Gene Regulation with O-GlcNAc Glycosylation by dCas9-OGT/OGA Fusion Proteins

Yu Shu Chou

A Thesis

Submitted to the Graduate Faculty of

the Department of Chemistry

in Partial Fulfillment

of the Requirements for the Degree of

Master of Science

Boston College
Morrissey College of Arts and Sciences
Graduate School

Jun 2022

Gene Regulation with O-GlcNAc Glycosylation by dCas9-OGT/OGA Fusion Proteins

Yu Shu Chou

Advisor: Jia Niu, PhD

Abstract

O-linked β-*N*-acetylglucosamine (O-GlcNAc) is a post-translational modification that plays important role in cellular network/diseases such as autophagy, transcriptional activity, and protein stability, phosphorylation modification competition, cancers, neurodegenerative disorders, etc. However, there are still lots of unknown roles in gene regulation especially while there is the existence of more than 1000 transcription factors and many other coactivators/repressors which might have an influence on expression level with or without O-GlcNAc modification.

In this project, we use CRISPR technology with catalytically inactive Cas9 from *Streptococcus pyogenes* fused with the only writer and eraser of O-GlcNAc to selectively target the DNA sequence of interest by guide RNA to see what the result of additional modification or cleavage of O-GlcNAc on proximal proteins in gene regulation is.

A cost and time-efficient way for guide RNA construction is developed in the project with one-piece PCR and Gibson assembly for 2 different backbones of guide RNA: 11 guide RNA for each. Other different cloning methods have also been used for future work. For future work, a second reporter will be introduced to normalize the luciferase signal. In addition, a new metabolic O-GlcNAc reporter and Y289L GalT/ UDP-GalNAz could be used to find out proteins that might be modified by dCas9-OGT/OGA with click reaction. Furthermore, a split OGA/Turbo ID system could also be used to reduce the background and find out proteins/modification sites that might

be important to the gene regulation for DNA-protein or protein-protein interaction.

TABLE OF CONTENTS

Table of Contents v

List of tables vi

List of figures vii

Acknowledgments viii

Introduction 1

Chapter 1. Introduction of O-GlcNAc and CRISPR 1

1.1 Introduction and importance of O-GlcNAc modification 1

1.2 CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats): introduction and variants 2

1.3. Experimental designs 4

Chapter 2. Results and Experiments 5

2.1 gRNA plasmid construction 5

2.2 Other cloning methods: two-piece Gibson, ligation and oligonucleotide annealing 11

2.2.1 Plasmid construction with ligation 11

2.2.2 Plasmid construction with two-piece Gibson assembly 12

2.2.3 Plasmid construction with oligonucleotide annealing 12

2.3 Mammalian cell culture and experiments 13

2.3.1. Trouble shooting of cell culture issue 13

2.3.2 Transient transfection 14

2.3.3 Western blot practice for future work for anti-O-GlcNAc with RL2 or CTD110.6 16

2.3.4 Luciferase assay 19

Chapter 3. Conclusions and Discussion 20

3.1 Summary of work 20

3.2 Experiment designs improvement 22

3.3 Future work 23

Chapter 4. Material and Methods 25

Chapter 5. Reference 27

Chapter 6. Sequencing Results 29

LIST OF TABLES

Table 1. Troubleshooting of cell culture issue ……... 14
Table 2. Transient transfection condition optimization ……. 15
Table 3. Plasmid elements cut by different types of IIS restriction enzymes …..... 21
Table 4. Dual-luciferase system used for experiment design improvement …….. 22
Table 5. Primer sets for SpCas9 Scaffold RNA: 46916_pU6-sgGAL4-4 backbone ..34
Table 6. Primer sets for SpCas9 Scaffold with Two MS2 Stem Loop: sgRNA(MS2) cloning backbone ……34

LIST OF FIGURES

Figure 1. CRISPR: SpCas9 components …… .3

Figure 2. Experiment designs of 3 plasmids ……... 4

Figure 3. Developed methods for guide RNA construction …… 6

Figure 4. Primers design for one-piece PCR for guide RNA …….. 7

Figure 5. PCR results for each set of primers …….. 8

Figure 6. Region of SV40 promoter targeted by 11 different guide RNA …….. 8

Figure 7. Miniprep results after condition optimization …….....9

Figure 8. Engineered sgRNA with 2 MS2 stem-loop and PCR results ……..……....10

Figure 9. Miniprep results with optimized condition ……..11

Figure 10. Backbone inserted with sequence cut by type IIS restriction enzyme …...13

Figure 11. Steps after gRNA prepared: 3 plasmids transient transfection ……..14

Figure 12. pU6-sgGAL4-4 transfection condition optimization …….16

Figure 13. dCas9-OGT transfection efficiency …….. 16

Figure 14. Western blot of GAPDH ……18

Figure 15. PVDF membrane image shown to exclude some possible errors made …19

Figure 16. Brief mechanism of luciferase chemiluminescence ……. 19

Figure 17. Luciferase result of guide RNA 1-5 ……..20

Figure 18. O-GlcNAc Chemoenzymatic labeling ……. 23

Figure 19. Newest O-GlcNAc metabolic reporter …….24

Figure 20. Future work: split OGA/Turbo ID system with sgRNA inserted with 2 different stem-loops ……24

ACKNOWLEDGMENTS

First, I would like to thank Professor Jia Niu for the opportunity to work in his laboratory on a very exciting project. Without this chance, I can't explore any faster chemical biology than I expect before. In addition, the standard, way, enthusiasm, and curiosity you own in science are always inspiring.

Second, I would like to thank my mentors at National Taiwan University. Any words can't express my thankfulness for your encouragement, kind words, and experience sharing.

Finally, I want to thank my parents, my pet, my family, and myself who always support me and hope for me to live a happy life. You are the main reason that I could still be on my way to my dream. I would never forget why I am here and wouldn't stop my growth in the following endless years.

Chapter 1: Introduction of O-GlcNAc and CRISPR

1.1 Introduction and importance of O-GlcNAc modification

O-GlcNAc (*O*-linked β-*N*-acetylglucosamine) is a post-translational modification that is found on serine and threonine which plays an important role in many cellular networks such as transcription, protein stability, phosphorylation competition, etc. Dysregulation of *O*-GlcNAc has been found in many pathologies: including neurodegenerative disorders, and cancers, and is implicated in increasing the phenomenon of autophagy. Contrary to its diverse role in the cellular network, the dynamic modification of O-GlcNAc is only controlled by one eraser and one writer, OGA and OGT, which is relatively simple than phosphorylation dynamics.[1,2,3,4]

Post-translational modification is a process the side chain of an amino acid is added or cleaved with modification after biosynthesis. Readers, writers, and erasers involve in the process and result in changes in protein structure, protein function, protein-protein interactions, and protein-DNA/RNA interaction. There are many modifications such as methylation, acetylation, phosphorylation, and ubiquitination found important in various biological processes. However, the knowledge of O-GlcNAc modification in the role of gene expression is still not well-known.

In this research, we are interested in the role and impact of O-GlcNAc modification. Here, we propose to construct the CRISPR (clustered, regularly interspaced, short palindromic repeats) technology with the only eraser and the writer for further discovery. The hypothesis is made that by selectively target to the sequence,

the CRISPR system fused with OGT/OGA will modify the proteins proximal to the DNA region and gene regulation will change because of the O-GlcNAc role which we are interested in for further discovery.

1.2 CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats): introduction and variants

There are two types of CRISPR systems: Class I and Class II. Class I system works with protein complexes that are not suitable for usage in gene editing or gene regulation approach. Hence, the CRISPR system we use in our project is one of Class II; SpCas9 which is from *Streptococcus pyogenes* and works by only a single protein with multi-subunit domains. Other Class II Cas9 proteins include SaCas9, NmeCas9, CjCas9, StCas9, and CdCas9 and they are from *Staphylococcus aureus, Neisseria meningitidis, Campylobacter jejuni, Streptococcus thermophilus,* and *Corynebacterium diphtheriae* separately. Cas9 proteins from different strains have different Cas9 protein structures (sequence), sequence of scaffold RNA, and the protospacer adjacent motif (PAM site) that is recognized by Cas9 protein and is directly 2-6 nucleotide downstream of the sequence targeted by guide RNA.[5,6]

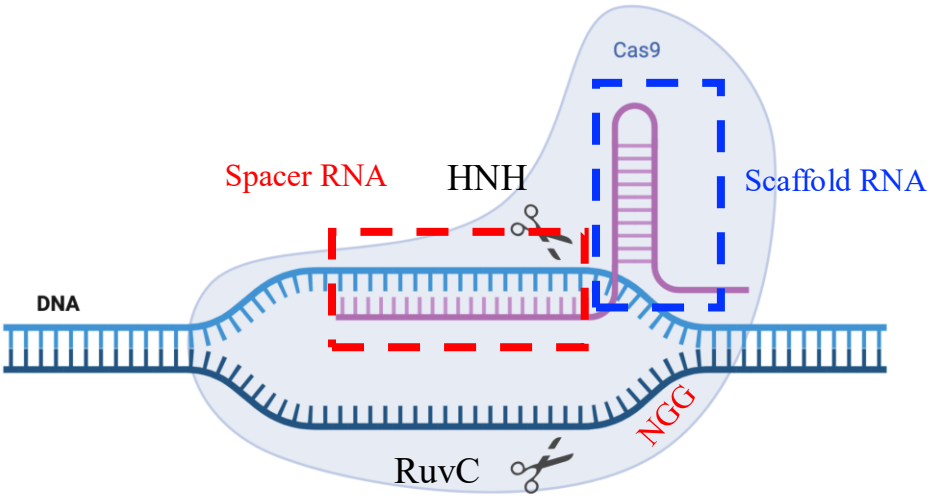


Figure 1. CRISPR: SpCas9 components

Native Cas9 protein and mutant Cas9 protein have been developed for different applications. In the beginning, native SpCas9 was found to make a blunt end double-strand break of DNA targeted sequence by the HNH nuclease domain and RuvC nuclease domain of Cas9. HNH nuclease cleaves the DNA strand complementary to the guide RNA and RuvC nuclease cleaves the opposite strand of DNA. The most popular mutants made for catalytic amino acids of the HNH domain are H840A and D10A for RuvC.

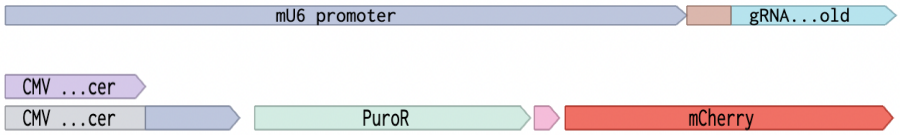
Nickase Cas9 (nCas9) is made by one of two mutants and dead Cas9 (dCas9) is made by double mutants. So far, nCas9 has been used for base editings such as A to G, C to T, and C to G, prime editing, enhanced selectivity of gene editing by homology-directed repair, etc.[7,8] However, dCas9 has been used in a different application even though it loses its nuclease ability while sustaining the ability to bring the Cas9 system to the target sequence. Catalytic inactive Cas protein is fused with many kinds of proteins such as transcription activator, transcription repressor, methyltransferase, fluorescence reporter, DNA polymerase, and kinases to regulate the cellular system,

label genomic locus, or find out the role of different kinds of modifications on protein, DNA or RNA in the cellular network.

