Norwegian University
of Life Sciences

Master's Thesis 2021 60 ECTS
Faculty for Biosciences

Cloning and characterization of Cas9 in IPEC-J2 cells for CRISPR screens against porcine E. coli pathogen

Jostein Arntsen
Master Biology

Acknowledgments

I would like to extend my deepest gratitude to my supervisors Thomas Nelson Harvey and Victor Boyartchuk for their help in the lab and insightful comments on the master thesis. I also thank Matthew Peter Kent for organizing weekly meetings that helped bring structure to the master and giving encouraging words. I also like to thank Mari Austad Brandt for her help in the cloning of Lentiguides. I very much appreciate the support from various people at Cigene with the questions they answered when I needed things in the lab.

Abstract

Background:

Pigs are a livestock animal widely used in agriculture, used for its meat. A large problem in the pig industry is diarrhea caused by various intestine bacteria, with enterotoxigenic *Escherichia coli* (ETEC) being one of the main contributors. It attaches to the intestine epithelia cells using fimbria and releases endotoxins causing diarrhea. Resistance can be tied to genetic differences between pigs. To determine which genes that play a role in resistance, a CRISPR screen can be used.

Results:

In this master thesis, the goal is to test the viability of a CRISPR screen of IPEC-J2 cells for resistance genes against F4 ETEC. CRISPR/Cas9 was delivered into the cells via lentivirus, made into cell lines and tested for expression of the Cas9. Guide RNA was designed to be used against known resistance genes against ETEC, MUC4, MUC13 and FUT1. Eight different transduced IPEC-J2 Cell lines with different levels of Cas9 expression were chosen and had one gRNA delivered via lentivirus to perform knockout of MUC4 gene. In addition, an experiment was done to evaluate using propidium iodide to test cell permeabilization of IPEC-J2 cells affected with ETEC.

Conclusion:

The transduction proved to be successful with little optimization. Creation of cell lines were difficult, but a change in method showed improved results. Both qPCR and western blot was used to evaluate mRNA and protein expression, which showed the transduced IPEC-J2 cell lines having a wide range of Cas9 expression. Transduced cells with both Cas9 and gRNA for MUC4 knockout showed indels in the MUC4 gene in all cell lines except one. This proves that IPEC-J2 cells can used for a CRISPR screen. Propidium iodide proved to be a bad tool for testing phenoptype in an eventual phenotype test for ETEC.

Sammendrag

Bakgrunn

Svin er ofte brukt husdyr i landbruket, brukt primært for dets kjøtt. Et ofte forekommende problem i svineindustrien er diare forårsaket av bakterier, hvor enterotoksinproduserende

Escherichia coli (ETEC) er en av de største bidragsyterene. Den fester seg til tarmveggens epithel celler via reseptorer på celleoverflaten med fimbriae og skiller ut enterotoksin som forårsaker diare.

Resultater

I denne master oppgaven så er målet å teste gjennomførbarheten av en CRISPR knockout skjerm av IPEC-J2 celler for resistens gener mot F4 ETEC. CRISPR/Cas9 ble levert inn i cellene av lentivirus, gjort om til celle linjer og testede for ekspresjon av Cas9. Guide RNA ble designet for å bli brukt mot kjente resistens gener mot ETEC: MUC4, MUC13 og FUT1. 8 Forskjellige transduserte IPEC-J2 cellelinjer med forskjellige nivåer av Cas9 utrykkelse ble valgt og hadde en gRNA levert via lentivirus for å utføre en deaktivering av MUC4 genet. I tillegg ble et eksperiment utført for å evaluere om propidiumjodid kan brukes for å teste endring i celle permeabiliteten til IPEC-J2 celler påvirket av ETEC.

Konklusjon

Transduskjonen viste seg å være vellykket med lite behov for optimalisering. Opprettelsen av celle linjer var utfordrene, men en endring i metode ga forbedrede resultater. Både qPCR og western blot ble brukt for å evaluere mRNA og protein produksjon nivåer, som viste at de transduserte IPEC-J2 cellene hadde en bred variasjon av Cas9 uttrykkelse. Dette viser at IPEC-J2 celler kan bli brukt for CRISPR skjerming. Propidiumjodid viste seg å være en dårlig verktøy for å teste fenotype i en eventuell fenotype test for ETEC effekt på IPEC-J2 celler.

Content

1 Introduction

1.1 E. coli pathogen in pigs

Pigs are an important animal in agriculture. Pork comprises 32,7% of all meat produced worldwide in 2019. (Food & Agriculture Organization of the United, 1997) As such, is any diseases that targets pigs can have a substantial economic cost. A pathogen oft contributing to economic losses related to disease is EnteroToxigenic *E. Coli* (ETEC), which causes pig diarrhea and can eventually lead to death. The cause for it can be manyfold, but a common cause is enterotoxigenic *Escherichia coli* (ETEC), which can cause diarrhea both before and after weaning. (Fairbrother et al., 2005)

ETEC bind mammalian cells by attaching its fimbriae to receptors and then excreting exotoxins, damaging the cells and causing diarrhea. Different fimbria bind to different receptors. As such ETEC can be classified according to their fimbriae, which include F4 , F5 , F6 , F18 and F41. The expression of target receptors are genetically determined by the host, so they play a role in resistance against ETEC. (Sinha et al., 2019).

Enterotoxin is a kind of toxin that targets cell in the intestine. ETEC has many well documented enterotoxins divided into two types, heat stable and heat liable (ST and LT respectively). Examples of enterotoxins produced in ETEC are STa, STI, ST_p, ST_h, ST_b ,LT-I, LT-II and ST_b (Sears & Kaper, 1996). The latter is of special interest in this study. ST_b interact with sulfatide present on the surface of intestinal epithelial cells in the porcine jejunum, casing a cascading effect with an iflux of Ca^{2+} into the cell. The increased Ca^{2+} levels in the cell affects calmodulin-dependent protein kinase II (CaMKII) through the Ca2+-calmodulin pathway and also protein kinase C (PKC)-mediated activation of CFTR, causing the cell to expel fluid into the intestine. ST_b also affects tight junctions in intestinal epithelial cells. Tight junctions are transmembrane protein complexes making a semi-permeable barrier across the cell membrane. The increased levels of Ca^{2+} caused by ST_b affects the tight junctions by redistributing a transmembrane protein crucial for tight junction integrity, claudin-1, from the plasma membrane to the cytosol. This causes an increased paracellular permeability. (Dubreuil, 2017)

1.2 CRISPR Technology

CRISPR, which is short for Clustered Regularly Interspaced Short Palindromic Repeats, refers to a set of molecules capable of performing targeted gene editing in cells. CRISPR specifically refers to a repeating pattern in DNA with short, non repeating sequences between them, which occurs in 50% of bacterial and 90 % of archaeal complete genomes. (Makarova et al., 2015) This was found out to be a library of sequences taken from invading bacteriophages and incorporating it in its own. These sequences would be translated and used together with CRISPR associated proteins (Cas) to find and cut sequences from other bacteriophages, as a form of learned immunity. (Ishino et al., 2018)

By editing the RNA sequence used to guide the Cas protein to a target sequence, the target of the Cas protein can be specified. This enables the CRISPR system to be used as a sequence specific gene editing tool. The most common used for this is Crisper II. Originally isolated from *Streptococcus pyogenes*, it consists of a Cas9 protein, CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA). When expressed, the crRNA and tracrRNA hybridizes and makes a RNA hybrid complex. This complex is bound to the Cas9 protein and guides it to a complementary DNA sequence where the Cas9 will make a double stranded cut. It is the crRNA sequence that decides where the Cas9 will cut. The tracrRNA functions as a scaffold so that the mRNA hybrid has the correct shape to fit into Cas9. By changing the crRNA, the target for the Cas9 can be specified. The crRNA and tracrRNA are separate in nature, but can be combined into a single sequence, called single guide RNA (sgRNA). This is what is used in this experiment and all references to gRNA hereafter refers to this type of sequence. After the cut, the cell machinery attempts to repair the damage caused by the Cas9. This can be done in two ways: homology directed repair (HDR) or non-homologous end joining (NHEJ). HDR uses a template containing sequences homologous to the regions flanking the cut site for repair while NHEJ does not. Because of the lack of a template, the NHEJ is prone to errors and introducing indels to the cut site. This is often used to perform knockout experiments, where the CRISPR system is used to introduce a cut in a gene of interest, which will be rendered non-functional due to indels caused by the NHEJ which shift the reading frame of the protein. Such a loss-of-function experiment can give insight into a genes function by contrasting phenotype of the cell with and without the edit. (Wang et al., 2016)

Even with the sequence specific targeting of CRISPR, it can have off-target effects. It is possible for a guide designed for a gene to also be able to target other parts of the genome that has the same or similar sequence as the target gene. In addition, the gRNA Does not necessarily need 100% overlap with a sequence to bind and facilitate a cut. So when designing guide RNA, is has to be compared to the target genome to avoid overlap. (Yip, 2020)

1.3 Lentiviral delivery method

For CRISPR to work in a cell, it first must be delivered to the cell. The gRNA and Cas9 delivery method influence the editing efficiency. The CRISPR components can be provided in the form of DNA (as a plasmid), RNA (mRNA of Cas9 and gRNA) and ribonucleoprotein complex (Cas9 protein with gRNA). Each form has its own advantages and disadvantages. Proteins work fastest, since the Cas9 protein can perform the edit soon after it enters the nucleus, but genome editing is transient as the protein will degrade after some time. An upside is that this reduces the chance of off target effects. Additional weakness in this approach to delivery is, since the protein comes from bacteria, that it can induce carryover of endotoxins from bacteria and trigger an immunologic response from the cell (You et al., 2019). RNA is faster than DNA, but slower than protein. mRNA is susceptible to be degraded by RNases, so it too will be transient, reducing the probability of off target effects (Wu et al., 2014). DNA, where the Cas9 gene is in a plasmid, is a good option but requires target cells to be easily transfectable. However, when CRISPR components are put on a viral plasmid, the Cas9 gene and gRNA can be packages into a virus and inserted into the genome of the cell, allowing it

to be continually expressed. The tools necessary to work with lentiviruses are available in a standard laboratory. However, it takes longer than both mRNA and protein. (Yip, 2020).

