UNIVERSITY OF CALIFORNIA SAN DIEGO

Optical Pooled Screens for Identifying Perturbations with In-Situ Sequencing

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Biology

by

Benjamin Lee

Committee in charge:

Professor Eugene Yeo, Chair
Professor Brenda Bloodgood, Co-Chair
Professor Ivonne Gonzalez-Gamboa

2024

The Thesis of Benjamin Lee is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Single Guide RNA	sgRNA
In-Situ Sequencing	ISS
Sequencing by Synthesis	SBS

ACKNOWLEDGEMENTS

I would like to acknowledge my committee for their invaluable time and constructive feedback. Their critiques and guidance will continue to shape my scientific endeavors. I owe a special gratitude to Dr. Eugene Yeo and the Yeo Lab for creating a supportive environment that made the realization of our research goals possible. Gene's pioneering spirit in research has left a lasting impact on my academic path.

Deep appreciation goes to Deepak Pant, Dr. Noorsher Ahmed, and Dylan Lam for their significant contributions to this thesis project. Their assistance was crucial in overcoming numerous obstacles, and their unwavering support provided the encouragement I needed. The guidance they provided will be essential in my future endeavors.

I also want to acknowledge my friends and family, whose support has been invaluable. Without their daily encouragement, none of this would have been possible. To my parents and Your love and support have been my foundation. I carry you in my heart always, and your belief in me makes everything easier. I hope this is just the beginning of many more moments of gratitude we will share together.

Finally, I would like to express my heartfelt thanks to everyone in the Yeo Lab. My time here has been filled with fond memories, and I will always cherish the experiences and knowledge I gained during this period.

ABSTRACT OF THE THESIS

Optical Pooled Screens for Identifying Perturbations with In-Situ Sequencing

by

Benjamin Lee

Master of Science in Biology

University of California San Diego, 2024

Professor Eugene Yeo, Chair
Professor Brenda Bloodgood, Co-Chair

Understanding and manipulating biological systems at a molecular level is crucial for therapeutic drug development. Perturbations introduced through various methods allow researchers to observe effects on cellular processes, providing insights for therapeutic intervention. Old school methods like chemical treatments, transfection, and genetic

modifications often face limitations in efficiency, precision, and cost. Advances such as RNA interference (RNAi) and CRISPR/Cas9 have revolutionized gene editing, with CRISPR/Cas9 enabling precise and permanent genetic modifications. These advances allowed for a pooled approach where sgRNAs can target multiple genes within a single sample. We propose to recreate an innovative optical pooled screen method with in-situ sequencing (ISS) to directly sequence guide RNAs (sgRNAs) using the CROPseq vector, eliminating barcode assignment errors and enhancing imaging throughput. Two major experiments were conducted: identification of rolling circle products to confirm amplification of sgRNA and achieving sequencing by synthesis. Both experiments have multiple iterations that target both an endogenous gene and a transduced sequence. This method surpasses previous techniques by efficiently linking genotype to a wide range of phenotypes, providing a powerful tool for high-resolution, high-throughput genetic screening.

INTRODUCTION

The discovery and development of therapeutic drugs are fundamentally anchored in our ability to understand and manipulate biological systems at a molecular level. Perturbations, whether chemical, genetic, or environmental, play a critical role in clarifying the complex interplay between various cellular processes and pathways. Through the introduction of specific perturbations, researchers can observe resultant effects on signaling pathways, cell function, differentiation, protein expression, and cellular metabolism, thereby gaining insights essential for therapeutic intervention.

Chemical libraries were used to treat cells with specific chemicals that could affect cellular processes but struggled with identifying direct effects of the perturbations[1]. Transfection allowed for foreign DNA to be introduced thus overexpressing or suppressing genes but can unintentionally disrupt cell physiology during integration[2]. Transgenic and knockout mice can introduce or delete a gene but is very cost effective and time-consuming[3]. Antisense oligonucleotides are also a way to suppress gene expression by binding to mRNA to stop translation but often suffer with stability and effective delivery (Figure 1A)[4].

Following the discovery of RNA interference (RNAi), perturbation studies took a significant leap forward (Figure 1B). RNAi enables precise mRNA targeting for degradation, achieving gene knockdown rather than complete knockout[5]. This method provides a distinct advantage over previous techniques by using small interfering RNA (siRNA) or small hairpin RNA (shRNA) to specifically target and degrade/inhibit mRNA, allowing temporary suppression of gene expression without permanently altering the genome. Additionally, RNAi developed

from primarily arrayed screens to pooled screens which cut time and cost while simultaneously testing different perturbations in a single well[6].

The most impactful system to date is CRISPR/Cas9, which revolutionized genome editing by allowing precise, efficient, and versatile manipulation of DNA (Figure 1C). Unlike RNAi, which provides temporary gene knockdown by targeting mRNA, CRISPR/Cas9 directly modifies the DNA, enabling permanent gene knockout or modification[7]. This system uses a single-guide RNA (sgRNA) to direct the Cas9 nuclease to specific DNA sequences, introducing double-strand breaks that the cell repairs, leading to targeted gene edits. CRISPR technology evolved as it could knockout, activate, and interfere with gene expression[8]. CRISPR/Cas9 offers unparalleled precision, minimal off-target effects, and can be applied to virtually any organism, making it superior to previous techniques.

Arrayed CRISPR screens became widely popular for their ability to investigate hundreds of perturbations simultaneously[9]. In these screens, cells are distributed into individual wells, each receiving different perturbations. After a selection process, readouts typically occur through non-sequencing phenotype assays such as microscopy, cell viability, proliferation, flow cytometry, or reporter gene assays. Despite their effectiveness, arrayed screens are labor-intensive, expensive, require complex experimental designs, and are susceptible to batch effects.

The advent of next-generation sequencing (NGS) transformed the landscape with the development of pooled CRISPR screens. Unlike arrayed screens, where each well has a designated perturbation and direct observations suffice, pooled CRISPR screens involve editing many genes across a mixed population of cells without pre-associating each edit with a specific cell[10]. NGS is crucial in this context, as it allows for the decoding of barcodes or unique identifiers associated with each guide RNA, linking them back to phenotypes[11, 12]. This

capability makes pooled screens more scalable and cost-effective for high-throughput genetic studies.

Optical pooled screens have been a growing subfield as they combine spatial information from high-content imaging with the high-throughput capabilities of pooled CRISPR screens. The most recent breakthrough in optical pooled screens is performing in-situ sequencing (ISS) to directly sequence the sgRNA itself or a barcode through creating Rolling Circle Amplification (RCA) products or rolling circle products (RCPs)[13, 14]. This method involves introducing perturbations through lentiviral delivery and then after fixation, obtaining an RCP from the mRNA that was transcribed. Imaging is used to obtain a wide variety of optical phenotypes, including cellular morphology, subcellular localization, dynamics, cell-cell interactions, stimuli response, organelle function, and signal transduction pathways. mRNA transcripts undergo reverse transcription (RT) into cDNA that hybridizes with padlock probes that extend and ligate themselves. RCA then occurs as the circular probe is used as a template for polymerases to add nucleotides onto the cDNA[15,16](Figure2B). Barcode sequences are associated with sgRNAs and decoded through rounds of fluorescence imaging. When sequencing the sgRNA itself, sequencing by synthesis (SBS) is performed through rounds of adding dye-labeled nucleotides that emit at different wavelengths, imaging, and cleavage of the nucleotides[17] (Figure 3).

To sequence the guide, cells must be transduced with lentiviral vectors that contain the sgRNA[18]. A popular dual vector system utilizes LentiCas9-Blast, which contains the Cas9 nuclease, and LentiGuide-Puro, which contains the sgRNA cassette. LentiCRISPRv2 is also a popular vector that contains the Cas9 endonuclease and sgRNA cassette in a single vector, allowing for fewer transductions and higher titer[19]. This paper will focus on the CROPseq vector, which is a unique vector in that it utilizes fixed hybridization sites for the padlock probe[20]. The

CROPseq vector was a modification of the LentiGuide-Puro vector to include a hU6-gRNA cassette within the 3' long terminal repeat (LTR) (Figure 4A). Typically, with LentiGuide-Puro, the sgRNA sequence is transcribed by RNA polymerase III, which then guides the Cas9 nuclease to its target site to edit the genome. With the addition of the hU6-sgRNA cassette into the 3'LTR, the CROPseq vector allows for RNA polymerase II to transcribe the sgRNA into a polyadenylated mRNA transcript that can be sequenced or detected with poly-A enrichment (Figure 4B). This eliminates error-prone barcode assignment that could occur during the transduction process.