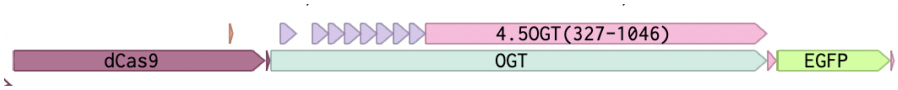
1.3 Experimental design

In the project, we use catalytic inactive Cas9 protein from *Streptococcus pyogenes* to fuse with OGT/OGA and target the sequence next to NGG to further discover the O-GlcNAc modification roles in gene regulation. We will transiently transfect HEK293T with three plasmids including guide RNA, dCas9-OGT/OGA, and luciferase reporter which the protein expression is driven by SV40 promoter. Several guide RNAs will be constructed to target the SV40 promoter and regulate the luciferase protein expression. T2A is installed between PuroR/mCherry and dCas9-OGT/EGFP.

Plasmid 1: pU6-sgGAL4-4



Plasmid 2: dCas9-OGT



Plasmid 3: pMLS-SV40-firefly luciferase

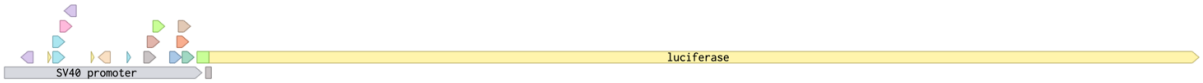


Figure 2. Experiment designs of 3 plasmids

Chapter 2: Results and Experiments

2.1 gRNA plasmid construction

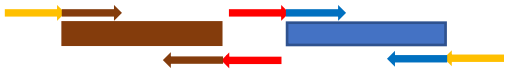
Since the development of CRISPR technology, different kinds of gRNA plasmid construction have been developed to full fill the need to target genome loci with gRNA libraries. In the system of SpCas9, spacer RNA is usually composed of 18-24 nucleotides of RNA which is complementary to the DNA that the Cas9 system targets. Two-piece Gibson assembly, long oligonucleotide annealing, and application of Type IIS restriction enzyme have been developed to reach out to the need. Among those methods, double recognition and cuts by Type IIS restriction enzyme is the way which is used most broadly. However, there is no double Type IIS in the backbone of the plasmid used in our project. To construct gRNA with the method, small oligonucleotide annealing inserted with double Type IIS recognition of AarI could be used to construct the backbone. On the other hand, two-piece Gibson assembly and long oligonucleotide annealing seem to be relatively not cost-efficient or time-efficient methods to construct gRNA.[9,10,11]

Hence, the possibility of a one-piece PCR/one-piece Gibson assembly has been developed in the project, and the reaction is optimized and applied to two gRNA plasmid backbones, one is for native gRNA of SpCas9, and another is for scaffold RNA which is inserted with double MS2 stem-loop. With the optimization, only one forward primer is needed for tons of different kinds of spacer, and we believe this kind of design logic could be applied to different kinds of Cas proteins that have different scaffold RNA.

Method 1

Cut by Type IIS restriction enzyme

Method 2



Two-piece Gibson assembly

Method 3

….

Long oligonucleotide annealing

Figure 3. Developed methods for guide RNA construction

Firstly, assemble junction is set up between forward and reverse primer over 30 nt. However, there is no PCR product according to the design even though several ways have been used such as adding DMSO into the PCR reaction, using more DNA polymerase, or diluting the concentration of the primers. There is still a faint band under the gel which is around 200 bp. The reason each PCR reaction doesn't work might be because forward and reverse primer bind together and interrupt their binding to the backbone. To avoid the situation, the junction is moved to the spacer RNA region, but there is a new problem which is a non-specific product derives from the design. The junction is then moved to the scaffold RNA region that expected size band is shown after diluting the concentration of the primers.

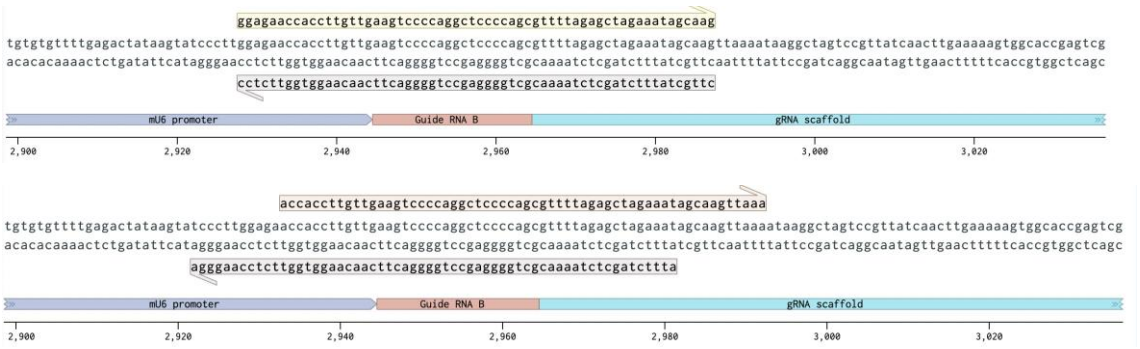
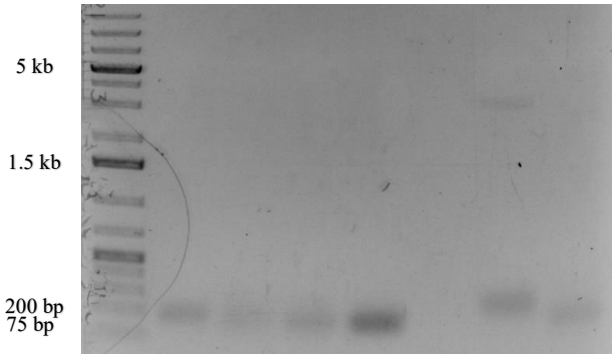


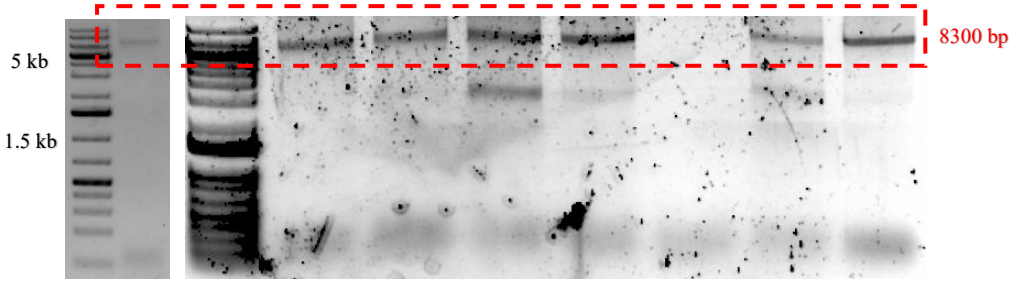
Figure 4. Primers designed for one-piece PCR for guide RNA



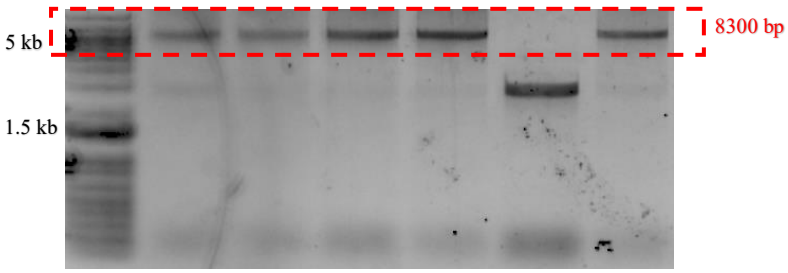
PCR A B C D B^a C^b

a: more Q5 polymerase

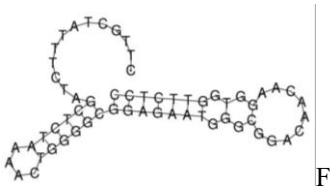
b: primer concentration dilution



PCR B A C D E F G H



PCR I J K L M N



F reverse primer: Free Energy (delta G) = -17.20 kcal/mol

Figure 5. PCR results for each set of primers

However, there is no PCR band shown on gRNA F, and a non-specific PCR shown gRNA M. The reason for the failure of gRNA F might be the reverse F primer will form a very stable hairpin that is not good for the PCR reaction. And the non-specific band from gRNA M might be solved by moving the junction 4 nucleotide left so the annealing affinity part to the target site is larger than the non-specific binding site.

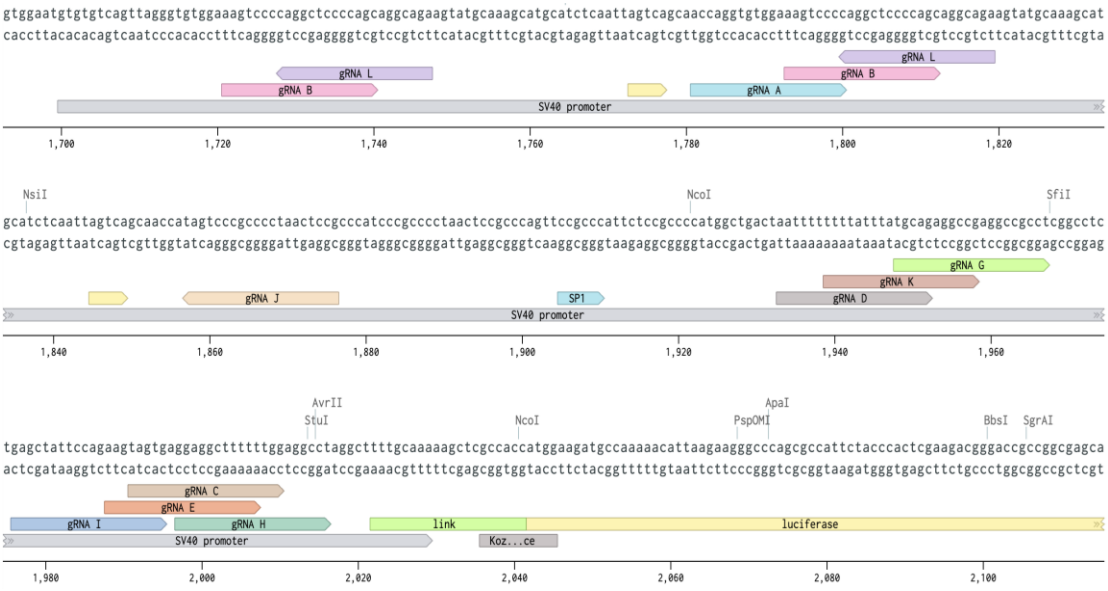


Figure 6. Region of SV40 promoter targeted by 11 different guide RNA

gRNAs are designed to target the SV40 promoter which drives luciferase protein expression. Each spacer RNA is 20 nucleotides in length and NGG is downstream of the region targeted by guide RNA.

The second problem comes after transformation/miniprep, there is almost no target sequence that the spacer RNA designed but most of them are templates with some

undesired bands. The reason might result from the low efficiency of the assembly and there is still a template plasmid in the assembly reaction before the transformation. To purify the PCR reaction which might benefit the efficiency of Gibson Assembly reaction or the possibility to get the newly constructed spacer RNA. DpnI (37 ºC, 1-4 h) is added to degrade the template plasmid which is amplified by TOP10 or Turbo strain of bacteria (both are Dam^+), and the spin column is used to purify the PCR solution by excluding protein that might interrupt the efficiency of Gibson Assembly. However, none of them works. Finally, gel extraction is performed which almost excludes the chance of getting a template but there is still some issue about the salt which might have an impact on Gibson assembly and transformation. While doing electroporation with assembly solution, arcing happens much of the time and there is still some undesired band after picking colonies/miniprep. The phenomenon is much excluded and optimized by increasing the ratio of deionized water and the time of assembly reaction from 15 minutes 60 ºC to an hour 60 ºC.

