To explore the potential of the CLERA method for cloning short fragments in a pooled manner, sgRNA vectors were cloned in pool using the CLERA3 method (Figure 6A). Previously, researchers used a multichannel pipette and PCR strip-cap tubes to clone many sgRNAs at the same time (Nageshwaran et al., 2018). However, an oligo duplex pool can be used to construct many sgRNAs in one reaction with the CELAR method. First, oligo pairs were independently annealed in PCR strip tubes and then mixed in equimolar amounts (Figure 6A). The resulting equimolar mixture of oligo duplexes was cloned into XhoI-digested lentiCRISPRv3 per the procedures of the CLEAR3 method (Figure 6A). After plasmid minipreps, Sanger sequencing was performed on the colonies to identify the inserts, which were then matched back to the original designs (Figure 6A). This minimizes labor requirements and results in significant cost savings in terms of both time and resources.

To further explore optimal conditions for pooled cloning with the CLEAR method, we conducted a Monte Carlo simulation in which n oligo duplexes were mixed and cloned and n colonies were picked and sequenced. For the sake of simplicity and generality, we made two assumptions: firstly, there are always 50 colonies in each experiment and all of them are positive. secondly, each oligo duplex has an equal probability to be cloned and picked. With n oligo duplexes mixed, n colonies will be sampled to calculate the coverage, defined as the percentage of unique sgRNAs among the sampled colonies. The sampling process was repeated 10,000 times for each given value of n. The mean coverages based on simulated sampling of 1 to 50 oligo duplexes cloned in pool were obtained (Figure 6B). The analysis reveals a positive correlation between the number of oligo duplexes (larger than 10) in the pool and the resulting coverage (Figure 6B). Notably, coverage exceeds 75% when the number of oligo duplexes is more than half of the total number of colonies (Figure 6B).

To experimentally validate the simulated sampling data, coverages were determined for different numbers of oligo duplexes cloned in pool (Figure 6C). The total number of colonies analyzed in each experiment was about 50 by adjusting the amount of competent cells used. The results demonstrated varying coverages depending on the pool size, with an increase in coverage observed as the number of oligo duplexes in the pool increased (Figure 6C). A similar trend was observed when the total number of colonies was controlled at approximately 25 in each experiment (Figure 6D). These data further confirmed the positive relationship between pool size and coverage, indicating the suitability and robustness of the CLERA3 method for cloning short fragments in a pooled manner.

To streamline data analysis of pool cloning, a custom Python script was developed (Figure 6E). The script effectively extracted DNA sequences and quality data from the ab1 files generated from Sanger sequencing. It accurately identified the inserts by confirming matching ends of the target sequences. Additionally, the script efficiently cross-referenced the inserts within a CSV dataset, providing information about the gene of interest and specific guide RNA numbers. In summary, the CLEAR method can clone short fragments in a pooled manner efficiency, with increasing coverage when the number of fragments is larger than half of the number of the total colonies in each experiment. A custom Python script streamlined data analysis by accurately identifying inserts and providing relevant gene and guide RNA information. Overall, the CLERA method demonstrated its effectiveness, scalability, and utility for molecular biology applications.

### Materials and Methods

### Monte Carlo simulation

We use a custom R script to conduct a Monte Carlo simulation. The simulation iterates over different sample sizes (n), ranging from 1 to N (26, 50, or 100 here), where N represents the total number of colonies one can get from each experiment. For each distinct sample size n, the simulation generates a pond of N samples with 1 to n. Each sample has an equal probability to be chosen into the pond. Then, n samples are picked out from the pond randomly. The proportion of unique elements in each sample are calculated. This process is repeated 10,000 times for each sample size n.