With all the various pooled genetic screen techniques in mind, here we propose to recreate the method of performing an optical pooled screen with ISS to directly sequence the guide. This allows for direct identification of the sgRNA on a vector level while achieving higher imaging throughput with low magnification. Since there are no barcodes involved, sgRNA associated with the wrong barcodes is not an issue. A perturbation library can be introduced, and universal padlock probes can be used to directly link genotype to a multitude of phenotypes with spatial positioning on a single-cell level.

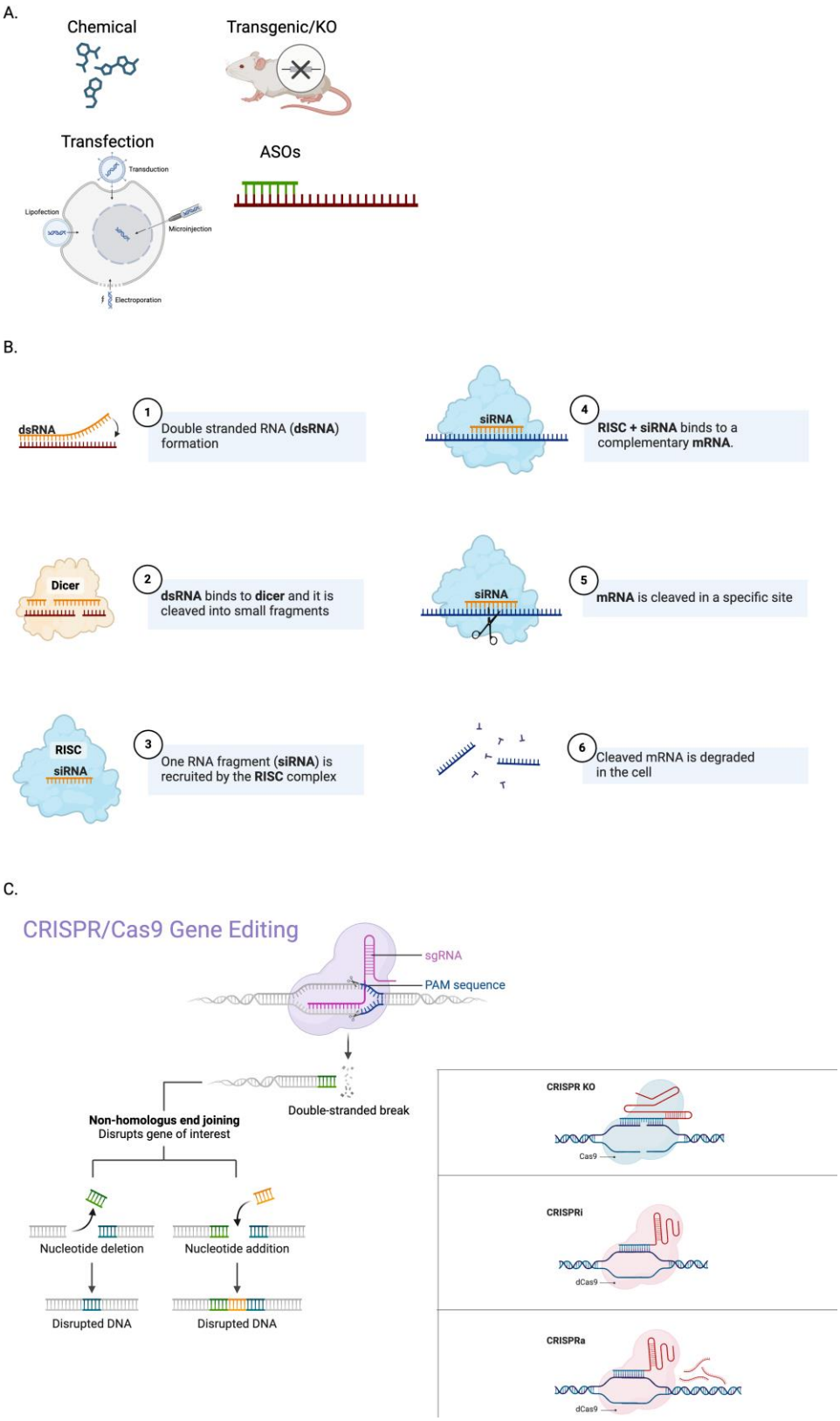


Figure 1: History of Perturbation Screens

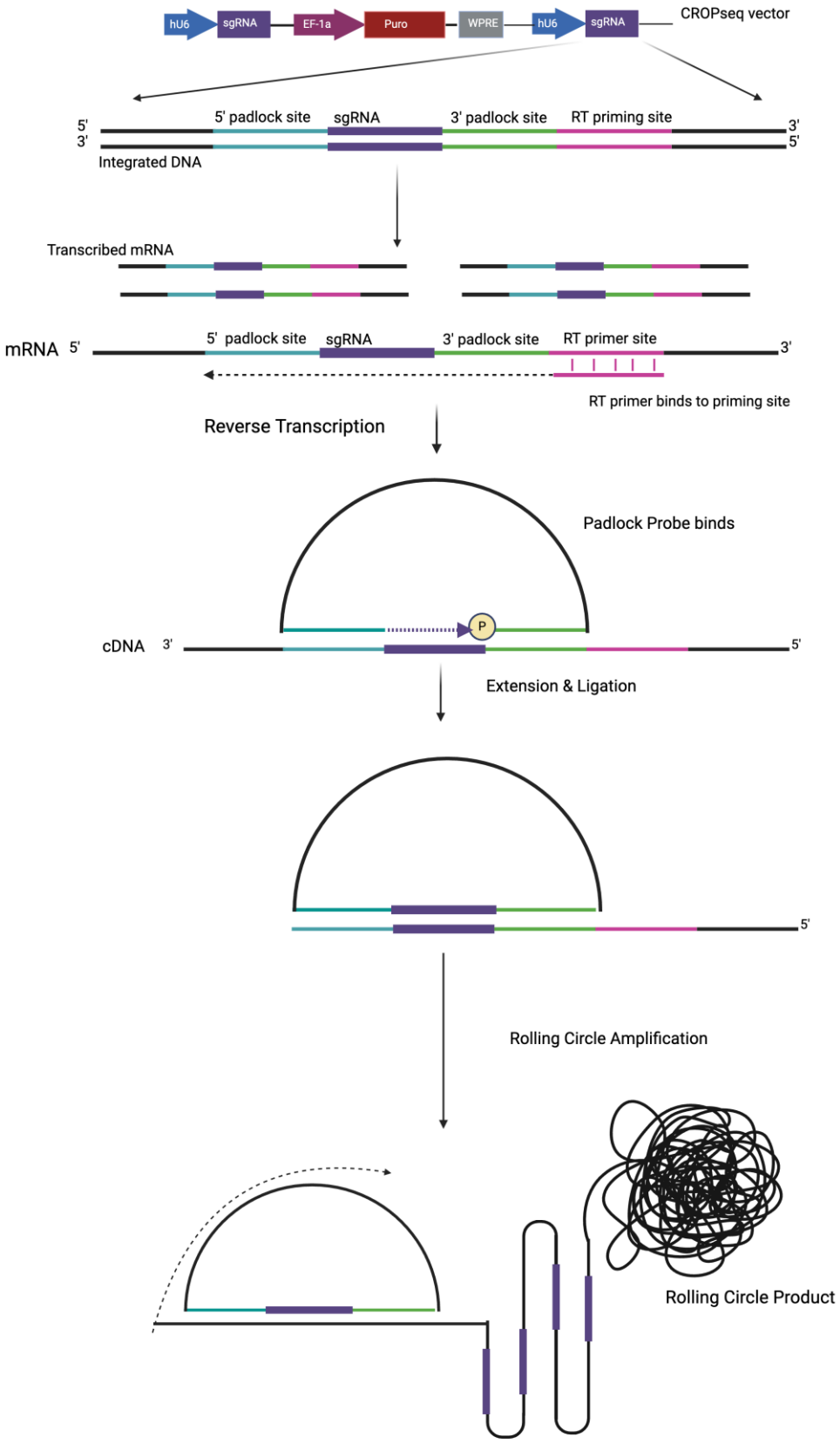


Figure 2: In-Situ Sequencing

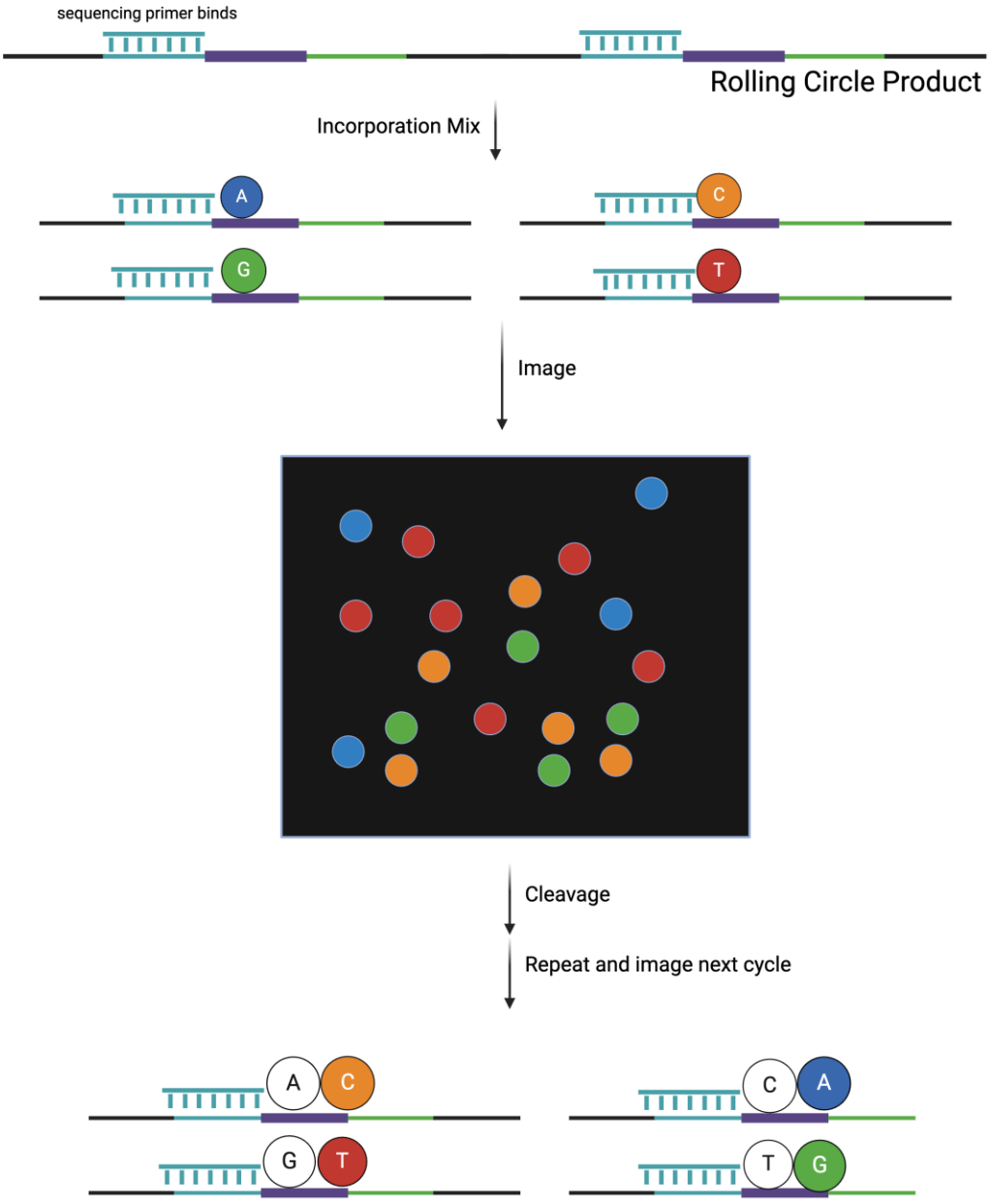


Figure 3: In-Situ Sequencing: Sequencing by Synthesis

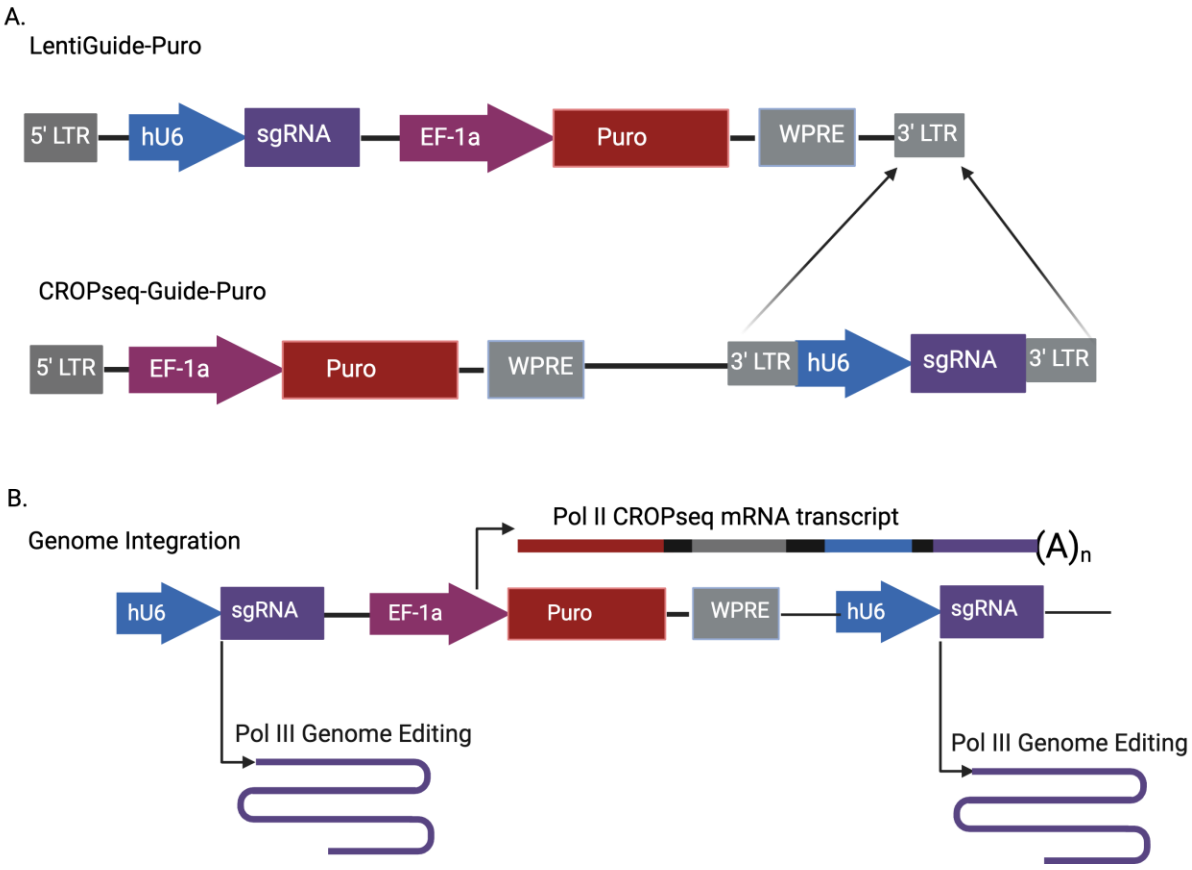


Figure 4: CROPseq Vector

MATERIALS & METHODS

sgRNA Library Preparation

After designing sgRNAs, single-stranded oligonucleotide sequences (oligos) for the top and bottom guide fragments were synthesized and ordered from a DNA synthesis vendor (Integrated DNA Technologies, IDT). The oligos were diluted to a starting concentration of 100 µM each. Equal volumes (1 µL) of the top and bottom oligos were mixed with 0.5 µL of T4 Polynucleotide Kinase (PNK) and 1 µL of 10x T4 ligase buffer in a total reaction volume of 10 µL, with the remaining volume made up by nuclease-free water. The reaction mixture was incubated on a thermocycler for 30 minutes at 37 °C, 5 minutes at 95 °C, and then held at 12 °C. During this process, the fragments were phosphorylated by T4 PNK and annealed together to form a single fragment (Figure 5A). The annealed and phosphorylated guide fragments were then used in a Golden Gate assembly reaction to digest the CROPseq vector/backbone with the restriction enzyme FastDigest Esp3I (BsmBI) (Figure 5B) and simultaneously ligate the guide fragments into the backbone with T4 DNA ligase(Figure 5C)[21].