For plasmid DNA delivery, there are a number of different methods that can be used. Examples include microinjection, gene gun, electroporation and virus. Injection of the plasmid into the cell is simple in concept but can be difficult in practice. Depending on the size of the cell, it would be hard to get a needle into it, which can also damage the cell. Gene gun works by loading the plasmid onto heavy and non-reactive particles, like gold, and then sending the plasmid coated particles at high speed at the cells. The particles penetrate the membrane and delivers the plasmid into the cell. This approach works even on cells with cell walls, like plants. Electroporation works by applying series of electrical pulses to the cell which stimulates temporary opening of pores in the membrane, where plasmids can enter. Many viruses have developed to be able to integrate its genome into a wide array of cells. By replacing the part cloned into the host by the virus with Cas9, the virus can be used as a delivery method for introducing the Cas9 gene into the genome of a cell. There are many different types of viruses that can be used for this. Examples are Adeno-associated viruses, Adenoviruses and lentiviruses. (Yip, 2020).

Adeno-associated viruses (AAV) have the advantage that they don't integrate into the genome, so the expression of their load is transient. They do have a low cloning capacity at less than 4,7 kb. Adenoviruses also do not integrate into host genome and can tranduce both dividing and non-dividing cells. The drawback is that they trigger a high immune response and require a lot of time and effort to make. Lentiviruses do integrate into the host genome and have a high cloning capacity at up to 8 kb. They are easier to make than AAV and adeno viruses. The drawback is that the integration into host cell genome is random, which can cause undesired effects like tumorigenesis. This makes lentiviruses a poor choice for in vivo experiments. (Yip, 2020).

1.4 CRISPR Screens

Genes affect a wide range of phenotypes. Some genes cause the phenotype directly by creating proteins that are integral to the task, other genes regulate other genes which in turn affect the phenotype. When investigating a phenotype, like survival in an environment, resistance to disease, effect of drugs or defining what affects certain cell states, it is reasonable to expect genes to have an effect. But determining which genes has an effect, and at what magnitude, can be difficult. One way of doing this is by a CRISPR screen. (Spencer, 2019)

A knockout CRISPR screen is a large scale loss-of-function experiment, where all genes are knocked out using CRISPR in individual cells to determine what specific genes play a role in determining the phenotype. Since the experiment is set up in such a way that only one gene is knocked out per cell, in the end there will be a population of many cells, each with one gene knocked out. The delivery is done with lentivirus since it integrates both Cas9 and gRNA into the genome. Even after the Cas9 has cut the target, the cell produces gRNA and the since the gRNA is integrated into the genome, successive generations of cells also have the gRNA. Because of this, mixed populations of cells can get their genomic DNA sequenced to determine which of gene knockouts allowed to survive to selection. If the target of the

screen is vital genes for survival, the cells that survived the initial knockout have their genomic DNA sequenced. In the case of testing a disease or drug, the cells are exposed to them and the surviving cells DNA sequenced. Based on the gRNA sequences found in the mixed population, which genes are present can be determined.

There are two kinds of CRISPR screens, positive and negative. I will use the example of screening for vulnerability/resistance to a drug. In a negative screen, it is genes that confer resistance that are being looked for. Cells with the resistance gene will survive, while cells that have had their resistance gene knocked out will die. When sequencing, it is those gRNA not present that shows which genes confer resistance. In a positive screen, it is genes that cause vulnerability that is being looked for. Cells with their vulnerability gene intact, will be killed by the drug. Cells with their vulnerability gene knocked out will survive. So when sequencing, it is the gRNAs present that represent what genes are relevant for vulnerability. (Poirier, 2017)

After the genes of interest are selected, which can include all genes in the genome if there is little knowledge of the physiological phenomenon to be explored, a CRISPR library needs to be made. A CRISPR library contains A collection of individual lentiviruses with different gRNA for each gene to be targeted in the screen. The reason to have multiple gRNA for each gene is to increase the chance that the CRISPR system knocks out the genes in the screen. Additionally, it reduces the uncertainty since many gRNAs missing or present gives more confidence in the result than one, which is more susceptible to random chance. A total of 6-8 guides per screen is the recommended amount.(Cuellar et al., 2018)

1.5 Intestinal porcine enterocyte cell line (IPEC-J2)

The intestinal porcine enterocyte cell line are secondary cells from epithelium cells isolated from the second part of the small intestine, the jejunum, in a neonatal piglet in 1989. The cells closely resemble the morphological and functional traits of epithelium cells *in vivo*. IPEC-J2 cells express F4 fibrial receptor, a receptor used by ETEC to attach to the cells, but not F18 receptor. (Vergauwen, 2015)

Mucins are genes expressed in epithelial cells in the surface of respiratory, gastrointestinal and reproductive tracts and create high-molecular-weight proteins that protect and lubricate the surface. The mucin 4 (MUC4) protein is membrane-bound O-glycoprotein and is shown to be linced to ETEC resistance. (Jin & Zhao, 2000; Peng et al., 2007; Sinha et al., 2019)

1.6 Goals of the master

The goal of the thesis is evaluating the feasibility of lentiviral deliver of CRISPR/Cas9 gene and gRNA into porcine IPEC-J2 cells, create monocultures of Cas9 transduced cells, and determine levels of mRNA and protein Cas9 expression in individual monocultures. A set of Cas9 expressing cells with varying levels of expression will then be transduced with gRNA targeting ETEC resistance genes and the frequency of mutations in ETEC resistance genes will be determined to identify those with optimal Cas9 levels for editing in porcine cells. Subsequently estimate the magnitude of resistance of generated mutant cells to infection with ETEC.

2 Methods

2.1 Cell lines and cell culture

IPEC-J2 cells were grown in Dulbecco's modified Eagle Medium – High glucose supplemented with 10% Fetal Bovine Serum and 1% penicillin-streptomycin (Thermo Fisher) Cells were incubated at 37°C and 5% CO2 in Nunc™ EasYFlask™ Cell Culture Flasks with. Cells were passaged by trypsinising the cells with 0,25% trypsin EDTA and diluted 1:10, which was done every 2-4 days. Before use, the cells were washed in PBS 1-2 times. Cells were counted by mixing with 0,4 % Trypan blue in 1:1 dilution and then using a Bio-Rad TC20™ Automatic Cell counter. To freeze the cells for long term storage, cells were first centrifuged at 125xg for 10 minutes and then removed the supernatant. The pellet was resuspended in DMEM with FBS and penicillin-streptomycin as described above, in addition with 5% DMSO. The amount of media depended on how many cells there were and how many separate cryotubes were needed. 1 ml of minimum $1*10^6$ cells (maximum $5*10^6$) were added to individual cryotubes and then put on a Mr. Frosty™ Freezing Container (Thermo-Fisher) and put in a -80°C freezer. After at least 24 hours the frozen cells were taken out of the container and put in a liquid nitrogen tank for long term storage.

HEK293T cells that were used for packaging lentiviruses were grown in Dulbecco's Modified Eagle Medium – High glucose supplemented with 10% Fetal Bovine Serum, 1% pencillin-streptomycin and 2mM L-glutamine. The cells were incubated at 37°C in 5% CO_2. Subculturing is done in the same way as IPEC-J2 cells.

2.2 Plasmids and cloning

2.2.1 Cas9 plasmid

lentiCas9-Blast plasmid was used to introduce Cas9 gene into IPEC-J2 cells. LentiCas9-Blast was a gift from Feng Zhang (Addgene plasmid # 52962 ; http://n2t.net/addgene:52962 ; RRID:Addgene_52962).

2.2.2 Design of guides

Sequences for the targeting guides were identified using the webpage based tool ChopChop (Labun et al., 2019). The genes targeted were MUC4, MUC13 and FUT1, in Sus Scrofa, with Cas9 doing a Knockout.

2.2.3 Cloning of gRNA

To create the guide RNA for Cas9 knockout, a set of primers were made based on the gRNA designed for MUC4, MUC13 and FUT1. One primer was the forward guide with bases CACCG added to the 5' end. The other primer, reverse, had a C base added to the 3' end and CAA added to the 5' end. See figure 1 for an illustration. The primers can be found in table 1.

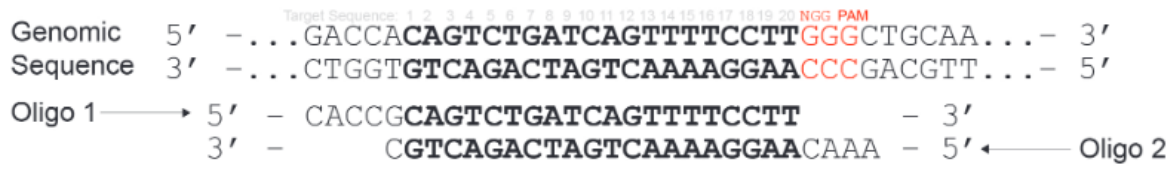


Figure 1: Illustration of what bases were added to guide sequence to be hybridised and inserted into lentiGuide-Puro. PAM sequence is not included in the primers.

5000 ng lentiGuide-Puro, 25 Units of BsmBI and 9,4 µl NE 3.1 buffer diluted in water to 50 µl total was incubated for 30 min at 37 ℃. 5 µl digested lentiGuide-Puro were then run on a gel to verify the digestion and then all remaining digested lentiGuide-Puro DNA was run on another gel and the larger lentiGuide-Puro backbone was extracted by cutting the band from the gel. The cut plasmid was purified by using a QIAquick Gel Extraction Kit (Qiagen). The primers used would target MUC4, MUC13 and FUT1 genes, with two guides for each gene. The primers are listed in Table 1.

Annealing of guide sequence primers were done by adding 100 µM of each complementary primer together with 1 µl 10X T4 Ligation Buffer (NEB) and diluted with ddH_2O to 10 µl final volume and incubated at 37 ℃ for 30 min and then 95 ℃ for 5 min and then ramped down to 25℃ at 5℃/min. Annealed oligos were then diluted 1/200.

For annealing reaction 50ng of BsmBI digested lentiGuide-puro, 1 µl annealed oligos and 5 µl 2x quick ligase buffer (NEB) and x µl ddH_2O to 10 µl was made, then 1µl T4 DNA ligase were added for a total 11µl and incubated for 10 min at room temperature.

Bacteria used for cloning were *E. coli* Stbl3 strain (Thermo-Fisher). The transformation was done following manufacturers protocol. 10 colonies were selected from each guide and streaked out on a plate, making large monocultures. To verify if the gRNA plasmid were correct size, several colonies were resuspended in a lysis buffer (20 % w/v sucrose, 200mM NaOH, 10 mM EDTA, 0,5% SDS, 120 mM KCl and Bromocresol green) and extract run a 1% agarose gel for 30 min at 100V. The three best monocultures, based on if they had correct fragment size and strong band intensity in the colony cracking , were chosen for each guide to be tested further.

