DEVELOPING NEW TOOLS FOR TRACKING SUBCELLULAR LOCALIZATION OF GENOME-EDITING ENZYMES

by

Anna-Marina MRAK

Marshallplan Scholarship Paper

Host Institution:

University of California, Berkeley

Innovative Genomics Institute

Wilson Lab

External Supervisors:

Ross Wilson, P.I.

Lorena DeOñate, PhD

Home Institution:

IMC University of Applied Sciences, Krems

Internal Supervisor:

Prof.(FH) Priv.Doz. Dr. Reinhard Klein

Marshallplan-Jubiläumsstiftung

Austrian Marshall Plan Foundation

Fostering Transatlantic Excellence

imc | FH KREMS UNIVERSITY OF APPLIED SCIENCES/AUSTRIA

Berkeley UNIVERSITY OF CALIFORNIA

1 Acknowledgements

I would like to express my gratitude to the Marshallplan Foundation for providing funding for my research semester at the University of California, Berkeley. Furthermore, I would like to thank the whole team of the Wilson lab and especially my P.I. Ross Wilson for making this internship possible and my supervisor Lorena DeOñate for being such a great teacher and mentor throughout my stay. Last but not least, I want to thank my family and friends for their encouragement and constant support.

2 Table of Contents

3 List of abbreviations

NGS	Next Generation Sequencing
°C	Degree Celsius
bp	Base pair
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
crRNA	CRISPR RNA
DNA	Deoxyribonucleic acid
gRNA	Guide RNA
hr	hour(s)
min	minute(s)
NHEJ	Non-homologous end joining
nt	nucleotide
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
sec	second(s)
sgRNA	Single guide RNA

4 Abstract

In-vivo delivery of genome-editing enzymes like CRISPR-Cas9 remains one of the biggest hurdles when it comes to using those powerful tools therapeutically. One possible solution to this problem is to engineer ribonucleoproteins (RNPs) to harbor cell type specific ligands and deliver them together with cell-penetrating peptides. In the present project, 2 reporter cell lines for intracellular localization of Cas9 are evaluated. It is shown that both cell lines are promising tools to track the uptake of Cas9 in HepG2 cells and determine efficiency of major steps of this pathway, which are endosomal escape and nuclear import. This makes them a suitable tool for monitoring this “receptor-mediated” uptake of Cas9.

5 Introduction

5.1 CRISPR-Cas9 mechanism and advantages compared to other gene editing technologies

The ability of introducing changes into genomes holds tremendous value for biomedical research and therapeutic development and after years of basic research, CRISPR-Cas9 (CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats) has emerged as an easy and efficient tool for genome editing and beyond.

With the discovery of restriction enzymes in the 1970s, used by bacteria for protection against phages, DNA manipulation and recombinant DNA technology was possible for the first time. [4] Followed by this major breakthrough in molecular biology, was the insight that double-strand breaks in mammalian cells are repaired by both, non-homologous end joining and homology directed repair upon presence of a linear DNA fragment and that integration of an exogenous DNA can thereby be induced by a DSB at the target site. [5] Even though scientists were able to mutate endonucleases to increase their DNA specificity, targets were still limited to the unique recognition sites of the restriction enzymes, making it hard to find one that targets exactly the needed locus. [6] This problem of variability was tackled by the discovery of so-called zinc finger proteins, which are zinc ion-regulated small protein motifs that can recognize a 3-bp DNA sequence. [7] By assembling 6-7 zing finger modules out of the unique 64-finger pool and fusing them to the domain of the Fok I endonuclease that cleaves the DNA, a specific 18-21 bp sequence within the genome could be targeted and edited. [8, 9] Similar to the zinc fingers, transcription activator-like effector (TALE) proteins derived from *Xanthomonas* bacteria, can recognize a single base instead of three and can also be fused to the Fok I DNA cleavage domain to form a target-specific nuclease. [10, 11] Even though discovery of those technologies increased the possibilities of editing different target sites of the genome, they were fairly difficult to engineer and were therefore never as prominent as the CRISPR technology.

Derived from the *Streptococcus pyogenes (Spy)* CRISPR locus, Cas9 is an enzyme capable of cleaving DNA by inducing double-strand breaks at locations with complementarity to the guide RNA (gRNA) in complex with Cas9, forming a ribonucleoprotein (RNP). As a result, cells will induce repair pathways like non-

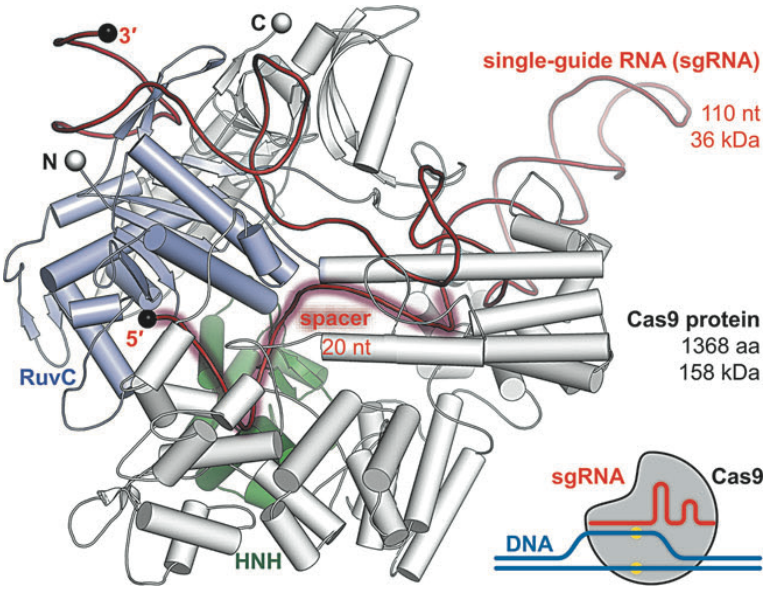


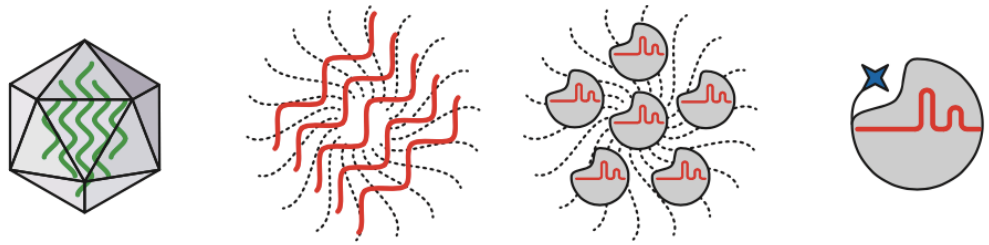
Figure 1: From [3]. SpyCas9 anatomy showing the Cas9 protein (white), sgRNA (red), and nuclease domains RuvC (blue) and HNH (green). Bottom left shows Cas9 binding to the complementary DNA strand and cleaving sites in yellow.

homologous (NHEJ) end joining that will most likely cause a small deletion or insertion (indel) at the location of the double-strand break or, in presence of donor DNA, incorporation via homologous recombination. Using this tool, almost any location of the genome can be targeted by simply forming a complex of Cas9 and the specific gRNA complementary to the region of interest. Simply having to synthesize a corresponding gRNA component to form a functional genome editing enzyme makes Cas9 such a rapid and inexpensive tool compared to previous genome editing techniques. [12, 13]

In its canonical form, SpyCas9 is a 1,368 amino acid big protein in complex with a ~75 nucleotide tracrRNA and a ~40 nucleotide crRNA. While the tracrRNA acts as a scaffold to form a stable RNP complex, the crRNA contains the 20 nt long region (spacer) that will allow Cas9 to find and bind to the complementary DNA sequence on the target strand. [14] The tracrRNA and crRNA form a double-stranded RNA helix by having a 24 nt complementarity to each other. To simplify the use of Cas9 even more, scientists fused this 2-piece gRNA together to form a single molecule by addition of a 4 nt loop that creates a so-called single-guide RNA (Figure 1). Cas9 carries two nuclease domains, RuvC and HNH, that are responsible for the formation of DSB in the complementary DNA strand. [12] In the early beginnings of using Cas9, other modifications were made to improve its function, including the addition of one or more nuclear localization signals (NLS) to promote nuclear import of the genome editing enzyme [15, 16] and the optimization for expression in human cells. One of the biggest concerns when it comes to using Cas9 is the potential off-target cleavage at DNA locations with a similar sequence as the target site and scientists are constantly working on engineering the enzyme in a way to improve specificity e.g. by mutating positively charged residues that could lead to non-specific binding by providing affinity to the negatively charged DNA. [17, 18] In addition to re-engineering Cas9 to increase specificity and reduce off-target binding, Cas9 has been engineered for use beyond genome editing. One examples for alternative applications of Cas9 include the use of a catalytically inactive enzyme (dCas9) which is not capable of cleaving the DNA anymore but can still target and bind to the desired DNA sequence. This variation of Cas9 has been reused for several applications like gene expression regulation due to tight binding of dCas9 to the DNA which interferes with other proteins like transcription factors or polymerases. [19]

5.2 Genome-editing enzyme delivery in vivo

As the aspect of delivery contributes largely to the possible therapeutic use of genome editing enzymes, several different strategies have been investigated (figure 2) in the past. While early applications in-vitro delivery were performed by transfecting cells



Approach	DNA	RNA	RNA & protein (RNP) enzyme	Engineered RNP enzyme
Example	Adeno-associated virus (AAV) packaged with DNA encoding Cas9 protein & sgRNA	Liposomes encapsulating Cas9 mRNA & sgRNA	Cas9 ribonucleoprotein associated with carriers, e.g. cationic nanoparticles	Cas9 ribonucleoprotein modified to add desirable properties
Size	20 nm	≥50 nm	≥50 nm	12 nm
Advantages	Extremely effective; prior use with traditional gene therapy	Straightforward to prepare; low immunogenicity; effective	Short half-life limits off-target editing; can be highly efficient	Short half-life limits off-target editing; simplified manufacture
Disadvantages	Risk of increased off-target cutting & genomic integrations; can be immunogenic; capacity for DNA storage can be limiting; non-trivial manufacture	Toxicity not fully characterized; tends to accumulate in the liver, thus limiting versatility	Unknown immunogenicity profile; co-carrier molecules may be toxic	Unknown immunogenicity profile; requires additional engineering, transduction reagents, or electroporation to enter cells
Tissue Specificity	Some tropism inherent to various strains; additional tropism can be engineered	Intravenous use edits the liver preferentially; molecular targeting is being developed	Shows great promise via local injection	No inherent cell-penetrating properties: molecular targeting can be precisely engineered