For the Golden Gate assembly reaction, the CROPseq plasmid (25 ng) was mixed with 1 µL of the insert, 2 µL of T4 DNA Ligase, 2 µL of 10X T4 DNA ligase buffer, and 1 µL of Esp3I to make a total volume of 20 µL with the remaining volume made up of nuclease-free water. The mixture was subjected to 15 cycles of digestion at 37 °C for 5 minutes followed by ligation at 20 °C for 5 minutes. Now we have our sgRNA plasmid that consists of the CROPseq backbone with the sgRNA insert.

Lentivirus Production & Titer

A 10 cm culture plate with HEK293FT cells at ~85% confluency with DMEM media + 10% FBS was prepared for optimal transfection. A lentivirus master mix was created using 1.5 µg pMD2.g (VSV-G envelope expressing plasmid), 3.5 µg psPAX2 (2nd generation lentiviral packaging plasmid), 5 µg of our sgRNA plasmid, 15 µl of the JetOptimus Reagent, and 100 µl of the JetOptimus Buffer. JetOptimus is used to introduce foreign DNA/RNA into cells with high transfection efficiency (Figure 5D). After incubating the mix at room temperature, the mixture was added dropwise to the 10 cm plate where the plate would incubate for two days, having the media changed in between. After two days of incubation, the lentivirus was harvested by collecting the supernatant and filtering it to remove cellular debris.

To determine lentiviral titer, U2OS cells were trypsinized and plated onto a 12-well plate with DMEM media + 10% FBS, aiming for ~90% confluency after spinfection. Different volumes of lentivirus were added to each well (0 µl, 1 µl, 5 µl, 10 µl, 25 µl) and spinfected through centrifuging the plate at 1,000 g for 2 hours at 33 °C. The cells were then incubated at 37 °C for two days, with the media being replaced in between. After incubating for two days, cells were trypsinized and diluted to plate on a 12-well plate with DMEM + 10% FBS. Each lentivirus condition was diluted to ¼ and 1/40 of the original volume of cells, and puromycin was added for selection to the ¼ dilution wells. The appropriate antibiotic selection marker, puromycin, was applied in accordance with lentiviral concentrations, but a standard 1:10,000 puromycin was added. Multiplicity of Infection (MOI) is used to measure the number of viral particles added per cell, and lentiviral titer refers to the concentration of lentivirus in a given volume. A high MOI would mean that multiple sgRNAs may be introduced to each cell, while a low MOI would ensure only one guide enters each cell. To determine low and high MOI, we selected the

populations with the highest and lowest lentiviral concentrations in which the cells survived, thus ensuring the guides were transduced due to the selection marker. This experiment is not interested in the combination of sgRNAs in individual cells, so we will be focusing more on the low MOI samples and aiming for a viral titer of 3000 transduced cells/µL. Transduction followed the same spinfection process as transfection (Figure 5E).

In-Situ Sequencing

Cells in DMEM media + 10% FBS are plated onto 18-well Glass Bottom Ibidi u-Slides, with each well having ~50,000 cells in a total volume of 100 µl. After incubating overnight, the cells are fixed to preserve cellular structure and prevent mRNA degradation with a solution consisting of 20% paraformaldehyde in 1X PBS + RNase Inhibitor (PBS*) for 30 minutes. After washing the cells three times with PBS*, the cells are permeabilized to allow for machinery to penetrate the cell membrane with 0.5% Triton X-100 in 1X PBS for 5 minutes. The cells are then washed five times with 0.05% vol/vol Tween-20 and 0.2% vol/vol RNase Inhibitor in 1X PBS (PBS-T), waiting 5 minutes for one of these washes. The reverse transcription (RT) mix is then assembled accordingly: RevertAid H Minus Reverse Transcriptase 4.8 units/µl, Revert Aid RT Buffer 1X, RNase Inhibitor 0.8 units/µl, BSA 0.2 mg/ml, RT primer 1 µM, and Deoxyribonucleotide Triphosphate (dNTPs) 250 µM in a total volume of 50 µl per well. The RT primer initializes the reverse transcription process by binding to the RT primer site that is located next to the 3' padlock site on the CROPseq vector. After binding, the RT enzyme uses the dNTPs to add nucleotides to create a mRNA/cDNA duplex. After adding the RT mix, the plate was subjected to 3 hours at 42 °C.

After returning the plate to room temperature, the plate is washed three times with PBS-T. Post-fixation is performed with a solution consisting of 3% PFA and 0.1% Glutaraldehyde in

1X PBS for 30 minutes to further ensure structural integrity, as glutaraldehyde provides more extensive cross-linking. The cells are then washed five times with PBS-T, and the Extension-Ligation (EL) mix is assembled accordingly: Taq DNA polymerase 0.02 units/µL, RNase H 0.4 units/µL, Ampligase 0.5 units/µL, Ampligase Buffer 1X, padlock probe 100 nM, dNTPs 50 nM each, BSA 0.2 mg/mL, in a total volume of 50 µL per well with the remaining being UltraPure Water. The plate is then subjected to 5 minutes at 37 °C and 90 minutes at 45 °C. The EL mix first hydrolyzes the phosphodiester bonds of mRNA bound to cDNA with RNase H, separating the duplex and freeing the cDNA. Then, as the temperature increases, the padlock probe binds to the 3' and 5' padlock sites. Following that, the polymerase enzyme adds dNTPs to the 3' end of the padlock probe and extends the sequence that is complementary to the sgRNA. It is essential to freshly dilute the dNTPs from the supplier's 10 mM concentration 1:1,000 in water to a 10 µM concentration. Overly concentrated dNTPs in the EL mix will cause errors in extension due to crowding of nucleotides. After the extension reaches the 5' end of the padlock probe that is already phosphorylated when ordering from the supplier (IDT), the ligase enzyme ligates the probe to be a circular sequence, creating a cDNA/padlock probe duplex.

The plate is washed three times with PBS-T, and the RCA mix is assembled accordingly: Phi29 DNA Polymerase 1 unit/µL, Phi29 buffer 1X, dNTPs 250 µM, BSA 0.2 mg/mL, and 5% Glycerol in a total volume of 50 µL per well with the remaining being UltraPure Water. Empty wells were filled with water to avoid evaporation, and the plate was incubated overnight at 30 °C. Phi29 polymerase expresses 3' to 5' exonuclease activity and thus digests the 3' overhang of the cDNA sequence up until the 5' padlock site[22]. The polymerase then initiates RCA by using dNTPs to extend the cDNA using the padlock probe as a template. After incubating overnight, plates were washed three times with PBS-T and could be stored for imaging at 4 °C.

During the experiments where a Cy5 probe was used to validate RCP formation, plates were first washed three times with 2X SSC for five minutes each. The fluorescent probe (Cy5) mix was then assembled accordingly: Cy5 probe 200 µM, Formamide 35%, 20X SSC 2X in a total volume of 100 µL with the remaining being UltraPure Water. The Cy5 mix was hybridized at room temperature for 15 minutes, protected from light. DAPI staining was then conducted by diluting a 1 mg/mL solution to 20 µg/mL in 2X SSC and incubating for 30 minutes, protected from light at room temperature. Plates were then washed six times with 2X SSC with a wait time of five minutes for four of the washes and ready to image.

Cy5 Detection & Sequencing by Synthesis

For the SBS experiments, plates were first washed twice with PBS-T. The sequencing primer, which was diluted to 1 µM in 2X SSC, was added, and plates were incubated for 30 minutes at 37 °C. After incubation, plates were washed three times with PBS-T, then three times with Incorporation (PR2) buffer. Incorporation mixes containing the dye-labeled nucleotides from the Miseq or Nextseq kits were thawed and 100 µL were added to each well. Plates were then incubated for 4 minutes at 60 °C, as it was crucial not to incubate longer than five minutes because the longer the incorporation goes, the background staining dramatically increases. We then washed the plates six times with the PR2 buffer in quick succession. The plate is then incubated five times for five minutes at 60 °C, being washed with fresh PR2 in between each incubation. With the incorporation mix being very concentrated, these wash steps help minimize background noise by removing any nucleotides that may have bound non-specifically to cells or any surface. Then DAPI solution was diluted from 1 mg/mL solution to 20 µg/mL in 2X SSC and added to plates where they incubated protected from light for 15 minutes at room

temperature. After washing with PR2 three times, plates were ready to image. After imaging and in between cycles of sequencing, plates can be stored at 4 °C for several weeks.

When ready to begin the next sequencing cycle, the cleavage mix is added, and the plate is incubated at 60 °C for 6 minutes. The plate is washed three times with PR2 in quick succession at room temperature. Then the plate is washed three times in PR2, incubating at 60 °C for one minute each. After completing three more PR2 washes, the plate is ready to have the incorporation mix added for the next sequencing round.