The guides chosen after colony cracking were amplified using Platinum™ II Hot-Start PCR Master Mix (2X) (Thermo-Fisher) according to manufacturer's protocol. Amplified guides were then cut using SalI-HF restriction enzyme and the digest was run on a gel at 100V for 20 mins to confirm that the guides had the gRNA insert by looking at size. Guides with correct size were then sent to Eurofins for sequencing.

2.2.4 Lentivirus packaging and transduction

The first step of transduction was to package lentiviruses withCas9 and our guides using HEK293T cells. This was done by first seeding the cells in a 6-well plate, $7*10^5$ cells per well and grown overnight. Next day, the transfection mix were made and added to the cells. The transfection mix was made of a plasmid mix (1000ng transfer plasmid, 750 ng psPAX2, 500 ng pMD2.G and 250 µl OptiMEM) and Lipofectamine mix (10 µl Lipofectamine 2000 and 250 µl OptiMEM).

Plasmid mix was added to the lipofectamine mix and incubated at room temperature for 20 min. 500 µl transfection mix were then added directly to the cells dropwise. Cells were then incubated at 37 ℃ 5% CO2 overnight. Media were changed next day. The day after that the supernatant were removed and filtered with a 0,45 µm filter and ready for use.

The IPEC-J2 cells were transduced with Cas9 blast viruses via reverse transduction. The virus was added to a 6 well plate in dilutions 1:10, 1:25, 1:50, 1:75 and 1:100 + one well without lentivirus. Then $5*10^4$ IPEC-J2 cells were addted to each of the wells and incubated at 37 °C, 5% CO2 for 2-3 days. Polybrene was also in the media, at a concentration of 10 ug/mL After the incubation, the media was replaced with 10 µg/ml blasticidin for selection of transformed cells.

Once all cells in the control well were dead, surviving cells would be made into single cells clones. The cells exposed to the lowest concentration of lentivirus were used, since the higher concentration of virus makes it more likely that the Cas9 gets integrated more than once into the genome. This increases the possibility of the Cas9 being integrated into important genes and affect the cell adversely. In addition, too high expression of Cas9 is toxic to cells (Ye et al., 2020) To do that, cells were suspended in media and diluted to 5 cells/ml and 100 µl were added to each well in a 96 well plate, meaning each well had 0,5 cells, and then incubated.

A different method was also used. All wells in a 96 well plate were filled with 100 µl media. $1x10^4$ cells in 100 µl of media were then seeded in a corner of a 96 well plate. Half of that well would then be moved to the well beneath it, mixed and then repeated. This would be done for all of the column, then the same procedure would be done the lateral direction. This way the number of cells in the well would decrease the further it was from the first well, eventually getting one cell per well.

Half of the media used when incubating cells were conditioned media. The conditioned media was media harvested from untransformed cells before they were split and filtered using 0,22 µm filter.

After single cell cultures were established, they were grown to near confluence, trypsinized and moved to a larger container. When a single cell culture had expanded enough to fill two flasks, it was frozen following the ATCC cryopreservation protocol.

2.4 Western blot

After transduced IPEC-J2 cell lines were frozen for long term storage, some were left to grow and be used for western blot and qPCR of mRNA. When they were ready to be harvested, cells for protein were trypsinized, washed in PBS 2x and then frozen as a pellet in a -18°C freezer. When all the cell lines had been frozen, with long-term storage and protein, the pellets were thawed and lysed using 2ml RIPA buffer (10mM Triss, HCl pH 8.0, 1mM EDTA, 1% triton x100, 0,1 % sodium deoxycholate, 0,1 % SDS, 140 mM and 1mM PMSF) per pellet.

The lysate was kept at constant agitation for 30 min at 4°C. Tubes were then spun at 16,000 xg for 20 min in 4°C and then moved to ice. 20 ul of lysate from each cell line were then used in a Bradford assay to determine protein concentration.

Western blot was done two times.

20 ul of lysate were taken from each sample and got an equal volume of 2x Laemmli sample buffer (Bio-rad). 20 µg of protein from each sample were then loaded on to a mini-protean TGX stain free gel (Bio-Rad) and run on 200 V for 20 min in running buffer (25mM Tris, 190mM glycine, 1%SDS, pH 8,3).

After electrophoresis, the gel was removed from the cassette and photographed on a Chemidoc MP imaging System (Application: Stain-free gel, Gel Activation time: 45 sec, Image activation: Optimal automatic exposure). The gel was then put back in running buffer.

Separated proteins were transferred to a nylon membrane via trans-blot Turbo transfer pack (Bio-Rad). The gel was placed on top of blot membrane in a western sandwich, had bubbles removed by rolling a pipette tip over the top of the sandwich and was run for 7 min with 1,3 A. Once the transfer was finished, both the gel and blot were put in a container with deionized water. They were imaged on Chemidoc MP imaging System (settings for gel were Application: Stain-free gel, Gel Activation time: None, Image activation: Optimal automatic exposure, Exposure time: same as the exposure time for pre-transfer image. Settings for blot were: Application:stain-free Blot and Optimal automatic exposure).

Membrane was then moved to container with Tris Buffered Saline with Tween 20 (TBST. 20mM Triss, 150mM NaCl and 0,1 % Tween 20). Then the blot was moved to a container with 1% Casein in TBS for 5 min with gentle agitation. After that primary antibody, Cas9 antibody (7A9-3A3), was added with 1:500 dilution and the blot was incubated at room temperature for 1 hour with gentle agitation.

After incubation with Cas9 antibody, the blot was rinsed 5 times for 5 minutes with TBST. Then the blot was incubated at room temperature with casein 1% blocking buffer with horseradish peroxidase-conjugated diluted at 1/10000 for 1 hour. The blot was then rinsed 6 times for 5 min with TBST.

After antibody incubation, another picture was taken of the blot on the ChemiDoc MP imaging system set to stain-free Blot and optimal automatic exposure. Then the blot was incubated for 5 min in 1:1 mix of Clarity Max Substrate (Bio-rad) components. The blot was then imaged on imaging system with chemiluminecent blot and optimal automatic exposure.

Normalization of bands was done by first dividing total protein in the lane and dividing it with the lane with the highest protein level. This gives a normalization factor for all the lanes. The Cas9 band was then normalized by multiplying it with the normalization factor and then standardized by dividing all samples on chosen sample.

2.5 Quantitative Real Time PCR

RNA was extracted using RNeasy plus mini kit (Qiagen). Extracted RNA was converted to cDNA using Bio-Rad iScript cDNA synthesis kit. For the qPCR, SSoAdvanced Universal SYBR Green Supermix (Bio-Rad) was used as reagents. The program for the qPCR was as follows: 95°C for 30 seconds, then a 40 cycle with 95°C denaturation and 50℃ annealing/extending and ends with 65°C for 5 seconds and then an increase of 0,5°C per second to 95°C.

qPCR was done three times, first to determine the best primers to use using WT and two samples, second and third time to test the remaining samples with the chosen primers. The reference genes used were ACTB, RPL4 and TBP. Primers for the Cas9 were Cas9, Cas9_1 and Cas9_2. In the first run, all primers were used. For the second and third, only ACTB, RPL4 and Cas9_1 primers were used. For sequences, see table 1. The positive control was lentivirus Cas9-Blast v2, the same plasmid used for the transduction of the cells. The negative control was a No Template Control (NTC).

The setup for the first qPCR was one standard curve series for each primer, a positive control for each Cas9 primer, a negative control for each primer and test for all primers. Each sample were run with all the primers. Replicates were three for each, except negative control which was only one for each. The standards were 3 wells and had a dilution of 1:2 for each well.

The second and third qPCR had a standard curve series, negative control, positive control and Cas9 test with the Cas9_1 primer. All had a triple replicates. The standard curve series were 4 wells and had a dilution of 1:10 in each.

The results were standardized by following the $2^{-\Delta\Delta CT}$ (Livak) Method (Livak & Schmittgen, 2001). The cell line P18B7 was used as a calibrator.

2.6 Validation of Cas9 gene knockout

The DNeasy blood and tissue kit (Quiagen) was used to extract DNA from 8 different transduced cells with both Cas9 and gRNA. Platinum™ II Hot-Start PCR Master Mix (2X) (Thermo-Fisher) was used on extracted DNA with primers (prMUC4_seq_g1g2_Forw and prMUC4_seq_g1g2_Rev) which targeted the KO site in MUC4. Program for PCR was 2min 94°C, then 40 cycles of 94°C for 15 sec, 51°C for 15 sec and 68°C for 7 sec. After cycles the samples were cooled down to 4°C and held indefinity. Ran the amplified product on 1% agarose gel at 50 V for 30 minutes. Used QIAquick Gel Extraction Kit (Qiagen) to purify the product.

Samples were labeled and sent to Eurofins for sequencing. Primes used were MUC4 sequence forward and reverse. See table 1 for sequence.

The site DECODR (https://decodr.org/) were used to deconvolute the sequence results. For all samples reverse primer generated sequences were used, with WT reverse sequence as reference.

2.7 *E. coli* adhesion assay experiment

To test ETEC resistance in IPEC-J2 cells, an adhesion assay was performed to determine how well *E. coli strains* with and without F4 gene would bind to stock IPEC-J2 cells.

2.7.1 Strains

E. coli strains used were 1598 and 1462. Strain 1598 expresses STb toxin and F4 fimbriae. Strain 1426 does not express of either STb toxin or F4. Bacteria was grown in LB overnight at 37°C.

2.7.2 Assay

All wells in a 96 well plate were seeded with $2x10^4$ cells IPEC-J2 cells. 100 µl bacteria was added to 5 ml of LB media and grown overnight. Concentration of bacteria determined by measuring opacity at 600 nm wavelength and calculated concentration by multiplying the OD with $1*10^9$. 1, 10 and 100 Multiplicity of infection (MOI) of bacteria were added to wells with IPEC-J2. Amount of *E. coli* cells added for MOI 1 was $2x10^4$ cells, $2x10^5$ for MOI 10 and $2x10^6$ for MOI 100. For positive control, Triton-100 0,25% and 0,5% were added to the positive control wells and nothing in the negative control wells. Plate was centrifuged at 1000g for 2 minutes to get the bacteria on the cells. Incubated at 37°C for 30 min. Then, washed the cells in PBS to remove bacteria and added Propidium iodide (PI) to the cells and examined them using a fluorescent microscope. One flask of unmodified IPEC-J2 cells were also stained with PI and checked in the microscope.

Positive control was Triton-100 0,25% and 0,05%. Negative control was without bacteria. Cells were spun 1000xg for 2 minutes. And then incubated at 37°C for 30 min.