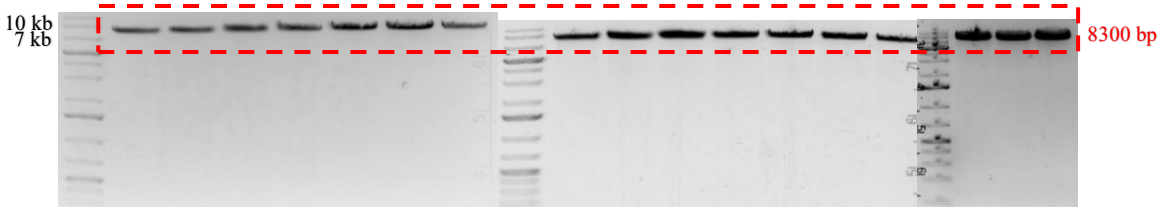


Figure 7. Miniprep results after condition optimization

After transformation and miniprep, the plasmid is digested with BamHI with one cut on the plasmid. There is no undesired band and 1 template after analytic gel or sequencing result.

The optimized primer designs and the procedure is used directly on another backbone for gRNA construction which is used in synergistic activation mediator (SAM). The gRNA difference from the native one is that two MS2 stem-loops which

are bound specifically by the coat protein of MS2 bacteriophage are engineered into scaffold RNA.[12]

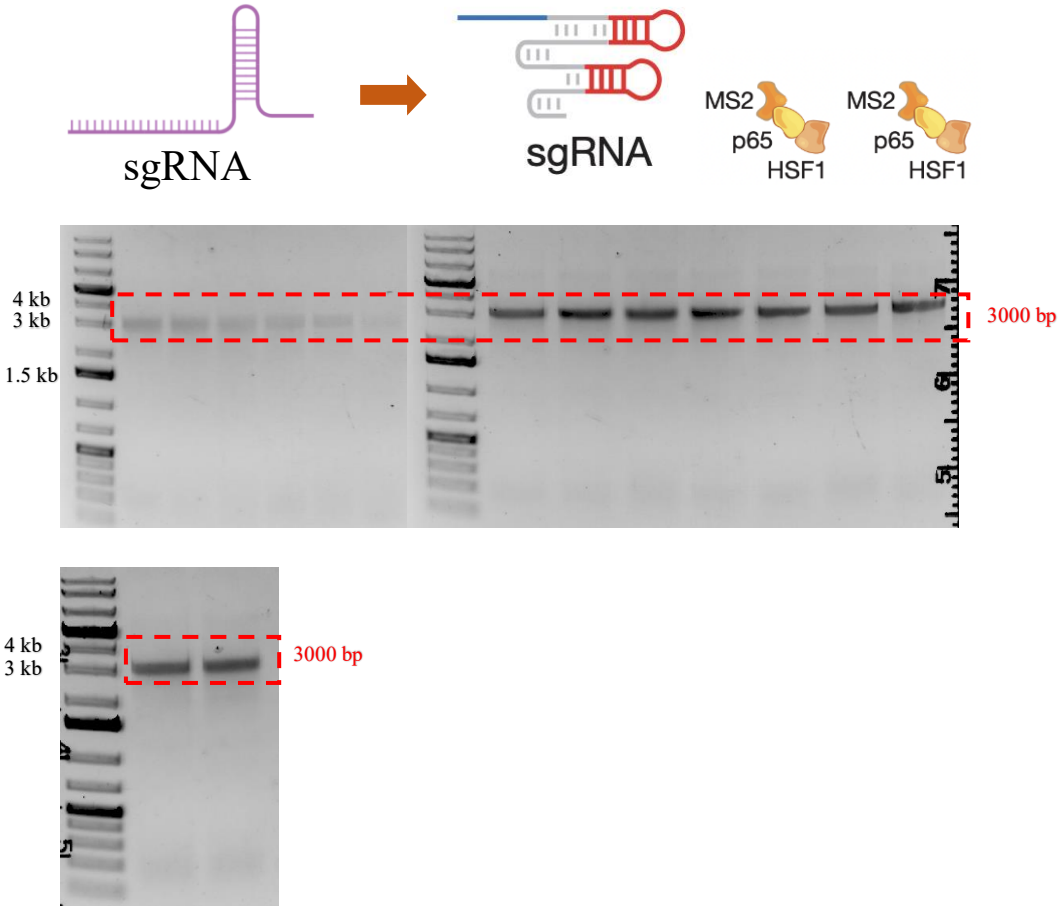


Figure 8. Engineered sgRNA with 2 MS2 stem-loop and PCR results

PCR product is shown above which includes all newly constructed spacers from the first backbone. There is no non-specific product shown in the gel but still with some faint primer dimer product around 75-200 bp. An interesting phenomenon is found that the band at the bottom is relatively faint compared to PCR the first backbone. The reason might be because the junction for Gibson assembly is moved 4 nucleotides left and the reverse primer could anneal much more than the original design.

After Gibson assembly and transformation, miniprep is performed and digested with one cut. There is no undesired band and there is almost no template. Correct

sequence results are obtained which only need to pick 2 colonies for each different spacer RNA.

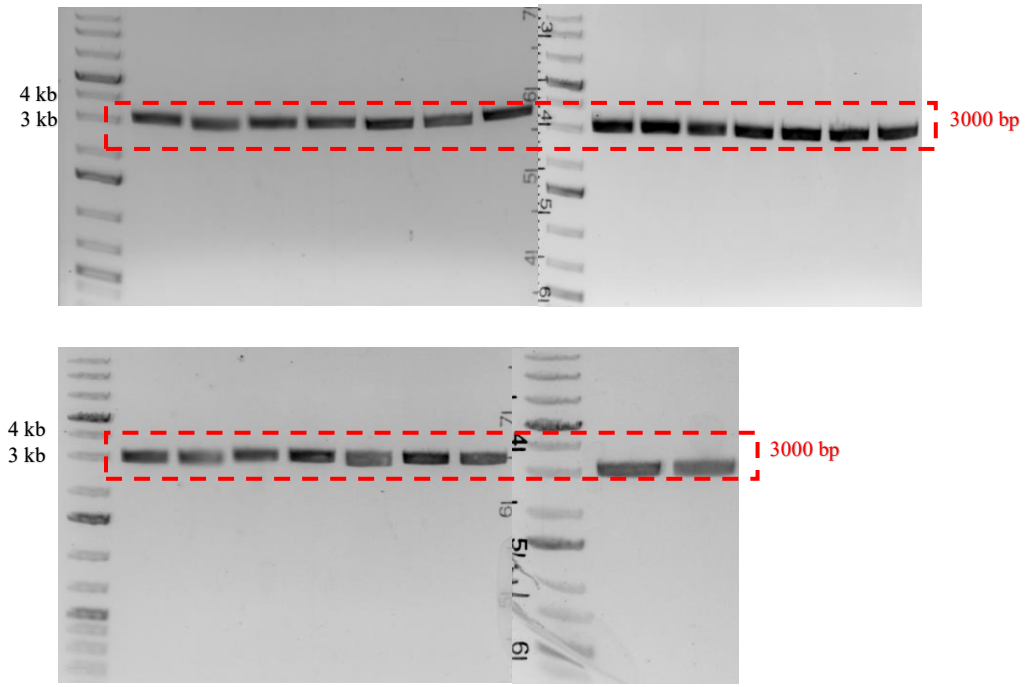


Figure 9. Miniprep results with the optimized condition

2.2 Other cloning methods: two-piece Gibson, ligation, and oligonucleotide annealing

2.2.1 Plasmid construction with ligation

Ligation is performed as a practice for the following oligonucleotide annealing. To construct plasmid with ligation, recognition site of restriction enzyme which causes sticky end is chosen and inserted in the 5'-end of forward and reverse primer with 5-7 nt added upstream of the recognition site. After PCR reaction and double restriction enzyme digestion for each piece, gel extraction (spin column is not good) is performed to exclude the rest of the primers and restriction enzymes which might interrupt the ligation reaction. Ligation reaction is conducted under 4 ºC 24 hours and heat inactivation is performed before electroporation. Transformation done in TOP10 has more colonies than Turbo in my case. However, because of further consideration about

if there is a frameshift, especially for translation fusion and the low reaction efficiency, a two-piece PCR combination is performed mainly by a two-piece Gibson Assembly.

2.2.2 Plasmid construction with two-piece Gibson Assembly

Two-piece Gibson Assembly is performed for a replacement of fluorescence reporter, chemiluminescent reporter (codon-optimized), and promoter (sgRNA target to the different promoter) for future work. Before assembly reaction, gel extraction which is performed to increase assembly efficiency is still needed. (46919_pMLS-SV40-EGFP to 46919_pMLS-SV40-mCherry, 46919_pMLS-SV40-EGFP to 46919_pMLS-CMV-EGFP and 46919_pMLS-SV40-luciferase to 46919_pMLS-SV40-luciferase codon-optimized).

2.2.3 Plasmid construction with oligonucleotide annealing

Oligonucleotide annealing is tried to insert one AarI recognition site into the backbone as a practice to see if the protocol and experiment design work. In my case, Type IIS restriction enzyme (BbsI, BsaI) could not be used because they will cut the regions outside the guide RNA. Hence, two pairs of oligonucleotides one of which contains an AarI recognition site are designed and successfully inserted into the backbone. Three pairs of oligonucleotides are designed to target the region targeted by gRNA F and M but finally, the results are not expected. The reason might because the ligation reaction is not efficient enough for 3 pairs of oligo ligation to work in only 15 minutes. Hence the following ligation with the vector doesn't work either. Vector is prepared by digestion with BstXI and BamHI separately in different buffers and is purified by gel extraction. Each pair of small oligonucleotides is annealed and

phosphorylated in different Eppendorf tubes. After 5'-end phosphorylation, two or three pairs of nucleotides are ligated together and then ligated with the vector.

The sequencing result shows the newly constructed backbone with the AarI recognition site is inserted and the design is shown as follows.

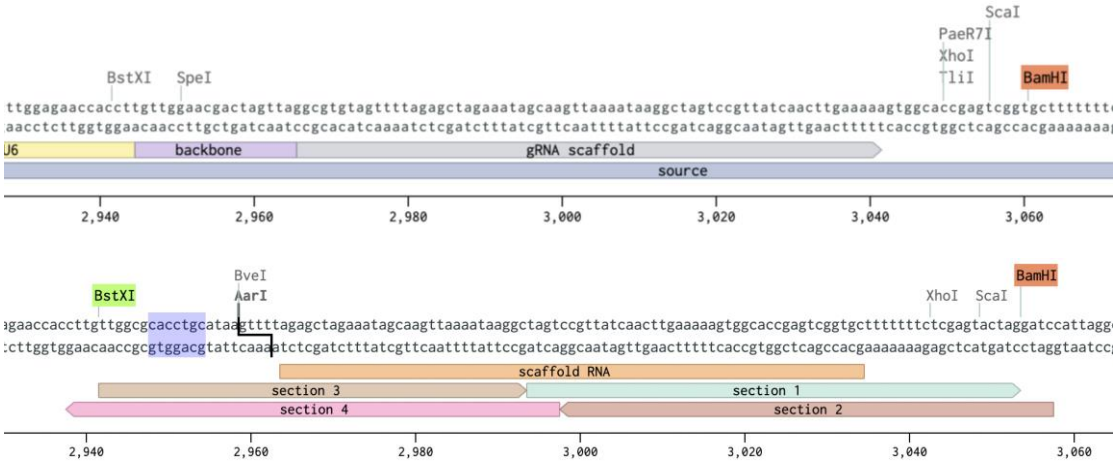


Figure 10. Backbone inserted with sequence cut by type IIS restriction enzyme

2.3 Mammalian Cell Culture and experiments

2.3.1 Troubleshooting of cell culture issue

There is a cell culture issue that happened in the lab which is after cell passages cells couldn't attach to the dish and die. The reason might be that we didn't add FBS in our DMEM, glutaMax medium. Even though glutaxMax supplement is already added in the commercially available medium which could improve cell viability and growth, attachment form cells such as the cell line used in the project, HEK293T, still need growth factors, hormones, etc, which are included in FBS. Without FBS, cells couldn't grow more than 3 passages and will die or lose their ability to attach to the dish in starvation conditions based on my observation.