Imaging

Imaging was done primarily using the Zeiss LSM 780 microscope, but the Keyence Fluorescence Microscope Bz-X Series was also used. DAPI staining was imaged at 405 nm, while Cy5 was imaged at 633 nm. While performing SBS, two kits were used: Illumina Miseq Reagent Kit v3 and Illumina Nextseq 1000/2000 Reagent Kit. Since the machines that these kits were originally meant for are not being used, we faced difficulty obtaining the correct kit and extracting the reagents. Due to this, some experiments were done with expired/old reagents which could have inactivated the incorporation mix. Using the numbered positions on the Miseq cartridge, extraction of reagents was done with position 1 being the incorporation mix containing the dye-labeled nucleotides and position 4 being the cleavage mix that contains the cleave enzyme. To extract Nextseq reagents, specific instructions were received from a collaborator.

Both kits differ in terms of their color chemistry as Nextseq has a two-color chemistry that uses 477 nm and 546 nm excitation lasers that fit with conventional filters such as EGFP and Cy3 (Figure 6A). This two-color chemistry means the G nucleotide will be a blank image with no fluorescent signal, while the A nucleotide will have both channels being fluorescent (Figure 6B). The other two nucleotides, T and C, will be fluorescent in EGFP and Cy3, respectively. The

Miseq kit, on the other hand, utilizes four-color chemistry with G and T being excited with a 546 nm laser while A and C are excited by a 637 nm laser with all four nucleotides expressing an emission signal at different wavelengths (Figure 6C, 6D). These unique emission spectra require specific filter sets that are unconventional in most microscopes. Although Nextseq is readily adaptable to more microscopes due to the standard filters, since obtaining a fresh Nextseq kit was difficult, we only did a few runs with the Nextseq and proceeded to move forward with only the Miseq kit. Though Miseq requires unconventional filter sets, fresh Miseq reagent kits were able to be purchased directly from the supplier.

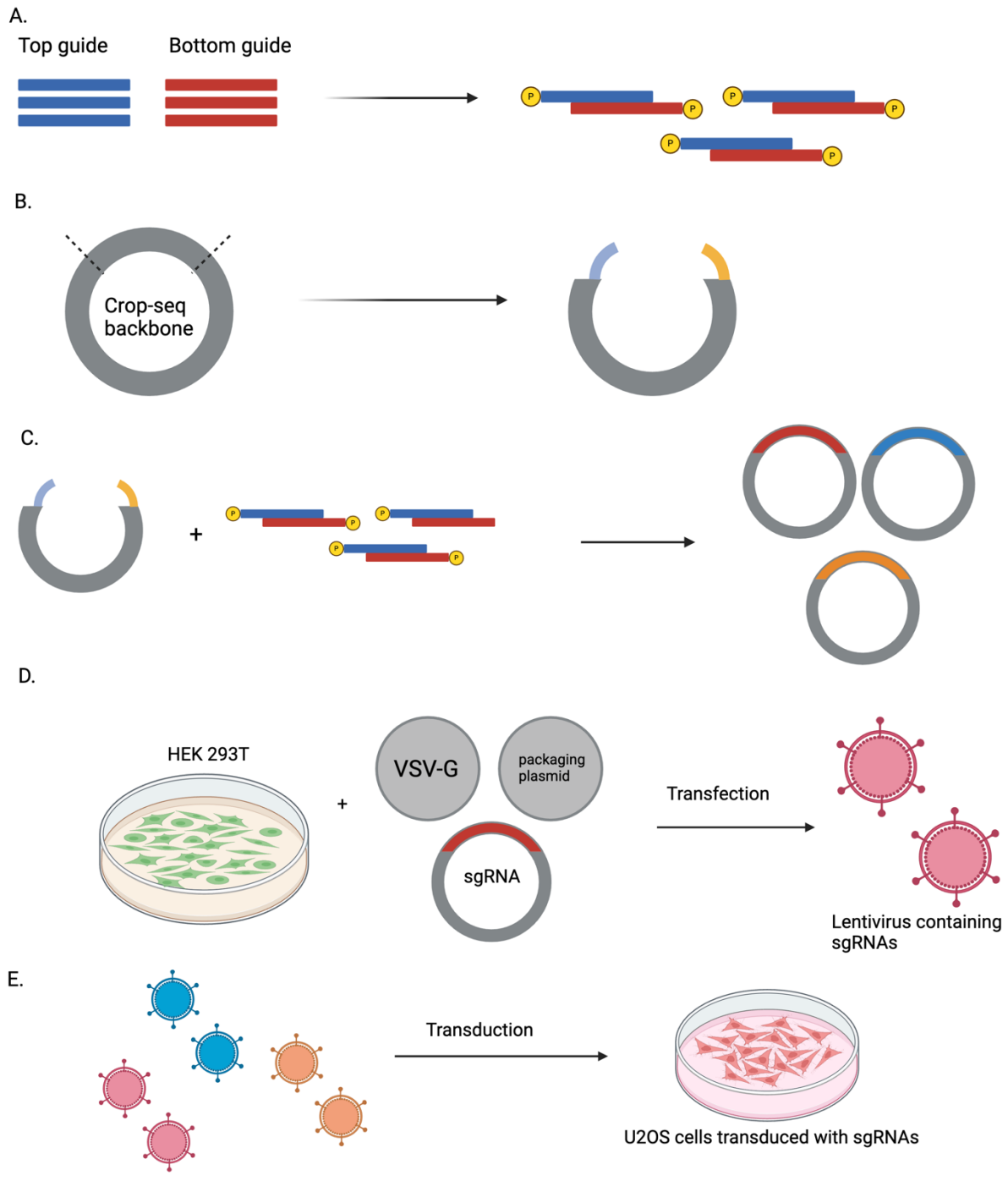
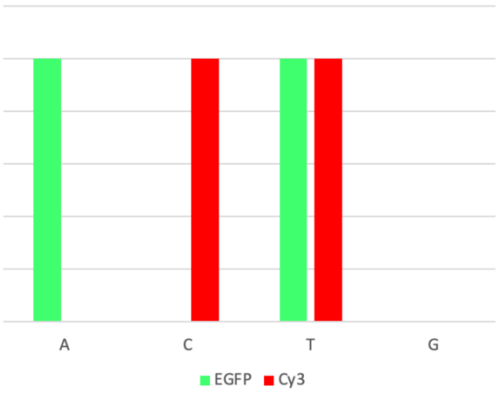


Figure 5: Golden Gate Assembly & Transduction

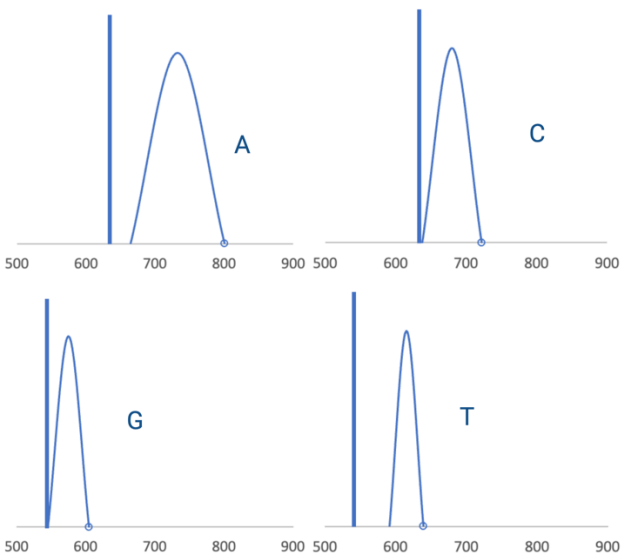
A.



B.

Base	Channel 1 – 477 nm - EGFP	Channel 2 – 546 nm – Cy3
A	X	
C		X
T	X	X
G		

C.



D.

Base	Excitation (nm)	Emission (nm)	Dye	
A	637	732/68	Cy4 (Cy5*)	
C	637	680/42	Cy5.5 (near infrared)	
T	546	615/24	Alexa Fluor 594	
G	546	575/30	Cy3	

Note*
Zeiss Microscope does not offer Cy4 dye so Cy5 was used instead.