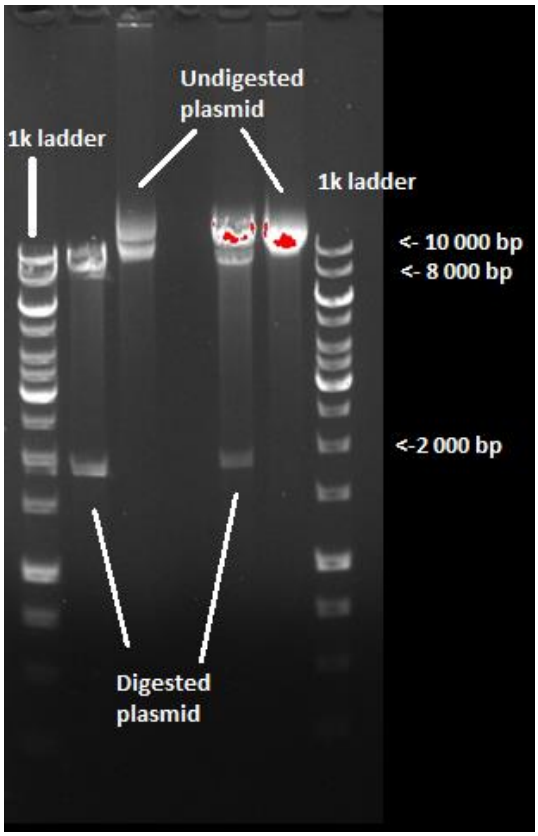
Table 1:Primers and their sequences used during the master. All the Cas9 primers are different primers targeting Cas9. The Cas9_2 primer uses the Cas9_1 reverse primer.

Primers	**Sequence**
prMUC4_g1_fwd	CACCGTGATTGTTCCTAGGCAGGCA
prMUC4_g1_rev	AAACTGCCTGCCTAGGAACAATCAC
prMUC4_g2_fwd	CACCGGTCCCTCCCAAGCCGAAGAA
prMUC4_g2_rev	AAACTTCTTCGGCTTGGGAGGGACC
prMUC13_g1_fwd	CACCGGTTGTAGTTGCTTCGATGGT
prMUC13_g1_rev	AAACACCATCGAAGCAACTACAACC
prMUC13_g2_fwd	CACCGTCGATGGTAGGAGTTGTAGT
prMUC13_g2_rev	AAACACTACAACTCCTACCATCGAC
prFUT1_g1_fw	CACCGCAAAGACGGGGGCTTCGTAT
prFUT1_g1_rev	AAACATACGAAGCCCCCGTCTTTGC
prFUT1_g2_fwd	CACCGGGTGTACCGGCGTGCCCGCC
prFUT1_g2_rev	AAACGGCGGGCACGCCGGTACACCC
ACTB Forward	CACGCCATCCTGCGTCTGGA
ACTB Reverse	AGCACCGTGTTGGCGTAGAG
RPL4 Forward	CAAGAGTAACTACAACCTTC
RPL4 Reverse	GAACTCTACGATGAATCTTC
TBP Forward	AACAGTTCAGTAGTTATGAGCCAGA
TBP Reverse	AGATGTTCTCAAACGCTTCG
Cas9 Forward	CCGAAGAGGTCGTGAAGAAG
Cas9 Reverse	GCCTTATCCAGTTCGCTCAG
Cas9_1 Forward	TGCGCGAGATCAACAACTAC
Cas9_1 Reverse	TCGCTTTCCAGCTTAGGGTA
Cas9_2 Forward	GTGCGCGAGATCAACAACTA
prMUC4_seq_g1g2_Forw	CTCTTTACCCCAGGTGGCAG
prMUC4_seq_g1g2_Rev	CCTCAGCACCCAACACTGAT

3 Results

3.1 Making gRNA

To make lentiGuide-puro clones with individual gRNA inserts, the first step was to digest the lentiGuide-Puro plasmid.



Figur 2 Gel of lentigudide puro digested with BsmBI.

The digestion of lentiGuide-Puro was done in parallel and the best digested sample was used for all subsequent steps. Figure 1 shows a gel with the digested plasmid and undigested plasmid as a control. The digested plasmid backbone was then purified by separating the products on a gel by electrophoresis and cutting out the correct band. Figure 1 shows a gel with the 5 µl of digested plasmid used to check if the plasmid had been cut. Digested plasmid backbone is 8298 bp, fragment is 1885 bp and undigested plasmid is 10183 bp. Both digested plasmids have two fragments of correct size. For the purification of the cut plasmid, the remaining digested plasmids were added to two wells and extracted by cutting out the cut plasmid band from the gel and run with a gel extraction kit (see 2.2.3 for details).

Annealed gRNA oligonucleotides were ligated into the lentiGuide-Puro backbone and then transformed into competent Stbl3 *E. coli* and grown on LB media containing Ampicillin. 10 colonies for each guide (MUC4 g1, MUC4 g2, FUT1 g1, FUT1 g2, MUC13 g1 and MUC13 g2) were streaked out on LB agar petri dishes. Cloning efficiency of the clones are showed in table 2.

Table 2 Cloning efficiency of lentiguides.

Target gene	Guide	Cloning efficiency Cfu/µg DNA
MUC4	G1	$9{,}3*10^6$
	G2	$8{,}73*10^6$
MUC13	G1	$1{,}8*10^7$
	G2	$1{,}53*10^7$
FUT1	G1	$1{,}77*10^7$
	G2	$2{,}7*10^7$

Colony cracking was done on plasmid containing colonies to test for the presence of inserts in lentiGuide-Puro. The colony cracking was done on each of the 10 individual colonies for each guide. Muc4 guides were done first and had been purified when the cracking of MUC13 and Fut1 guides were to be done, so a good Muc4 guide was used as a control. The result of the cracking of the MUC13 colonies are shown in figure 2.

After the best plasmid candidates were chosen, inserts from each clone were amplified with PCR. In addtion some of the resulting plasmid was digested with SalI-HF restriction enzyme and run on a gel. The Muc4 guide result is shown in figure 3, where all of the guides were correct size. The guides which were the correct size were sent for sequencing.

Sequencing results showed that for all genes there was at least one proper gRNA sequence insert. Figure 4 shows an alignment of the lentiguides with MUC4 gRNA insert compared to the MUC4 gRNA sequence template.

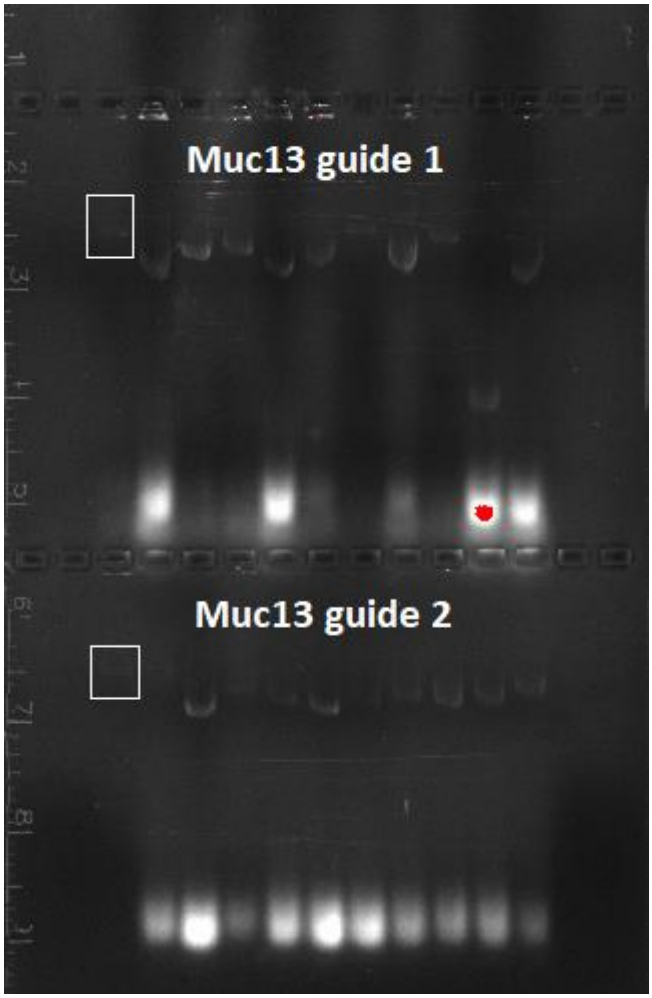


Figure 2: Result of Muc13 colony cracking. The white squares contains a good, if faint, MUC4 guide as a reference for the other guides. Guides closest resembeling the muc4 guide was chosen for further experiments.

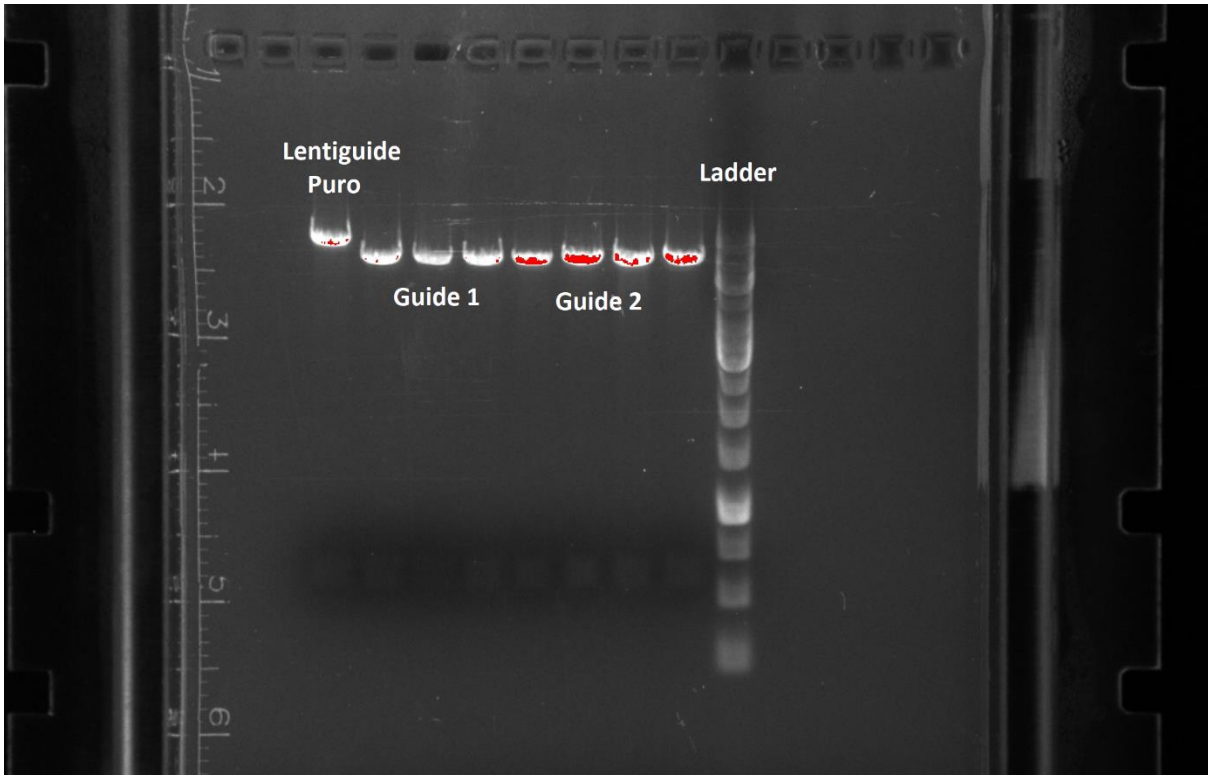


Figure 3. Gel electrophoresis of SalI-HF digestion of lentiGuide-Puro plasmids. LentiGuide-Puro at the left is unmodified, while the guides were digested by BsmBI and then got an insert, resulting in a netto loss of 1860 basepairs. All the guides were the correct size. Ladder used is 1k plus.