A Series of troubleshooting has been made to fix or find out the problem which is shown in the table

Entry	Improvement	Reason
1	Disperse cells in the middle of dish	Avoid clamp while dispending cell
2	Disperse cells evenly in dish	Avoid clamp while dispending cell
3	Add Penicillin-Streptomycin	Anti-bacteria, negative and positive gram
4	Pipette smoothly	Pipette too strong will kill the cell
5	Pipette strongly	Suspend cell into single cell, not clamp
6	Finish cell passage in 5 minutes	Avoid death of cell from external pressure, temperature, % of CO_2, possible contamination
7	Trypsin time longer	Make cell into single cell
8	Trypsin time shorter	Over-trypsinization kill cell or cell can't attach
9	Lab coat contamination	Specific lab coat for cell culture room
10	Medium contamination	Open a new bottle of medium
11	Technical issues	Not crossing over dishes, not possible contact contamination of medium, trypsin or PBS.

Table 1. Troubleshooting of cell culture issue

2.3.2 Transient transfection

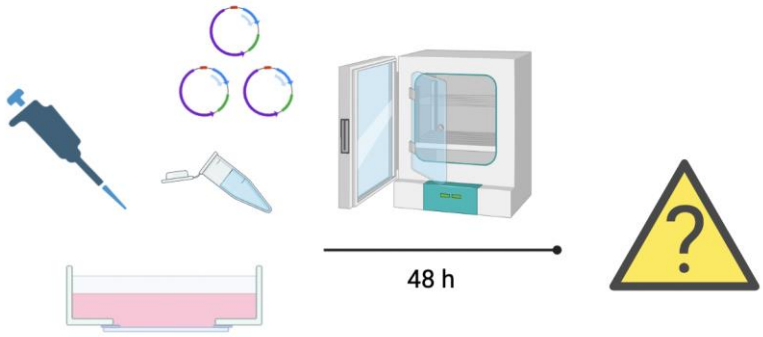


Figure 11. Steps after gRNA prepared: 3 plasmids transient transfection

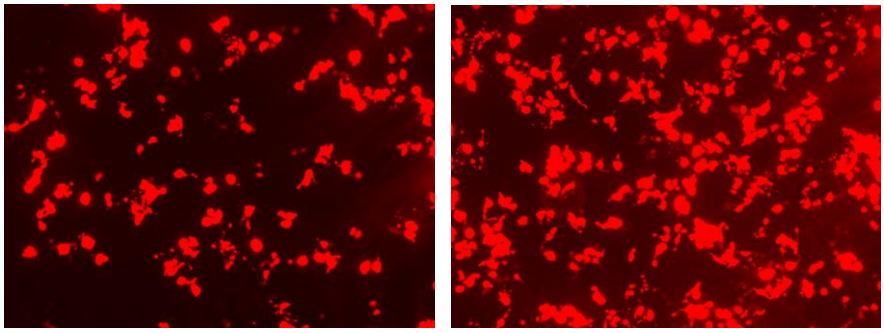
5 μg of each plasmid is transiently transfected into HEK293T cell. The ratio of plasmid/PEI (1 mg/mL) is tested at 1:3, 1:4, or 1:5 during the incubation. Less Opti-MEM for incubation is found to have higher transfection efficiency which might result

from a higher interaction rate of PEI and plasmid so more plasmids could be bound with PEI with a positive charge and enter the cell with endocytose pathway. The difference in transfection efficiency is found between the cell just revived and the cell already passaged 25 times. Furthermore, transfection efficiency is found to be higher when the experiment is performed while the cell is at 60-70% confluency than at 80% confluency or above.

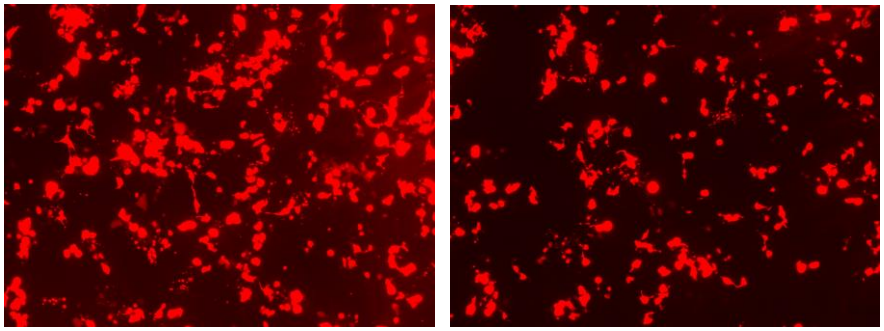
To further increase transfection efficiency which might be needed in the future, more amount of plasmid could be transfected into a 10-cm dish from 5 µg to 10 or more and the ratio of transfection reagent and plasmid should be tried further. Different kinds of methods could be used such as lipid-based transfection reagent, lipofectamine 2000, or lipofectamine 3000. Bright-field should be presented overlapping with fluorescence reporter imaging, but the data can't be read for no reason.

Entry	Amount of plasmid µg	Volume of 1 mg/mL PEI (µL)	Opti-MEM (µL)	Incubation time (min)
1	5	15	125	15
2	5	20	125	15
3	5	25	125	15
4	5	25	200	15

Table 2. Transient transfection condition optimization



Entry 1 Entry 2



Entry 3 Entry 4

Figure 12. pU6-sgGAL4-4 transfection condition optimization

Condition of entry 2 is applied to another plasmid with eGFP and the image is presented as follows. The transient transfection of eGFP plasmid is replicated and the result of transfection efficiency seems to be reliable and fine.

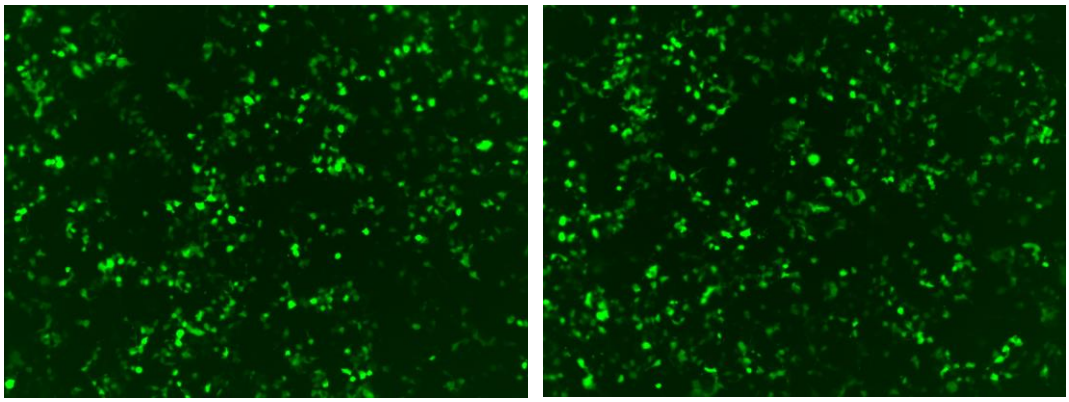


Figure 13. dCas9-OGT transfection efficiency

2.3.3 Western blot practice for future work for anti-O-GlcNAc with RL2 or CTD110.6.

To perform western blot, the cell medium is removed, and the cell is washed with PBS (could be +/+ Ca^{2+}/ Mg^{2+}) 2-3 times to make sure there is no rest of the medium with a little pink color which might interrupt the western blot result. The cell is then collected with 2 mL PBS in a 15 mL tube and incubated in ice for 15-20 minutes before

ultrasonication for protein extract. This step is to lower the overall temperature so cavitation bubbles from ultrasonication couldn't reach out to high temperatures and cook the proteins in lysates.

The cell is ultrasonicated until there are no pale white particles (cell) precipitate at the bottom of the tube then the protein extract could be used directly for the following western blot procedure or storage at -20 ºC or -80 ºC temperature. Repeat freezing and thawing should be avoided so 2 mL (or a little bit more because of the cell) could be aliquot into 40 (less or more) 0.5 mL tubes with 50 µL protein extract for each depending on the expected repeat time of western blot or storage capacity.

15 µL protein extract is mixed with a 5 µL blue ladder (4X to 1X) and incubated under a condition of 90 ºC for 5 minutes. Stacking gel (4%) solution is added to the stack and then add about 1 mL of isopropanol on the top evenly to make the end of the curve straight. After 30 minutes of polymerization, isopropanol is removed. Resolving gel (12%) is added to the specific height and the cap is inserted. After 30 minutes of polymerization, the cap is carefully removed, and samples are loaded into each well that the tip should be deep enough to make sure samples do not disperse/diffuse into the running buffer solution but do not destroy the gel. The protein ladder is loaded aside.

In the beginning, 120 V 1 hour is performed to separate the protein, but it does not work well. Several trials have been made to troubleshoot the problem if there is an error made in buffer preparation or gel preparation. Finally, the condition is changed to 180V for 1 hour and it works well which means it is not any problem with preparation in advance.

After protein separation, the gel is carefully removed from the stack and is cut into a shorter piece that can fit better into the sandwich for protein transfer to the PVDF membrane. 10% of methanol contained in transfer buffer is found to have better transfer ability. PVDF membrane is incubated into methanol for 15 sec for activation and is

found to have better transfer ability too. The first time, 20 mA 2 hours is performed for the transfer condition, however, it doesn't work well for the rest of the imaging detection or even the protein ladder transfer. Hence, 150 mA 1 hour is performed (sometimes need 10-15 minutes more) to transfer the protein.

Membrane blocking is performed at room temperature for 4 hours to overnight but finally found 1 hour is enough and over blocking under room temperature conditions is not good. The membrane is then washed with wash buffer and the following primary antibody incubation; secondary antibody incubation is performed. After each time of incubation, the membrane is washed with wash buffer 3 times, 5 minutes for each. Secondary antibody incubation should be less than 2 hours. Horseradish peroxidase is conjugated with a secondary antibody so Clarity^{TM} Western ECL substrate, 200 mL is added to the membrane for detection. 1 mL of each luminol reagent and peroxide reagent is added to the membrane with an area of around 10 cm^2.

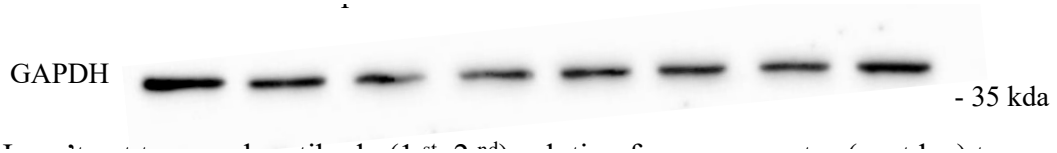


Figure 14. Western blot of GAPDH

While performing primary antibody incubation, 10 mL solution is found out that it couldn't cover all the membrane in the smallest tips covered for incubation procedure by the rocker. The middle of the membrane is found sometimes expose to the air while rocking. This could be fixed by using more primary antibody solution or a shaking mode of incubation.