Figure 2: From [3]. Approaches for Cas9 delivery either as DNA, RNA or RNP with different carriers.

chemically or via electroporation with a plasmid encoding Cas9 and the gRNA [20] for proof of concept, in-vivo delivery requires more complex technologies for cellular uptake. The entry of the large ribonucleoprotein complex is well-prevented by cells and existing delivery platforms like viral vectors or lipid nanoparticles burden disadvantages like increased risks of off-target editing or weak cell-specificity. [21, 22]

Initial gene therapies used adeno-associated virus (AAV) as a tool for efficient and relatively safe delivery of single-stranded DNA, and this system was adapted for packaging the genetic information for the Cas9 protein and the sgRNA. [23] AAV is a small, non-pathogenic ssDNA virus that is weakly immunogenic. Upon infection, it remains mainly in an extrachromosomal state but can also integrate in the human genome, most likely at the AAVS1 locus. Due to this low occurrence of integrations or integrations at a neutral site in the genome, it is considered to be relatively safe. [24] Furthermore, AAVs can be used to encode both, CRISPR-Cas9 and a homology repair donor DNA for knock-in mutations. [25] One of the major downsides of this delivery

method is the small genome size of the virus, which limits the use of *S.pyogenes* Cas9, which is relatively big and therefore cannot be encoded together on the same vector with a sgRNA and two separate AAVs need to infect at the same time, which reduces therapeutic potential. [26] While other virus systems that also have a low immunogenicity like the lentivirus do not have this problem of size limitation and can encode both, *S.pyogenes* Cas9 and a sgRNA in a single virus, they will reverse transcribe their viral RNA to DNA and randomly integrate it into the host genome. This leads to a high risk of off-target effects that might lead to undesired mutations.

Another strategy of delivery is to use lipid nanoparticles (LNPs), which are polymers that can encapsulate mRNA encoding Cas9 and a sgRNA, enabling translation and subsequent RNP complex formation within the cells. [27] LNPs cannot only be used for the delivery of Cas9 and sgRNA mRNA but also for direct delivery of Cas9 RNP. One example for this approach is the delivery of Cas9 RNP by using cationic lipids which can interact with nucleic acids and therefore with the negative charges at the sgRNA portion of the RNP. This method achieved a two-fold editing increase compared to the delivery of the Cas9 plasmid. [28]

A possible solution to this problem could be Cas9 delivery via receptor-mediated uptake by engineering RNP to harbor receptor ligands that result in internalization by cells expressing the corresponding receptor. In particular, bearing of asialoglycoprotein receptor ligand (ASGPrL) showed uptake of Cas9 into liver-derived cell lines (like HepG2) expressing the asialoglycoprotein receptor. [29]

This approach relies on three steps (Figure 1): endocytosis of the RNP complex after ligand-receptor binding – escape from the endosome before cargo is degraded in the lysosome – nuclear import of the complex mediated by one or more nuclear localization signals (NLS) flanking Cas9. All three steps have to be

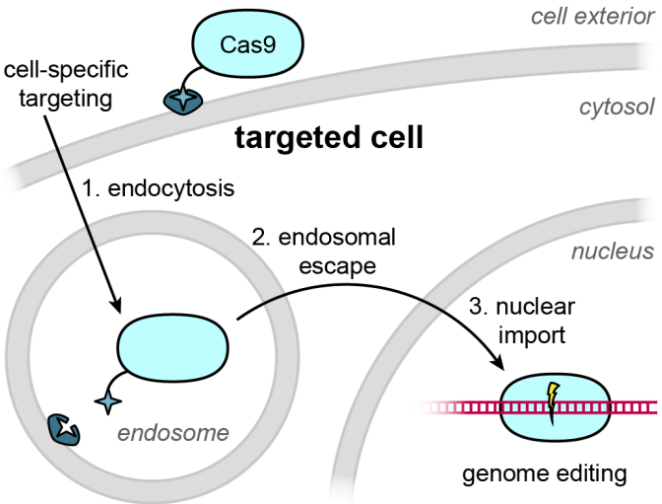


Figure 3: Receptor-mediated uptake of Cas9 to the nucleus.

successful in order for Cas9 to edit the genome and while it is easy to monitor cellular uptake, it is unclear whether endosomal escape or nuclear import is the bottleneck of transport efficiency. [29] Having tools to track effectiveness of each step would accelerate the optimization of delivery technologies and help understanding the hurdles of subcellular transport of genome-editing enzymes.

6 Aim and research question

Purpose of this project is to evaluate two reporter systems to track either endosomal escape and nuclear localization through genome editing in cultured cells. Those systems should be optimized for use in HepG2 cells and subsequent flow cytometry or confocal microscopy. The final goal of this project will be to have a quick, cost-efficient and reliable tool to track genome editing pathways that can be used for high throughput screenings to improve nuclear entry of CRISPR-Cas9. Finally, the effect of different amounts of nuclear localization signals (NLS) on the Cas9 protein on cytosolic and nuclear transport is determined using the above reporters.

More in detail, the research question tackles several methods for the evaluation of two reporter cell lines. The first reporter is intended to be used to assess efficiency of nuclear import. As described in the paper of Nguyen, Miyaoka [1] a genetic element, called reporter cassette (Figure 4), consisting of an array of a red fluorescent mCherry gene, a target site to which the Cas9 RNP is supposed to bind and induce a double-strand break and an eGFP (enhanced green fluorescent protein) gene. Located between the two fluorescent genes is a stop codon, so the eGFP downstream is not translated under normal conditions. Additionally, the eGFP gene is out of frame, meaning the start codon triplet is not within the reading frame of the ribosome but shifted 1 base pair downstream. [1] When Cas9 binds to the target site and induces a double-strand break, the two ends are joined back together through a NHEJ repair mechanism, which can result in insertions or deletions causing a

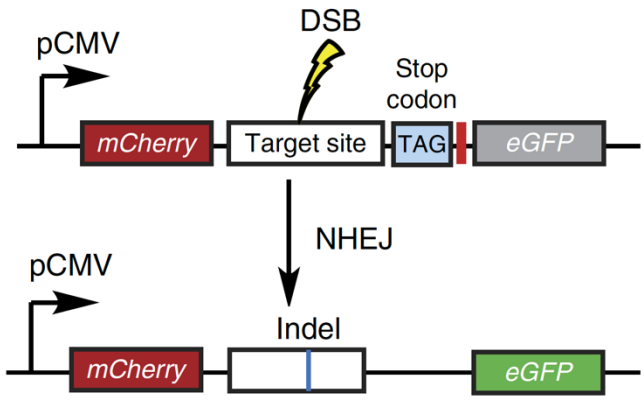


Figure 4: From [1] fluorescence-based genome editing reporter cassette.

frameshift mutation. [30, 31] If the right frameshift of either -2 bp or +1 bp is achieved, the stop codon will fall out of frame and translation can continue. Furthermore, eGFP gets shifted into frame and can be translated so that subsequently green fluorescence can be detected under the microscope and using flow cytometry. [1] It is anticipated that the more eGFP signal is detected by flow cytometry, the more editing/nuclear localization took place, thereby enabling comparison of different modifications of Cas9 to optimize for an increased nuclear import.

The other reporter that is to be evaluated is concerned about the second important step of Cas9 delivery: endosomal escape. Performing straight microscopy by labeling Cas9 with a fluorescent antibody to track its localization inside the endosomes vs. in the cytosol is complicated due to the signal-to noise ratio of the highly inefficient cargo escape (0.1-5%), revealing essentially all of the Cas9 signal to be located inside the endosome. [32]

Utilizing a so-called “split-GFP” approach, like in the papers of Lonn, Kacsinta [33] and Kamiyama, Sekine [34], to detect endosomal escape solves this problem. A reporter GFP, consisting of 11 β strands, is split into two parts that only give a fluorescent signal when encountering each other. The large β1-10 fragment is constantly expressed by the reporter cells, whereas the smaller β11 fragment gets genetically fused to a Cas9 protein (figure 6). As soon as both fragments encounter each other, meaning when Cas9 escapes from the endosome into the cytosol, a fully active β1-11 GFP barrel can form and green fluorescence can be detected under the microscope and using flow cytometry. The localization of the GFP signal within the cell is therefore also an indicator of the location of the Cas9 RNP.

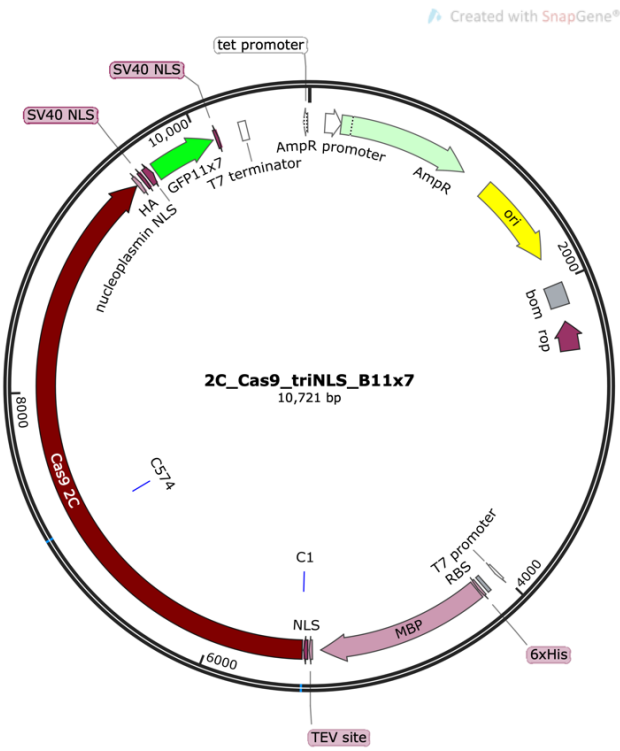


Figure 6: Plasmid of Cas9 harboring three different NLS and 7 GFP B11 fragments that can bind to the GFP B1-10 peptide and form an active GFP protein.