Figure 4: Alignment of MUC4 guides with MUC4 guide template. Both guides being compared are the same guide, just from different colony

3.2 Transduction with Cas9 lenti and single cell cloning

Because the concentration of lentivirus was unknown, transductions were done in a 6-well plate with one control well without lentivirus and 5 wells with dilutions from 1:10 to 1:100. The amount of surviving cells after selection began decrease higher than 1:50 dilution, meaning they had a low amount of lentivirus present to transduce the cells. Cells with lower dilution than 1:50 had a high survivability after selection, with similar number of cells compared to the control well before selection.

When creating single cell culture, the blasticidin resistant cells obtained with the lowest concentration of virus was chosen. This is because the Cas9 gene can be inserted multiple times in the same cell. Since the insertion is random, a greater number of insertions increases the probability of the Cas9 gene being inserted into important parts of the genome and having an adverse effect on the cell. So by choosing to use the cells exposed to

the lowest concentration of lentiCas9-blast, that had managed to survive, the less likely it would be that Cas9 was inserted more than once into the cell.

Creating single cell cultures was difficult because cells usually did not survive when alone. The number of monocultures per 96 well plate were 0-2. The rest were either empty, the cells died or had more than one surviving cell. The first method of seeding 0,5 cells per proved to be too inefficient, with many 96 plates having no growing cells. A different method was then used to make single cell cultures , based on dilution series instead. It had more success in creating surviving cells, even though the total number of wells that could hold single cells was lower. Cell lines were named after the plate and position on the plate they were grown on. For example, sample P3E9 is from plate 3, position E9 on the 96 well plate. Samples with double digit plate number are done with dilution series method, plates with single digit is done with 0,5 cell/well method.

Conditioned media became used halfway through monoculture production to increase the success chance. It did not show any notable improvement over regular media but was used for all monocultures after.

3.3 Characterization of Cas9 cell lines

After IPEC-J2 cells were transduced with Cas9-blast and split into cell lines, they were tested to measure expression levels of mRNA and protein of Cas9.

3.3.1 Results of qPCR on mRNA in transduced cells

qPCR on Cas9 mRNA was done three times. The first qPCR experiment was done to determine which Cas9 primer would be used and evaluate what housekeeping gene should be used as a reference gene. There were three different Cas9 primer sets and three housekeeping primer sets. The reason for using so many different primers for Cas9 was to find the most efficient version. The second and third qPCR were done on all cell line samples to test mRNA expression using one Cas9 primer and two housekeeping genes per cell line.

The result of the first qPCR showed that the Cas9 primers were not significantly different from each other, but the Cas9_1 primer was chosen because it was somewhat higher than the other two primers. For reference genes, RPL4 and ACTB were chosen to be used because TBP gene gave a

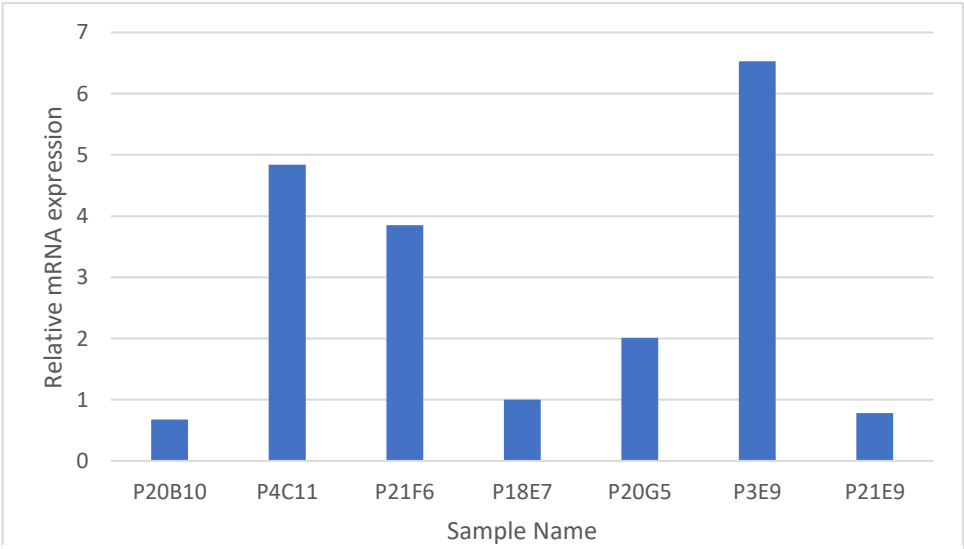


Figure 6 Relative expression of Cas9 mRNA for different transduced samples. All samples are standarized on P18E7, so the numbers are relative to P18E7. For example, P20G5 have double fold the expression of Cas9 mRNA as P18E7.

lot of variation between samples.

The standard curves were unsuccessful in the first two qPCR experiments. The first qPCR had a standard curve for each primer, but none of the standard curves showed no linear progression. The second qPCR had one standard curve for the Cas9 primer, but there was no amplification in any of the wells. The third qPCR had standard curve with three replicates of the Cas9 primer. This showed a more linear trend. Figure 7 shows the standard curve of the Cas9 primer made by average of all the replicate wells. This shows an PCR efficiency of 112% and a R^2 of 0,98.

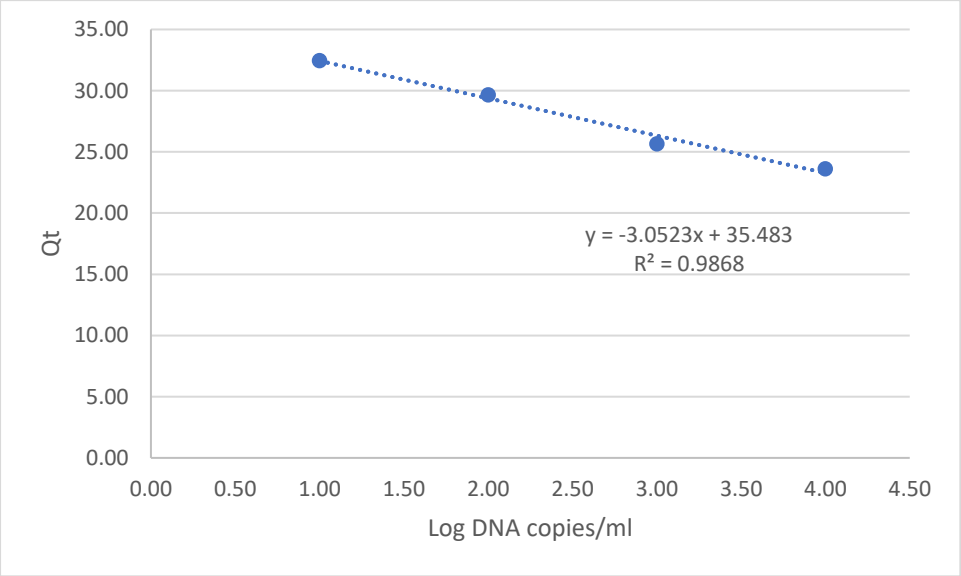


Figure 5: Standard curve of Cas9_1 primer from the third qPCR.

The positive control did not get any amplification during the last two qPCR. The last qPCR had amplification in two of its three NTC. With a Cr of 36,5 and 32,07, it was higher than all of the other samples, exept for some standard curve wells.

Figure 5 shows the relative expression of Cas9 in the samples. Only the samples used for Cas9 knockout are shown here. For the rest of the samples, see the appendix. As shown, mRNA expression varied a lot between the different cell lines.

3.3.2 Western results

Two separate western blots were done, the first on all cell lines to select transduced IPEC-J2 cell lines to be used for KO experiment and the second only on the cell lines to be used for CRISPR KO. Both followed the procedure listed in materials, except for a few differences:

Because the imaging of chemiluminescent blot took longer than expected, one blot was incubated in Clarity max Substrate for 20 min instead of 5 min. In the second western blot, the blot was incubated in primary blocker overnight at 4°C instead of 1 hour in room temperature. The second western blot was also incubated in clarity substrate for 10 min instead of 5 min.

Cas9 protein expression between the different samples were standardized using Image Lab 6.1 (Bio-Rad) software. Figure 8 shows the relative expression of Cas9 in the different cell

lines in relation to P18B7.