Buffer preparation, gel preparation, and the procedure of membrane transfer seem not to be an issue in the picture shown below. The picture is filmed after membrane transfer and before blocking/primary antibody incubation. There is no uneven distribution of bands between each well ranging from small to large proteins and the

line of each well of the protein ladder is straight.

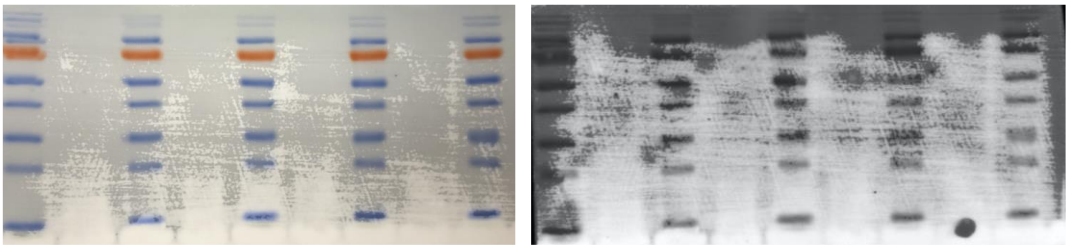


Figure 15. PVDF membrane image is shown to exclude some possible errors made

2.3.4 luciferase assay

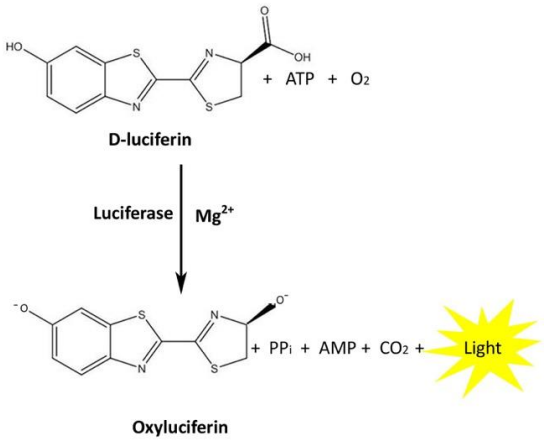


Figure 16. Brief mechanism of luciferase chemiluminescence

Luciferase, guide RNA, and the dCas9-OGT plasmid are transfected into HEK293T with PEI and the signal is detected after 48 hours of transfection. Before cell lysis and the following procedure to do the luciferase assay, the medium is removed, and the cell is washed with PBS 2-3 times until there is no pink color in the solution (medium). To make the assay result more reliable, the cell must be cultured with FBS because once the cell is cultured in starvation conditions, the cell is extremely easy to come off by PBS which is very hard to avoid losing some cells. The deviation of some part of the signal is huge and might also come from each transient transfection

efficiency or after cell lysis with the kit procedure, cell lysates from some of the sample aggregate easily even with smooth/harsh pipetting.

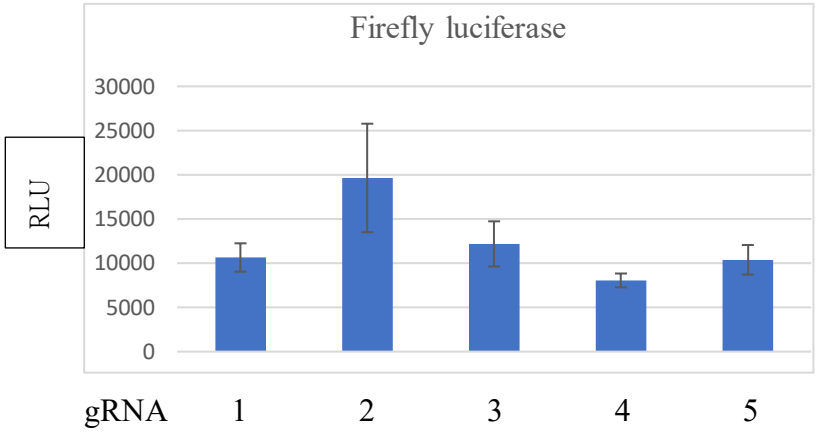


Figure 17. Luciferase result of guide RNA 1-5

Those five guide RNA are constructed by our collaborator with a long oligonucleotide annealing method to construct the gRNA. Each of the oligo pairs is over 100 bp. According to the result, different gRNA designed to target the SV40 promoter results in the different protein expression levels of firefly luciferase, and the rest of the gRNA constructed by one-piece Gibson assembly would be tested and tried to find out the role of O-GlcNAc among the signaling pathway.

To improve the reliability of the assay result, further improvements to the experiment designs should be applied and written in future work.

Chapter 3: Conclusions and Discussion

3.1 Summary of work

One-piece PCR and Gibson Assembly are developed for gRNA plasmid construction which avoids using Type IIS restriction enzyme and are more feasible to

construct plasmid with less limitation. In addition, it is also a relatively cost and time-efficient gRNA plasmid construction developed by optimized primer designs and the following procedure to exclude the template, and undesired bands and increase the assembly efficiency. To construct gRNA, only one forward primer is needed for different reverse primers and is confirmed to work on two backbones. The table below shows the Type IIS restriction enzyme cut site in some elements while designing plasmids.

NEB/ThermoFisher	Cut Site
BbsI/BpiI	WPRE, mCherry, VSV-G, NmCas9, BGH polyA, luciferase*,
BsaI/Eco31I	Puro, Amp, 5'-LTR, VSV-G, hGH polyA,
AlWI/BspPI	Amp, 5'-LTR, WPRE, NeoR/KanR,
BfuAI/BveI	SaCas9, PP7CP, chimeric intron, HIV-pol, nucleoplasmin NLS, NmCas9, luciferase*, IRES, NeoR/KanR, SaABE8e, SpCas9,
PaqCI/AarI	HIV-pol, chimeric intron, NmCas9, luciferase*, IRES,
BsmBI/Esp3I	Puro, gentamicin resistance, hGH polyA,

Table 3. Plasmid elements cut by different types of IIS restriction enzymes

Other cloning methods have been used to construct different plasmids for different needs including ligation, two-piece Gibson, and oligonucleotide annealing. All of them are successfully performed so there are several alternative ways to construct recombinant DNA if one has an issue. Experiments such as western blot, luciferase, transient transfection, and cell culture issues have been performed and the reasonable/possible improvement or the issue which is already confirmed is pointed out.

3.2 Experiment designs improvement

To develop a reliable luciferase assay, cells should be cultured with FBS, and experiment designs should include a second bioluminescence reporter for normalization. The second reporter should not be constructed on the same plasmid as the first luciferase or use self-cleavage peptides to construct them together. Luciferase of pMLS-SV40-Luc is replaced with luciferase PLR1. CMV-Puro-T2A-mCherry construct with gRNA is cloned into EF1 alpha-luciferase PLG3-T2A-mCherry by two-piece Gibson Assembly.[13]

In addition, experiments would be performed with a 96-well plate with control PLR1/PLG3 only transfected. After 48 hours of transfection, the cell is lysed and both substrates are added to each well at the same time, and data is collected by the plate reader.[14,15]

	Name	Substrate	RLU wavelength
Luciferase 1	PLR1	LH2	635 ± 16 nm
Luciferase 2	PLG3	BtLH2	516 ± 10 nm

Table 4. Dual-luciferase system used for experiment design improvement

The cell line used in the project is HEK293T but not HEK293 which might be more complicated to discuss because of a non-natural protein in a mammalian cell.

3.3 Future Work

Once all gRNAs are tested triplicate in 96-well with a second reporter, the candidates with the most and least signal from PLR1 after normalization would be further tested with one or some of the following methods to try to find out which proteins are modified or less modified by O-GlcNAc.

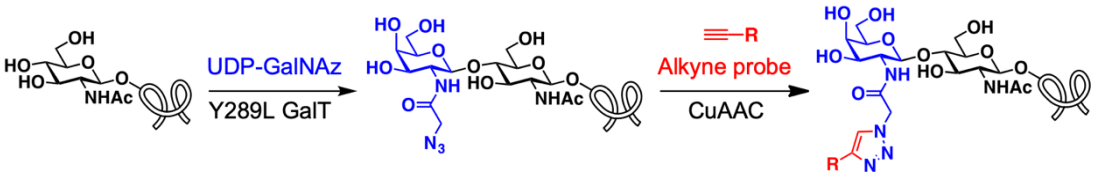


Figure 18. O-GlcNAc Chemoenzymatic labeling

3.3.1 After 48 hours of transfection with three plasmids, candidates are then added with expressed and purified Y289L GalT/ UDP-GalNAz at the end of O-GlcNAc and N-glycosidase F (PNGase F) is treated to remove GlcNAc-containing N-linked glycans. CuAAc (Copper-catalyzed azide-alkyne cycloaddition) reaction is performed to click the azide and Biotin/TAMRA conjugated alkyne for following in-gel fluorescence or biotin/streptavidin pull-down for protein identification.[4] However, there should be more other proteins O-GlcNAcylated by dCas9-OGT after biosynthesis but are not related to the gene regulation of luciferase.

3.3.2 To lower the signal-to-noise ratio, split turbo ID and split OGA could be used to specifically find out which proteins are further modified.[16] However, the scaffold RNA should be designed to insert two different stem-loop such as MS2 and PP7 which their corresponding coat proteins bound specifically. And for example, plasmid encodes PCP-Turbo(N) and MCP-Turbo(C) should be transfected at the same time. Split-OGA system developed by Woo's lab could also be used in the project to deregulate

GlcNAcylation of the proteins proximal to the genome locus.[17] Selective metabolic O-GlcNAc reporter ($Ac_3$6AzGlcNAc) could also be used in the project for easier discovery of proteins that are O-GlcNAcylated or not.[18]

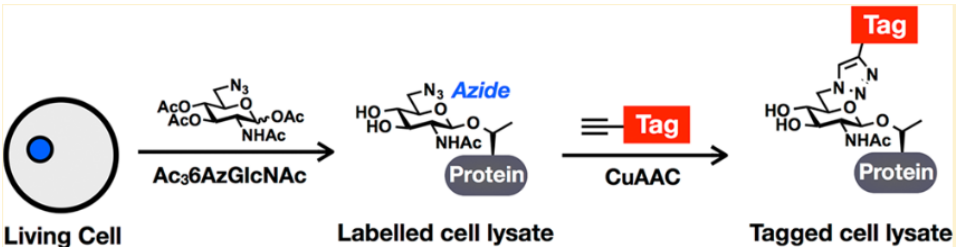
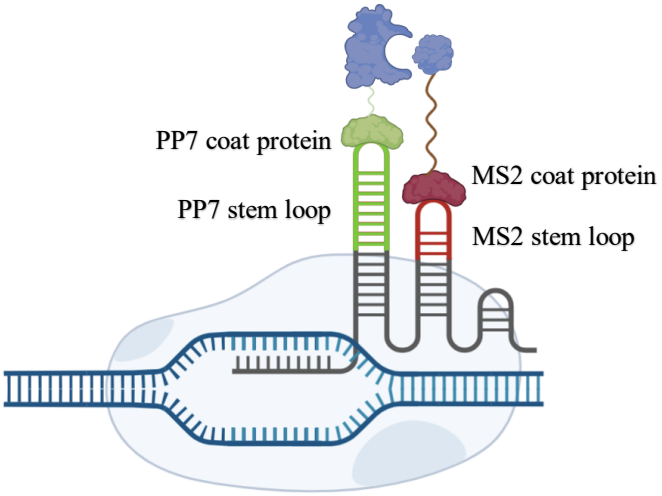


Figure 19. Newest O-GlcNAc metabolic reporter



PCP is fused with one part of turbo ID or OGA.