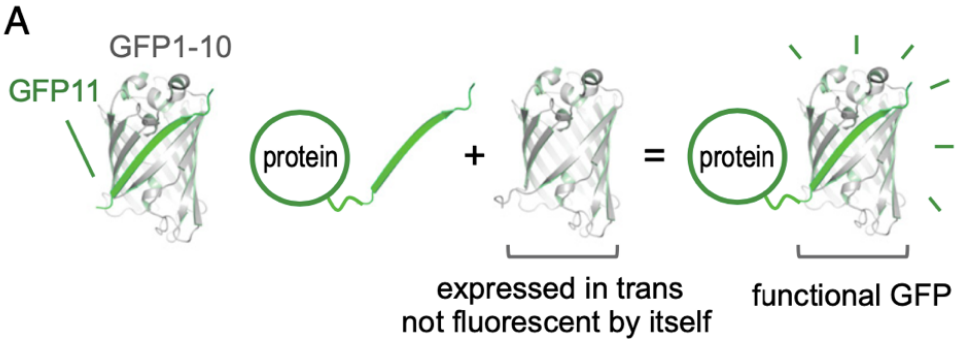


Figure 5: From [2]. Split-GFP strategy. A protein (Cas9) is fused to the missing B11 strand of the GFP barrel, which is needed to form an active fluorescing GFP protein.

7 Scientific Approach

Nuclear Entry Reporter

In the beginning of evaluating the generated clones, different gRNAs are screened to determine which RNP or multiguide (two RNPs cutting in close proximity to “cut out” a fragment of DNA) gives the highest editing rate. The cells are electroporated, meaning the cell membrane will be permeabilized by applying and electric field, allowing the RNPs to enter the nucleus. Editing efficiencies can be compared by sequencing the genomic DNA of the treated cells after 48 hours. At the same time, those editing rates are compared to the visual output of the cells seen under the fluorescent microscope and using flow cytometry. As explained earlier, the more cells are edited efficiently, the more GFP is expected to be visible.

As a more gentle method, a so called “co-incubation” is performed after a successful first electroporation, using the same RNPs but instead of “shooting” the RNP into the nucleus with an electric pulse, the cells are incubated in a mix of cell-penetrating peptide (CPP) and Cas9-RNP to transfer the genome editing enzymes into the cells via endocytosis. [3, 35] This will be the method of choice to collect relevant data. Again, the results from sequencing the DNA and flow cytometry are compared to see if they correlate.

In addition to the overall percentage of editing, which is determined by next-generation sequencing of the genomic DNA, the indel distributions, meaning the prevalence of distinct insertion or deletion patterns of each gRNA, have to be analyzed since only a certain indel will give the right frame shift to mediate translation of the eGFP gene.

While some gRNAs can give high editing rates, they might not be compatible for this reporter if they don't give the right frame-shift to turn on GFP, thereby not being able to use flow cytometry for analysis.

The spacer that has the highest percentage of generated indels in addition to the correct insertion or deletion of base pairs resulting in expression of eGFP can be chosen as a standard for all upcoming screenings.

As a next step, different variations of nuclear localization signals (NLS) flanking Cas9 are compared to see if nuclear import can be improved by the number or configuration of those NLS. As an example, the editing efficiencies of Cas9 with 1xNLS are compared to Cas9 with three different NLS.

Endosomal Escape Reporter

The plasmid of Cas9 fused with the GFP β11 peptide has to be cloned and subsequently expressed to be used with the reporter cells. Again, effectiveness of this reporter is first tested with a simple electroporation of the cells targeting a known site on the human EMX1 gene. If successful, it is expected to see green fluorescent protein in the cytosol of the cells due to binding of the GFP β11 to the β1-10 fragment forming an active GFP barrel. Then the same experiment has to be repeated in a co-incubation set-up as explained earlier.

In order to see how fast and efficient endosomal escape takes place, a time-course experiment has to be performed. This means that every hour after the first addition of the RNP to the cells, it has to be checked if green fluorescence is visible by microscopy or flow cytometry. The first time-point where GFP is visible is supposed to be when endosomal escape of Cas9 first takes place. At a later timepoint, it is expected to also see GFP in the nucleus, since Cas9 is supposed to be imported by nuclear transport together with the GFP barrel.

As a result, it is expected to use this reporter in the lab to compare endosomal escape capabilities of Cas9 modifications by seeing more or less GFP in the cytoplasm. The more green is detected by microscopy or flow cytometry, the more effective is endosomal escape.

8 Materials

8.1 Equipment

Equipment	Manufacturer
Countness automated cell counter	Invitrogen
4D-Nucleofector Core Unit	Lonza
EVOS FL Cell imaging system	Thermo Fisher Scientific
Isotemp CO_2 incubator	Fisher Scientific
SterilGARD class II type A2 biosafety cabinet	BAKER
NanoDrop 8000 spectrophotometer	Thermo Fisher Scientific
5424 Centrifuge	Eppendorf
Countess Automated Cell Counter	Invitrogen

8.2 Buffers and media

Buffer	Composition/Manufacturer
Flow Cytometry Buffer	1xDPBS (-$MgCl_2$,-$CaCl_2$) + 10% FBS + 2mM EDTA
10x Folding Buffer	100 mM HEPES-NaOH pH 7,5 / 1.5 M NaCl
Gel Filtration (GF) Buffer	20 mM HEPES-NaOH pH 7.5 / 150 mM NaCl / 10% Gylcerol
S.O.C. medium	Invitrogen
LB liquid medium	Innovative Genomics Institute
LB/Amp plates	Innovative Genomics Institute

8.3 Chemicals

Chemical	Manufacturer
1xDPBS (-$MgCl_2$,-$CaCl_2$)	Fisher Scientific
0.5M EDTA, pH 8.0	Fisher Scientific
1xFBS	Invitrogen
Ampicillin	VWR Stanley
Penicillin/Streptomycin	Fisher Scientific
6x DNA Gel Loading Dye, purple	New England Biolabs
100bp GeneRuler	Fisher Scientific
1kb GeneRuler	Fisher Scientific
dNTP mix, 10μM each	TaKaRa
100% Glycerol	Fisher Scientific
25mM $MgCl_2$	
SybrSafe DNA gel stain	Invitrogen
UltraPure Agarose	Fisher Scientific
0.05% Trypsin-EDTA	Fisher Scientific

0.25% Trypsin-EDTA	Fisher Scientific
TryPLE	Invitrogen
SOC medium	Fisher Scientific
Trypan Blue Stain (0.4%)	Invitrogen
100% Ethanol	VWR

8.4 Primers

Primer name	Sequence
Fw InFusion Backbone annealing n6	GGCAGCAGCGGATCCCGCGCCACCTGGTG
Rev InFusion Backbone annealing (all) NP n7	CTCGAGCTTTTTCTTTTTTGCCTGGC
Fw InFusion annealing B11x7 n8	AAGAAAAAGCTCGAGGGCAGCGGCTCCCG
Rev InFusion annealing B11x7 n9	GGATCCGCTGCTGCCGT
Frame shift GFP NGS F1	GCGCCTACAACGTCAACATC
Frame shift GFP NGS R1	CTTCTTCAAGTCCGCCATGC

8.5 Enzymes

Enzyme name	Manufacturer
triNLS 2C-Cas9	Home-made
1xNLS 2C-Cas9	Home-made
2xNLS B11-Cas9	Home-made
1xNLS B11-Cas9	Home-made
PrimeStar GxL DNA Polymerase	Clontech/TaKaRa
Infusion Enzyme	TaKaRa
QuickExtract DNA Extraction Solution	Lucigen

8.6 Kits

Kit	Manufacturer
QIAprep Spin Miniprep Kit	QIAGEN
QIAquick Gel Extraction Kit	QIAGEN
PCR Cleanup Beads	UC Berkeley DNA Sequencing Facility

8.7 Bacterial Strains

Bacterial Strain	Manufacturer
E.Coli XL-1 blue competent cells	Home-made

8.8 Cell lines

Cell line	Manufacturer
HepG2	ATCC

9 Methods

9.1 Cultivation of HepG2 cells

HepG2 cells were cultivated in T75 or T25 cell culture flasks in EMEM+10%FBS+1%Penicillin/Streptomycin with media changes every 2-3 days. As soon as cells reached ~80% confluency, they were split by first washing 2x with 1xDPBS (-$MgCl_2$,-$CaCl_2$), addition of 1:2 TryPLE:0.05% Trypsin-EDTA, incubation at 37°C for ~5 minutes until all cells were detached, addition of fresh cell culture media to stop trypsinization followed by centrifugation at 200xg for 4 minutes and resuspension in fresh medium. Cells were then split at a maximum ratio of 1:8 to provide an optimal seeding density and were plated in new flasks which were filled up with fresh medium.

9.2 Generation of HepG2 reporter cells via lentiviral transduction

The genetic elements for both, the expression of GFP ß1-10 peptides for the endosomal escape cells as-well as the reporter cassette for the genome editing reporter were integrated into the genome of HepG2 cells via lentiviral transduction. This was done by the Wilson lab prior to the beginning of this project. The lentivirus was produced in HEK 293 cells and harvested from media of the transduced cells. After transduction, cells were sorted by fluorescence-activated cell sorting (FACS) to select single clones, which were then expanded, frozen and stored in liquid nitrogen for use in this project.

9.3 PCR and gel electrophoresis of Cas9 tri-NLS backbone and B11 insert

In order to clone the GFP B11 insert into Cas9, minipreps of triNLS Cas9 and of the 7xB11 plasmid were amplified in a PCR. For the backbone primers Fw InFusion Backbone annealing n6 and Rev InFusion Backbone annealing (all) NP n7 were used, for the insert primers Fw InFusion annealing B11x7 n8 and Rev InFusion annealing B11x7 n9.