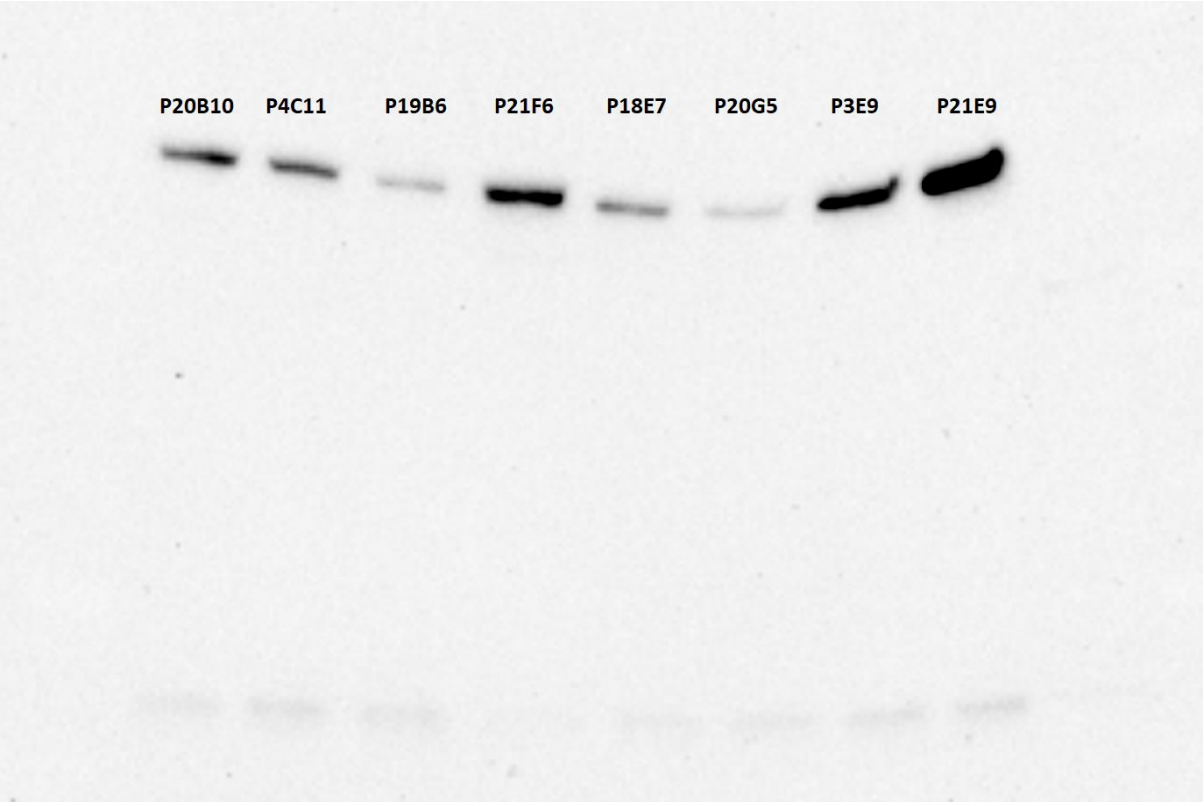


Figure 8:Chemiluminesent western blot of Cas9 proteins from 8 different samples. All samples were 20 ng, so variation in intensity is because of differing amount of Cas9 in the samples.

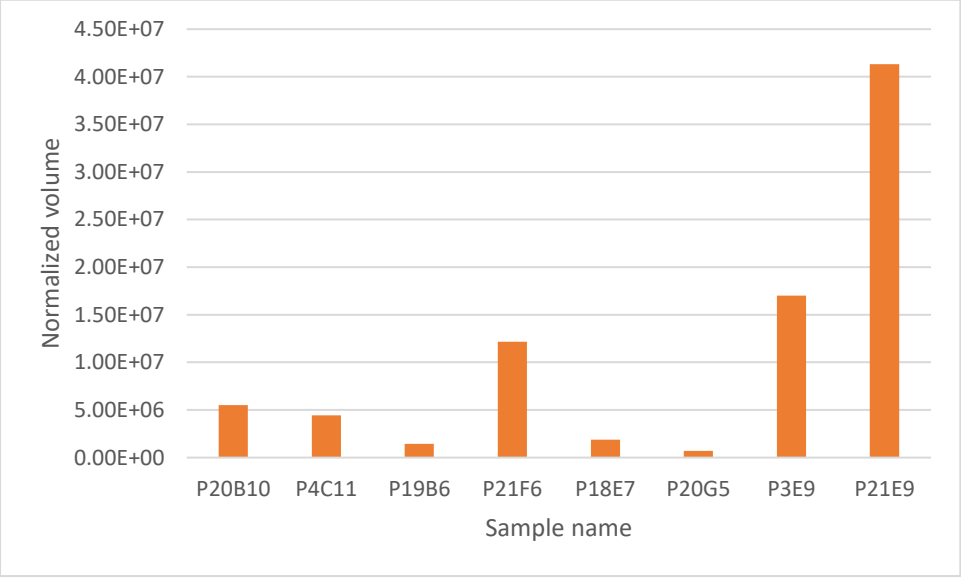


Figure 9: Standardized Cas9 protein levels in different samples.

Protein expression also varies between the different samples. The highest expressing cell line is P21E9, with over twice as much Cas9 protein expression as the second highest sample.

3.3.3 Correlation of mRNA and protein

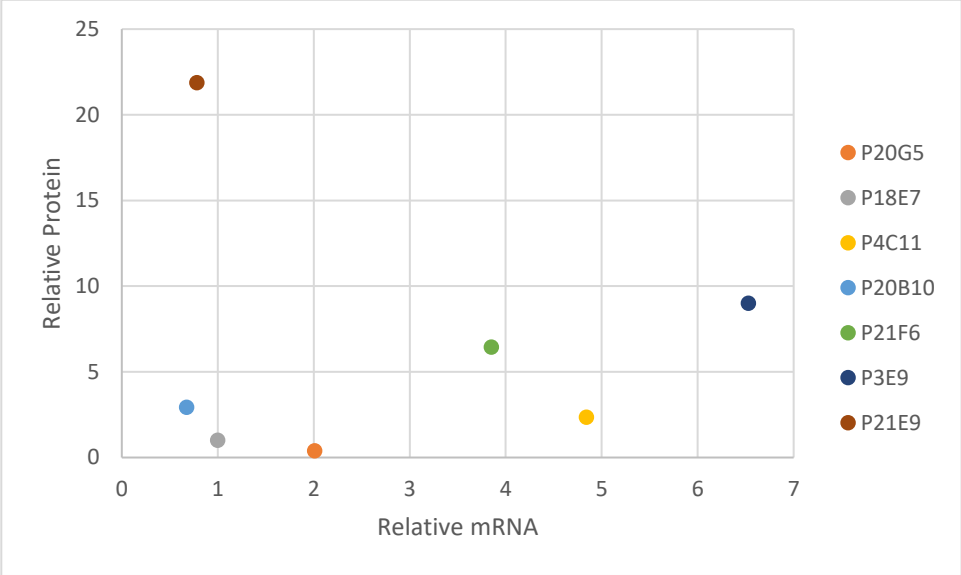


Figure 10: Point graph showing the relationship between Cas9 mRNA and Cas9 protein in different monocultures of transformed IPEC-J2 cells. The values are all relative to P18E7.

There is a correlation between mRNA and protein of Cas9 if the outlier P21E9 sample is excluded. A trend line were plotted without P21E9, showing a R^2 of 0,56.

3.4 Validation of Cas9 activity

3.4.1 Knockout edit efficiency in various transduced IPEC-J2 strains

Sanger sequencing was done to determine sequence of IPEC-J2 KO cells. After sequencing, the result was deconvoluted using Decodr (https://decodr.org/). Both forward and reverse were checked. The forward results had generally poor quality, so for the graphs, only the reverse sanger sequencing has been used. P3E3 did not get good enough sanger sequencing reverse results to be able to be deconvoluted, so it is ignored.

The knockout efficiency of MUC4 gene varied between the different cell lines, as illustrated in figure 11. One cell line has no knockouts, while three cell lines has a high knockout at above 80% knockout efficiency. These are good results, since only one guide was used.

The indels range from 1 to 100, with the most prevalent sort of indels being small, under 10 bp. The majority of indels are frameshift, so they will cause knockout of genes. The spread of indels are shown in table 3. Figure 13 shows a graphic representation of the spread of indels in the P4C11 sample as an illustration of indel distribution.

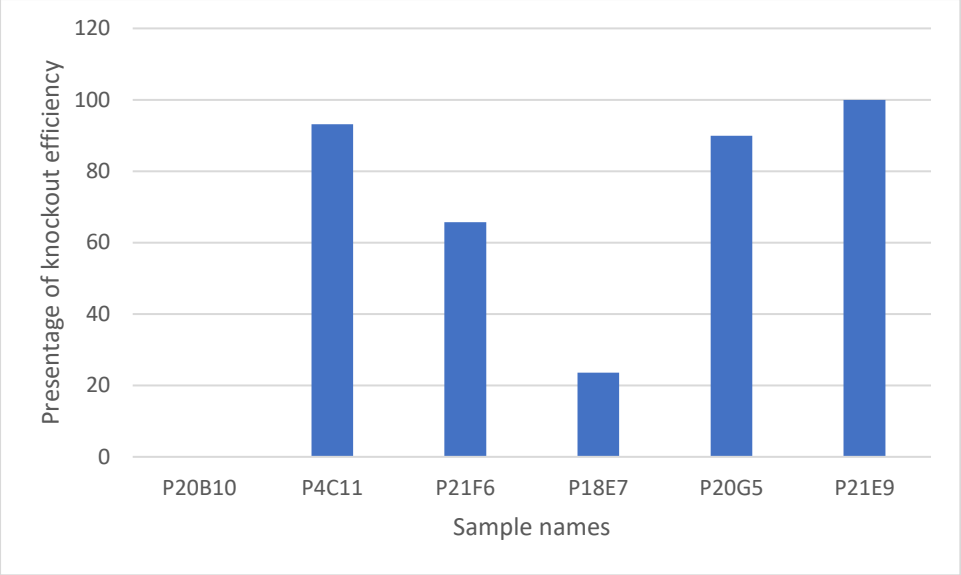


Figure 11: Graph for Knockout efficiency of MUC4

Comparing knockout efficiency of MUC4 with Cas9 protein expressed in the cells shows little correlation between them. This is clearly illustrated in figure 12. Sample P21E9 has both high knockout efficiency and protein expression, but P20G5 does have high knockout efficiency and the lowest protein expression of the samples.

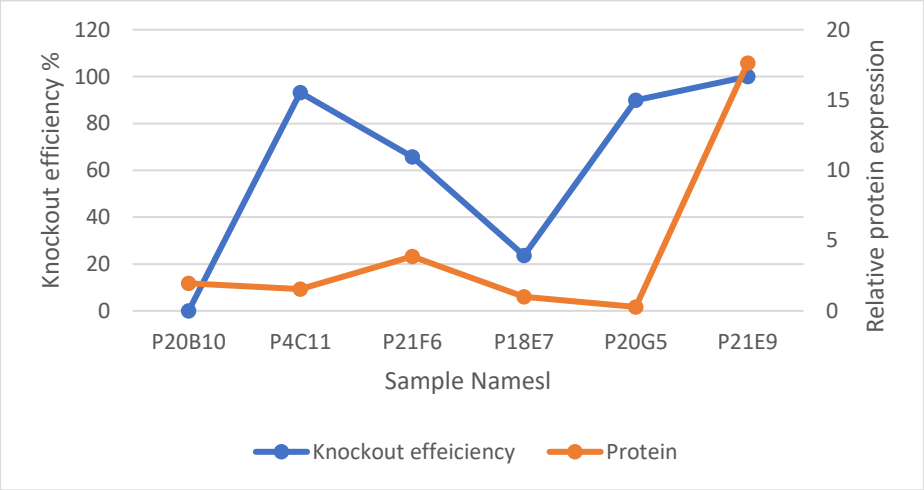


Figure 6: Knockout efficiency in MUC4 gene and Cas9 protein expression levels in different cell lines.