MCP is fused with another part of turbo ID or OGA.

Dead SpCas9 is still needed to localize genome loci.

Scaffold RNA is engineered with one PP7 stem-loop and an MS2 stem-loop.[19]

Figure 20. Future work: split OGA/Turbo ID system with sgRNA inserted with 2 different stem-loops

Chapter 4: Material and Methods

Gels and western blot were imaged with the BioRad ChemiDoc imaging system. Absorbances were taken using the Thermo Fisher Scientific NanoDrop One Micro-UV/Vis Spectrophotometer. Cells were monitored for viability with EVOS XL Core imaging system from Thermo Fisher Scientific and imaged with Zeiss Axio Observer A1 X-Cite Series 120 Q combination microscope.

To clone the plasmids of interest using Gibson assembly, the backbones and inserts were amplified with PCR. A 50 µL reaction of 1X Q5 reaction buffer (NEB), 200 µM dNTPs (NEB), 0.25-0.5 µM forward primer (Genewiz), 0.25-0.5 µM reverse primer (Genewiz), and 0.02 U/µL Q5 polymerase (NEB) was subjected to the thermal cycler with the following program: [98 °C for 30 s] + [98 °C for 15 s + annealing temperature for 30 s + 72 °C for (# k length + 20-50 s depends on target size)] * 30 + [72 °C for 2 min] + [4 °C hold]. The amplified product was purified by PCR purification using the QIAGEN QIAquick PCR purification kit according to the manufacturer's protocol, and the resulting DNA template concentration was quantified by UV spectroscopy at Abs_{260}. DNA integrity was analyzed by agarose gel electrophoresis with a 1% agarose gel, alongside the GeneRuler 1 kb Plus DNA ladder (ThermoFisher Scientific) in 1X TAE buffer at 120 V for 45 minutes stained with 0.33 µg/mL EtBr solution. Gel extraction is performed based on the protocol of the QlAquick Gel Extraction Kit. Once the PCR products were purified and quantified, an assembly reaction comprising 1X Gibson assembly master mix (NEB), 0.03 pmol vector, and 0.06 pmol insert(s) was diluted to 10 µL with nuclease-free water. The reaction was subjected to the thermal cycler with the following program: [50 °C for 15 min] + [4 °C hold]. For one-piece Gibson assembly for sgRNA, the solution is diluted, and time is increased from 15 min to 1 h.

After the assembly reaction, 1 µL of the product was transformed into NEB Turbo or other strain of electrocompetent *E. coli* cells using a benchtop electroporator, and the transformed cells were incubated in 1 mL SOC media (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 20 mM glucose) for recovery at 37 °C for 1 hour. After recovery, 100 µL recovered cells were cultured on Luria broth agar plates containing 100 µg/mL ampicillin and incubated at 37 °C for 12-16 hours for Turbo, 18-24 hours for TOP10. The following day, several single colonies were inoculated in Luria broth supplemented with 100 µg/mL ampicillin and grown at 37 °C for 12-14 hours for Turbo, 16-24 for TOP10 at 250 rpm. The overnight cultures were then isolated and purified using the GeneJET Plasmid Miniprep kit (ThermoFisher Scientific), and the resulting plasmid concentration was quantified by UV spectroscopy at Abs_{260}. To confirm that the isolated plasmid was correct, restriction enzyme digests were performed following NEB's standardized protocols, and the samples were sent out for Sanger sequencing (Azenta/Genewiz) to confirm the DNA sequence.

HEK293T cells were treated in 100-mm culture dishes (Corning). Cells were transfected at approximately 70% confluency with 1 µg/µL or 1 mg/mL PEI (Sigma Aldrich) and 5 µg of the purified plasmid of interest. The cell viability was measured every 24 hours post-transfection using the bright-field microscope, and fluorescence was visualized with the fluorescent microscope after 48 hours to visualize mCherry, eGFP, and gene expression levels. HEK293T was seeded in a 10 mm dish and transfected with luciferase. Luciferase assay is detected with the protocol Pierce Firefly Luciferase Glow Assay Kit under BioTek synergy Neo-2, Gen5 Image Prime 3.10. Stacking gel (4%): 1.98 mL 30% Acrylamide/bis, 3.78 mL 0.5M Tris-HCl, pH 6.8, 9 mL diH_2O, 150 µL 10% SDS, 15 µL TEMED, 75 µL 10% APS. % of the Resolving gel is adjusted based on the size of the protein of interest. Resolving gel (12%): 6.0 mL 30% Acrylamide/bis, 3.75 mL 1.5M Tris-HCl, pH 8.8. 5.0 mL diH_2O, 150 µL 10%

SDS, 7.5 µL TEMED, 75 µL 10% APS. SDS-PAGE running buffer: 1g SDS, 3 g Tris-base, 14.4 g glycine. Transfer buffer: 3 g Tris-base, 14.4 g glycine, 900 mL diH_2O, 100 mL methanol (% of methanol in the solution could be adjusted). Transfer membrane PVDF. Blocking buffer: 1 g blotting grade blocker non-fat dry milk, 200 µL Tween 20, 50 mL 1X PBS. Wash buffer: 50 mL 1XPBS, 200 µL Tween 20.

Chapter 5: Reference

1. Parker MP, Peterson KR, Slawson C. O-GlcNAcylation, and O-GlcNAc Cycling Regulate Gene Transcription: Emerging Roles in Cancer. Cancers. 2021; 13(7): 1666.
2. Rahman MA, Hwang H, Cho Y, et al. Modulation of O-GlcNAcylation Regulates Autophagy in Cortical Astrocytes. Oxidative Medicine and Cellular Longevity. 2019, 1-13.
3. Brimble S, Wollaston-Hayden E, Teo C, et al. The Role of the O-GlcNAc Modification in Regulating Eukaryotic Gene Expression. Current Signal Transduction Therapy. 2010; 5(1): 12-24.
4. Kim EJ. Chemical Arsenal for the Study of O-GlcNAc. Molecules. 2011, 16(3): 1987-2022.
5. Kim E, Koo T, Park SW, et al. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nature Communications. 2017; 8(1): 14500.
6. Hirano S, Abudayyeh OO, Gootenberg JS, et al. Structural basis for the promiscuous PAM recognition by Corynebacterium diphtheriae Cas9. Nature Communications. 2019; 10(1): 1968.
7. Kurt IC, Zhou R, Iyer S, et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. Nature Biotechnology. 2020; *39*(1): 41–46.
8. Anzalone AV, Koblan LW, Liu DR. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nature Biotechnology. 2020; *38*(7): 824–844.
9. https://www.addgene.org/crispr/zhang/

10. Kim H, Colaiácovo MP. CRISPR-Cas9-Guided Genome Engineering in *Caenorhabditis elegans*. Current Protocols in Molecular Biology. 2019; 129(1): e106.

11. Wang T, Lander ES, Sabatini DM. Single Guide RNA Library Design and Construction. Cold Spring Harbor Protocols. 2016; 2016(3): pdb.

12. Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature. 2014; 517(7536): 583–588.

13. Wang X, Xu Z, Tian Z, et al. The EF-1α promoter maintains high-level transgene expression from episomal vectors in transfected CHO-K1 cells. Journal of Cellular and Molecular Medicine. 2017; 21(11): 3044–3054.

14. Branchini BR, Southworth TL, Fontaine DM, et al. A Firefly Luciferase Dual Color Bioluminescence Reporter Assay Using Two Substrates to Simultaneously Monitor Two Gene Expression Events. Scientific Reports. 2018b; 8(1): 5990.

15. Woodroofe CC, Meisenheimer PL, Klaubert DH, et al. Novel Heterocyclic Analogues of Firefly Luciferin. Biochemistry. 2012b; 51(49): 9807–9813.

16. Cho KF, Branon TC, Udeshi ND, et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. Nature Protocols. 2020b; 15(12): 3971–3999.

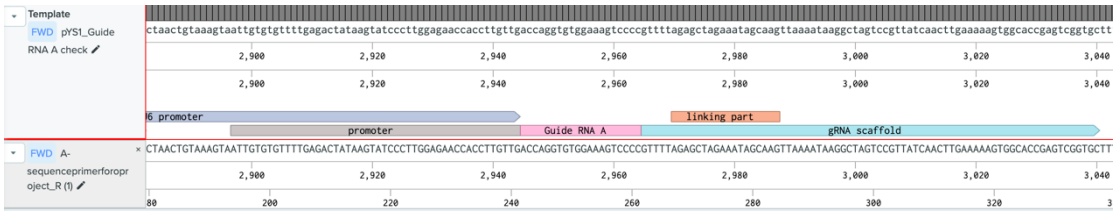
17. Ge Y, Ramirez DH, Yang B, et al. Target protein deglycosylation in living cells by a nanobody-fused split O-GlcNAcase. Nature Chemical Biology. 2021b; 17(5): 593–600.

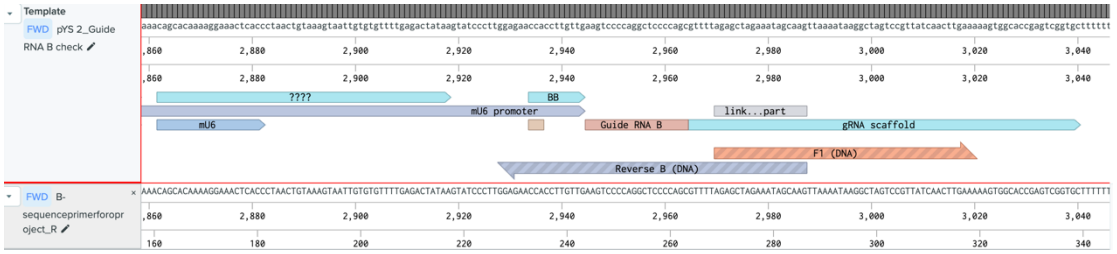
18. Chuh KN, Zaro BW, Piller F, et al. Changes in Metabolic Chemical Reporter Structure Yield a Selective Probe of O-GlcNAc Modification. Journal of the American Chemical Society. 2014; 136(35): 12283–12295.

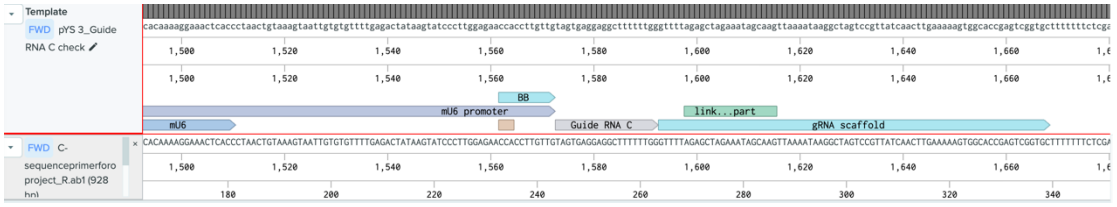
19. Ma H, Tu LC, Naseri A, et al. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nature Biotechnology. 2016b; 34(5): 528–530.

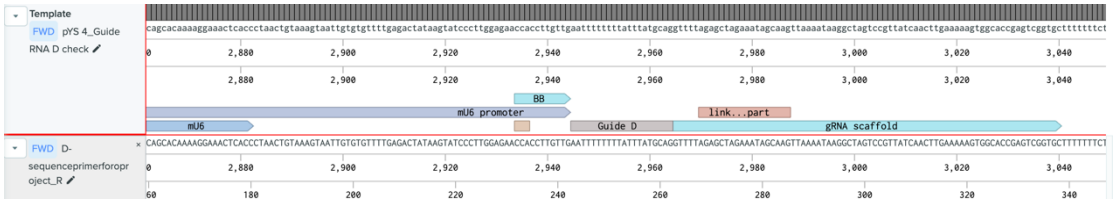
Chapter 6: Sequencing Results

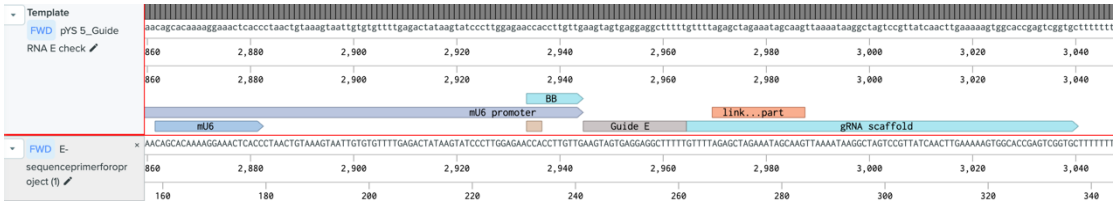
46916_pU6-sgGAL4-4 Guide RNA: A, B, C, D, E, G, H, I, J, K, L

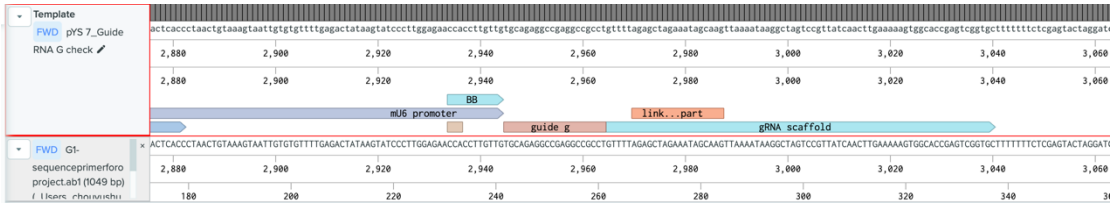


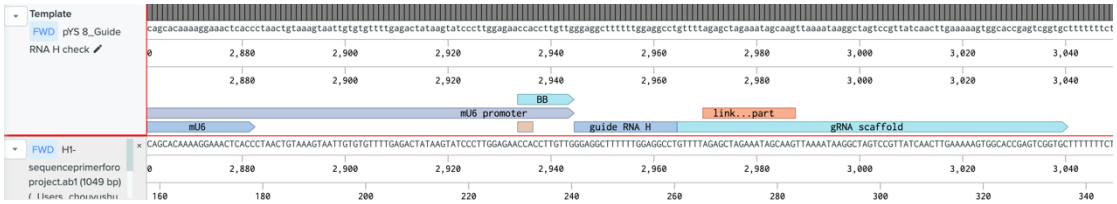




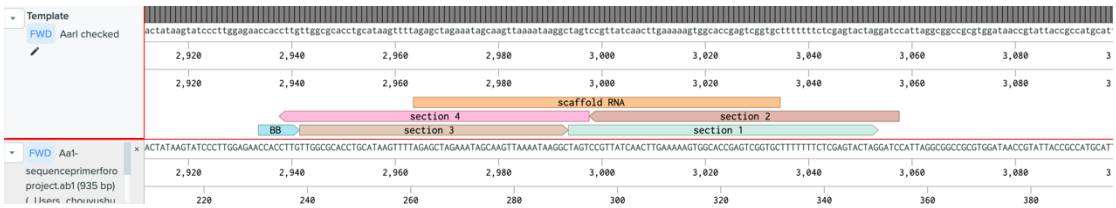




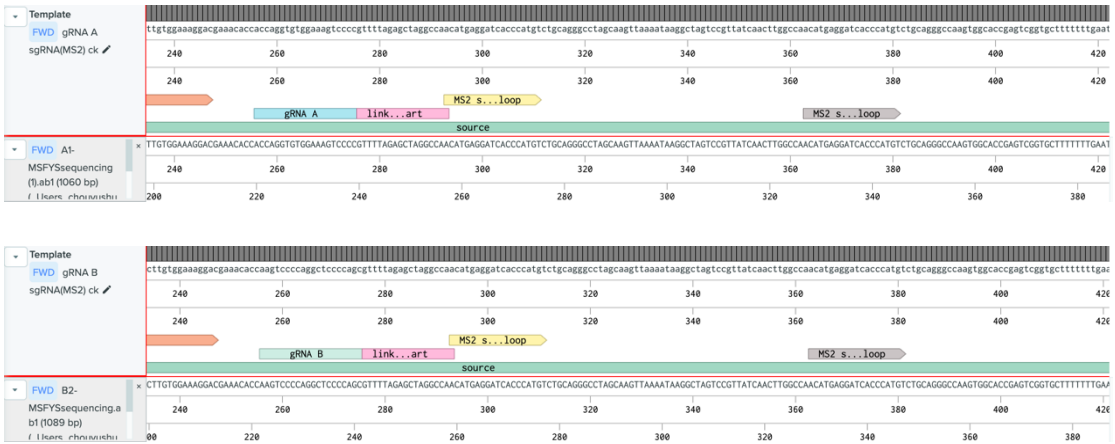


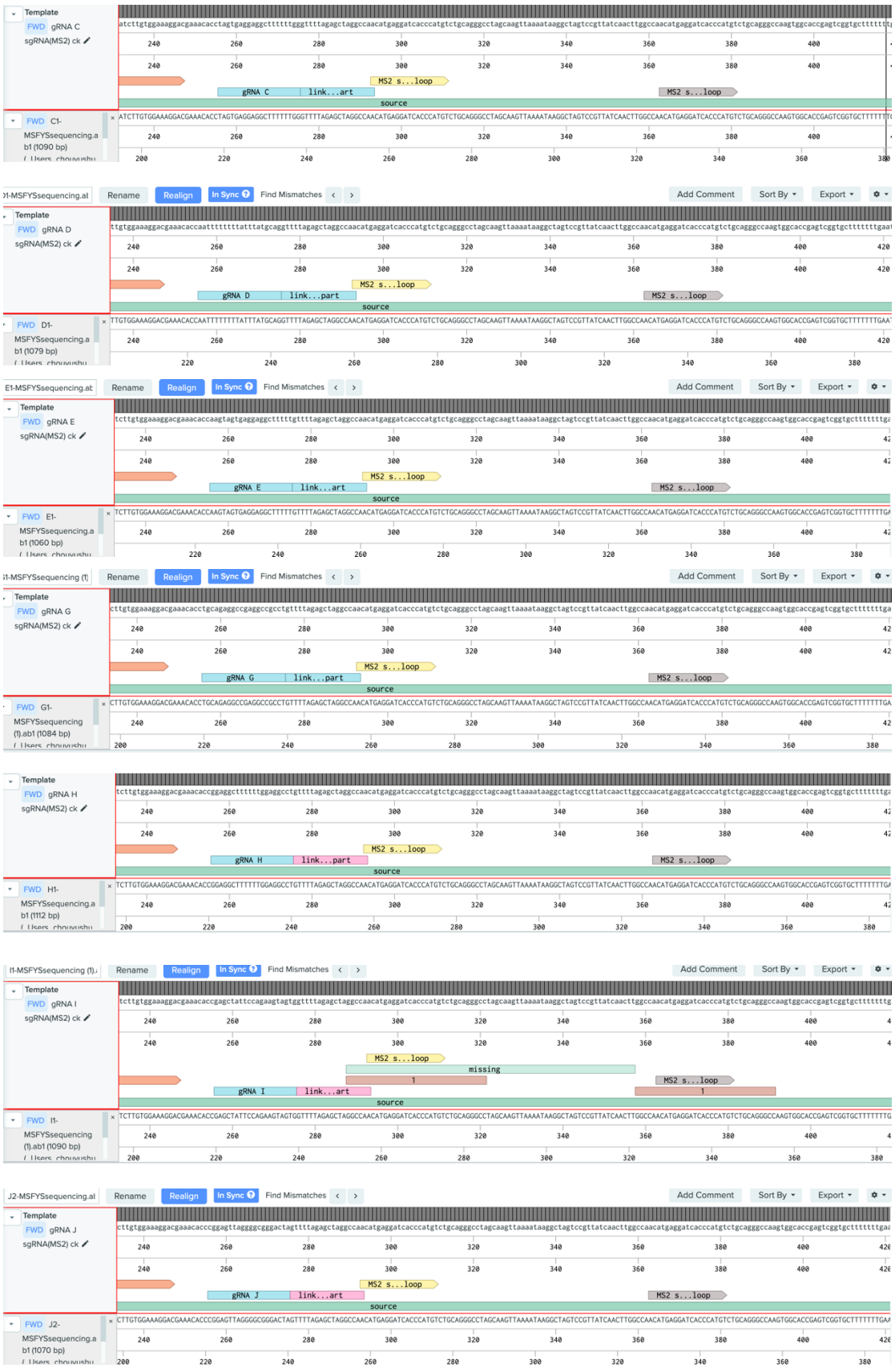


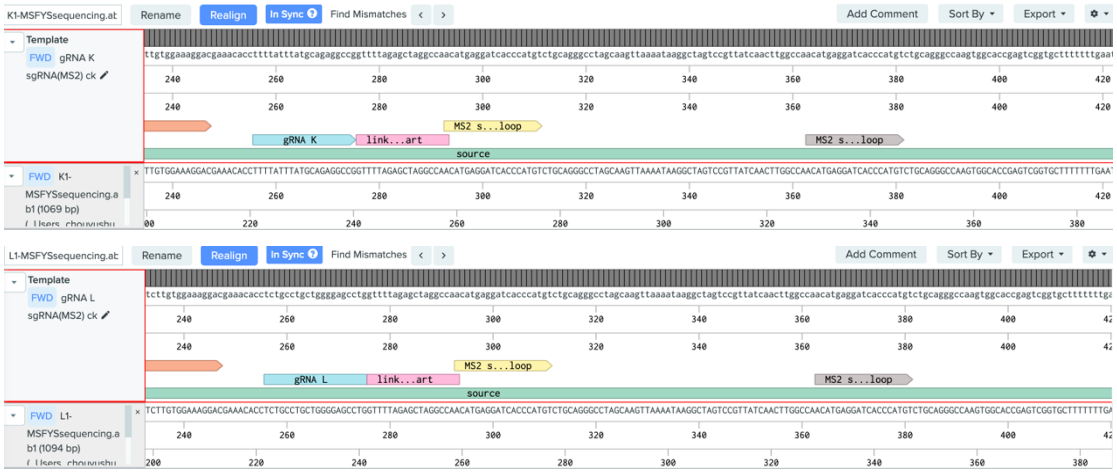
46916_pU6-sgGAL4-4 Guide RNA (1 AarI inserted)



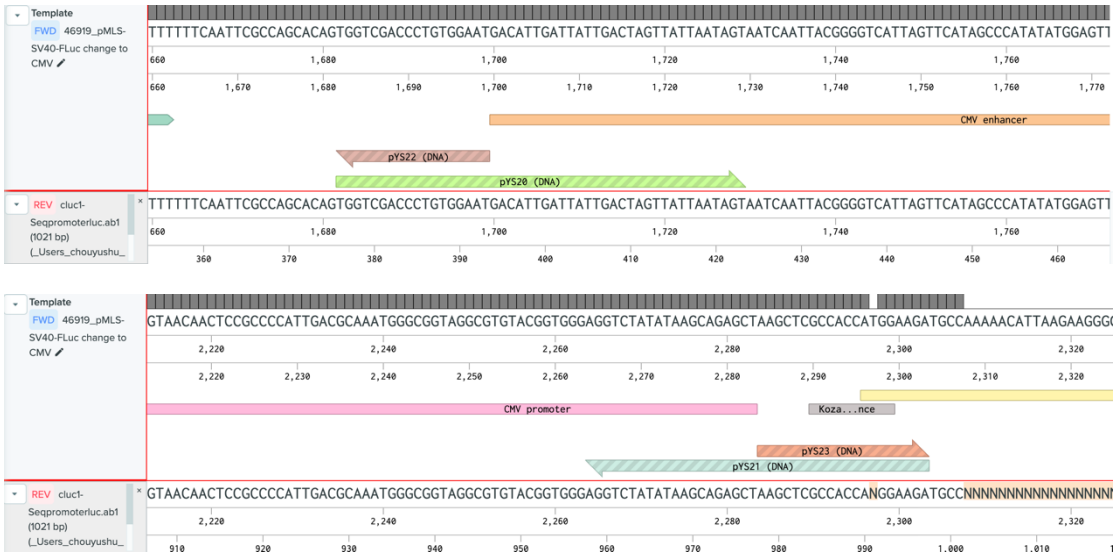
gRNA sgRNA (MS2): A, B, C, D, E, G, H, I, J, K, L



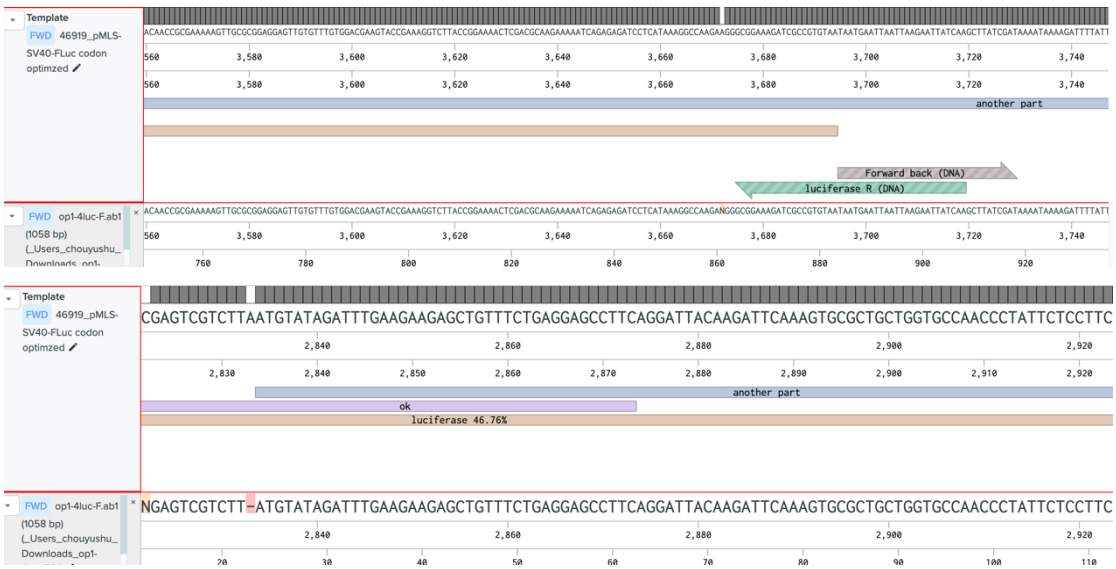

Template
FWD gRNA C
sgRNA(MS2) ck
gRNA C
link...art
MS2 s...loop
source
FWD C1-MSFYSsequencing.ab1 (1090 bp)
Rename
Realign
In Sync
Find Mismatches
Add Comment
Sort By
Export
Template
FWD gRNA D
sgRNA(MS2) ck
gRNA D
link...part
FWD D1-MSFYSsequencing.ab1 (1079 bp)
Template
FWD gRNA E
sgRNA(MS2) ck
gRNA E
FWD E1-MSFYSsequencing.ab1 (1060 bp)
Template
FWD gRNA G
sgRNA(MS2) ck
gRNA G
FWD G1-MSFYSsequencing (1).ab1 (1084 bp)
Template
FWD gRNA H
sgRNA(MS2) ck
gRNA H
FWD H1-MSFYSsequencing.ab1 (1112 bp)
Template
FWD gRNA I
sgRNA(MS2) ck
gRNA I
missing
FWD I1-MSFYSsequencing (1).ab1 (1090 bp)
Template
FWD gRNA J
sgRNA(MS2) ck
gRNA J
FWD J2-MSFYSsequencing.ab1 (1070 bp)



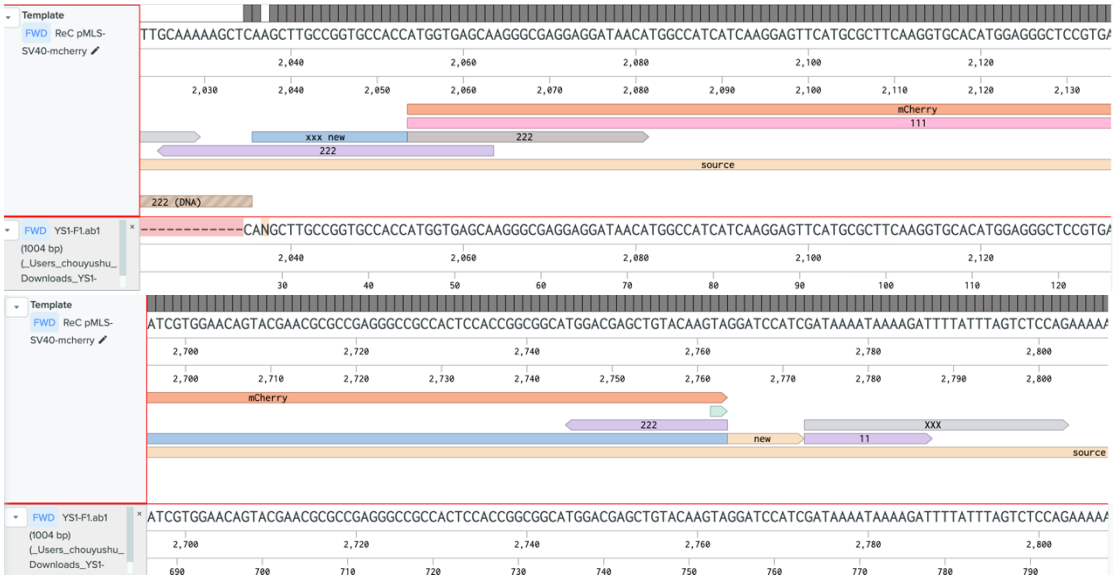
46919_pMLS-SV40-FLuc->46919_pMLS-CMV-FLuc



46919_pMLS-SV40-FLuc to 46919_pMLS-SV40-FLuc codon-optimized



46919_pMLS-SV40-EGFP -> 46919_pMLS-SV40-mCherry



Target	Forward Primer	Reverse Primer
gRNA A	AGAGCTAGAAATA GCAAGTTAAAAT AAGGCTAGTCCGT TATCAACTTGAAA	CTTGCTATTTCTAGCTCTAAAACGGGGACTTTCCACACCTGGTCAACAAGGTGGTTCTCC
gRNA B		CTTGCTATTTCTAGCTCTAAAACGCTGGGGAGCCTGGGGACTTCAACAAGGTGGTTCTCC
gRNA C		CTTGCTATTTCTAGCTCTAAAACCCAAAAAAGCCTCCTCACTACAACAAGGTGGTTCTCC
gRNA D		CTTGCTATTTCTAGCTCTAAAACCTGCATAAATAAAAAAAATTCAACAAGGTGGTTCTCC
gRNA E		CTTGCTATTTCTAGCTCTAAAACAAAAAGCCTCCTCACTACTTCAACAAGGTGGTTCTCC
gRNA G		CTTGCTATTTCTAGCTCTAAAACAGGCGGCCTCGGCCTCTGCACAACAAGGTGGTTCTCC
gRNA H		CTTGCTATTTCTAGCTCTAAAACAGGCCTCCAAAAAAGCCTCCCAACAAGGTGGTTCTCC
gRNA I		CTTGCTATTTCTAGCTCTAAAACCACTACTTCTGGAATAGCTCCAACAAGGTGGTTCTCC
gRNA J		CTTGCTATTTCTAGCTCTAAAACTAGTCCCGCCCCTAACTCCGCAACAAGGTGGTTCTCC
gRNA K		CTTGCTATTTCTAGCTCTAAAACCGGCCTCTGCATAAATAAAACAACAAGGTGGTTCTCC
gRNA L		CTTGCTATTTCTAGCTCTAAAACCAGGCTCCCCAGCAGGCAGACAACAAGGTGGTTCTCC

Table 5. Primer Sets for SpCas9 Scaffold RNA: 46916_pU6-sgGAL4-4 backbone

Target	Forward Primer	Reverse Primer
gRNA A	GTTTTAGAGCTAG GCCAACATGAGGA TCACCCATGTCTG CAGGGCCTAGC	TTGGCCTAGCTCTAAAACGGGGACTTTCCACACCTGGTGGTGTTTCGTCCTTTCC
gRNA B		TTGGCCTAGCTCTAAAACGCTGGGGAGCCTGGGGACTTGGTGTTTCGTCCTTTCC
gRNA C		TTGGCCTAGCTCTAAAACCCAAAAAAGCCTCCTCACTAGGTGTTTCGTCCTTTCC
gRNA D		TTGGCCTAGCTCTAAAACCTGCATAAATAAAAAAAAATTGGTGTTTCGTCCTTTCC
gRNA E		TTGGCCTAGCTCTAAAACAAAAAGCCTCCTCACTACTTGGTGTTTCGTCCTTTCC
gRNA G		TTGGCCTAGCTCTAAAACAGGCGGCCTCGGCCTCTGCAGGTGTTTCGTCCTTTCC
gRNA H		TTGGCCTAGCTCTAAAACAGGCCTCCAAAAAAGCCTCCGGTGTTTCGTCCTTTCC
gRNA I		TTGGCCTAGCTCTAAAACCACTACTTCTGGAATAGCTCGGTGTTTCGTCCTTTCC
gRNA J		TTGGCCTAGCTCTAAAACTAGTCCCGCCCCTAACTCCGGGTGTTTCGTCCTTTCC
gRNA K		TTGGCCTAGCTCTAAAACCGGCCTCTGCATAAATAAAAGGTGTTTCGTCCTTTCC
gRNA L		TTGGCCTAGCTCTAAAACCAGGCTCCCCAGCAGGCAGAGGTGTTTCGTCCTTTCC

Table 6. Primer sets for SpCas9 Scaffold with Two MS2 Stem Loop: sgRNA (MS2) cloning backbone