Table 1: 1x PCR Mix for insert and backbone	
DEPC H_2O	33.5 µl
5x KAPA HiFi Buffer for GC rich sequences	10 µl
Takara dNTP Mix 2.5 mM each	1.5 µl
10 µM forward primer	1.5 µl
10 µM reverse primer	1.5 µl
KAPA HiFi HotStart DNA Polymerase	1 µl
Plasmid template [1 ng/µl]	1 µl

Table 1: PCR conditions B11 insert			
	Denaturing	95°C	3 min
30 cycles	Denaturing	98°C	15 sec
	Annealing	62°C	15 sec
	Extension	72°C	30 sec
	Final extension	72°C	1 min
		4°C	∞

Table 3: PCR conditions Cas9 backbone			
	Denaturing	95°C	3 min
30 cycles	Denaturing	98°C	15 sec
	Annealing	62°C	15 sec
	Extension	72°C	5 min
	Final extension	72°C	10 min
		4°C	∞

Samples were then analyzed using gel electrophoresis by diluting 5µL of PCR product in 5µL of DEPC H2O and addition of 2µL 6x Loading Dye and were run on a 1% Agarose + 1xSybrSafe gel at 120V for ~20 min. The target DNA fragments were cut out directly from the agarose gel and purified using the QUIAGEN QIAquick Gel Extraction Kit. To do so, target fragments from agarose gel were cut out using a clean razor blade and transferred into separate microcentrifuge tubes. The weight of each gel slice now had to be determined. Each tube was filled up to 1.5 mL with buffer QG and incubated at 50°C for ~10 min until the gel was completely dissolved. If the color of the mixture was orange or violet, 10µL of 3 M sodium acetate, pH 5.0 had to be added until it turned yellow. To each dissolved fragment, 1x gel volume (gel weight=gel volume) of isopropanol was added and mixed. This mixture was then transferred into a QIA quick spin column on top of a 2 mL collection tube and the columns were centrifuged at 17,900 x g for 1 min. Flow-through was discarded and columns were placed back into the collection tubes. Now, 750 µL of Buffer PE was added to each column and centrifuged at 17,900 x g for 1 min. Flow-through was again discarded and columns were centrifuged again at 17,900 x g for 1 min. To elute DNA, columns were placed on clean 1.5 mL microcentrifuge tubes and 30 µL of DEPC H_2O were added to the center of each column membrane, incubated for ~2 min and then centrifuged at 17,000 x g for 1 min. Concentration of purified DNA could now be measured using the Nanodrop.

9.4 Infusion reaction of tri-NLS B11 Cas9 plasmid and transformation of *E.Coli XL-1 blue* competent cells

After successful DNA gel extraction, an infusion reaction was done to clone the B11 insert into the Cas9 backbone. A molar ratio of 5:1 insert:backbone was chosen for the reaction, which was performed using the Takara In-Fusion Kit and according to the manufacturers protocol. An In-Fusion mix was prepared, incubated for 15 min at 50°C and then placed on ice until the transformation.

Table 4: In-Fusion mix B11 insert + Cas9 backbone

	B11 insert (1:10 dilution) [0.0153 mol/µL]	Cas9 backbone [0.00371 mol/µL]	Takara Infusion mix	DEPC H_2O
Total volume = 10 µL	2.42 µL	2 µL	2 µL	3.58 µL

For the transformation, competent *E.Coli XL-1 blue* cells were thawed on ice right before use. As soon as they were completely thawed, 2.5 µL of the In-Fusion mix were added to 50 µL of competent cells and incubated on ice for 30 min. Then, cells were heat shocked for 45 sec at 42°C and placed on ice for 2 min. 450 µL of S.O.C. medium were added to each tube of transformed cells and tubes were incubated shaking for 1 hr at 37°C. After incubation, 1/5 of each transformation reaction was plated on a separate LB/Ampicillin plate and incubated overnight at 37°C. As a positive control, cells were transformed with the intact backbone plasmid, as a negative control, with no plasmid at all.

9.5 Overnight cultures, glycerol stocks and miniprep

5 colonies (clones#1 - #5) from the plate of the transformed bacteria were picked and inoculated in separate tubes with 5 mL LB/Amp liquid medium and put on the shaker at 37°C overnight. On the next day, minipreps were performed using the QIAGEN miniprep kit, according to the manufacturers protocol. The liquid cultures were centrifuged at 3,500 x g for 3 min, the supernatant was removed and each pellet was resuspended in 250 µL of Buffer P1 and transferred into a microcentrifuge tube. To each tube, 250 µL of Buffer P2 was added and the tubes were inverted 5 times to mix thoroughly. The mix then turned blue and the lysis reaction should not proceed for more than 5 min. To stop the reaction, 350 µL of Buffer N3 were added to each tube and mixed by inverting 5 times. Tubes were centrifuged for 10 min at 17,900 x g in a table-top microcentrifuge and 800 µL of the supernatant were transferred to QIAprep 2.0 spin columns by pipetting without disturbing the white pellet. Spin columns were centrifuged for 1 min at 17,900 x g and flow-through was discarded. Columns were washed by adding 750 µL of Buffer PE and centrifugation for 1 min at 17,900 x g. Flow through was again discarded and columns were dried by centrifuging again for 1 min at 17,900 x g. Now, columns were placed in clean 1.5 mL microcentrifuge tubes and DNA was eluted by addition of 40 µL DEPC H_2O to the center of each column, incubation for ~1 min and centrifugation for 1 min at 17,900 x g.

The DNA concentration of each miniprep was then measured using the Nanodrop and the samples were sequence verified by Sanger sequencing. Additionally, glycerol stocks of each overnight culture were prepared by mixing 500 µL of the liquid culture with 500 µL of 40% Glycerol and storage at -80°C in case more plasmid DNA is needed after an extended time.

9.6 Ribonucleoprotein formation

In order to form a complex between the guide RNA and the Cas9 protein, gRNAs from Synthego [100µM] were diluted in DEPC H_2O and 10x folding buffer (200 mM HEPES-NaOH pH 7,5 / 1,5 M NaCl2) to reach a final concentration of 15 µM and were placed on the heating block at 95°C for 5 min. After incubation, they were placed on the bench to reach room temperature and glycerol was added. Cas9 protein was diluted to a final concentration of 12.5 µM in 1x GF buffer (20 mM HEPES-NaOH pH 7.5 / 150 mM NaCl / 10 % glycerol) and $MgCl_2$ [20mM], which was added right before RNP formation. To initiate RNP complex formation, Cas9 was pipetted slowly to the gRNA in a 1:1 volume:volume ratio and the tubes were placed into the heating block at 37°C for 15 min. In case of precipitation at any step, samples had to be spun down and supernatant had to be recovered. Samples were stored at room temperature until electroporation or Co-Incubation or at -80°C for long-term storage.

9.7 Transfection by nucleofection/electroporation

HepG2 cells had to be ~80% confluent to perform an electroporation. RNP was prepared right before Nucleofection. For each nucleofection, 200.000 cells and 50 pmol RNP were needed. A 24-well plate with 1 mL culture media (EMEM+10%FBS+1% Pen/Strep) in each well was pre-warmed at 37°C and 5% CO_2. Cells were washed 2x with 1xDPBS (-$MgCl_2$-$CaCl_2$) and detached by addition of 2:1 TryPLE:0.25% Trypsin-EDTA and subsequent incubation at 37°C for ~5 minutes. Trypsinization was stopped by addition of fresh media and cells were centrifuged at 200xg for 4 min. Supernatant was removed and cells were resuspended in 10mL 1xDPBS (-$MgCl_2$-$CaCl_2$) and the cell concentration was determined with a Countess Automated Cell Counter from Invitrogen by mixing 20 µL of the cell suspension with 20 µL Trypan Blue and adding 10 µL of this mix to a disposable cell counting chamber slide for the cell counter. Now, the needed amount of cell suspension was transferred into a fresh 15 mL conical tube and spun down at 200xg for 4 min. In the meantime, Lonza SF solution was prepared by mixing 82% SF buffer and 18% supplement up to a total volume of 20 µL per reaction. Now, 50 pmol of RNP was added to each of the Lonza 16-well Nucleocuvette Strip wells. The cell pellet had to be resuspended in the SF solution to reach a final concentration of 200,000 cells/ 20µL and 20µL of the mix were added to each RNP containing well of the nucleofection cassette, which was then placed into the Lonza 4D-Nucleofector and electroporated choosing the SF method

on the machine. Samples were recovered by quickly adding ~80µL of the pre-warmed media to each nucleofection mix and transferring everything back to the plate into the corresponding well. Cells were incubated at 37°C for 48 hr and checked under the microscope after 24 and 48 hr for successful transfection.

9.8 Transfection by co-Incubation with cell-penetrating peptide (CPP)

Cells had to be plated in 96-well plates 24h prior to performing the Co-Incubation at a density of 15,000-20,000 cells/well in order to actively grow while performing the procedure. The old media had to be removed and wells were filled with 75 µL of fresh EMEM+10%FBS+1% Pen/Strep. RNP was prepared right before the Co-Incubation. Cell-penetrating peptide stock [10,000µM] was thawed at room temperature and diluted with DEPC H_2O to a working concentration of 100 µM. From this dilution, a master mix with 3 µL of CPP and 11µL OPTIMEM per samples was prepared. For each reaction, 1.6µL of RNP [6.25 µM] were diluted in 11 µL OPTIMEM and 14 µL of the CPP master mix were added. In case of a multiguide, 0.8µL of one RNP + 0.8µL of the other RNP were used. This mix was incubated for max. 10 min and then added to the HepG2 cells in the 96-well plate without disturbing the cell layer. Plates were incubated for 48 hr and checked under the microscope after 24 and 48 hr for successful transfection.