Figure 6: Miseq & Nextseq Imaging Channels

RESULTS

Identification of Rolling Circle Products

To achieve ISS, we first confirmed the identity of rolling circle products (RCPs) to ensure successful amplification of the padlock probe sequence containing the sgRNA. The initial experiment targeted the endogenous Actin-Beta (ACTB) gene in human osteosarcoma cells (U2OS). This experiment had two negative control wells of U2OS: one without the RT mix and the other without the RCA mix. We selected the ACTB gene because it is constantly expressed in U2OS cells due to its essential role in maintaining basic cellular functions such as cell shape, cytoskeleton, and cell division[23]. A primer sequence targeting the ACTB gene initiated reverse transcription of mRNA into cDNA. After binding the padlock probe and performing RCA, a Cy5-labeled oligo was used to bind the RCP. The Cy5 probe did not target the extended and ligated region of the padlock probe to minimize variables. The experiment had two versions: v1, where the padlock probe was only ligated, and v2, where the padlock probe was extended by four nucleotides before ligation. Corresponding Cy5 probes were designed to bind the RCPs. Imaging was performed using a Zeiss LSM 780 microscope with a 20x objective, showing positive signals for both v1 (Figure 7A) and v2 (Figure 7B), with no signals in the negative controls. Positive signals were confirmed through histogram analysis and visual inspection.

After proving RCP formation for an endogenous gene, we transduced U2OS cells with sgRNA sequences targeting the FMR1 gene using lentivirus at both high and low multiplicity of infection (MOI). FMR1 codes for the fragile X mental retardation protein (FMRP), crucial for cognitive functions and neural development[24]. Cells were selected with puromycin (1:10,000) in DMEM + 10% FBS and processed to achieve RCPs. Positive signals were observed in both high and low MOI cells (Figures 7C).

Seeing some success, we wanted to streamline the protocol to go from three to two days of wet-lab work. Originally, the reverse transcription step was incubated at 37°C overnight, but we found that RevertAid H Minus Reverse Transcriptase's enzyme was sufficiently active at a temperature of 42°C for 1 and 3 hours (Figure 8A, 8B). Wash times were also shortened as positive signals were extremely prevalent due to the efficiency of the CROPseq vector.

SBS & Analysis

With confirmed RCP formation, our next experiment focused on sequencing by synthesis (SBS) to sequence the sgRNA. To test SBS, we used both the Illumina Nextseq P2000 and Illumina Miseq v3 reagent kits. Due to the difficulty of obtaining both reagent kits, unorthodox measures were taken which could explain some of the failed experiments. All the SBS experiments were designed so that only the first nucleotide was sequenced to eliminate the cleavage process.

We decided to design four new padlock probes targeting the ACTB gene, each using one sequencing primer. These four padlock probes all have the same 3' and 5' binding sites that were used in the last experiment, but in the middle of the padlock probe where the primer binds to initiate sequencing, there is a four-nucleotide barcode. This way all four nucleotides can be sequenced with one primer on a single plate. Although this experiment was designed to eliminate the issue of sequencing primer design and binding, no positive signal was seen with the Nextseq and Miseq kits. To identify if the SBS primer is binding properly, we tested two padlock probes with a Cy5 oligo attached and saw positive signals (Figure 8A). Although we confirmed that the SBS primer is binding, the sequences for the ACTB gene were received from a collaborator who

had not validated their efficiency, unlike the previously used FMR1 sequences. As a result, this experiment was ultimately abandoned.

The next experiment focused on the transduced FMR1 cell line where three additional sequencing primers were designed so that all primers would attach to the RCP and the first cycle of SBS would produce a different nucleotide. Similarly to the previous SBS Actin experiment, all four nucleotides were able to be tested on a single plate. The three additional primers designed are shifted from the original sequence with different starting points. The first iteration with the Miseq kit failed, most likely due to expired incorporation mix, as the liquid lacked color. However, positive signals were captured with the Nextseq kit, sequencing the "C" nucleotide and observing Cy3 fluorescence (Figure 9A). This experiment was repeated and yielded the same results.

After acquiring fresh Miseq kits, a new iteration (v2) was attempted with a different sequencing primer that was now extended from the original sequence. Four new primer sets were designed to test all the nucleotides and all primers still contained the same sequence as the original sequencing primer but just had an extended length. This was done as there was suspicion that the v1 sequences, which were shifted up to eight nucleotides, were not binding properly to the RCP or not adding the correct nucleotide. Our v2 iteration yielded no positive results with the Nextseq kit, but the Miseq kit showed fluorescent signal with the "T" nucleotide (Figure 9B).

Our past experiments inspired us to no longer target the ACTB gene and focus on the FMR1 gene while also excluding Nextseq reagent kits to build upon previously successful experiments. Four U2OS cell lines were transduced with different CROPseq vectors designed to test all nucleotides, using validated primers and sequences from the successful FMR1 experiments. Each cell line had a different four-nucleotide sgRNA sequence reminiscent of our

SBS Actin experiment. This design allowed the four cell lines to use the same validated primers from past experiments, hoping to eliminate issues regarding downstream binding of sequences and nucleotides during ISS.

