rev primers provide unique barcodes (bold) for distinguishing different sgRNA libraries (e.g., experimental and control conditions for four screening bioreps) in a pooled sequencing run, and the number of NGS-Lib-Rev primers needed depends on the number of different sgRNA libraries to be sequenced together in a pooled run. As the sgRNA backbone is different between the GeCKO and SAM libraries, separate NGS-Lib-KO-Rev and NGS-Lib-SAM-Rev primers have been provided for each library.

**Table 20**

| **Component** | **Amount per reaction (μl)** | **Final concentration** |
| --- | --- | --- |
| NEBNext High Fidelity PCR Master Mix, 2× | 25 | 1× |
| Pooled sgRNA library template from Step 31 | 1 | 0.4 ng μl−1 |
| NGS-Lib-Fwd primer (unique) | 1.25 | 0.25 μM |
| NGS-Lib-KO-Rev or NGS-Lib-SAM-Rev primer (barcode) | 1.25 | 0.25 μM |
| UltraPure water | 21.5 |  |
| Total | 50 |  |

**CRITICAL STEP**

Using a different reverse primer with a unique barcode for each library allows for pooling and sequencing of different libraries in a single NextSeq or HiSeq run. This is more efficient and cost-effective than running the same number of libraries on multiple MiSeq runs.

**CRITICAL STEP**

To minimize error in amplifying sgRNAs, it is important to use a high-fidelity polymerase such as NEBNext. Other high-fidelity polymerases, such as PfuUltra II (Agilent) or Kapa HiFi (Kapa Biosystems), may be used as a substitute.

33 Perform a PCR using the following cycling conditions:

**Table 21**

| **Cycle number** | **Denature** | **Anneal** | **Extend** |
| --- | --- | --- | --- |
| 1 | 98 °C, 3 min |  |  |
| 2–23 | 98 °C, 10 s | 63 °C, 10 s | 72 °C, 25 s |
| 24 |  |  | 72 °C, 2 min |

34 After the reaction is complete, pool the PCR reactions and purify the PCR product by using the QIAquick PCR Purification Kit according to the manufacturer's directions.

35 Quantify the purified PCR product and run 2 μg of the product on a 2% (wt/vol) agarose gel. Successful reactions should yield a ∼260- to 270-bp product for the knockout library and a ∼270- to 280-bp product for the activation library. Perform gel extraction using the QIAquick Gel Extraction Kit according to the manufacturer's directions.

**PAUSE POINT**

Gel-extracted samples can be stored at −20 °C for several months.

36 Quantify the gel-extracted samples using the Qubit dsDNA HS Assay Kit according to the manufacturer's instructions.

37 Sequence the samples on the Illumina MiSeq or NextSeq according to the Illumina user manual, with 80 cycles of read 1 (forward) and 8 cycles of index 1. We recommend sequencing with a 5% PhiX control on the MiSeq or a 20% PhiX control on the NextSeq to improve library diversity; we recommend aiming for a coverage of >100 reads per sgRNA in the library.

38 *Analyze sequencing data with count\_spacers.py*. We provide a Python script for analyzing the NGS results for the sgRNA representation ([Supplementary Data 3](https://www.nature.com/articles/nprot.2017.016#MOESM436)). Install Python 2.7 (<https://www.python.org/downloads/>) and biopython (<http://biopython.org/DIST/docs/install/Installation.html>). Prepare a .csv file containing the guide spacer sequences, with each line corresponding to one sequence.

39 To determine the spacer distribution, run python count\_spacers.py with the following optional parameters:

**Table 22**

| **Flag** | **Description** | **Default** |
| --- | --- | --- |
| '-f' | .fastq file containing NGS data for analysis | NGS.fastq |
| '-o' | Output .csv file with guide spacer sequences in the first column and respective read counts in the second column | library\_count.csv |
| '-i'′ | Input .csv file with guide spacer sequences | library\_sequences.csv |
| '-no-g' | Indicates the absence of a guanine before the guide spacer sequence | Guanine is present |

After running count\_spacers.py, spacer read counts will be written to an output .csv file. Relevant statistics, including the number of perfect guide matches, nonperfect guide matches, sequencing reads without key, number of reads processed, percentage of perfectly matching guides, percentage of undetected guides, and skew ratio will be written to statistics.txt. An ideal sgRNA library should have >70% perfectly matching guides, <0.5% undetected guides, and a skew ratio of less than 10.

**CRITICAL STEP**

The human SAM libraries do not have a guanine before the guide spacer sequence, so make sure to run the script with the parameter -no-g when analyzing those libraries.

| **Primer** | **Sequence (5′–3′)** | **Purpose** |
| --- | --- | --- |
| NGS-Lib-Fwd-1 | AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GATCTTAAGTAGAGGCTTTATATATCT TGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-2 | AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GATCTATCATGCTTAGCTTTATATATC TTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-3 | AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC GATCTGATGCACATCTGCTTTATATAT CTTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-4 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTCGATTGCTCGACGCTTTATATA TCTTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-5 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTTCGATAGCAATTCGCTTTATAT ATCTTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-6 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTATCGATAGTTGCTTGCTTTATA TATCTTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-7 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTGATCGATCCAGTTAGGCTTTAT ATATCTTGTGGAAAGGACGAAACACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-8 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTCGATCGATTTGAGCCTGCTTTA TATATCTTGTGGAAAGGACGAAACAC C | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-9 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTACGATCGATACACGATCGCTTT ATATATCTTGTGGAAAGGACGAAACA CC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-Fwd-10 | AATGATACGGCGACCACCGAGATCTA CACTCTTTCCCTACACGACGCTCTTCC GATCTTACGATCGATGGTCCAGAGCTT TATATATCTTGTGGAAAGGACGAAAC ACC | GeCKO or SAM sgRNA library NGS |
| NGS-Lib-KO-Rev- 1 | CAAGCAGAAGACGGCATACGAGAT**TC GCCTTG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 2 | CAAGCAGAAGACGGCATACGAGAT**AT AGCGTC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 3 | CAAGCAGAAGACGGCATACGAGAT**GA AGAAGT**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 4 | CAAGCAGAAGACGGCATACGAGAT**AT** **TCTAGG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 5 | CAAGCAGAAGACGGCATACGAGAT**CG TTACCA**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 6 | CAAGCAGAAGACGGCATACGAGAT**GT CTGATG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 7 | CAAGCAGAAGACGGCATACGAGAT**TT ACGCAC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-KO-Rev- 8 | CAAGCAGAAGACGGCATACGAGAT**TT GAATAG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGTGCC ACTTTTTCAA | GeCKO sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-1 | CAAGCAGAAGACGGCATACGAGAT**TC GCCTTG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-2 | CAAGCAGAAGACGGCATACGAGAT**AT AGCGTC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-3 | CAAGCAGAAGACGGCATACGAGAT**GA AGAAGT**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-4 | CAAGCAGAAGACGGCATACGAGAT**AT TCTAGG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-5 | CAAGCAGAAGACGGCATACGAGAT**CG TTACCA**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-6 | CAAGCAGAAGACGGCATACGAGAT**GT CTGATG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-7 | CAAGCAGAAGACGGCATACGAGAT**TT ACGCAC**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |
| NGS-Lib-SAM-Rev-8 | CAAGCAGAAGACGGCATACGAGAT**TT GAATAG**GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTGCCAAGTTGATAA CGGACTAGCCTT | SAM sgRNA library NGS and barcode (bold) |