9.9 Flow cytometry of HepG2 cells

HepG2 cells in 96-well plates were washed 2x with 1xDPBS(-$MgCl_2$,-$CaCl_2$) and detached by addition of 2:1 TryPLE:0.25% Trypsin-EDTA and subsequent incubation at 37°C for ~5 minutes. Trypsinization was stopped by addition of fresh EMEM+10%FBS+1% Pen/Strep to each well and content of wells was transferred into V-bottom 96-well plates. After centrifugation at 200xg for 4 min, the supernatant was removed and pellets were resuspended in flow buffer (1xDPBS(-$MgCl_2$,-$CaCl_2$) + 10% FBS + 2mM EDTA). This wash step was performed twice. Plate had to be kept on ice until it was analyzed with the Attune NxT Flow Cytometer.

In case of the mCherry-GFP reporter, a laser compensation had to be performed prior to analyzing the samples to prevent overlap between fluorophores and subsequent false positive results. This was done by running the compensation program on the Attune and using only GFP positive, only mCherry positive and GFP/mCherry negative cells.

9.10 Genomic DNA extraction of mCherry-GFP reporter cells

Genomic DNA was extracted by washing HepG2 cells 2x with 1xDPBS (-$MgCl_2$,-$CaCl_2$), addition of Lucigen QuickExtract onto the cells and incubation at RT for ~15 min while shaking. After incubation, the gel-like substance was transferred into PCR tubes and placed in the thermocycler (65°C for 40 min – 95°C for 20 min – 4°C ∞) for DNA extraction. Samples were stored at -20°C until further usage.

9.11 NGS preparation of mCherry-GFP reporter gDNA

The mCherry-GFP target region of the HepG2 gDNA had to be amplified in a PCR with modified NGS primers in order to submit it for Next Generation Sequencing and determine the percentage of editing in a cell population.

Table 5: 1x PCR Mix for NGS samples	
DEPC H_2O	33.5 µl
Takara 5x PrimeStar GXL Buffer	10 µl
Takara dNTP Mix 2.5 mM each	1.5 µl
10 µM forward primer (Frame shift GFP NGS F1)	1.5 µl
10 µM reverse primer (Frame shift GFP NGS R1)	1.5 µl
Takara PrimeStar GXL DNA Polymerase	1 µl
Plasmid template [1 ng/µl]	1 µl

Table 6: PCR conditions gDNA for NGS			
	Denaturing	95°C	3 min
30 cycles	Denaturing	98°C	10 sec
	Annealing	62°C	15 sec
	Extension	72°C	30 sec
	Final extension	72°C	1 min
		4°C	∞

PCR products were then analyzed using gel electrophoresis by diluting 5µL of PCR product in 5µL of DEPC H_2O and addition of 2µL 6x Purple Loading Dye and were run on a 1% Agarose + 1xSybrSafe gel at 120V for ~20 min. If all bands showed the expected length, PCR samples were cleaned using magnetic PCR cleanup beads. Magnetic bead particles were resuspended by vortexing for ~20 sec. To each 45 µL PCR sample, 81 µL of the magnetic bead solution were added and mixed by pipetting up and down 10 times. This mix was then incubated for 5 min to allow DNA binding to the beads. Samples were transferred onto a magnetic rack for ~ 10 min to separate

beads from the supernatant and clear solution was then retained from the tubes and discarded. To each sample, 200 µL of 70% ethanol was added and incubated for ~30 sec. This step was repeated for a total of two washes. Finally, the reaction tubes were removed from the magnetic rack and air dried completely on the benchtop for a minimum of 20 min or until the ethanol is evaporated completely. To retain the DNA, samples are resuspended in 40 µL of DEPC H_2O, placed on the magnetic rack and transferred away from the beads into fresh tubes. The DNA concentration of each sample had to be determined using the nanodrop and samples could be submitted for NGS.

10 Results

10.1 triNLS B11 Cas9 cloning

As seen in figure 7, amplification of the backbone and the insert was successful as there are bands visible at the expected size of ~10,000 bp (backbone) and ~500 bp(insert). Lanes 1 and 2 represent the amplified backbone and lanes 3 and 4 the amplified insert. On the very left, a 1kb DNA ladder can be seen and on the right side a 100 bp DNA ladder. Furthermore, insertion of the 7xB11 fragment of clone #4 was verified by Sanger sequencing and plasmid DNA from this clone was subsequently used to express and purify the 2C-B11 Cas9 construct in house, which was then used for all endosomal escape reporter experiments.

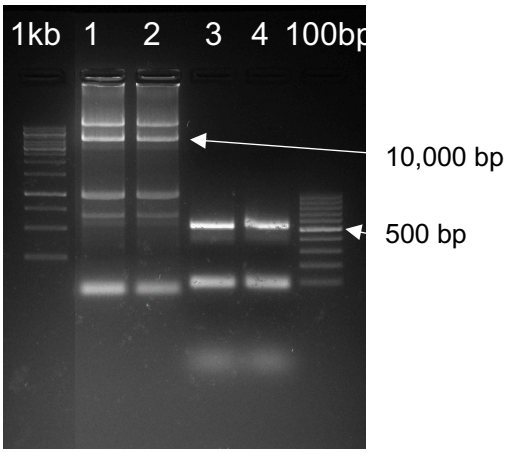
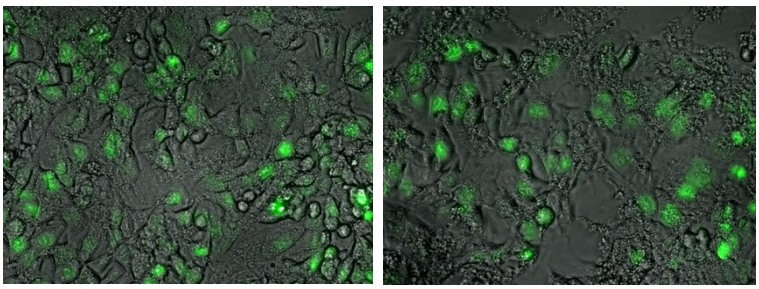


Figure 7: Gel electrophoresis of triNLS B11 Cas9 backbone and insert

10.2 Split-GFP reporter

Picture 8 shows transfected HepG2 endosomal reporter cells 30 hr after the Co-Incubation with B11-Cas9 and cell-penetrating peptide. Interestingly, almost all of the green signal was located in the nucleus after this time period and there is no significant difference in intensity or amount of signal visible by eye when comparing 2xNLS B11 Cas9 and triNLS B11 Cas9 and keeping in mind that the 2xNLS cells are a little less confluent in this image. This statement is also verified by flow cytometry, which showed very similar results for both Cas9 variants, with green fluorescent signal detected at around 35% of the cells.



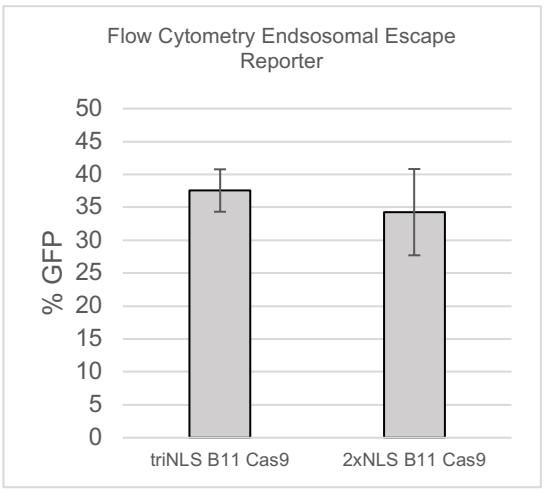


Figure 8: Top shows HepG2 cells under the fluorescent microscope with 20x magnification 30 hr after co-incubation with CPP and 2xNLS B11 Cas9 (left) and triNLS B11 Cas9 (right). Bar graph on the right shows the corresponding percentages of GFP signal detected by flow cytometry.