Table 3. Percentage of different indel types in MUC4 gene for all KO samples

Sample name	Percentage	Indel
P21E9	84.5	-4
	10.1	1
	5.5	-8
P4C11	33.5	-5
	17.1	11
	13.4	-4
	8.7	-37

	7.3	-26
	6.9	-13
	6.8	0
	6.3	-39
P20F6	35.9	-4
	34.3	0
	8.2	1
	8.1	-100
	7.9	-37
	5.6	-5
P20B10	100	0
P20G5	43	-4
	23.2	1
	13.1	-86
	10.7	-85
	10.1	0
P18E7	76.4	0
	15.2	-4
	8.3	-9

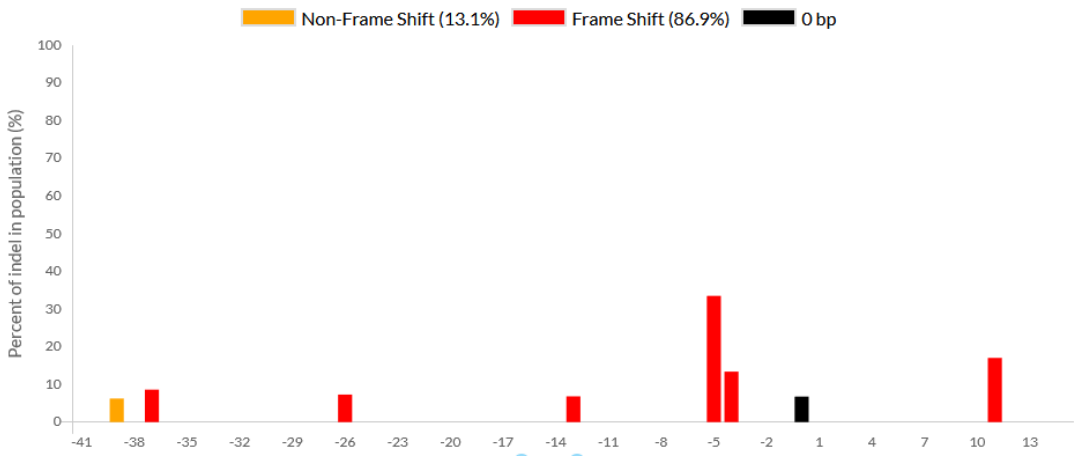


Figure 13: The size and amount of indel mutations in P4C11 knockout sample.

3.7 *E. coli* binding experiment

The result of the *E. coli* binding experiment was that there were no significant differences between the different ETEC isolate strains or the negative control. The positive control did have more stained cells, but not significantly more. The unmodified cells had the same number of PIstained cells compared to the ones exposed to toxin producing *E. coli*. In Figure 11 there is not a large difference between cells exposed to *E. coli* and cells not exposed to *E. coli*

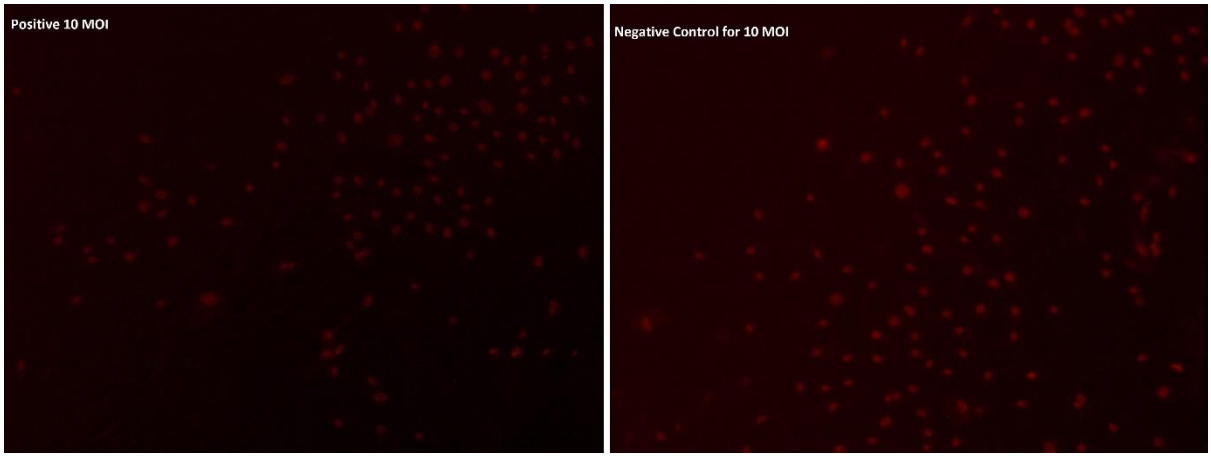


Figure 14: Two pictures of fluorecent BI stained E. coli bacteria. The picture on the left is from a 10 MOI e. coli expressing STB and F4. The image of the right is the negative control of the same series. There is no significant difference between the amount of permeated cells between the different images.

4 Discussion

The primary goal of the master thesis was to evaluate the viability of lentiviral delivery of Cas9 gene delivery in IPEC-J2 cells and then determining the resulting editing efficiency in *E. coli* resistance genes. The secondary goal war to test the magnitude of resistance of generated mutant cells to *E. coli.*

This section will be divided into five parts. The first will discuss the transduction of IPEC-J2 cells and generating cell lines, the second the characterization of Cas9 expressing IPEC-J2 cells, the third dicusses MUC4 KO, fourth the test of an *E. coli* phenotype to evaluate *E. coli* resistance in IPEC-J2 and finally what other steps would be done in the master if more time had been available.

4.1 Transduction of IPEC-J2 and creating cell lines

No non-targeting control were used in the transductions for Knockout. A non targeting control would have a gRNA sequence that did not target anywhere in the genome. It could show if Cas9 made cuts without a target. Off target effects were not checked in this study, but If they had been, the lack of a non targeting control would make It difficult to verify if off target effects were caused by gRNA or Cas9.

Non targeting controls are available commercially, so one could have been ordered. A non targeting control could also have been made. To do so, a lentiguide-puro backbone could have gotten inserted a guide sequence with no targets in the IPEC-J2 cells, in the same manner as the other guide sequences were inserted.

4.2 Characterization of transduced cell lines

The expression of mRNA and protein varied wildly between cell lines, with no clear correlation between them. This could be because of many factors. Examples are bad proteins in western blot, non linear amplification in qPCR and technical errors.

The western blot had to be redone multiple times in order to be done successfully. Almost every time, the extracted protein was thawed and used. Thawing and freezing proteins has an adverse affect on the quality of protein. If that had been the case for this experiment, it would have shown up as a smear on the western blot. This is not the case, so degrading Cas9 proteins are not a factor for the western blot results.

The final dilution curve of the qPCR gives an PCR efficiency of 112%. The ideal PCR efficiency is 100%, since it means the primer target is doubled for each cycle. Possible causes for too high efficiency are inhibitors in the template and amplification of unspesific products. The more dilute the sample, the less the inhibitors can affect the PCR. This increases the Ct of samples with high dilutions compared to samples with low dilution since the inhibitors are more concentrated and reduces Ct. The result is an increased PCR efficiency. Testing for amplification of unspesific products can be done by running the PCR result on a gel. If there were unspesific amplification, there would be more than one band or a smear, showing the unspesific amplicons. It should be noted that the PCR efficiency are only calculated for the last qPCR since the two former did not succeed on making any standard curve. This can be a sign of poor execution of all the qPCR experiments.

Another factor is where the Cas9 gene were incorporated. The insertion is random and is almost certainly different between all the cell lines. Depending on where the Cas9 was incorporated, it could affect not just the expression of Cas9, but also other genes. Unfortunately, time did not permit checking where Cas9 were incorporated.

Another possible cause for the poor correlation between mRNA and protein is technical error in all the different steps. Inexperience increases the risk of making mistakes that could affect the results. In the case of the qPCR, what could have been done is do the mRNA extraction again and run the qPCR on that one to validate the results.

It is also factors within the cells themselves that affect the mRNA and protein correlation. It is known that that there are a lot of variables affecting the mRNA translation into protein, such that there isn't always a clear correlation between them.

The bad positive control in the qPCR can be because the positive control was not diluted before being used in the qPCR, even though the concentration of the positive control was much higher than the other samples. As such, the signal would be so high initially to exceed the maximum Ct threshold, making the qPCR machine not able to see any change in DNA present.

4.3 Knockout Efficiency of MUC4 gene

Three cell lines show a great degree of knockout. However, since the IPEC-J2 cells are diploid, that does not necessary mean a complete knockout of MUC4 in all of the celsl. There are two copies of the MUC4 gene, one on each chromosome. For the gene to no longer be expressed, there must be an indel in each of the chromosomes. So even if editing efficiency of a cell line were 90%, the actual efficiency would be lower since the sanger sequencing is not of both chromosomes in the same cell, but many different cells.

Since the CRISPR and gRNA are both incorporated into the genome of the IPEC-J2 cells, it will continually express both the Cas9 protein and gRNA. The Cas9 protein was shown expressed with a Western blot. The gRNA was produced since the transduced cells did have a knockout. As long as there would be a sequence for the gRNA to attach to, the Cas9 would bind to it and cut. So over time, it could be that the gene editing efficiency in the transduced IPEC-J2 cells would increase. However, continued expression of Cas9/gRNA could also increase off target effects. If the off target effects negatively affected the transduced cells, cells with non-functioning Cas9 or gRNA could have a fitness advantage over cells with working Cas9/gRNA. This could then lead to lower editing efficiency in the population.

The small sample size makes the correlation between mRNA, protein and knockout efficiency less reliable. If more samples had been done, it could be that there would have been shown a stronger correlation, since it would reduce the effect of individual variation between the cells.