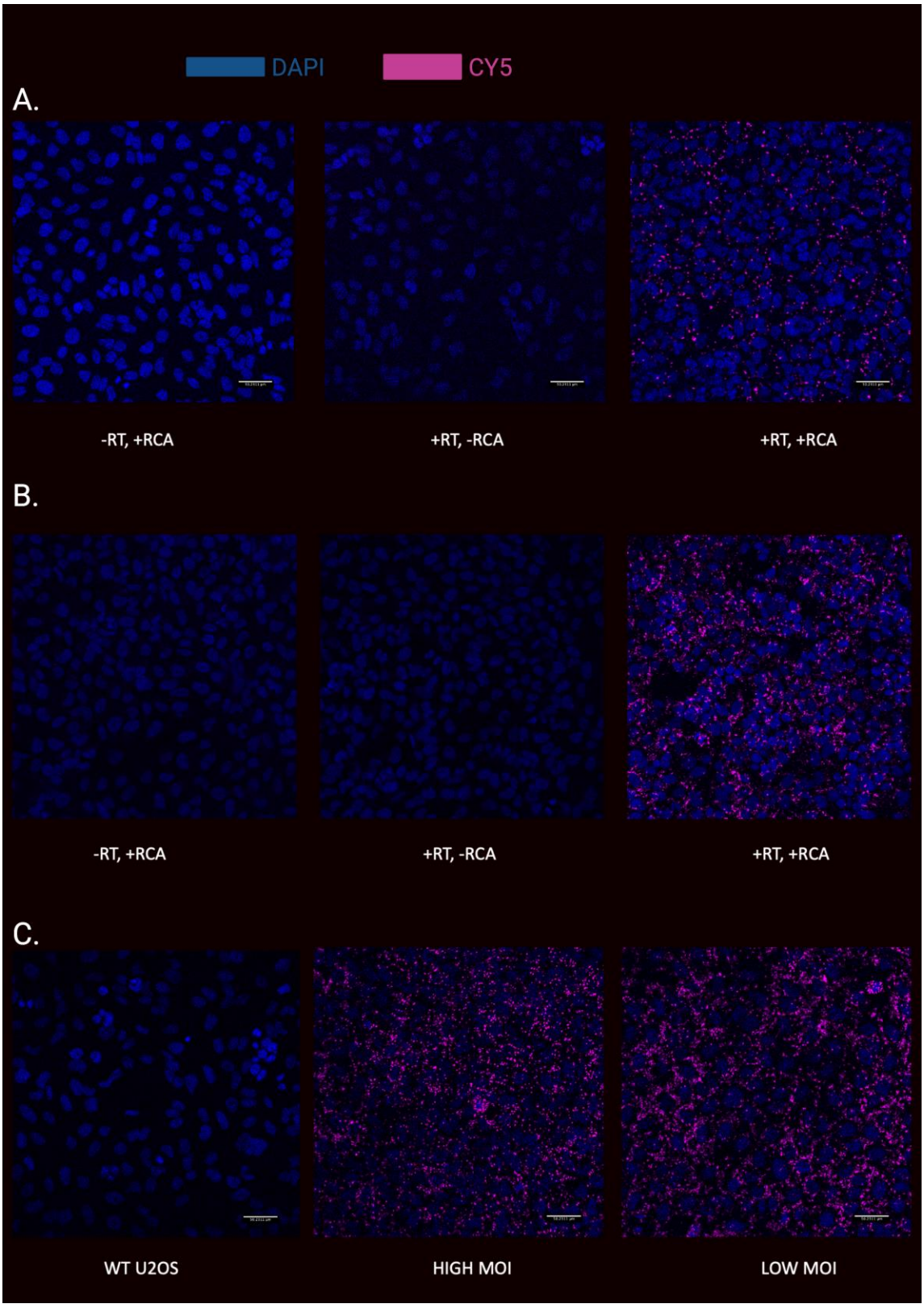


Figure 7: Identification of Rolling Circle Products Experiment Results

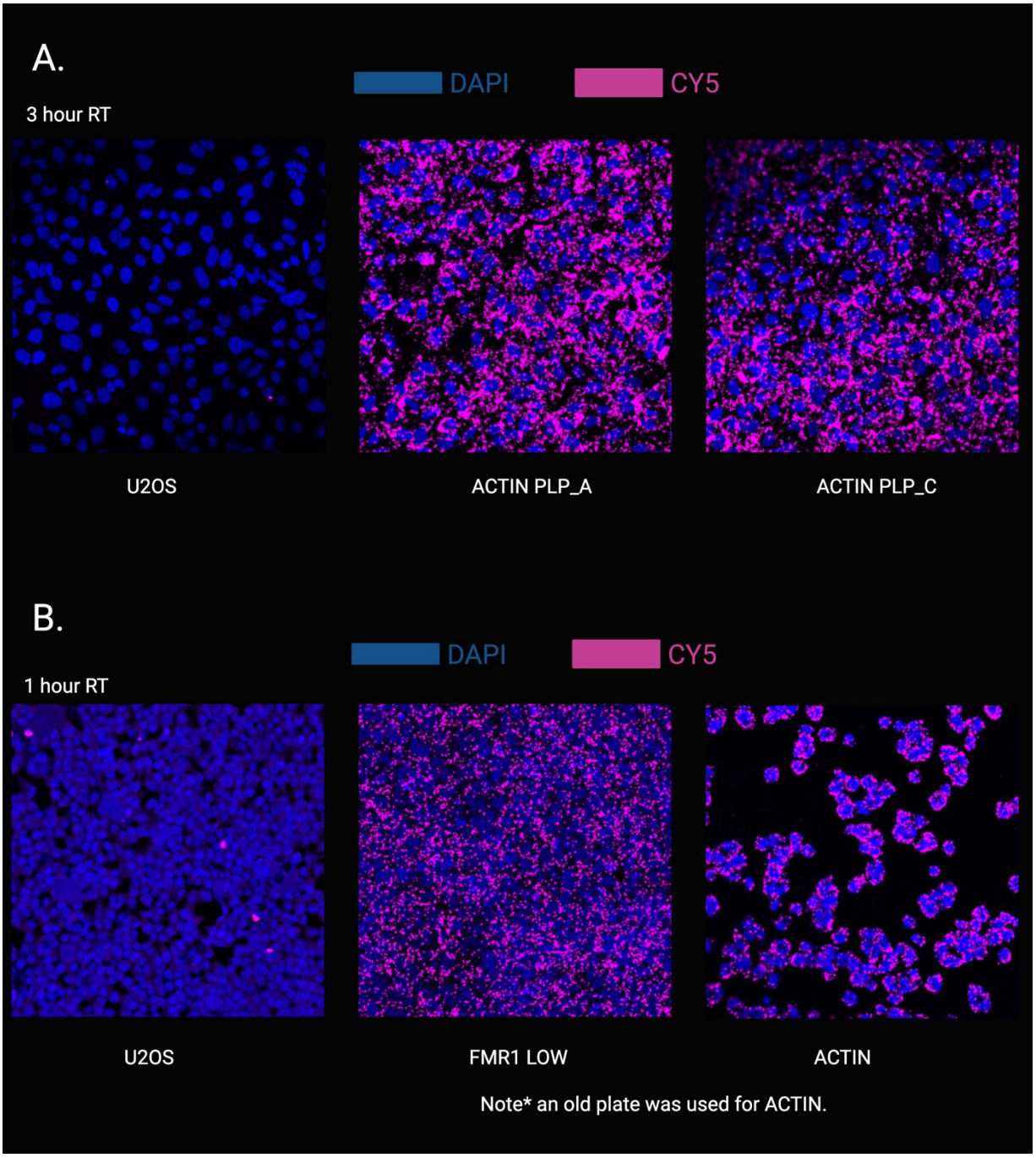


Figure 8: Validation of Actin SBS primer & Revised RT

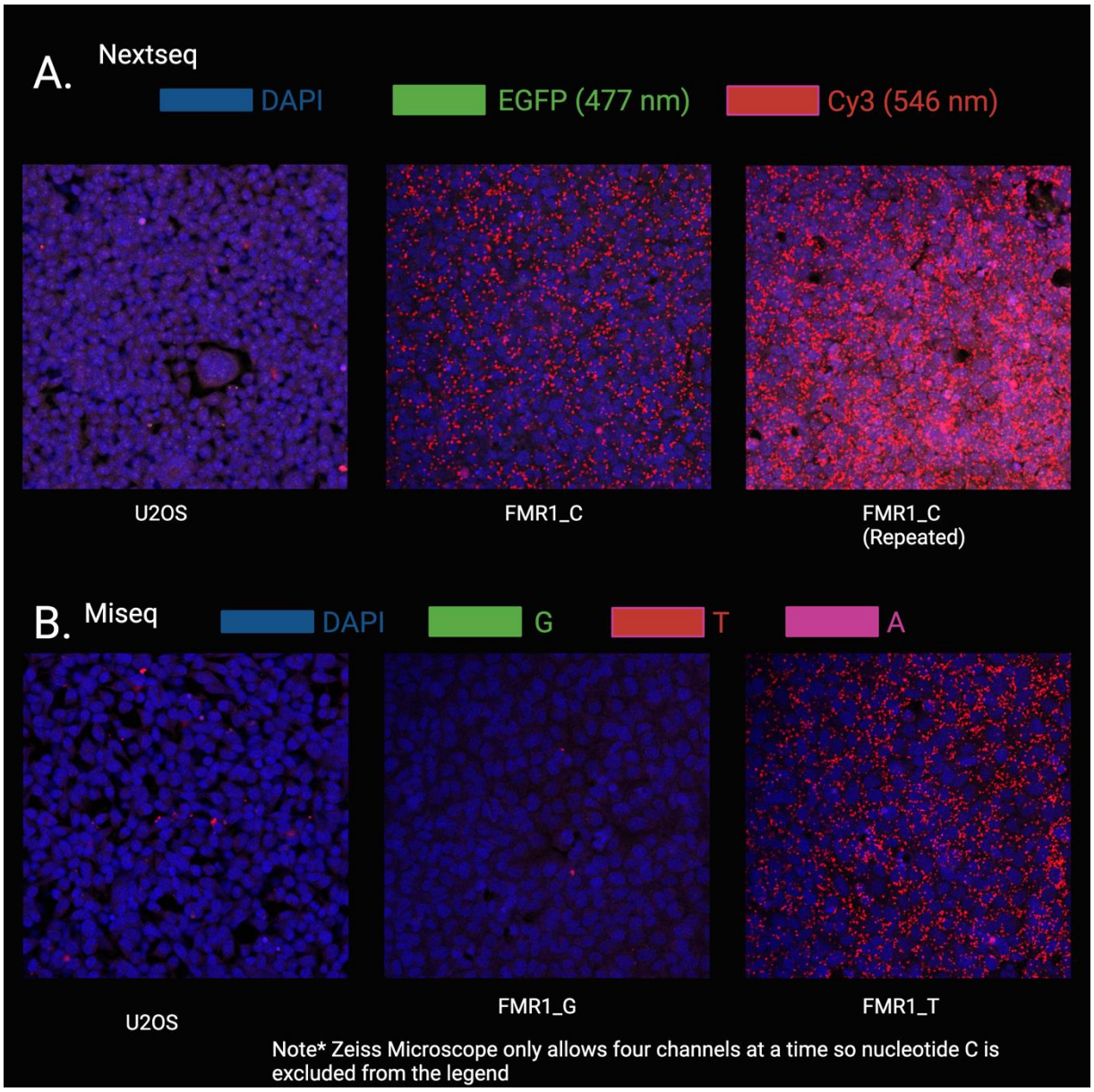


Figure 9: SBS Experiment Results

DISCUSSION

Optical pooled screens with ISS allow for perturbation identities to be sequenced in-situ, directly linking genotype to a wide variety of phenotypes such as cell-cell interactions, morphology, dynamic behaviors, and subcellular localization of RNA/protein. This method enables sgRNAs to be sequenced with low magnification, bypassing barcode assignment and achieving higher throughput images. This capability is due to the CROPseq vector's unique insertion of the hU6-sgRNA cassette, which produces a detectable mRNA transcript that contains the sgRNA. This vector, in combination with ISS, has proven to increase the activity of CRISPR knockout, interference, and activation.

Although our attempt at the full ISS fell short due to issues with SBS, we believe that RCP formation is robust with this protocol. With the major bottleneck in RCP formation being enzyme stability and dNTP concentrations, SBS faces difficulties with incorporation mixes and imaging platforms. We expect to successfully recreate the full ISS method, as there is suspicion that the issue in our experiment is the incorporation mix being inactive or the nucleotides not binding to the correct sequence.

When attempting to perform the full screen, further experiments will need to be conducted for the phenotype imaging process. As our experiments fell short, we have yet to explore the best time to perform immunofluorescent reporter assays to image the phenotype. It is recommended for staining to occur before the first cycle of SBS, but imaging before RT, after RT, and after RCA needs to be investigated. The full screen will also require a computational sequencing pipeline that can convert image files to base-called coordinates. Images need to undergo proper alignment, background correction, segmentation, peak detection, and base calling.

The field of optical pooled screens is growing quickly, with various methods being built upon ISS. To achieve multiple sgRNAs per cell would usually require a high MOI lentiviral delivery, but CROPseq-multi was recently created to multiplex perturbations in a single vector[25]. One CROPseq-multi vector can target a combination of genes or have two different sgRNAs that target the same gene. This overcomes random perturbation combinations that may be introduced during transduction with high MOI lentivirus in combinatorial screens.

Two major limitations that come with our ISS method are the permeabilization of cells and the instability of mRNA in the cytosol. Since we permeabilize the cells to allow our RT mix to be incorporated, cell-cell boundaries are damaged, which interferes with cell segmentation. The cytoplasmic mRNA is unstable and limits experiments to cell types that have high transcriptional activity to ensure sufficient signal. NIS-seq employs a modified CROPseq vector that utilizes a T7 polymerase to transcribe the gRNA sequences directly within the nucleus, ensuring all signals are retained there and avoiding issues with mRNA degradation[26].

The foundational work for optical pooled screens has been laid out, and the continued refinement of this ISS technique seems inevitable. With improvements already being seen, the backbone of this method holds promise for unlocking new insights into cellular function and disease, ultimately contributing to advances in biomedical research and drug therapy.

REFERENCES

1. Bunin, B. A., Plunkett, M. J., & Ellman, J. A. (1996). Synthesis and evaluation of 1,4-benzodiazepine libraries. *Methods in enzymology*, *267*, 448–465. https://doi.org/10.1016/s0076-6879(96)67028-1

2. Fus-Kujawa, A., Prus, P., Bajdak-Rusinek, K., Teper, P., Gawron, K., Kowalczuk, A., & Sieron, A. L. (2021). An Overview of Methods and Tools for Transfection of Eukaryotic Cells *in vitro*. *Frontiers in bioengineering and biotechnology*, *9*, 701031. https://doi.org/10.3389/fbioe.2021.701031

3. Shakweer, W. M. E., Krivoruchko, A. Y., Dessouki, S. M., & Khattab, A. A. (2023). A review of transgenic animal techniques and their applications. *Journal, genetic engineering & biotechnology*, *21*(1), 55. https://doi.org/10.1186/s43141-023-00502-z

4. Crooke, S.T., Baker, B.F., Crooke, R.M. Antisense technology: an overview and prospectus. *Nat Rev Drug Discov* **20**, 427–453 (2021). https://doi.org/10.1038/s41573-021-00162-z

5. Kim, D., & Rossi, J. (2008). RNAi mechanisms and applications. *BioTechniques*, *44*(5), 613–616. https://doi.org/10.2144/000112792

6. Wucherpfennig, K. W., & Cartwright, A. N. (2016). Genetic screens to study the immune system in cancer. *Current opinion in immunology*, *41*, 55–61. https://doi.org/10.1016/j.coi.2016.05.007

7. Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 157(6), 1262-1278. https://doi.org/10.1016/j.cell.2014.05.010

8. Kampmann M. (2018). CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS chemical biology*, *13*(2), 406–416. https://doi.org/10.1021/acschembio.7b00657

9. Proietti, L., Manhart, G., Heyes, E., Troester, S., & Grebien, F. (2022). Arrayed CRISPR/Cas9 Screening for the Functional Validation of Cancer Genetic Dependencies. *Bio-protocol*, *12*(24), e4577. https://doi.org/10.21769/BioProtoc.4577

10. Bock, C., Datlinger, P., Chardon, F. High-content CRISPR screening. *Nat Rev Methods Primers* **2**, 8 (2022). https://doi.org/10.1038/s43586-021-00093-4

11. Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human immunology*, *82*(11), 801–811. https://doi.org/10.1016/j.humimm.2021.02.012

12. Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L., & Trees, E. (2018). Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, *24*(4), 335–341. https://doi.org/10.1016/j.cmi.2017.10.013

13. Feldman, D., Singh, A., Schmid-Burgk, J. L., Carlson, R. J., Mezger, A., Garrity, A. J., Blainey, P. C. (2019). Optical Pooled Screens in Human Cells. *Cell*, 179(3), 787-799.e17. https://doi.org/10.1016/j.cell.2019.09.016

14. Feldman, D., Funk, L., Le, A., Carlson, R. J., Leiken, M. D., Tsai, F. N., Blainey, P. C. (2022). Pooled genetic perturbation screens with image-based phenotypes. *Nature Protocols*, 17(2), 476-512. https://doi.org/10.1038/s41596-021-00653-8

15. Mohsen, M. G., & Kool, E. T. (2016). The Discovery of Rolling Circle Amplification and Rolling Circle Transcription. *Accounts of chemical research*, *49*(11), 2540–2550. https://doi.org/10.1021/acs.accounts.6b00417

16. Xu, L., Duan, J., Chen, J., Ding, S., & Cheng, W. (2021). Recent advances in rolling circle amplification-based biosensing strategies-A review. *Analytica chimica acta*, *1148*, 238187. https://doi.org/10.1016/j.aca.2020.12.062

17. Guo, J., Yu, L., Turro, N. J., & Ju, J. (2010). An integrated system for DNA sequencing by synthesis using novel nucleotide analogues. *Accounts of chemical research*, *43*(4), 551–563. https://doi.org/10.1021/ar900255c

18. Dong, W., & Kantor, B. (2021). Lentiviral Vectors for Delivery of Gene-Editing Systems Based on CRISPR/Cas: Current State and Perspectives. *Viruses*, *13*(7), 1288. https://doi.org/10.3390/v13071288

19. Sanjana, N. E., Shalem, O., & Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nature methods*, *11*(8), 783–784. https://doi.org/10.1038/nmeth.3047

20. Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L. C., Kuchler, A., Alpar, D., & Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. *Nature methods*, *14*(3), 297–301. https://doi.org/10.1038/nmeth.4177

21. Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE*, 3(11): e3647. https://doi.org/10.1371/journal.pone.0003647

22. Del Prado, A., Rodríguez, I., Lázaro, J. M., Moreno-Morcillo, M., de Vega, M., & Salas, M. (2019). New insights into the coordination between the polymerization and 3'-5'

exonuclease activities in ϕ29 DNA polymerase. *Scientific reports*, *9*(1), 923. https://doi.org/10.1038/s41598-018-37513-7

23. Ke, R., Mignardi, M., Pacureanu, A. *In situ* sequencing for RNA analysis in preserved tissue and cells.*Nat Methods***10**, 857860 (2013). https://doi.org/10.1038/nmeth.2563

24. Sitzmann, A. F., Hagelstrom, R. T., Tassone, F., Hagerman, R. J., & Butler, M. G. (2018). Rare FMR1 gene mutations causing fragile X syndrome: A review. *American journal of medical genetics. Part A*, *176*(1), 11–18. https://doi.org/10.1002/ajmg.a.38504

25. Fandrey, C. I., Konopka, P., Jentzsch, M., Zackria, A. A., Maasewerd, S., Latz, E., & Schmid-Burgk, J. L. (2024). Cell Type-Agnostic Optical Perturbation Screening Using Nuclear In-Situ Sequencing (NIS-Seq). *bioRxiv*. https://doi.org/10.1101/2024.01.18.576210

26. Walton, R. T., Qin, Y., & Blainey, P. C. (2024). CROPseq-multi: a versatile solution for multiplexed perturbation and decoding in pooled CRISPR screens. *bioRxiv*. https://doi.org/10.1101/2024.03.17.585235 (BioRxiv)

27. Biorender.com