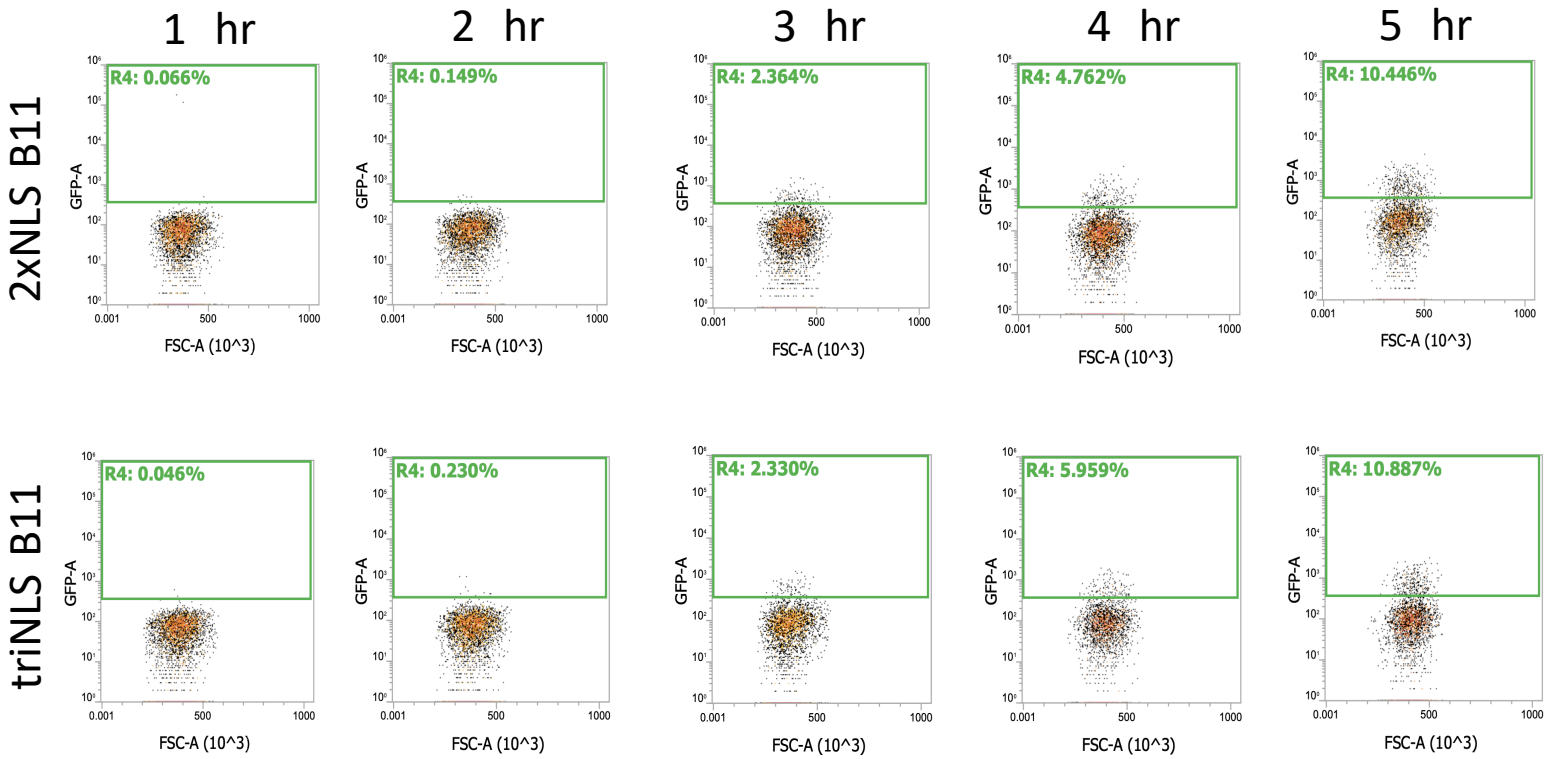


Figure 9: Time-course experiment of split-GFP reporter cells. Flow cytometry was performed every hour up to 5 hours post co-incubation with CPP and 2xNLS B11 Cas9 (top) and triNLS B11 Cas9 (bottom). Plots show the intensity of the GFP signal on the y-axis and the forward scatter on the x-axis.

Furthermore, a time-course experiment was performed to determine when endosomal escape first takes place. As seen in figure 9, first evidence of endosomal escape, indicated by detection of a green fluorescent signal when analyzing the cell samples on the Attune Nxt Flow Cytometer, was at 2-3 hours after the co-incubation of the endosomal escape reporter cells with B11-Cas9 and cell-penetrating peptide. The signal at timepoint 1 hr is not significant, since it is equal to the signal detected at the negative controls (around 0.05%), where no RNP was added. It can be seen that the GFP signal increases steadily as time passes, and there is a clear shift of the population upwards on the y-axis, which represents the fluorescent intensity. There is no significant difference in

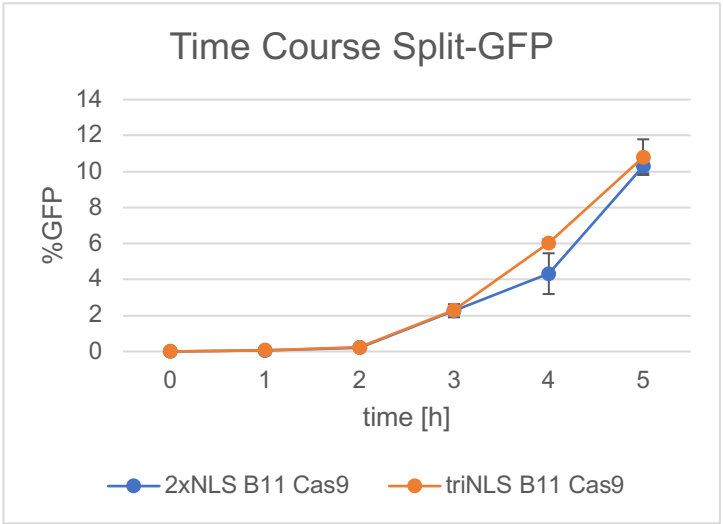


Figure 10: Graph summarizes time-course experiment results of all samples. GFP signal detected by flow cytometry on the y-axis and time-point after the co-incubation on the x-axis.

fluorescent intensity or speed of increasing signal when comparing triNLS B11 Cas9 and 2xNLS B11 Cas9. This experiment was conducted in triplicates and the results are summarized in figure 10, showing the percentage of detected GFP signal by flow cytometry at the y-axis. Again, the first detection of GFP, indicating endosomal escape, is at 2-3 hours with a steady increase up until 5 hr. Comparing all transfected samples, there was no significant difference between 2xNLS Cas9 and triNLS Cas9 in terms of endosomal escape, as both graphs mostly overlap or are very close to each other at almost all time-points.

10.3 mCherry-GFP reporter

In order to determine a standard gRNA for the mCherry-GFP editing reporter, different sgRNAs were tested for editing efficiency by nucleofection with RNP. Both, using a single RNP and a multi-guide approach (cutting out a small fragment of DNA in-between the fluorescent genes by using two RNPs with DNA targets in close proximity) were tried. Furthermore, Cas9 with 1 NLS was compared to Cas9 with three different NLS to see whether the number of NLS influences the nuclear import and thereby the editing efficiency. In figure 11, the blue bars represent the percentage of editing / non-

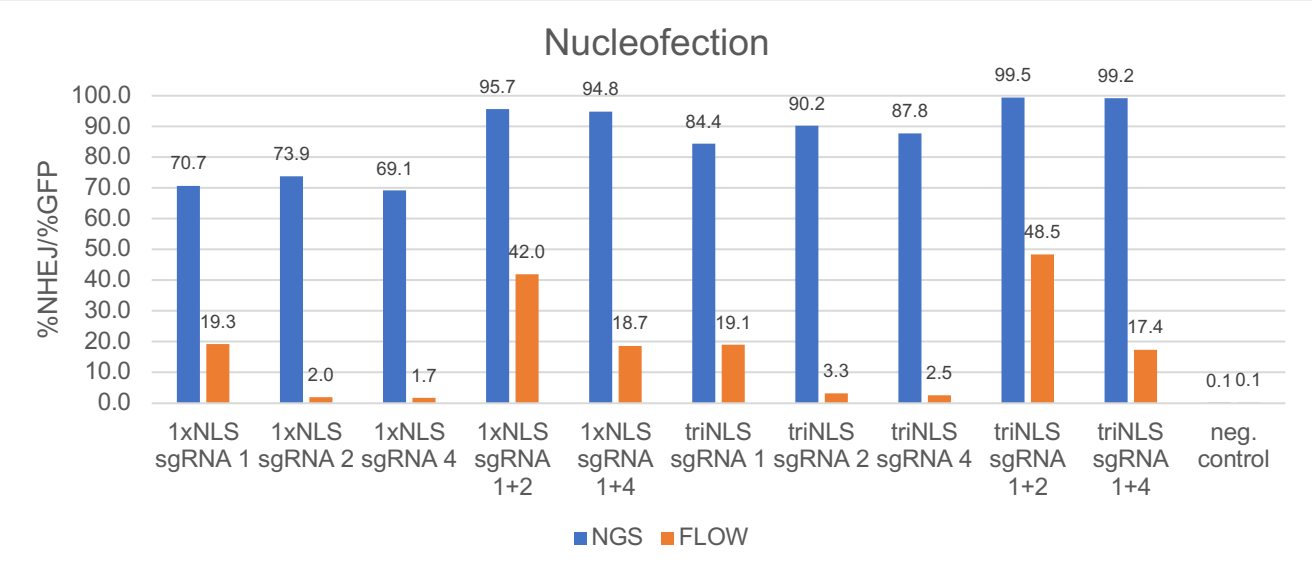


Figure 11: Screening of different gRNAs or combinations by nucleofection of mCherry-GFP reporter cells and subsequent NGS and flow cytometry.

homologous end joining that was detected by next generation sequencing and the orange bars represent the percentage of green fluorescent signal that was detected when analyzing the samples by flow cytometry. As a result, the multiguide approach with sgRNA 1+2 was chosen as the standard with the best editing and flow cytometry

results in terms of % of non-homologous end joining and % of GFP signal. This combination of sgRNAs gave the best editing results in both Cas9 variants and by both analysis methods – NGS and flow cytometry. The second-best candidate when looking at the NGS results was the other multi-guide (sgRNA 1+4), showing almost identical editing results, close to 100%, just like the chosen standard. The reason why it was not used in the end was because of the poor flow cytometry outcome, which was underreporting very strongly and did not show an increase in GFP signal compared to just using e.g. sgRNA 1. This was explained by looking at the position of the target of sgRNA 1 and sgRNA 4 on the plasmid map. The most probable cutting site of both gRNAs, which is 3 bp after the PAM sequence would excise a fragment of 42 bp, which is a multiple of 3 and would not result in a frameshift mutation that would lead to the turn-on of the GFP gene.

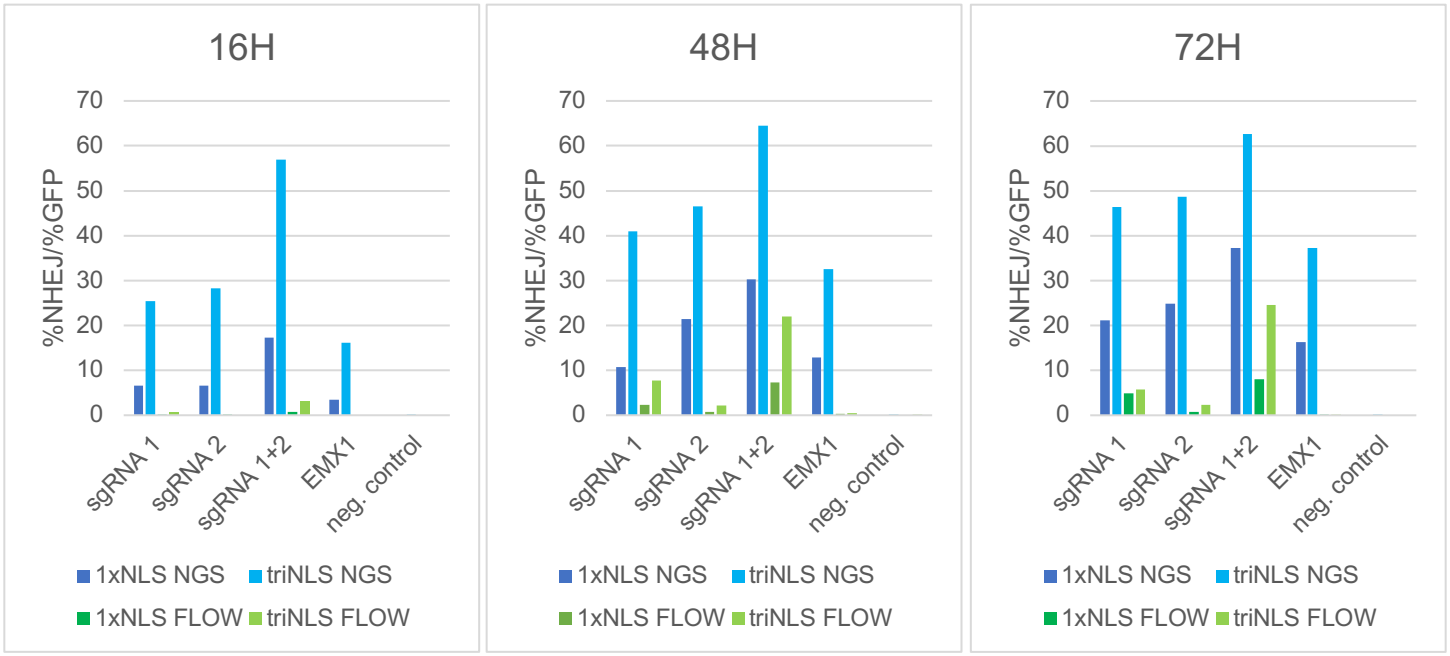


Figure 12: Comparing results from NGS and flow cytometry at 16, 48 and 72 hr after co-incubation with CPP and 1xNLS Cas9 or triNLS Cas9.

After excluding multi-guide 1+4 from the editing experiments, a co-incubation with CPP and multi-guide 1+2, sgRNA1 and sgRNA 2 just like a control with an RNP targeting the human EMX1 gene and a control with no RNP was performed. Samples were analyzed by flow cytometry and gDNA was harvested and sequenced via NGS after 16, 48 and 72 hr. As seen in figure 12, dark and light blue bars represent percentage of editing by NGS and dark and light green bars represent percentage of GFP signal detected by flow cytometry. Again, 1xNLS Cas9 was compared to triNLS

Cas9. It can clearly be seen that there is a significant increase of both, NGS and flow cytometry editing percentages going from 16 to 72 hr. This difference is especially visible for the flow cytometry samples, which gave almost no detectable GFP signal after 16 hr due to delay of protein expression compared to DNA editing, which can be detected earlier by NGS. Furthermore, findings from figure 11 were confirmed, showing the same trend and correlations of the used RNPs in terms of efficiency, meaning that multi-guide 1+2 gave the best editing results analyzed by NGS and flow cytometry in all cases. Also, like in the previous experiment, using sgRNA 2 gave better slightly better results by NGS but worse results by flow cytometry when compared to using sgRNA 1. Looking at the EMX1 controls, it can be seen that there was no GFP signal detected by flow cytometry, but very well indels detected by NGS, which proves that turning GFP on does not happen randomly in all reporter cells but only in those targeted by the right RNP. Also, targeting the EMX1 gene is a well-used control to check for success of editing in the Wilson lab and those results confirm that the co-incubation with CPP worked successfully and can be used as a method of transfection just like nucleofection. Comparing the samples transfected with 1xNLS Cas9 and triNLS Cas9 it is clearly visible that triNLS Cas9 gave significantly better editing results with both analysis methods. In almost all cases, NHEJ/GFP percentages were around 2-fold higher than with the 1xNLS Cas9. This is again the same trend that was already detected in the nucleofection experiment (figure 11)

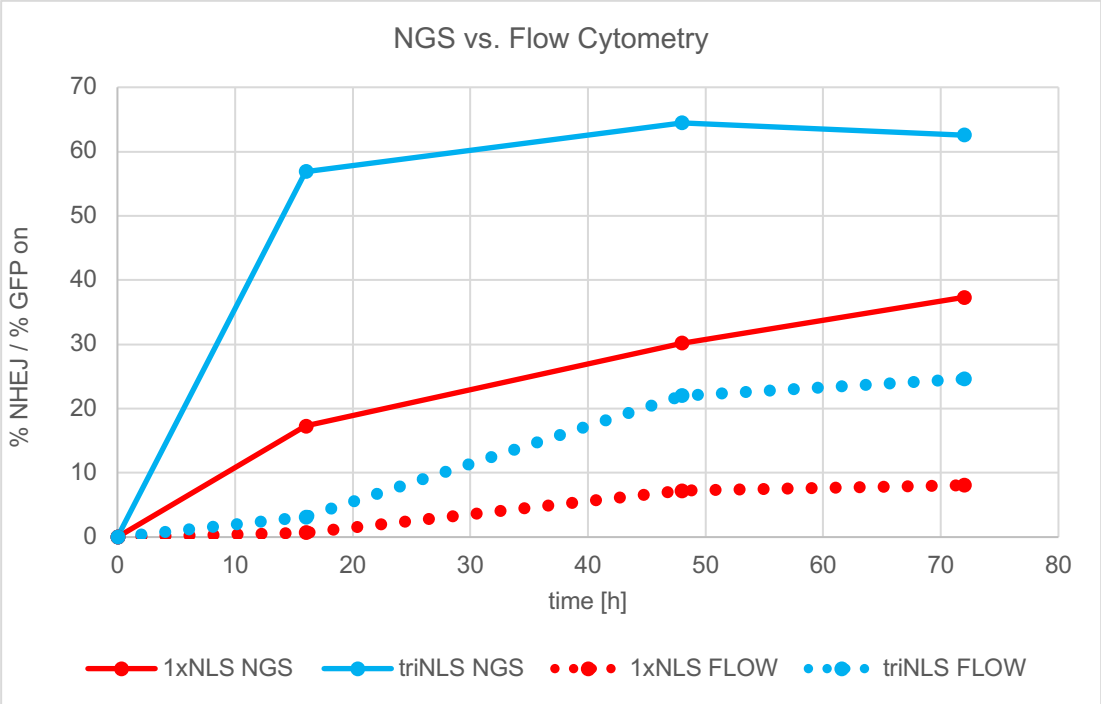


Figure 13: Comparing editing results from NGS and flow cytometry after 16, 24 and 48 hr post co-incubation with CPP and 1xNLS Cas9 or triNLS Cas9. Standard multi-guide RNA (sgRNA 1+2) was used in all samples.

Figure 13 outlines the data from figure 12, comparing only the samples with the standard multi-guide 1+2 using 1xNLS Cas9 and triNLS Cas9. Dotted lines represent flow cytometry results in terms of detected GFP signal and solid lines represent percentage of editing detected by NGS. When looking at the NGS results, the triNLS Cas9 samples in blue do not show a very big increase in terms of editing going from 16 to 72 hr, in contrast to the 1xNLS Cas9 samples, where editing clearly increases with time. Furthermore, it is even more visible in this graph that GFP expression in the reporter cell lines takes up to 48 hr and can thereby not be detected by flow cytometry after 16 hr, even though the DNA was most likely already edited as indicated by the NGS results. 72 hr ultimately turned out to be the best time-point to conduct both, flow cytometry and harvesting of gDNA for NGS. Also in this graph, the 2-fold increase when going from 1xNLS Cas9 to triNLS Cas9 is clearly visible.

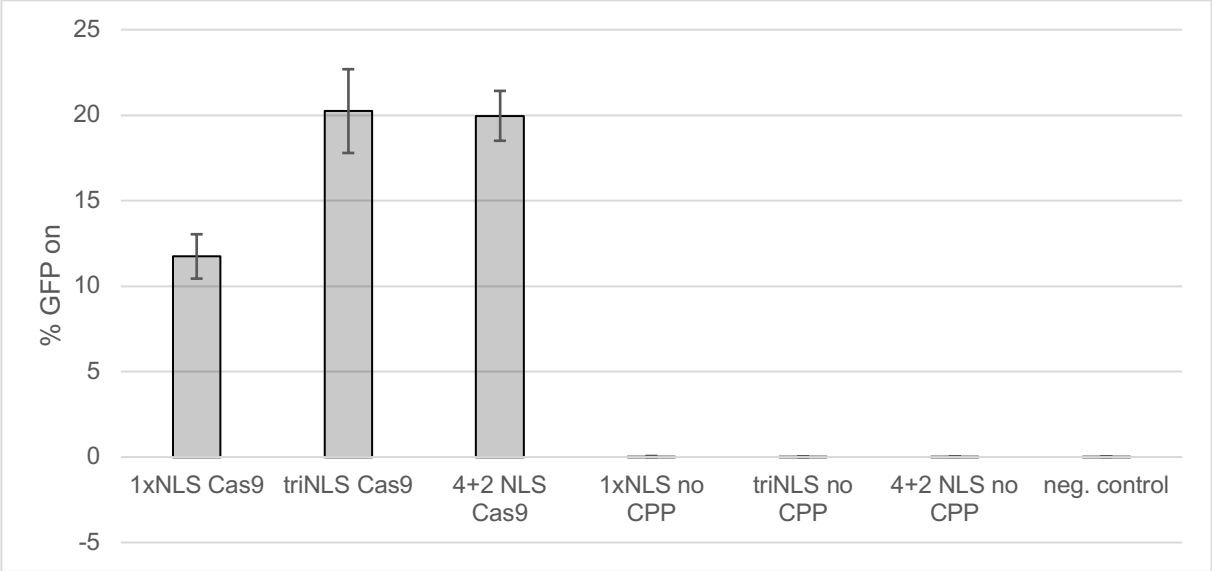


Figure 14: GFP signal detected by flow cytometry 72 hr after co-incubation with three different Cas9 variants with different amounts of NLS and with or without CPP.

Ultimately, the standard RNP was used to perform another co-incubation with CPP to compare 1xNLS, triNLS and 4+2 NLS Cas9 (Cas9 with 4 SV40 NLS at the N terminus and 2 SV40 NLS at the C terminus). This was done in triplicates. Results in figure 14 show the same trend of a 2-fold increase going from 1xNLS to triNLS Cas9, as already seen in previous experiments. Nevertheless, comparing triNLS and 4+2 NLS Cas9 does not show a significant difference in GFP signal despite having more NLS attached. Additionally, the cell-penetrating capacities of NLS alone were tested by co-incobation of HepG2 cells with RNP but without CPP. There was no editing detected

in any of the samples that were incubated with RNP only, indicating that even with more NLS attached, Cas9 cannot enter the cells unspecifically.

11 Discussion and Conclusion

General findings and remarks

Conclusively, it can be said that this project showed that both reporter cell lines are promising tools for further use as a standard analysis method for efficiency of cellular uptake of genome-editing enzymes in the Wilson lab. As expected, there was green signal detected in the endosomal escape reporter cells upon endosomal escape due to formation of an active GFP barrel as-well as in the editing reporter cells upon nuclear entry/DNA editing due to activation of the GFP gene and subsequent expression of GFP protein. In both cases, the signal increased according to the endosomal escape or editing efficiency as hypothesized, meaning the better Cas9 could escape the endosome or edit the DNA, the more green signal was detected. This enables comparison of different variations of Cas9 using the reporter cell lines which make them a possible tool to compare the influence of properties like addition of NLS to Cas9 on endosomal escape and nuclear import.

It should be mentioned that there were two different methods of transfection of Cas9 used in this project, nucleofection/electroporation and incubation with CPPs. As it can be seen when comparing figure 11 and 12, transfection rates and thereby editing efficiencies of nucleofections will always be higher because RNPs are “shot” directly into the nucleus of the cells with electric impulses, while co-incubation with CPPs is a little more gentle and allows uptake via an endocytic pathway. [36]

Split-GFP reporter for endosomal escape

Looking at the results from the split-GFP reporter, it is interesting to see that after 30 hr of incubation, all of the GFP was accumulated in the nucleus. This leads to the hypothesis that all of Cas9 is imported into the nucleus at this time-point or, the less desirable explanation, that the GFP protein is simply “chopped off” from Cas9 and is translocated because of the SV40 NLS on the C-terminus of the Cas9-B11 plasmid (figure 6), right after the 7xB11 fragments. To find out if Cas9 is still attached to the GFP proteins, immunostaining of Cas9 with subsequent confocal microscopy could be

performed. If the Cas9 signal is detected in the same locations as the GFP, they are still fused, if they are in different locations, it was most likely removed from Cas9.

mCherry-GFP reporter for genome editing/nuclear import

One thing that has to be considered when looking at the results from the DNA editing reporter is that flow cytometry analysis will always underreport compared to the NGS results. This can be explained due to the fact that with NGS, every kind of insertion or deletion after the DSB repair can be detected, while for flow cytometry, only a frameshift mutation of -2 bp or +1 bp leads to the right frameshift that activates the GFP transcription. Therefore, all other mutations stay “hidden” with this reporter. This also explains why for some of the tested gRNAs, editing by NGS was extremely high but unproportionally low when looking at the flow cytometry results simply because they induced the wrong frameshift. Nevertheless, it was shown by NGS that for the standard RNP (sgRNA 1+2), the most common indel did indeed give the right frameshift and due to robust indel profiles, meaning reproducibility of indel distributions, flow cytometry can be used with this multi-guide to compare ratios of editing efficiencies. This was one of the most important finding of this project – flow cytometry always gave the same overall results as NGS, making it a reliable method to do first big screenings of RNPs to compare editing efficiencies.

Effect of nuclear localization signals on endosomal escape and nuclear import

Another very interesting outcome was the comparison of Cas9 with different NLS attached and how they behave in terms of endosomal escape and nuclear import. As expected, there was no difference visible for endosomal escape of Cas9 with 2 and 3 NLS when looking at figure 8. Seeing figure 14 on the other hand, there was a big difference for nuclear import between Cas9 with 1 and 3 NLS. This clearly showed that having more NLS can improve nuclear import. Interestingly, there was no big difference when comparing Cas9 with 3 and 6 NLS attached. This might be due to the reason that triNLS Cas9 contains 3 different NLS (nucleoplasmin, SV40 and C-myc), while 4+2 NLS Cas9 contains of 6 of the same NLS (SV40). Therefore, variability and not only number of NLS might also have a positive effect on nuclear import. Another positive finding was that Cas9 alone was not able to enter the cells (figure 14), even with many NLS attached, meaning there was no non-specific uptake and subsequent editing. Due to their positive charges, some NLS, including the 4+2 NLS Cas9, showed

cell-penetrating properties [37-39], which would be highly unwanted in the context of target-specific delivery of genome editing enzymes.

Finally, it has to mentioned that all experiments were conducted with a relatively low sample size (1-3 replicates per sample). This is due to the fact that for the first “big-screenings”, sample variability was too simply too high for more replicates. This project therefore was a first successful baseline for further screenings and optimizations.

12 References

1. Nguyen, D.P., et al., *Ligand-binding domains of nuclear receptors facilitate tight control of split CRISPR activity.* Nat Commun, 2016. **7**: p. 12009.
2. Leonetti, M.D., et al., *A scalable strategy for high-throughput GFP tagging of endogenous human proteins.* Proc Natl Acad Sci U S A, 2016. **113**(25): p. E3501-8.
3. Rouet, R., et al., *Engineering CRISPR-Cas9 RNA-Protein Complexes for Improved Function and Delivery.* CRISPR J, 2018. **1**(6): p. 367-378.
4. Danna, K. and D. Nathans, *Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae.* Proc Natl Acad Sci U S A, 1971. **68**(12): p. 2913-7.
5. Rouet, P., F. Smih, and M. Jasin, *Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease.* Mol Cell Biol, 1994. **14**(12): p. 8096-106.
6. Seligman, L.M., et al., *Mutations altering the cleavage specificity of a homing endonuclease.* Nucleic Acids Res, 2002. **30**(17): p. 3870-9.
7. Klug, A. and D. Rhodes, *Zinc fingers: a novel protein fold for nucleic acid recognition.* Cold Spring Harb Symp Quant Biol, 1987. **52**: p. 473-82.
8. Kim, Y.G., J. Cha, and S. Chandrasegaran, *Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1156-60.
9. Urnov, F.D., et al., *Genome editing with engineered zinc finger nucleases.* Nat Rev Genet, 2010. **11**(9): p. 636-46.
10. Boch, J., et al., *Breaking the code of DNA binding specificity of TAL-type III effectors.* Science, 2009. **326**(5959): p. 1509-12.
11. Christian, M., et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
12. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
13. Finn, J.D., et al., *A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing.* Cell Rep, 2018. **22**(9): p. 2227-2235.
14. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
15. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.
16. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
17. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity.* Science, 2016. **351**(6268): p. 84-8.
18. Kleinstiver, B.P., et al., *High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects.* Nature, 2016. **529**(7587): p. 490-5.
19. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-83.
20. Jinek, M., et al., *RNA-programmed genome editing in human cells.* Elife, 2013. **2**: p. e00471.

21. Kim, S., et al., *Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins.* Genome Res, 2014. **24**(6): p. 1012-9.
22. Wilson, R.C. and L.A. Gilbert, *The Promise and Challenge of In Vivo Delivery for Genome Therapeutics.* ACS Chem Biol, 2018. **13**(2): p. 376-382.
23. Naso, M.F., et al., *Adeno-Associated Virus (AAV) as a Vector for Gene Therapy.* BioDrugs, 2017. **31**(4): p. 317-334.
24. Hamilton, H., et al., *Adeno-associated virus site-specific integration and AAVS1 disruption.* J Virol, 2004. **78**(15): p. 7874-82.
25. Gaj, T., et al., *Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery.* Nucleic Acids Res, 2017. **45**(11): p. e98.
26. Swiech, L., et al., *In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9.* Nat Biotechnol, 2015. **33**(1): p. 102-6.
27. Miller, J.B., et al., *Non-Viral CRISPR/Cas Gene Editing In Vitro and In Vivo Enabled by Synthetic Nanoparticle Co-Delivery of Cas9 mRNA and sgRNA.* Angew Chem Int Ed Engl, 2017. **56**(4): p. 1059-1063.
28. Zuris, J.A., et al., *Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo.* Nat Biotechnol, 2015. **33**(1): p. 73-80.
29. Rouet, R., et al., *Receptor-Mediated Delivery of CRISPR-Cas9 Endonuclease for Cell-Type-Specific Gene Editing.* J Am Chem Soc, 2018. **140**(21): p. 6596-6603.
30. Bibikova, M., et al., *Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases.* Genetics, 2002. **161**(3): p. 1169-75.
31. Adli, M., *The CRISPR tool kit for genome editing and beyond.* Nat Commun, 2018. **9**(1): p. 1911.
32. Kim, J.S., et al., *Engineering of a tumor cell-specific, cytosol-penetrating antibody with high endosomal escape efficacy.* Biochem Biophys Res Commun, 2018. **503**(4): p. 2510-2516.
33. Lonn, P., et al., *Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics.* Sci Rep, 2016. **6**: p. 32301.
34. Kamiyama, D., et al., *Versatile protein tagging in cells with split fluorescent protein.* Nat Commun, 2016. **7**: p. 11046.
35. Ramakrishna, S., et al., *Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA.* Genome Res, 2014. **24**(6): p. 1020-7.
36. Jiao, C.Y., et al., *Translocation and endocytosis for cell-penetrating peptide internalization.* J Biol Chem, 2009. **284**(49): p. 33957-65.
37. Yu, W., et al., *Highly efficient cellular uptake of a cell-penetrating peptide (CPP) derived from the capsid protein of porcine circovirus type 2.* J Biol Chem, 2018. **293**(39): p. 15221-15232.
38. Ragin, A.D., R.A. Morgan, and J. Chmielewski, *Cellular import mediated by nuclear localization signal Peptide sequences.* Chem Biol, 2002. **9**(8): p. 943-8.
39. Staahl, B.T., et al., *Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes.* Nat Biotechnol, 2017. **35**(5): p. 431-434.