4.4 Adhesive assay of ETEC on IPEC-J2 cells

Propidium Iodide cannot enter cells with intact cell membrane. In the experiment, the assumption was that the cells not attached with ETEC would not show any or very low

fluorescence while cells attacked by ETEC would have high fluorescence. But there could be other causes for cells to be permeable aside from toxins from ETEC. First is that some cells could be dead before the experiment started. The cells were seeded not long before the experiment. Some dead cells could have been transferred with the live cells and skewed the results. Another possibility is that some cells died during the experiment. Before incubation with bacteria, the cells were run on 1000xg for 2 min to get the *E. coli* spun down onto the cells. This speed could have killed cells, affecting the result.

The bacteria used were also in PBS during the experiment. This could adversely affect their ability to deliver toxin.

Another issue is that the results were checked manually using a fluorescent microscope. It is possible to see significant differences, but not small ones. A 10 % difference in number of dead cells would be difficult to see with the naked eye. It might have been better if the results were checked more objectively using a machine, like a flow cytometer. Here the results would be less susceptible to subjective bias. On the other hand if the difference is not large enough to be seen in a microscope, it might not be large enough to be of significant value.

4.5 Possibilities if more time had been available

Since the Cas9-blast and lentiGuide both have incorporated randomly in the genome, it could be useful to know where they were incorporated. One way of finding out is to sequence the whole genome and run BLAST to find the Cas9 gene and gRNA. Another, less resource intensive method is inverse PCR. The first step of inverse PCR is to select an restriction enzyme. It must create sticky ends and should have a recogniction sequence of 6 or more bases while not targeting within the known DNA. In this case within the inserted sequences: Cas9, S. pyogenes sgRNA cassette, Blasticidin and pyromycin resistance genes, and the gRNA sequence. It would then be used to chop up the DNA of a transduced cell line creating many long strands of DNA with sticky ends. Using ligase, we can ligate the ends together and make plasmids. Since we have a known in the plasmid, we can design primers that go away from the known sequence, amplify the plasmid using PCR and sequence the amplicon. Knowing the flanking sequence of the insert, we can then place it in the genome. (Ochman et al., 1988)

Another thing I would have done is to test the mRNA again. Both to get the standard curve right and to get another result of Cas9 mRNA levels. Another result of Cas9 mRNA levels could give insight into whether or not the original qPCR gave the correct numbers. If the numbers matched, it was done right and we have an added replicate. If it did not, we would have a number to contrast it with. It might also correlate better with the proteins.

5. Conclusion

A CRISPR screen needs a reliable way of delivering CRISPR/Cas9 components into cells at a large scale in order to determine a correlation between a gene and a phenotype. This study has shown that lentivirus is a viable option for delivery of both Cas9 gene and guide RNA

into IPEC-J2 cells. The creation of multiple IPEC-J2 cell lines expressing Cas9 at different levels also enables easier testing of knockouts with only delivery of guide RNA. However, further testing is needed to determine off target effects and gene insert sites, in addition to long term effects of Cas9 expression in the cells.

E. coli adhesion experiment shows that propidium iodide is a poor choice to test permeability of IPEC-J2 as a readout for resistance to ETEC infection.

6. Citations

Cuellar, T. L., Callow, M., Lau, T., Watanabe, C., Bainer, R., Costa, M. & Haley, B. (2018). Application of CRISPR for Pooled, Vector-based Functional Genomic Screening in Mammalian Cell Lines. In Appasani, K. (ed.) *Genome Editing and Engineering: From TALENs, ZFNs and CRISPRs to Molecular Surgery*, pp. 209-222. Cambridge: Cambridge University Press.

Dubreuil, J. D. (2017). Enterotoxigenic Escherichia coli targeting intestinal epithelial tight junctions: An effective way to alter the barrier integrity. *Microb Pathog*, 113: 129-134. doi: 10.1016/j.micpath.2017.10.037.

Fairbrother, J. M., Nadeau, É. & Gyles, C. L. (2005). Escherichia coli in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Animal Health Research Reviews*, 6 (1): 17-39. doi: 10.1079/AHR2005105.

Food & Agriculture Organization of the United, N. (1997). *FAOSTAT statistical database*: [Rome] : FAO, c1997-. Available at: https://search.library.wisc.edu/catalog/999890171702121.

Ishino, Y., Krupovic, M. & Forterre, P. (2018). History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *Journal of bacteriology*, 200 (7): e00580-17. doi: 10.1128/JB.00580-17.

Jin, L. Z. & Zhao, X. (2000). Intestinal receptors for adhesive fimbriae of enterotoxigenic Escherichia coli (ETEC) K88 in swine--a review. *Appl Microbiol Biotechnol*, 54 (3): 311-8. doi: 10.1007/s002530000404.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, 25 (4): 402-408. doi: https://doi.org/10.1006/meth.2001.1262.

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, 13 (11): 722-736. doi: 10.1038/nrmicro3569.

Ochman, H., Gerber, A. S. & Hartl, D. L. (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics*, 120 (3): 621-623.

Peng, Q. L., Ren, J., Yan, X. M., Huang, X., Tang, H., Wang, Y. Z., Zhang, B. & Huang, L. S. (2007). The g.243A>G mutation in intron 17 of MUC4 is significantly associated with susceptibility/resistance to ETEC F4ab/ac infection in pigs. *Anim Genet*, 38 (4): 397-400. doi: 10.1111/j.1365-2052.2007.01608.x.

Poirier, J. T. (2017). Chapter Four - CRISPR Libraries and Screening. In Torres-Ruiz, R. & Rodriguez-Perales, S. (eds) vol. 152 *Progress in Molecular Biology and Translational Science*, pp. 69-82: Academic Press.

Sears, C. L. & Kaper, J. B. (1996). Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. *Microbiol Rev*, 60 (1): 167-215. doi: 10.1128/mr.60.1.167-215.1996.

Sinha, R., Sahoo, N. R., Shrivastava, K., Kumar, P., Qureshi, S., De, U. K., Kumar, A., Kumar, G. V. P. P. S. R. & Bhushan, B. (2019). Resistance to ETEC F4/F18–mediated piglet diarrhoea: opening the gene black box. *Tropical Animal Health and Production*, 51 (6): 1307-1320. doi: 10.1007/s11250-019-01934-x.

Spencer, N. Y. (2019). *Overview: What is CRISPR screening?* Available at: https://eu.idtdna.com/pages/education/decoded/article/overview-what-is-crispr-screening (accessed: 27.5).

Vergauwen, H. (2015). The IPEC-J2 Cell Line. In Verhoeckx, K., Cotter, P., López-Expósito, I., Kleiveland, C., Lea, T., Mackie, A., Requena, T., Swiatecka, D. & Wichers, H. (eds) *The Impact of Food Bioactives on Health: in vitro and ex vivo models*, pp. 125-134. Cham: Springer International Publishing.

Wang, H., Russa, M. L. & Qi, L. S. (2016). CRISPR/Cas9 in Genome Editing and Beyond. *Annual Review of Biochemistry*, 85 (1): 227-264. doi: 10.1146/annurev-biochem-060815-014607.

Wu, X., Kriz, A. J. & Sharp, P. A. (2014). Target specificity of the CRISPR-Cas9 system. *Quant Biol*, 2 (2): 59-70. doi: 10.1007/s40484-014-0030-x.

Ye, S., Enghiad, B., Zhao, H. & Takano, E. (2020). Fine-tuning the regulation of Cas9 expression levels for efficient CRISPR-Cas9 mediated recombination in Streptomyces. *Journal of industrial microbiology & biotechnology*, 47 (4-5): 413-423. doi: 10.1007/s10295-020-02277-5.

Yip, B. H. (2020). Recent Advances in CRISPR/Cas9 Delivery Strategies. *Biomolecules*, 10 (6): 839. doi: 10.3390/biom10060839.

You, L., Tong, R., Li, M., Liu, Y., Xue, J. & Lu, Y. (2019). Advancements and Obstacles of CRISPR-Cas9 Technology in Translational Research. *Mol Ther Methods Clin Dev*, 13: 359-370. doi: 10.1016/j.omtm.2019.02.008.

7. Appendix

Table 4. Results from first western blot. It consisted of two blots, with one sample, P20G5, shared between them to standardize.

Western blot 1				
Cas9 Bands	Adj. Volume	Normalized protein	Cas9 to total lane ratio	Relative to P20G5
P18D5	139460	1248370.84	0.0159	1.60
P18D7	484960	4341100.85	0.0430	4.33
P18E7	86800	776986.87	0.0095	0.95
P19B2	89760	803483.20	0.0215	2.17
P19B6	179600	1607682.52	0.0155	1.57
P20G4	0	0.00	0.0000	0.00
P20G5	113260	1013842.55	0.0099	1.00
P21B10	744980	6668659.91	0.0522	5.26
Western blot 1				
P21D8	0	0	0.0000	0.00
P21E9	83120	83120.00	0.0037	0.10
P21F6	909680	909680.00	0.0661	1.73
P3E9	189680	189680.00	0.0145	0.38
P3G9	223860	223860.00	0.0201	0.53
P4C11	3430180	3430180.00	0.2883	7.53
P6D6	44580	44580.00	0.0066	0.17
P20G5	624180	624180.00	0.0383	1.00

Table 5. mRNA expressed for all of the cell lines.

Cell line	**Relative mRNA expression**
P20G5	2.01
P19B2	0.11
P18E7	1
P4C11	4.84
P20G4	0.80
P21B10	0.68
P21F6	3.85
P18D5	2.87
P3E9	6.53
P21D8	4.11
P21E9	0.78
P3G9	9.33
P6D6	15.33

Table 6 Protein expressed in different cell lines from second western blot

Sample	Norm. Vol.	Relative
P20B10	5.52E+06	2.92
P4C11	4.44E+06	2.35
P19B6	1.45E+06	0.77
P21F6	1.22E+07	6.45
P18E7	1.89E+06	1
P20G5	7.15E+05	0.38
P3E9	17001484	9.00
P21E9	41315616	21.88

Table 8. Percentage of different indel types in MUC4 gene for all KO samples

Sample name	Presentage	Indel
P21E9	84.5	-4
	10.1	1
	5.5	-8
P4C11	33.5	-5
	17.1	11
	13.4	-4
	8.7	-37
	7.3	-26
	6.9	-13
	6.8	0
	6.3	-39
P20F6	35.9	-4
	34.3	0
	8.2	1
	8.1	-100
	7.9	-37
	5.6	-5
P20B10	100	0
P20G5	43	-4
	23.2	1
	13.1	-86
	10.7	-85
	10.1	0

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway