

**Facultad de Ciencias Experimentales**  
**Grado Biotecnología**  
**Trabajo Fin de Grado**

**Implications Of Hippo Signalling Pathway Transcription Factors  
Involved In Idiopathic Pulmonary Fibrosis**

**Conditional CRISPR Cas9 A549 Cell Line Generation Employing TALEN Technology**

**Pablo Ovejero Romero**

**Under the supervision of Hani N. Alsafadi**

**Wagner Lab Lung Bioengineering and Regeneration**

“The adventure of life is to learn. The purpose of life is to grow. The nature of life is to change. The challenge of life is to overcome. The essence of life is to care. The secret of life is to dare. The spice of life is to befriend. The beauty of life is to give.”

***William Arthur Ward***

## **Acknowledgements**

I would like to thank whole LBR lab for kindly hosting me during this period. I want to show special gratitude to Darcy and Hani for supporting me and helping me in the beginnings in my developing as a competitive scientist.

I want also to express my sincere gratitude to all people that have accompany and supported me during this captivating university-journey. However, I do not have enough gratitude words to thank my mother and my father for giving me unconditionally love no matter what, because, love is, in my not so humble opinion, our most inexhaustible source of vitality and strength.

## Table Of Content

<b>1. ABSTRACT</b>	3
<b>2. INTRODUCTION</b>	4
2.1. Idiopathic Pulmonary Fibrosis (IPF)	4
2.2. Alveolar Epithelium	4
2.3. IPF Pathogenesis	5
2.4. Hippo Signalling Pathway	6
2.5. Genome Editing Technologies	7
2.5.1. TALEN Technology	7
2.5.2. CRISPR-Cas	8
2.6. Plasmids	9
<b>3. OBJECTIVES</b>	10
<b>4. MATERIALS AND METHODS</b>	11
4.1. Cell culture	11
4.2. Subculturing of the cells	11
4.3. iCRISPR A549 Cell Line Generation	11
4.3.1. Plasmid Isolation and Purification	11
4.3.2. Optimization of Antibiotic Selection	13
4.3.3. Transfection: Electroporation	14
4.3.4. Transfection: Lipofection	15
<b>5. RESULTS</b>	17
<b>6. DISCUSSION AND CONCLUSIONS</b>	29
<b>7. ANNEXES</b>	31
<b>8. BIBLIOGRAPHY</b>	34

# Implications Of Hippo Signalling Pathway Transcription Factors Involved In Idiopathic Pulmonary Fibrosis

---

## Conditional CRISPR Cas9 A549 Cell Line Generation Employing TALEN Technology

### 1. ABSTRACT

Idiopathic Pulmonary Fibrosis (IPF) is a devastating non-curable disease with increasing incidence rate. IPF manifests itself by uncontrolled tissue scarring and extracellular matrix (ECM) deposition leading to lung failure with no possibility of recovery once diagnosed and no available treatment can stop or reverse disease progression. Therefore, there is a real need for finding effective therapies for IPF. Several developmental signalling pathways are re-activated during fibrosis; one of which is the Hippo signalling, known for its role in differentiation, proliferation, and organ size. Hippo signalling functions by controlling the activity of the co-transcriptional activators, YAP and TAZ. The activity of YAP and TAZ was found to be increased during fibrosis by mediating release of pro-fibrotic molecules leading to other ECM production in the alveolar interstitium. YAP/TAZ exert their effect by binding to different transcription factors (TFs), some of which are known, most are not. To this end, a screening study was performed to identify YAP/TAZ interacting TFs. The aim of this project is to elucidate the effect of screened YAP/TAZ TFs binding partners on IPF through the generation of an A549 inducible CRISPR (iCRISPR) Cas9 cell line using the genome editing TALEN technology. The generation of this cell line is based on insertion of Cas9 and M2rtTA cassettes in both loci of the AAVS1 gene; a process in which each step needs to be optimized. The plasmid isolation and purification from bacterial stabs was successful with a remarkable efficiency. Antibiotics are used for selecting cells with successful transfection. However, the antibiotic concentration and treatment duration are crucial to appropriately select the desired cells. Thus, we tested a range of concentrations of the 2 antibiotics, G418 and Puromycin over the course of 12 days. We found that the G418 and puromycin perform best at concentrations of 1000 µg/ml and 2 µg/ml respectively based on morphology analysis and cytotoxicity LDH assay. Thus, we have optimized the steps necessary to perform plasmid transfection using electroporation or lipofection. The generation of this multi-target knockout tool will aid in the identification of pathologic TF complex interaction downstream of the Hippo pathway in IPF.

**Key words:** Idiopathic Pulmonary Fibrosis (IPF), YAP/TAZ, ECM, Hippo Signalling Pathway, Transcription Factors, inducible CRISPR-Cas9, TALEN.

## 2. INTRODUCTION

### 2.1. Idiopathic Pulmonary Fibrosis (IPF)

Idiopathic Pulmonary Fibrosis (IPF) refers to a whole subgroup of illnesses within the major group of Interstitial Lung Diseases (ILDs). ILDs are a type of respiratory diseases characterized by affecting the lung interstitium which correspond to the anatomic space between the alveolar epithelium and the endothelial epithelium. IPF is characterised by excessive deposition of extracellular matrix components in the distal lung leading to progressive injury to the lung epithelium. In IPF the fibrosis phenomenon leads to a marred gas exchange and progressive dyspnea that ends up in lung failure compromising patient's life. (1) IPF has a survival ratio between 3 to 5 years and there is no available treatment that can stop or reverse disease phenotype. The exact causes for IPF are unknown but molecular mechanisms driving disease progression are being studied. (2) The main driving theory of disease onset is the repetitive microinjuries that happen in the distal lung epithelium that lead to activation of repair mechanisms which are then deranged due to the repetitive nature of the injuries. (3)

The incidence and prevalence of IPF remains unknown due to the lack of diagnosis tools and codes for the proper identification of the molecular pathologies. Despite this fact, we can no longer assume that it is a rare disease since it is exponentially increasing worldwide.

### 2.2. Alveolar Epithelium

As IPF affects the gas exchange area of the lungs we have to focus on the architecture of alveolar epithelium. Alveolar cells known as Type I (AT1s) or Type II (AT2s) inhabit, in around a 96% and 4% respectively, the surface of this epithelium. AT2s are cuboidal cells that play an important role in injury repair. Evidence has shown that these cells have a great plasticity and can be considered as multipotent cells capable of renewing alveolar epithelium. (4, 5) Thanks to a transdifferentiation process in which lamellar bodies are lost the nucleus relocated and the cuboidal shape modified AT2 cells can promote to AT1 cells. (Figure 1) (6)

IPF development involves severe injury in alveolar epithelium causing an enormous AT1 and AT2 cell death. Repairing systems are affected during the course of the illness and AT2 cells lose the capacity to regenerate and repair lung epithelium.

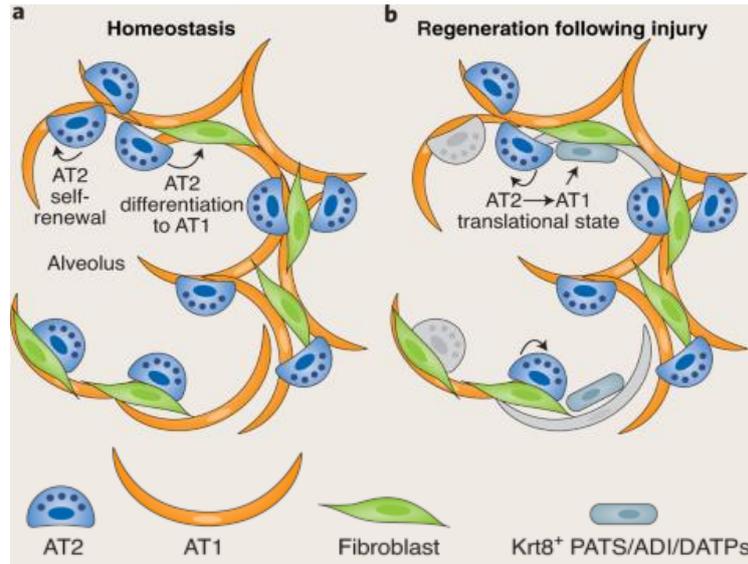


Figure 1. Regeneration of alveolar epithelium by AT2 cells

### 2.3. IPF Pathogenesis

The etiology of IPF remains unknown. However, it has been studied that there are some factors that may increase the susceptibility of developing IPF. Some gene mutations affecting surfactant proteins or mucins aggravate the fibrosis phenomenon. Environmental exposure to cigarettes, vapours and contamination also plays an important role as they can alter the lung architecture and promote the differentiation of fibroblast into myofibroblasts thus, being more likely to induce the expression of pro-fibrotic and pro-inflammatory cytokines.

Several studies draw the molecular pipeline in which subsequent events occur, finally reaching IPF. Persistent injury affecting the lung epithelium ends up in a defective wound healing cascade in addition to the induction of pro-fibrotic chemokines by epithelial cells. Fibroblasts and fibrocytes are, then, recruited to the wound focus and they are fostered to differentiate into myofibroblasts which are the main culprits of ECM and collagen accumulation. (7)

It has been shown that some routes are activated during fibrotic progression (8). Hippo signalling pathway was shown to mediate release of pro-fibrotic cytokines in lung epithelium. Hippo functions by regulating its main effectors, the Yes Associated protein (YAP) and the transcriptional coregulator (TAZ), which are transcriptional co-factors that bind to other transcription factors to exert their function. YAP and TAZ were found to be dysregulated in the fibrotic alveolar epithelium inducing gene expression that

leads to an excess of Extracellular Matrix (ECM) deposition contributing to IPF pathogenesis.(9)

## 2.4. Hippo Signalling Pathway

The Hippo signalling Pathway is a potent tissue mass and homeostasis regulator as it controls cell proliferation, differentiation, and survival. Their main effectors YAP/TAZ are targeted for phosphorylation involving a signalling cascade in which other proteins play their role such as the adaptor protein Salvador and the families of kinases MST and LATS. When phosphorylated these proteins are translocated into the cytoplasm where they are marked for degradation. However, loss of function of the phosphorylation cascade leads in nuclear translocation of YAP and TAZ. Inside the nucleus YAP/TAZ cannot bind directly to the DNA, thus they bind to other several types of Transcription factors (TFs) or even with TFs complexes to exert their function and induce gene expression. (Figure 2) (9, 10) The most studied YAP/TAZ interacting TF are the TAE domain Transcription factors (TEAD1-4) which are most known for their role in cancer and tumour development. Several other TFs have been explored as interactors of YAP or TAZ such as p53, Tbx5, Runx TFs, and Smad TFs. However, most of these interactions are studies within the context of cancer and the exact TF interactions within the lung epithelial cells is not known yet.

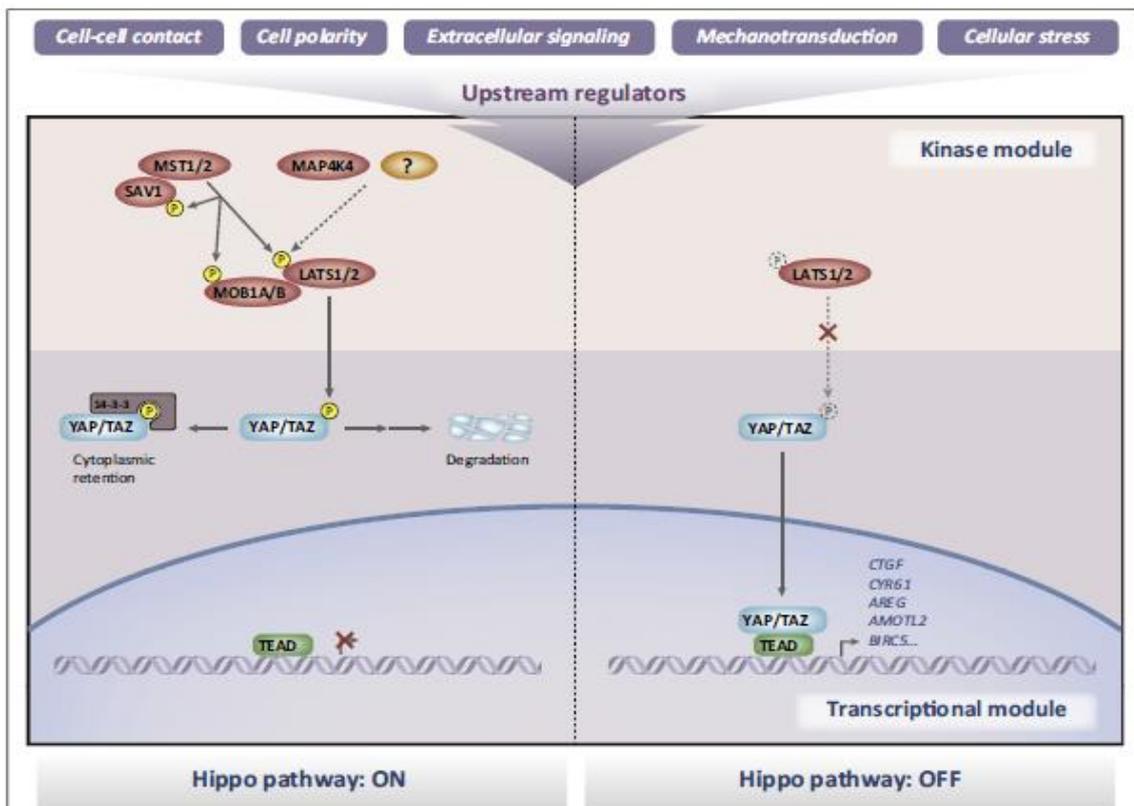


Figure 2. Hippo Signalling Pathway

As YAP/TAZ were found to be active in IPF, a proof-of-concept study was performed at the Wagner lab to explore the effect of pharmacologically targeting YAP/TAZ on fibrotic progression. Verteporfin, an FDA approved drug for photodynamic therapy, was found to inhibit the YAP/TAZ-TEAD interaction.(11) Thus, it was tested to treat fibrotic injury in mice induced by bleomycin treatment. Verteporfin was able to improve survival and attenuated the fibrotic injury (unpublished data). However, Verteporfin is highly toxic and it is not clear which interactions of YAP/TAZ it inhibits, it could potentially inhibit repair promoting interactions. To this end, a chromatin Immunoprecipitation study was performed in order to identify relevant TFs for fibrotic progression. Such studies produce a large number of hits and not all could be relevant. Thus, experimental validation is required to examine the TF identified. Genome editing to target these TFs alongside YAP and TAZ may aid in identifying their role in fibrotic development. However, a stable genomic tool is required to perform such multi-targeted editing.

## **2.5. Genome Editing Technologies**

### **2.5.1. TALEN Technology**

TALEN (Transcription activator-like effector nuclease) is a genome editing technology capable of performing a DSB in a desired point of DNA. If no DNA template is added the DSB will be repaired by NHEJ (Non-Homologous End Joining) introducing indel mutations. Adding a DNA template containing homology arms will end up in the insertion of it by HDR (Homology Directed Repair).

Looking into the molecular basis of this genome editing tool we can see that it comprises 2 major components: the TALE proteins and the Fok1 nuclease domain to which the TALE proteins are fused. (Figure 3) (12)

TALE proteins are a type of proteins that bind DNA specifically which were first found in plant-pathogenic bacteria of the genus of *Xanthomonas*. (13) These proteins consist of several Repeat Variable Domains (RVD) which have the same sequence of 34 amino acids except for the ones in positions 12 and 13. These residues will determine the nucleotide-specificity to each RVD. The TALEN protein library creation is simple and efficient because each nucleotide is complementary with one protein. Thus, a 4-protein library can be assembled. This raises a great advantage in comparison with other genome editing tools like ZFN (Zinc Finger Nucleases) in which each zinc domain binds to 3 nucleotides.

As said before, TALE proteins are fused to a nuclease domain that will be in charge of performing the DSB. However, dimerization of the FokI domain is required to exert its function so 2 TALEN arms are needed.

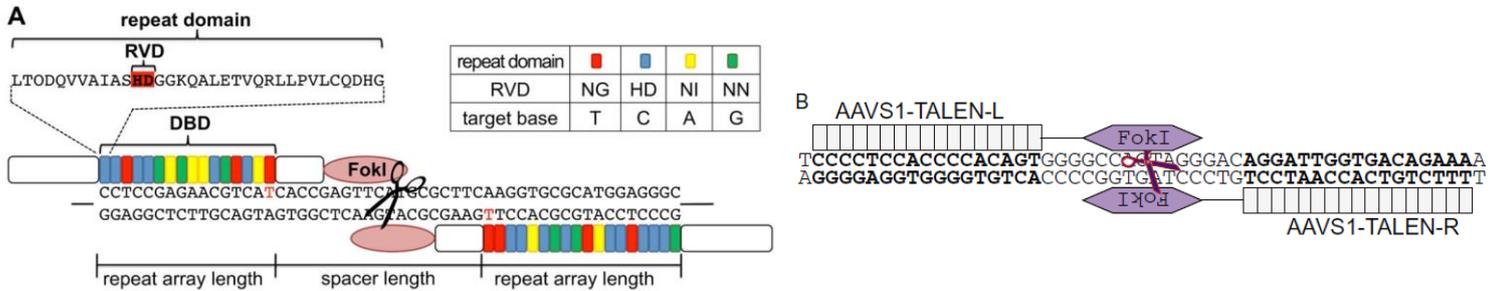


Figure 3. TALEN Technology

### 2.5.2. CRISPR-Cas

CRISPR-Cas is an adaptative immune system found in some prokaryotic organisms. Thanks to a small type of RNAs they are able to be protected among the invasion of foreign nucleic acids. (14) (15, 16)

CRISPR-Cas was adapted to function as an RNA-programable technology capable of targeting a desired sequence in vitro with high accuracy and in a simple way.

To exert its function as a genome editing tool it is necessary to have the components it comprises:

- Short-guide RNA (sgRNA) is the duplex formed between the crRNA and the trans-activating crRNA (tracrRNA) that will direct the DNA cleavage.
- Cas9 is a type of endonuclease guided to the target sequence or protospacer that employs the HNH and RuvC-like domains to perform a DSB. It is strictly necessary for Cas9 recognition that a PAM sequence (Protospacer Adjacent Motif) is present at the 3'-end of the protospacer.

Our target locus for Cas9 insertion is AAVS1 which is a model locus within the PPP1R12C gene that allows strong expression of transgenes.

## 2.6. Plasmids

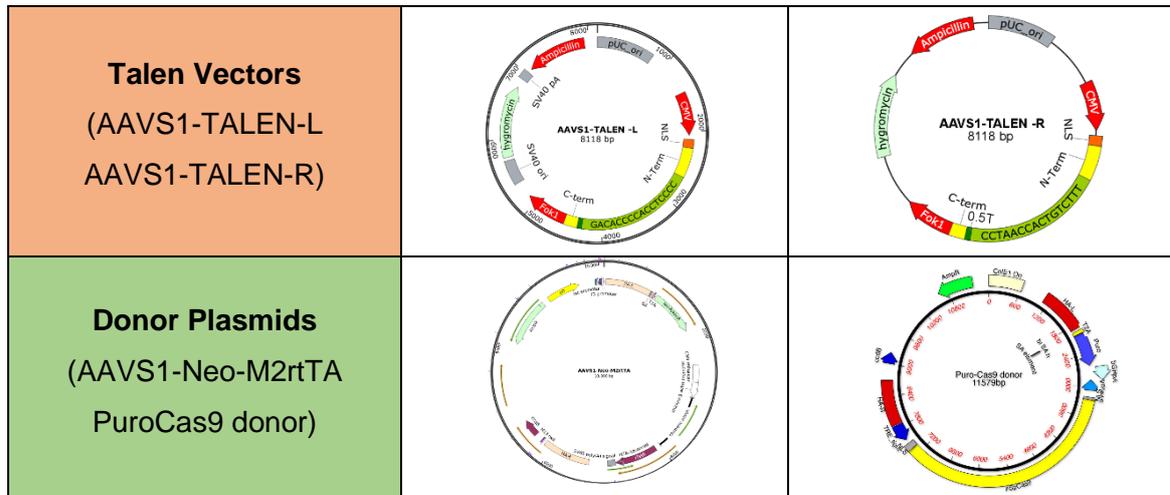


Figure 4. TALEN Vectors and Donor Plasmids Genomic Maps. (Bigger pictures can be found in Annex 1)

For the generation of the iCRISPR A549 cell line, 4 plasmids need to be delivered. The first pair contain the TALEN technology (Figure 4) (17) with which is possible to produce a Double Strand Break (DSB) in AAVS1 locus thanks to the base complementarity of the arms. They also contain an ampicillin resistance cassette for cloning purposes. Both plasmids have a molecular weight of 8,118 bp.

On the other hand, the Donor Plasmids (Figure 4) (18, 19) carry the Cas9 protein and the doxycycline Tet-On induction system. These plasmids serve as DNA templates to repair the DSB by Homology Directed Repair (HDR). Cas9 protein under the Tet Response Element (TRE) will be introduced in one of the strands and the reverse Tet Repressor in the other strand. Neo-M2rtTA and PuroCas9 donor plasmids weigh 10.080 bp and 11.579 bp respectively.

### 3. OBJECTIVES

The main objective of this project is the development of a doxycycline inducible CRISPR A549 stable cell line. This tool will allow us to perform multitarget Knockouts with the purpose of underlining any pathologic TFs complexes involved in IPF.

Specific objectives were set at the start of the project since there was time limitation incompatible with completing the whole proposal:

- Establish a clear and well-structured work pipeline.
- Optimize and set a stable A549 cell line.
- Isolate and Purify TALEN and Donor plasmids from bacterial stabs.
  - o Compare Plasmid growth at 30°C and 37°C.
- Optimize the Antibiotic Selection protocol with G418 and Puromycin.

## 4. MATERIALS AND METHODS

### 4.1. Cell culture

The cells selected for the generation of an iCRISPR cell line were A549 which are epithelial, adherent lung cells that came from a carcinomatous tissue. A549 cells were provided from a frozen cryovial by Hani N. Alsafadi, LBR Wargner Lab (ATCC, CCL-185). DMEM/F12 complete growth medium (Gibco™, 11554546) supplemented with 10% Fetal Bovine Serum (FBS) (Gibco™, 11573397) and 1% Penicillin/Streptomycin (Pen/Strep 10.000 U/ml) (Thermo Fisher Scientific, 15140122) was employed to culture the cells. The culture conditions applied in the incubator were 5% (CO<sub>2</sub>) and 37°C. The cells were cultured in T-75 flasks (Sarstedt, 83.3911.002).

### 4.2. Subculturing of the cells

Phosphate Buffered Saline 1X (PBS) (Nordic Biolabs, sh30256.01) was used to perform the washing step prior to the enzymatic cell disaggregation with Trypsin 0.05% (1X) 1:250/L in HBSS EDTA (Fisher Scientific GTF AB, 10011472). When the incubation time, around 5 minutes, has finished, DMEM/F12 was added to inhibit Trypsin and cells were seeded at a ratio of 1/10 in a new flask.

### 4.3. iCRISPR A549 Cell Line Generation

To fulfil this objective, several steps were required:

#### 4.3.1. Plasmid Isolation and Purification

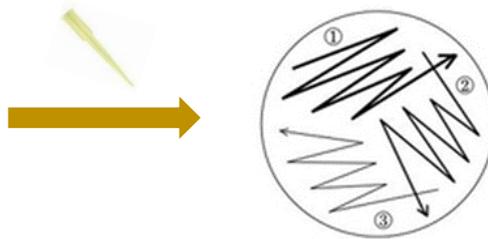
The TALEN and Donor plasmids were provided as bacterial *Escherichia coli* (*E. coli*) *Stbl3* stabs. AAVS1-TALEN-L, AAVS1-TALEN-R and Puro-Cas9 donor were a gift from Danwei Huangfu (Addgene plasmid #59025) (Addgene plasmid #59026) (Addgene plasmid #58409). AAVS1-Neo-M2rtTA was a gift from Rudolf Jaenisch (Addgene plasmid #60843).

To perform the bacteria streaking for single colony isolation Luria-Bertani broth and agar plates were prepared. 25g/L liquid media was employed made with LB broth base (Thermo Fisher Scientific, 12780052) and supplemented, after autoclaving, with Ampicillin Sodium Salt (Fisher Scientific, 10193433) at a final concentration of 100µg/ml. For the LB agar plates, 25g/L of LB broth was mixed with 20g/L of Select Agar, Powder (Thermo Fisher Scientific, 30391023). Once autoclaved it, Ampicillin at 100µg/ml was added when the mix was at ~50°C and the Petri Dishes 60x15 mm (Sarstedt, 821194500) were poured.

The streaking was carried out by touching with a sterile pipette tip the punctured area of the bacterial stab and drawing each streak changing tips in between.



AAVS1-TALEN-L Bacterial Stab



Addgene Streaking Diagram ([See Addgene Streaking Protocol](#))

Figure 5. Streaking method of *stb13* stabs

Plates containing the different plasmids were incubated at 37°C overnight. Neo-M2rtTA plasmid is highly recommended to be incubated at 30°C as it is prone to recombine. (12). In the first isolation assay all plasmids were grown at 37°C and then, a second assay was performed with Neo-M2rtTA and Puro-cas9 donor plasmids at 30°C. Thus, we were able to compare the different growth conditions. Once isolated single colonies, they were selected and, with a sterile tip, they were transferred into 5ml liquid LB cultures with ampicillin to grow overnight.

The following day, plasmids were ready to be purified but, so as to storage them for long-term, glycerol stocks were made by adding 600µl of culture to 400µl of 50% Glycerol in 2ml cryovials. The 50% Glycerol can be made by diluting 100% Glycerol in  $dH_2O$ .

The purification of the plasmids was carried by using the commercial kit Thermo Scientific™ GeneJET Plasmid Miniprep Kit (Fisher Scientific, 10319699). As all the vectors were high-copy plasmids 5ml of culture were employed as recommended. To analyse the efficiency of the purification process, DNA concentration and purity were measured employing Thermo Scientific NanoDrop 1000.

Finally, to examine the DNA integrity and calculate the molecular weight of the plasmids a 1% agarose electrophoresis (100V) was carried using as ladder 1KB Bio Rad molecular Ruler (Bio Rad, #1708355). Moreover, a second electrophoresis was carried but with a percentage of agarose of 0,5% to separate better the upper bands.

A final electrophoresis was performed to compare the plasmids grown at 37°C and the plasmids grown at 30°C.

The molecular weight of the plasmids were calculated by means of plotting a calibration curve with the logarithm of the molecular weight of the markers and its Retention factors ( $RF = Z/Z_0$ ; being  $Z_0$  the distance from the origin to the solvent

front and Z the distance from the origin to each respective band). Once drawn the tendency line, plasmids RF were interpolated within the linear equation given.

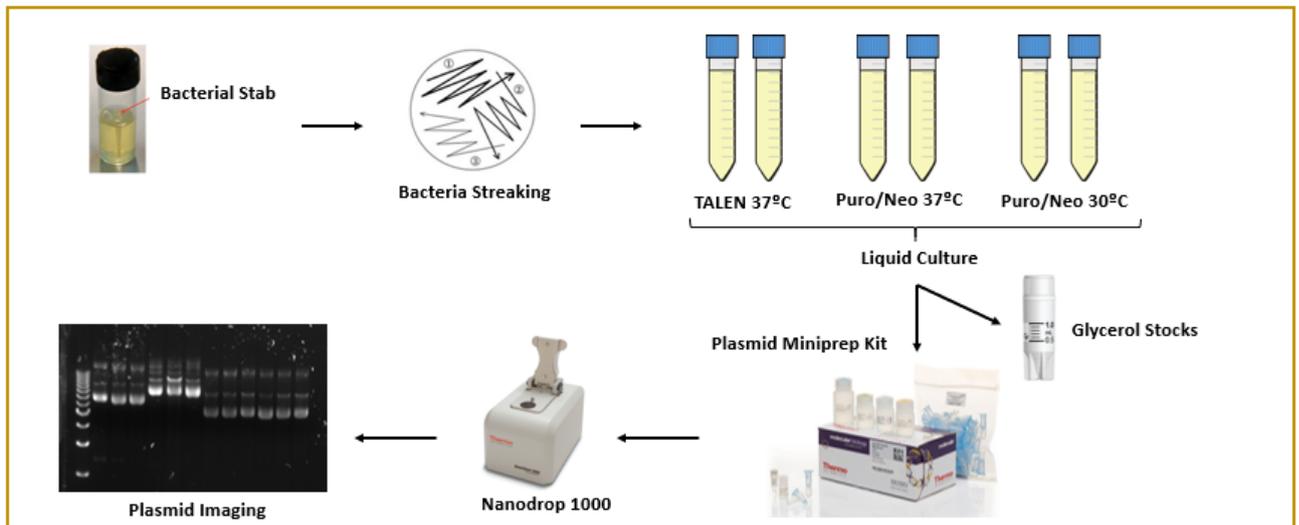


Figure 6. Plasmid Isolation/Purification workflow scheme

#### 4.3.2. Optimization of Antibiotic Selection

Prior to performing the transfection, the antibiotic selection was honed so only the desired cells would survive. The antibiotics employed for cell selection were Gibco™ Geneticin™ Selective Antibiotic (G418 Sulfate), Powder (Fisher Scientific, 10463982) and Puromycin dihydrochloride from *Streptomyces alboniger* (Sigma-Aldrich, P8833-10MG). Several concentrations and a control with no antibiotics were tested in 12-well plates for a course of 12 days:

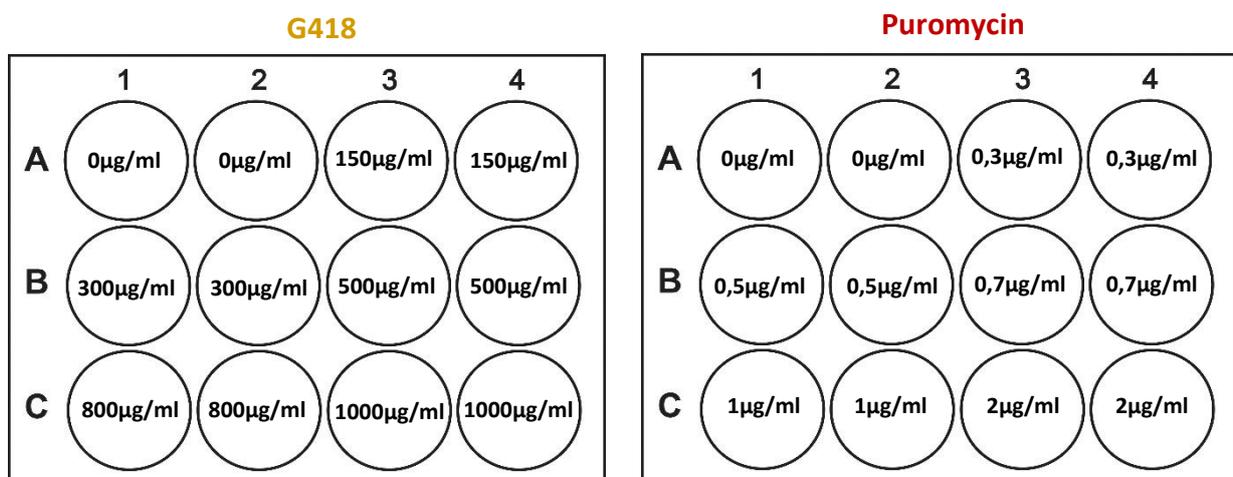


Figure 7. G418 and Puromycin tested Concentrations in 12-well plates.

Media was changed every 2-3 days and pictures were daily taken to keep a qualitative track of cell death. To perform a quantitative assay of cell viability a cytotoxicity test was carried on measuring the release of Lactate Dehydrogenase enzyme (LDH) with a commercial kit (Sigma-Aldrich, 11644793001). It consists in redox coupled reactions (Figure 8) (20) that lead in the formation of a coloured substance that absorbs light between 490-492 nm.

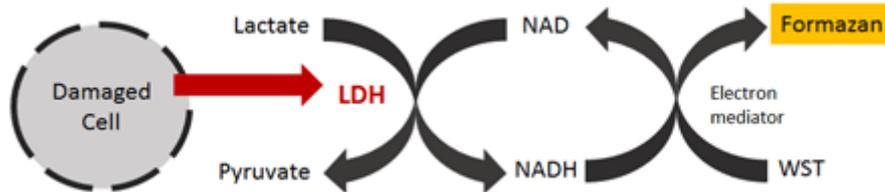


Figure 8. LDH Cell Cytotoxicity Assay Coupled Reactions

#### 4.3.3. Transfection: Electroporation

Electroporation was chosen among other transfection methods due to the high efficiency of the technique and to the cancerous nature of A549 cells which will oppose the inherent high mortality during the procedure.

One day before performing the electroporation cells were resuspended in DMEM/F12 media containing ROCK inhibitor (Cayman Chemical Company, 10005583) at  $5\mu\text{m}$  to enhance the survivability when transfecting the plasmids.

The electroporation was carried on with Bio Rad Gene Pulser® II Electroporation System. Cells were resuspended in electroporation buffer (Annex 2) and  $800\mu\text{l}$  containing approximately  $5 \cdot 10^6$  cells were added to each electroporation cuvette. Three conditions were established:  $37^\circ\text{C}$  plasmids,  $30^\circ\text{C}$  plasmids and negative control without any plasmid. The amount of plasmid employed is shown below.

Plasmid	Amount (µg)
AAVS1-TALEN-L	2,5
AAVS1-TALEN-R	2,5
Puro-Cas9 Donor	20
AAVS1-Neo-M2rtTA	20

Figure 9. Plasmid Quantities employed for Electroporation.

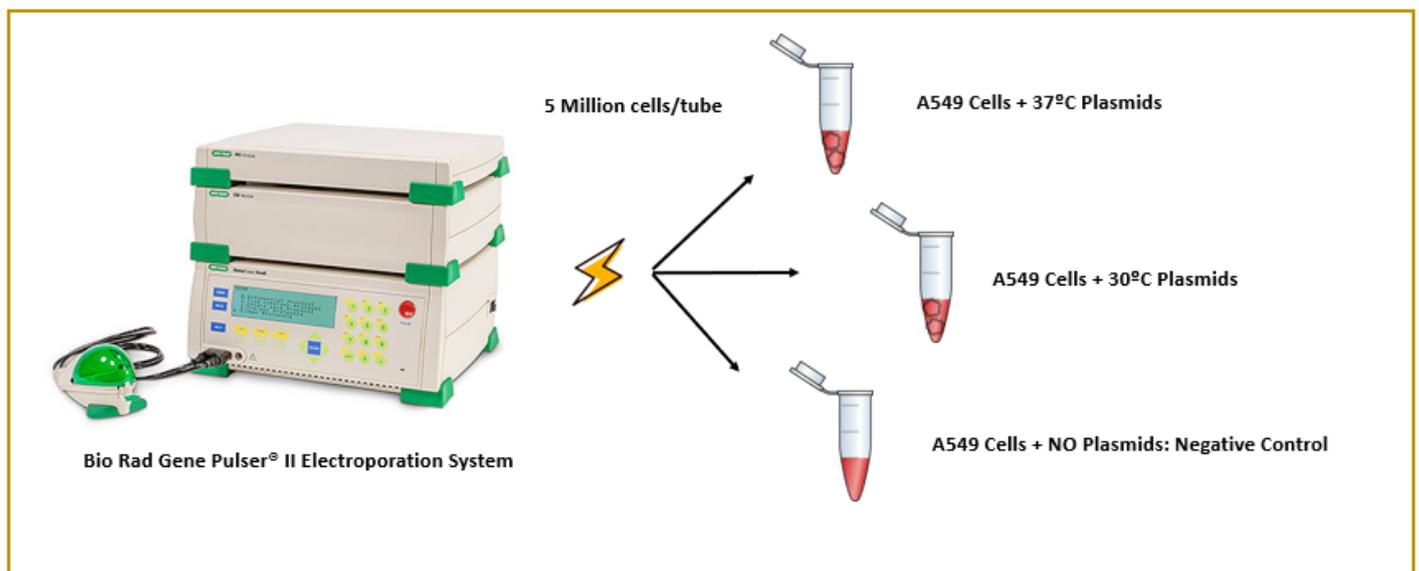


Figure 10. Electroporation Workflow

#### 4.3.4. Transfection: Lipofection

As an alternative an to compare delivery methods Lipofection was performed. As well as with the electroporation the two plasmid conditions and a negative control were evaluated. A first solution was prepared to prepare the plasmids before entering the liposomes. The solution contained the plasmids, Gibco™ Opti-MEM™ I Reduced Serum Medium (Fisher Scientific, 10149832) and P3000 reagent of Invitrogen™ Lipofectamine™ 3000 Transfection Kit (Fisher Scientific, 15292465). A second solution was prepared with Lipofectamine 3000 from the kit

above and Opti-MEM. After a 15-minute incubation the solutions are mixed and then added to the cells for liposome delivery.

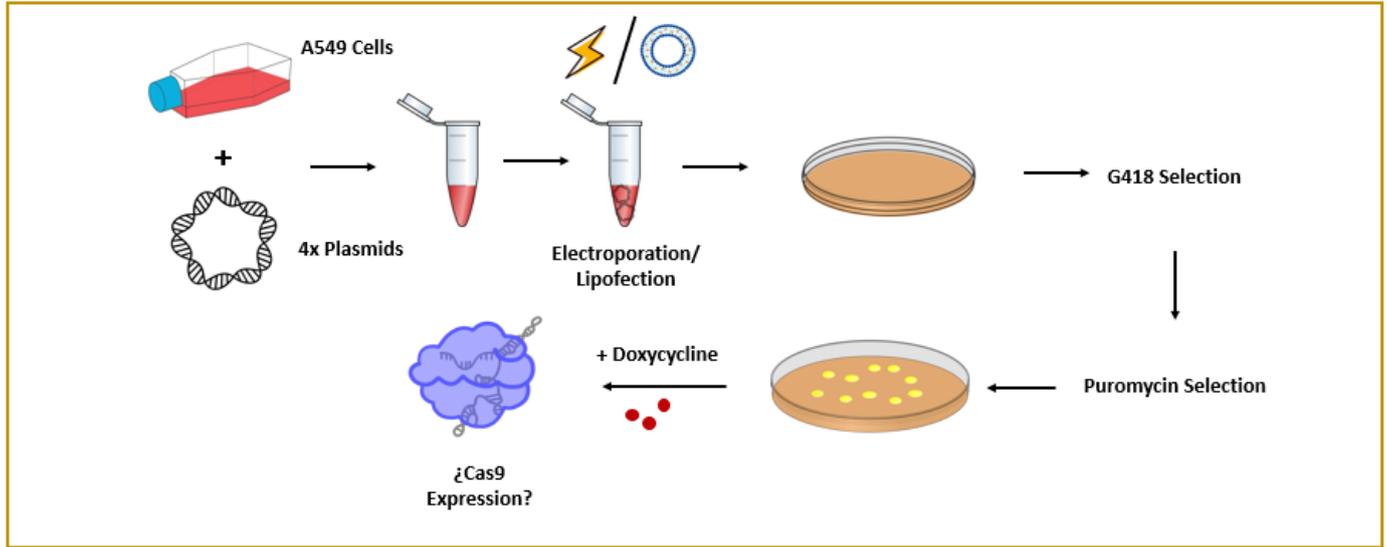


Figure 11. Transfection and Antibiotic Selection workflow scheme

## 5. RESULTS

### A. Required plasmids were successfully isolated and purified from Bacterial stabs at high efficiency

The streaking of the 4 plasmids was successfully performed being able to isolate single colonies. After the incubation in liquid cultures the concentration and the purity were measured with the Nanodrop 1000. Total amount of DNA was calculated by multiplying the concentration obtained by the 50 $\mu$ l of the elution buffer employed. Remarkable purification efficiency (maximum the kit retains is 20 $\mu$ g) higher than 50% was obtained, even reaching values around 100% with some plasmids like PuroCas9 at 37 $^{\circ}$ C and Neo-M2rtTA at 30 $^{\circ}$ C (Figure 11).

Isolation and Purification was first done of Neo-M2rtTA and PuroCas9 plasmids grown at 37 $^{\circ}$ C and then, repeated with the same plasmids grown at 30 $^{\circ}$ C so as make a comparison. Concentration values show that Neo-M2rtTA plasmid must be grown at 30 $^{\circ}$ C as advised. (12) However, despite Addgene recommendation, PuroCas9 donor exhibit better purification results at 37 $^{\circ}$ C.

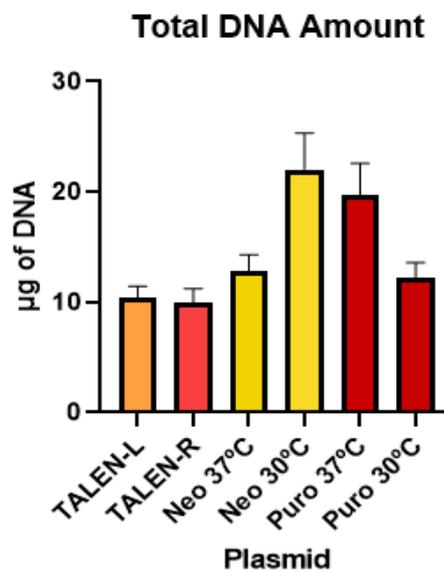


Figure 12. Total DNA amount isolated with miniprep.

Plasmid integrity was analysed by means of a 1% Agarose Electrophoresis. The percentage of agarose was selected according to the resolution, between 500 to 10.000 bp approximately for 1% agarose gel.

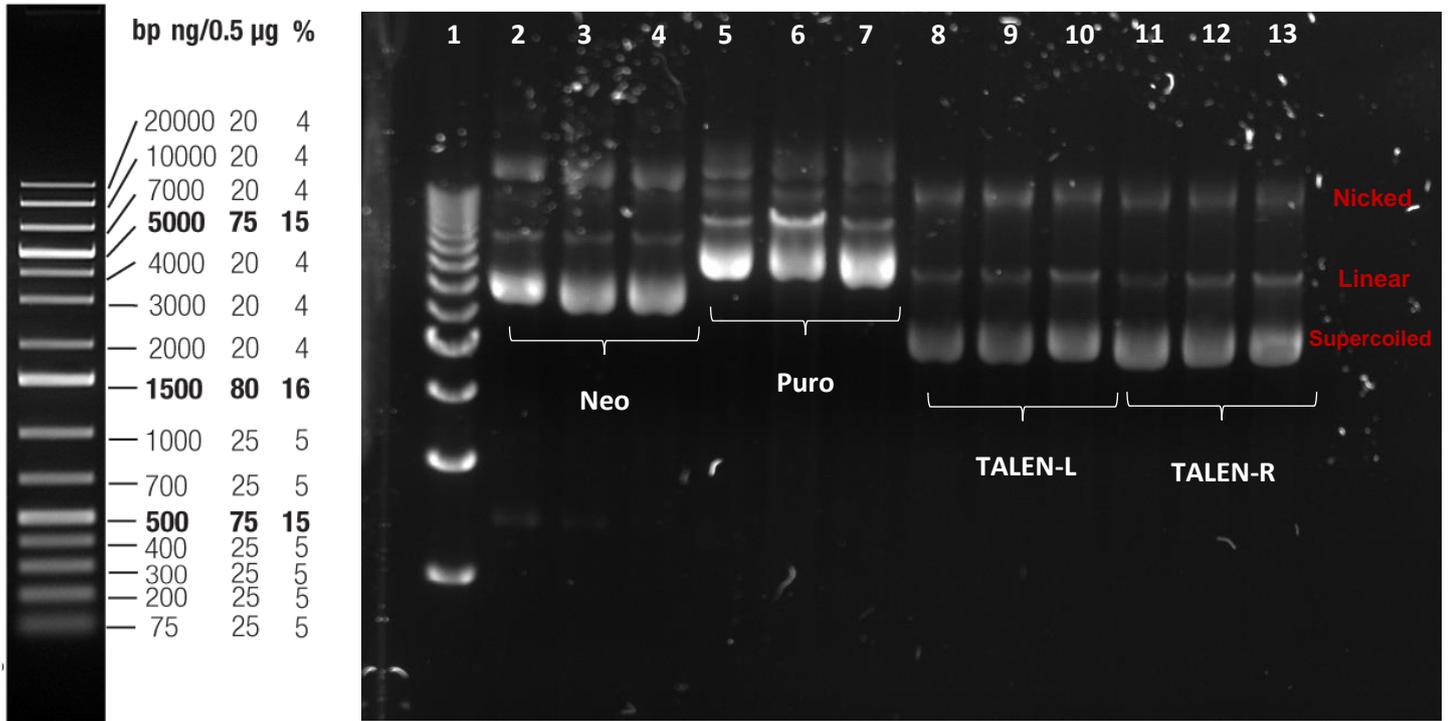


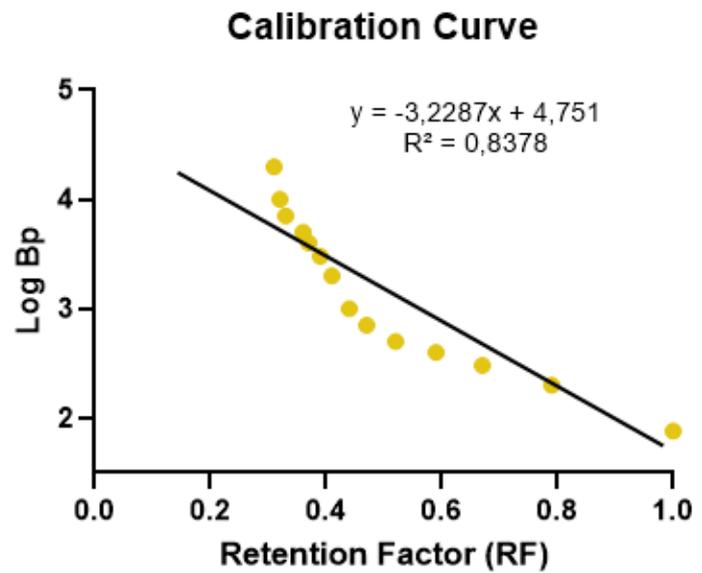
Figure 13. Plasmid Imaging: 1% Agarose Electrophoresis. Well 1: 1Kb Bio Rad molecular ruler. Wells 2-4: Plasmid Neo-M2rtTA (10.080 bp). Well 5-7: Plasmid PuroCas9 (11.579 bp). Well 8-10: Plasmid TALEN-L (8.118 bp). Well 11-13: TALEN-R (8.118 bp)

The calibration curve was built plotting the RF of each molecular marker and the base pair logarithm. The coefficient of determination ( $R^2$ ) obtained was 0,8378 which is a good value since it is close to 1. However, it can be observed that the upper bands were not perfectly separated thus, leading in a poor-linear region represented in the plot. Molecular weights of the plasmids were calculated by interpolating the RF in the equation of the tendency line. The values obtained match with the expected weights of each plasmid. The little deviation in weigh is caused by the non-linear region of the calibration curve.

**A**

Retention Factor ( $Z/Z_0$ )	Log Bp
1,00	1,88
0,79	2,30
0,67	2,48
0,59	2,60
0,52	2,70
0,47	2,85
0,44	3,00
0,41	3,30
0,39	3,48
0,37	3,60
0,36	3,70
0,33	3,85
0,32	4,00
0,31	4,30

**B**



**C**

	RF Plasmids	Log Bp	Bp
Neo-M2rtTA	0,23	4,01	10.203,99
Puro-Cas9	0,21	4,06	11.599,45
TALEN-L	0,26	3,90	7.896,52
TALEN-R	0,26	3,90	7.896,52

Figure 14. Plasmid Molecular weight Calculation: Table A shows molecular marker RF calculations and the logarithm of the Bp plotted in Graphic B. Tendency line and its linear equation is shown in the graphic as well as the  $R^2$ . Plasmid weighs calculations are shown in Table C.

In order to enhance the results obtained, a second electrophoresis was carried out. Agarose percentage was reduced to 0,5% to ensure better separation of bands.

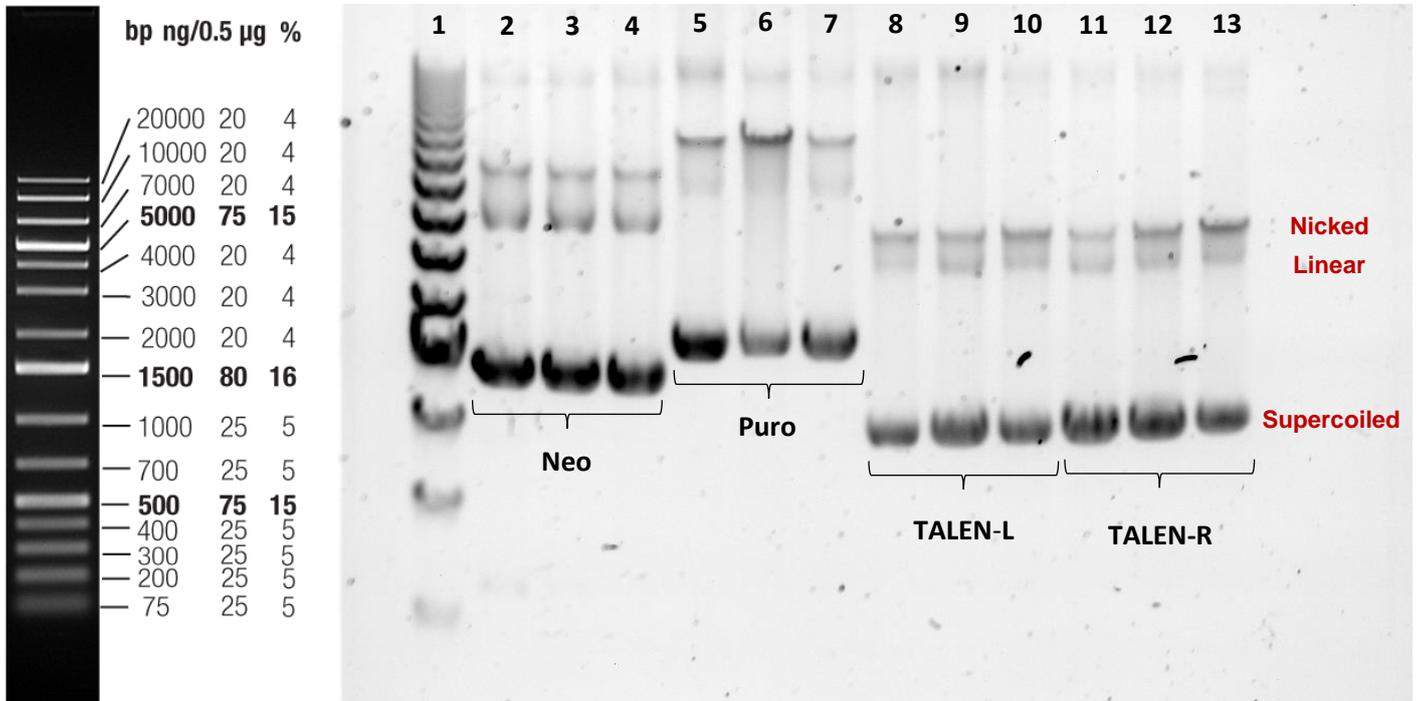


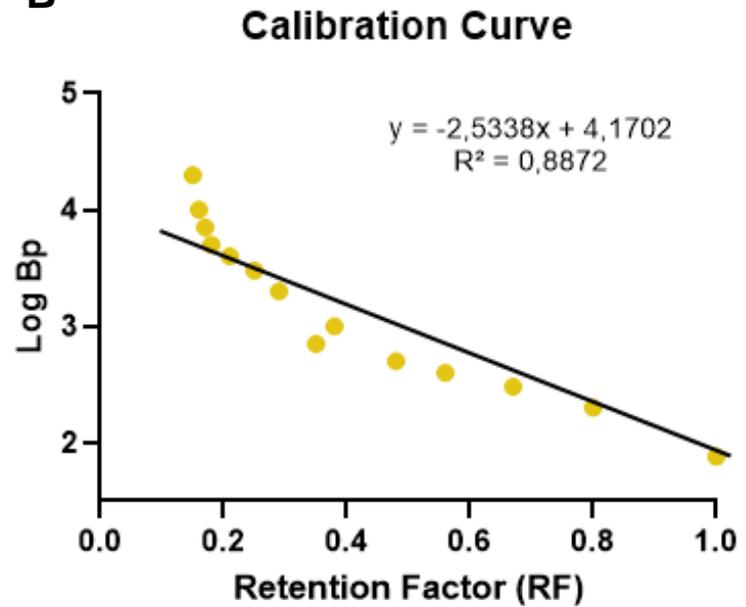
Figure 15. Plasmid Imaging: 0,5% Agarose Electrophoresis. Well 1: 1Kb Bio Rad molecular ruler. Wells 2-4: Plasmid Neo-M2rtTA (10.080 bp). Well 5-7: Plasmid PuroCas9 (11.579 bp). Well 8-10: Plasmid TALEN-L (8.118 bp). Well 11-13: TALEN-R (8.118 bp)

An increase in the  $R^2$  (0,8872) was obtained, proving that a slightly better separation was achieved. Therefore, gel shows that upper bands are still not perfectly separated. Lowering agarose percentage is not possible because it will compromise the gel polymerization. Not taking this into account, reliable results have been achieved correlating with expectancies.

**A**

Retention Factor ( $Z/Z_0$ )	Log Bp
1,00	1,88
0,80	2,30
0,67	2,48
0,56	2,60
0,48	2,70
0,35	2,85
0,38	3,00
0,29	3,30
0,25	3,48
0,21	3,60
0,18	3,70
0,17	3,85
0,16	4,00
0,15	4,30

**B**



**C**

	RF Plasmids	Log Bp	Bp
Neo-M2rtTA	0,07	3,99	9.795,86
Puro-Cas9	0,04	4,07	11.690,26
TALEN-L	0,11	3,89	7.738,68
TALEN-R	0,11	3,89	7.738,68

Figure 16. Plasmid Molecular weight Calculation: Table A shows molecular marker RF calculations and the logarithm of the Bp plotted in Graphic B. Tendency line and its linear equation is shown in the graphic as well as the  $R^2$ . Plasmid weighs calculations are shown in Table C.

A last electrophoresis was performed to compare plasmids grown at 30°C and 37°C. No visual differences were found between the 2 conditions.

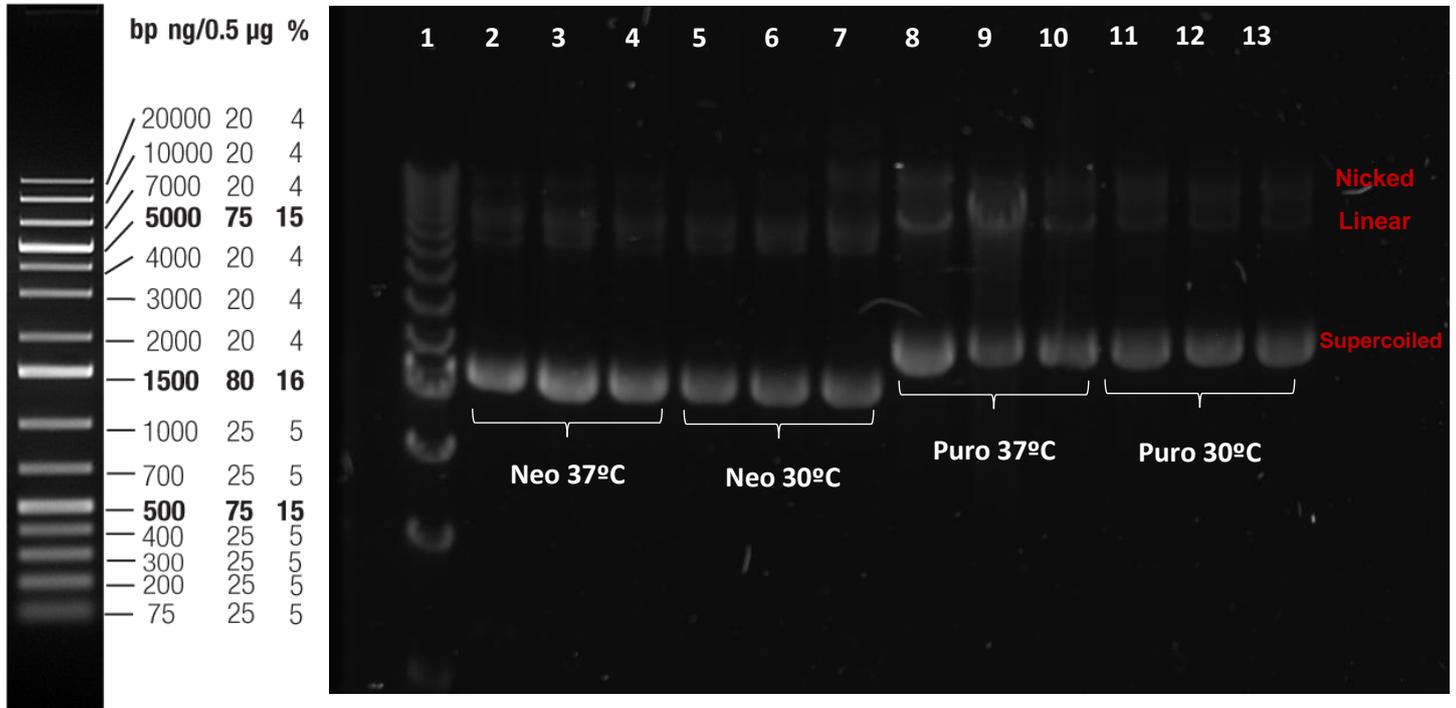


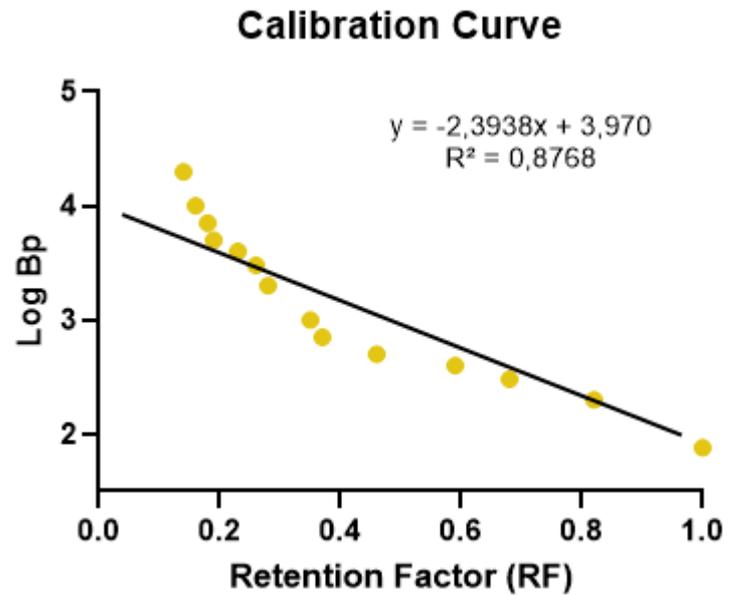
Figure 17. Plasmid Imaging: 0,75% Agarose Electrophoresis. Well 1: 1Kb Bio Rad molecular ruler. Wells 2-4: Plasmid Neo-M2rtTA 37°C (10.080 bp). Well 5-7: Plasmid Neo-M2rtTA 30°C (10.080 bp). Well 8-10: Plasmid PuroCas9 37°C (11.579 bp). Plasmid PuroCas9 30°C (11.579 bp).

Similar  $R^2$  was obtained with a 0,75% agarose gel. Once calculated the molecular weigh of plasmids of each condition, no difference was found either in Neo-M2rtTA plasmids or in PuroCas9 plasmid.

**A**

Retention Factor ( $Z/Z_0$ )	Log Bp
1	1,88
0,82	2,30
0,68	2,48
0,59	2,60
0,46	2,70
0,37	2,85
0,35	3,00
0,28	3,30
0,26	3,48
0,23	3,60
0,19	3,70
0,18	3,85
0,16	4,00
0,14	4,30

**B**



**C**

	RF Plasmids	Log Bp	Bp
Neo-M2rtTA 37°C	0,07	4,07	9.975,86
Neo-M2rtTA 30°C	0,07	4,07	9.975,86
Puro-Cas9 37°C	0,11	3,89	11.356,89
Puro-Cas9 30°C	0,11	3,89	11.356,89

Figure 18. Plasmid Molecular weight Calculation: Table A shows molecular marker RF calculations and the logarithm of the Bp plotted in Graphic B. Tendency line and its linear equation is shown in the graphic as well as the  $R^2$ . Plasmid weights calculations are shown in Table C.

## **B. Selective antibiotics showed best efficiency at high concentrations within 5 days of treatment**

To elucidate the correct antibiotic dose, different concentrations were tested over the course of 12 days. A quantification LDH colorimetric assay was employed so as to determine the precise concentration of each antibiotic.

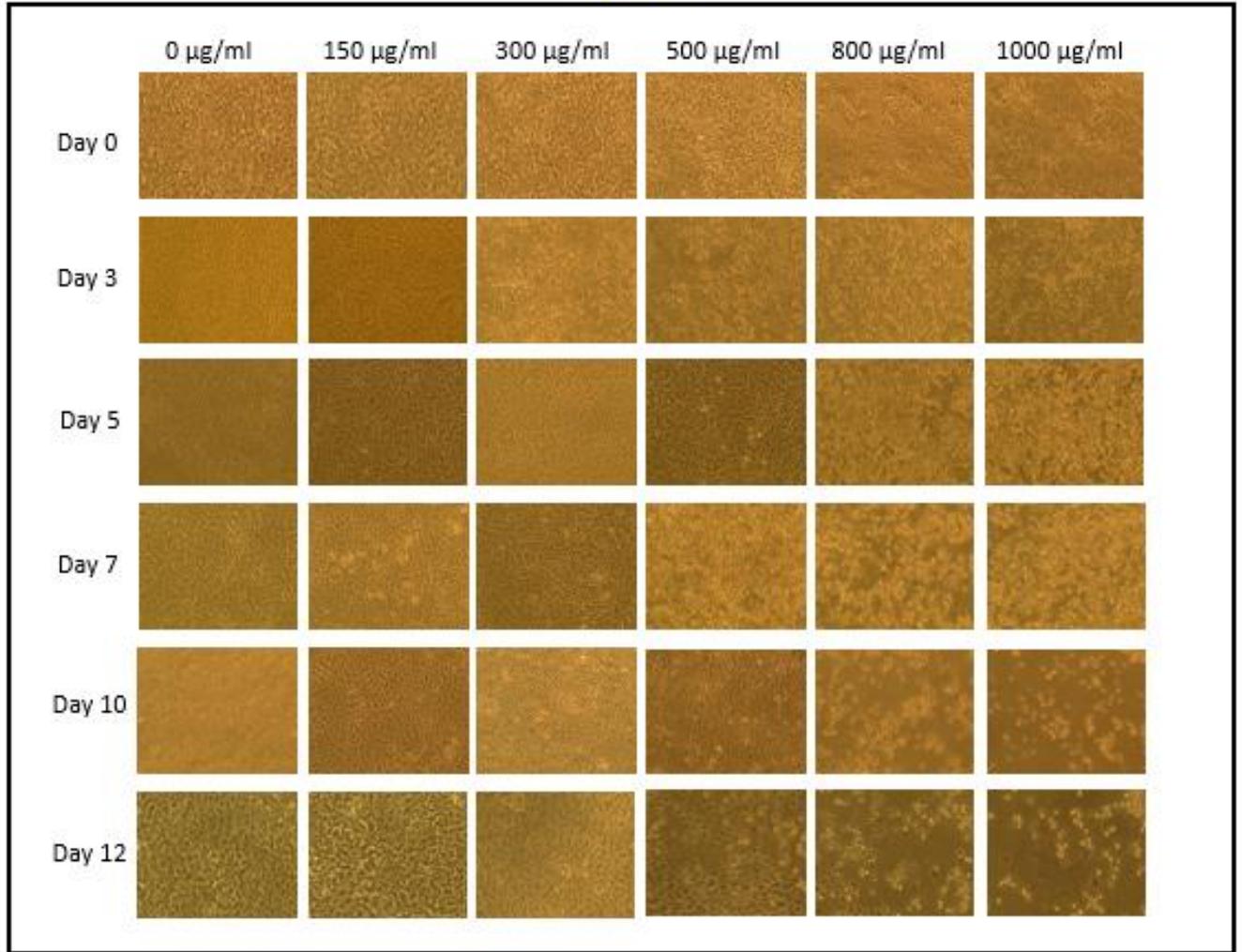
For G418 5 concentrations (0, 150, 300, 500, 800, 1000 $\mu$ g/ml) were evaluated as well as a control condition with no antibiotic. These concentrations were selected to surpass the concentrations suggested by the manufacturer. In figure 18 A it can be clearly observed that cell death increases diagonally at higher antibiotic concentrations and more days of treatment. Images show that between day 3 and day 5 at maximum concentration cell death reaches a peak since the majority of cells are not attached anymore.

The LDH cytotoxicity assay was performed at the same time with all the media samples recollected. Absorbance at 490-492 which represents LDH release was plotted among each timepoint. Day 3 shows a peak in LDH release. However, as said before, it cannot be compared as higher cell death was reached due to longer treatment time. If we pay attention to Day 5 it can be plainly seen how concentrations appear coherently ordered from the lowest to the highest dose. Maximum cell death was reached within this day at 1000 $\mu$ g/ml of G418 (Figure 18 B). It can be observed that at Day 7 cell death drops to minimum meaning that few cells are still alive.

Based on the data obtained, the 1000 $\mu$ g/ml dose of G418 has been chosen for the selection treatment over a 5-day period.

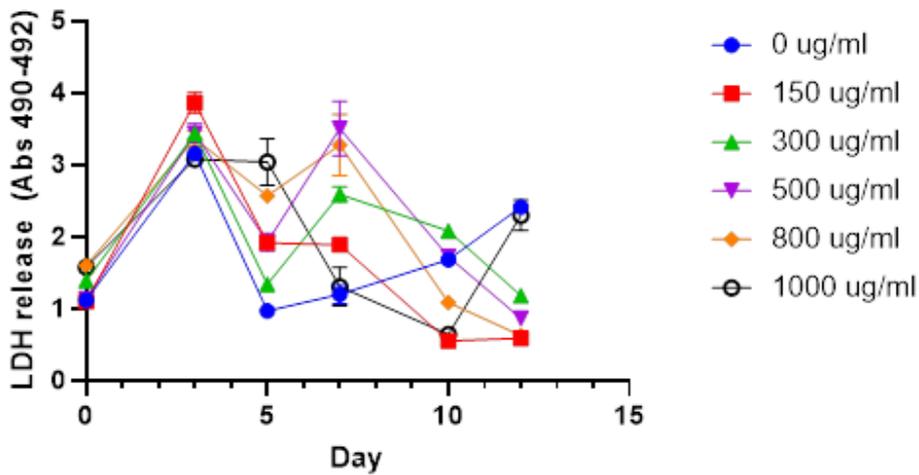
**A**

**G418**



**B**

**XY: G418**



**C**

**Day 5 G418**

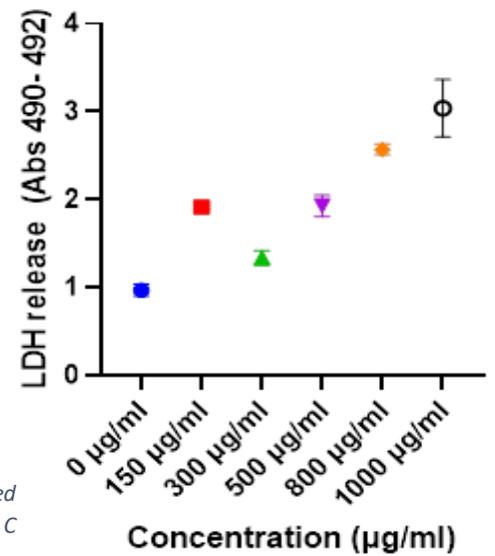
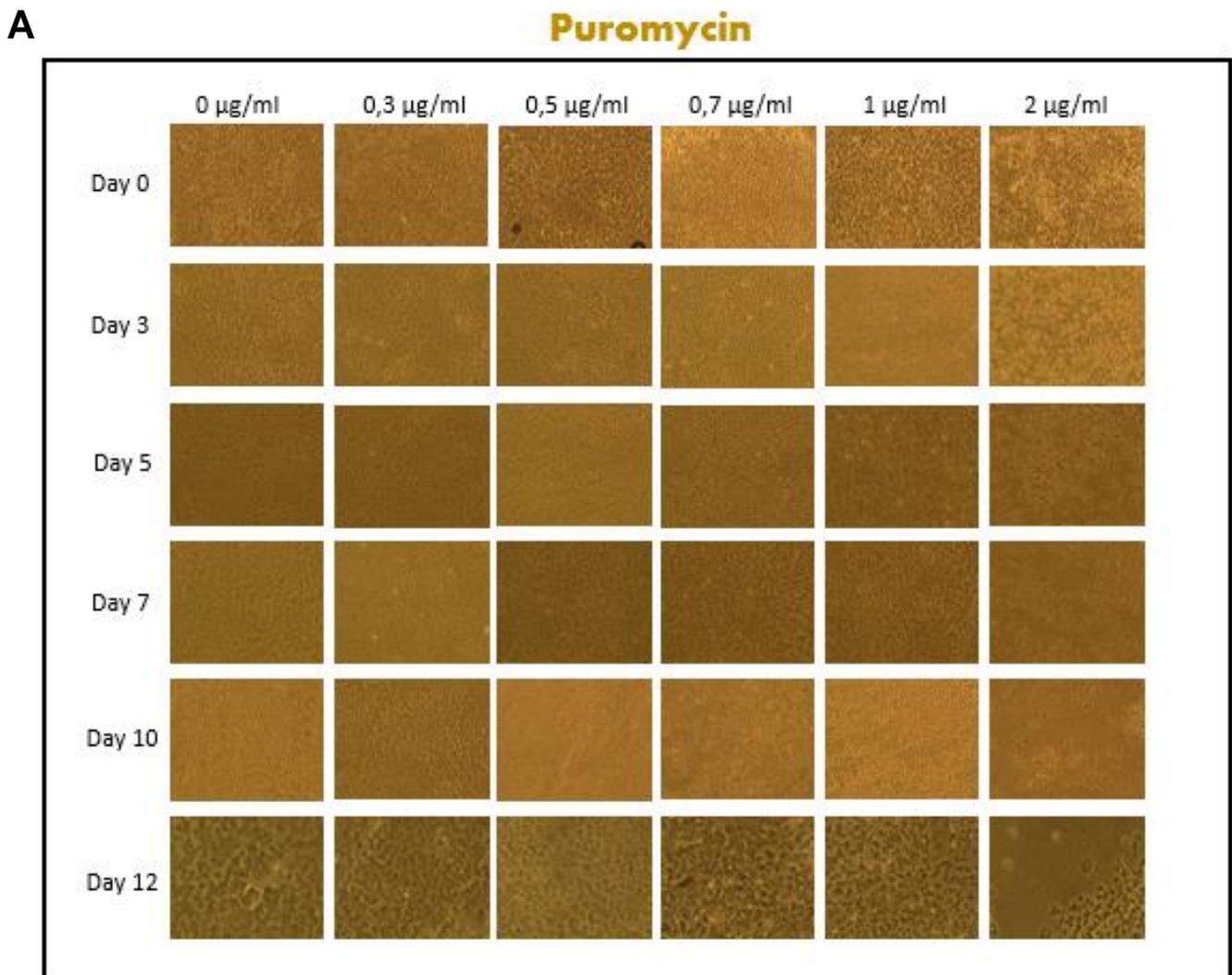


Figure 19. G418 Selection Optimization. Figure A shows pictures taken of each condition tested over a 12-day course. Graph in B represents LDH release tested in collected media and Graph C exhibits data on Day 5 specifically.

Looking at Puromycin antibiotic selection same results have been obtained as G418. At Day 5 when applied 2 $\mu$ g/ml dose maximum cell death was achieved. This concentration over a course of 5 days was selected as the optimal one for selection after the transfection.



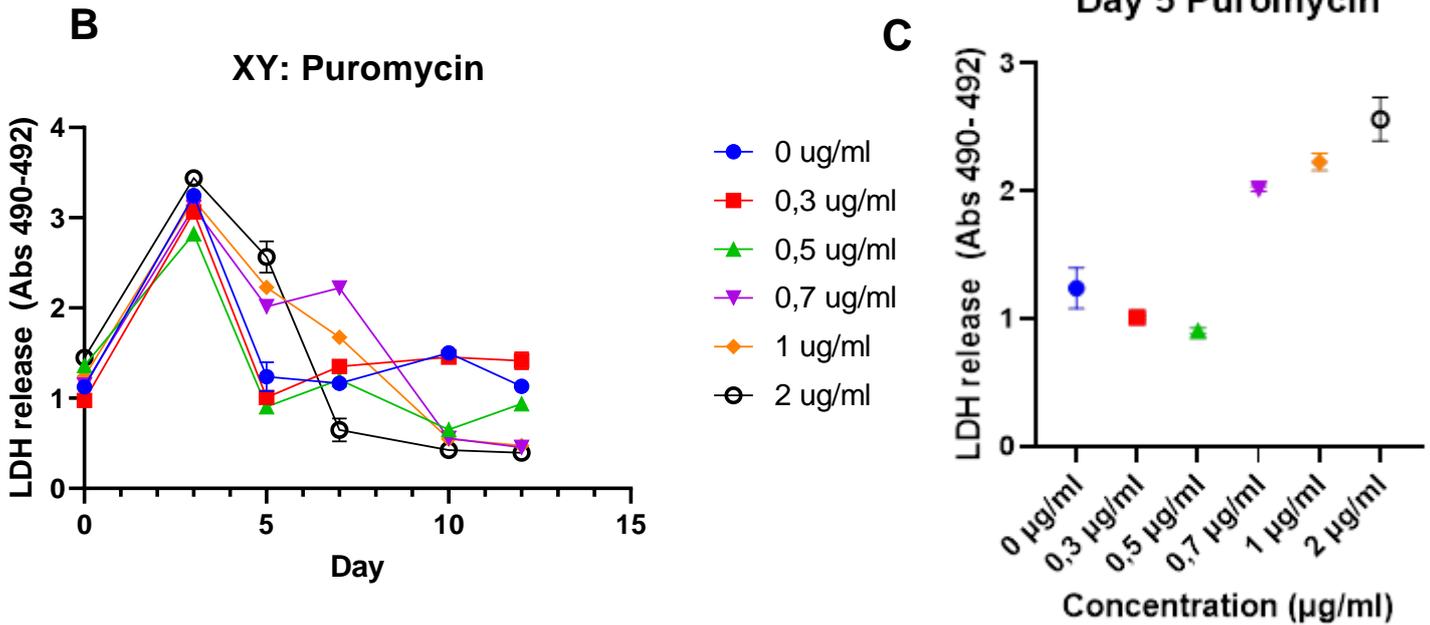


Figure 20. Puromycin Selection Optimization. Figure A shows pictures taken of each condition tested over a 12-day course. Graph in B represents LDH release tested in collected media and Graph C exhibits data on Day 5 specifically.

### C. Electroporation did not show remarkable results so lipofection was successfully performed

Firstly, electroporation was tried for plasmid delivery with no success as no cell colonies could be seen after the selection meaning that plasmids were not incorporated.

## Electroporation

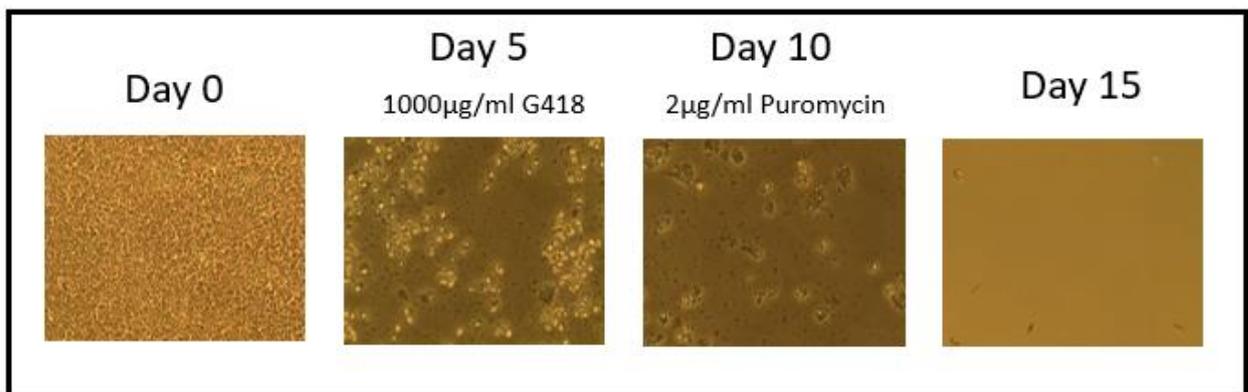


Figure 21. Electroporation trial. Pictures were taken from the 37°C plasmid condition.

On the other hand, lipofection was carried successfully as cell colonies can be observed at the final stage of the antibiotic selection. However, further experimental evidence is needed to validate the results obtained.

## Lipofection

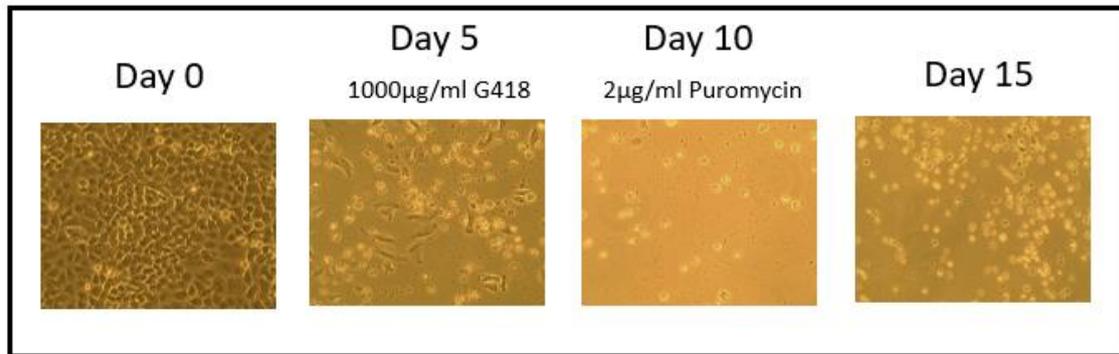


Figure 22. Lipofection trial. Pictures were taken from the 37°C plasmid condition.

## 6. DISCUSSION AND CONCLUSIONS

IPF has proved to be a dreadful disease which rampant increment is raising global awareness. Furthermore, the only available drugs can only slow down disease progression but they cannot stop or reverse it, which leaves lung transplantation as the only effective option. Effective. However, the number of available donor lungs is limited and their success rate remains at approximately 50% after 5 years of transplantation. Thus, new therapeutical approaches are needed focusing in targeting the root of the pathologic factors involved in IPF.

In order to elucidate the effect of identified TFs, we must examine the fibrotic changes during their knock-out in addition to YAP/TAZ knockout. A stable genome editing tool is required to precisely multitarget the TFs. To fulfil this goal, CRISPR/Cas9 was employed due to its versatility and simple way to work with. The main problem regarding this technique is the possible Off-targets that can be generated. Taking this into account, the iCRISPR cell line was created by means of the TALEN technology which will ensure the precise location of the cassettes. While TALEN produces stable edits in the genome, it requires lots of time and optimization, which makes it less suitable as the technology to be utilized in the knock-out cell line. This approach has been successfully used in generating an iCRISPR line in primary human pluripotent stem cells.

While many primary cells can be isolated and propagated, the lung alveolar epithelial cells are difficult to isolate and, more importantly, impossible to passage. Since IPF main site of injury is the lung epithelium, we ought to use a cell line that most closely mimics the lung alveolar epithelium. To this end, we have used A549 cells, which are human lung adenocarcinoma cells. While A549 cells are tumor cells, they are widely used in the lung field to examine certain mechanisms. In all cases, all results from A549 experiments are usually validated in primary cells. Additionally, unlike alveolar epithelial cells, bronchial epithelial cells are easy to isolate and propagate. Thus, generation of a stable bronchial iCRISPR line would potentially work after optimization is complete.

Cell survival in response to antibiotics treatment is highly variable among different cell types. Thus, we needed to establish the correct concentrations for the used antibiotics to ensure their selectivity for resistant cells. Optimization of the selection process with antibiotic was needed for A549 cells as, to our knowledge, no known concentrations are published for geneticin and puromycin. We successfully were able to elucidate the optimal

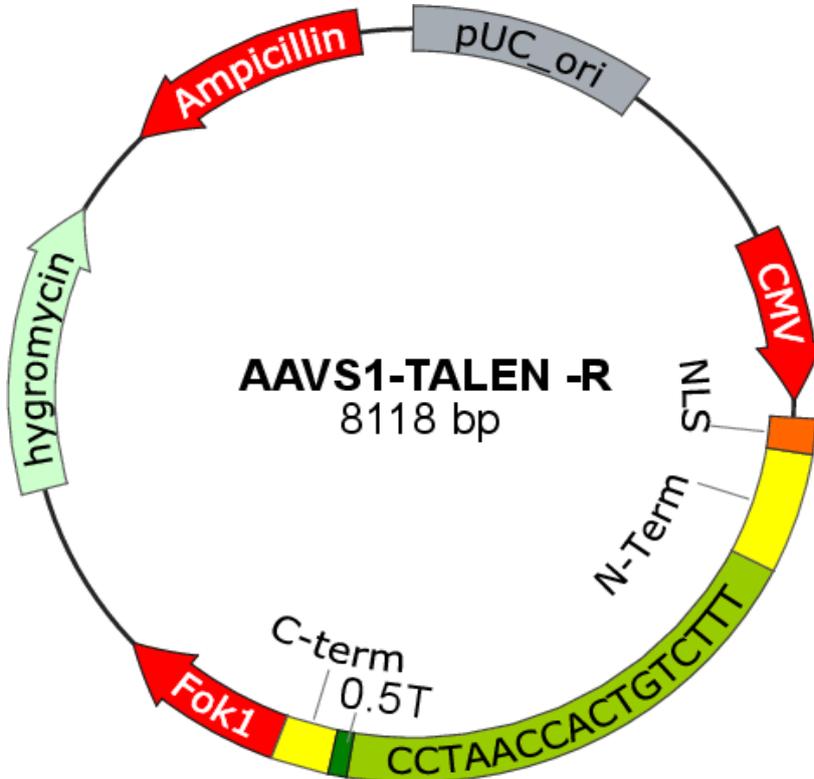
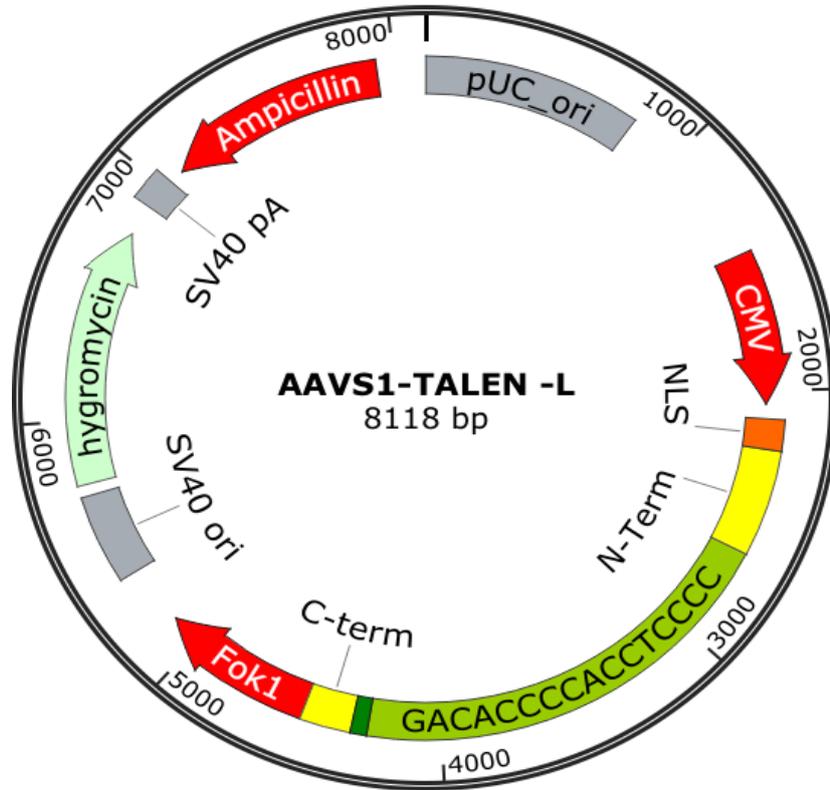
dose of antibiotic needed of each antibiotic in order to select only the desired cells when the transfection of the plasmids is done.

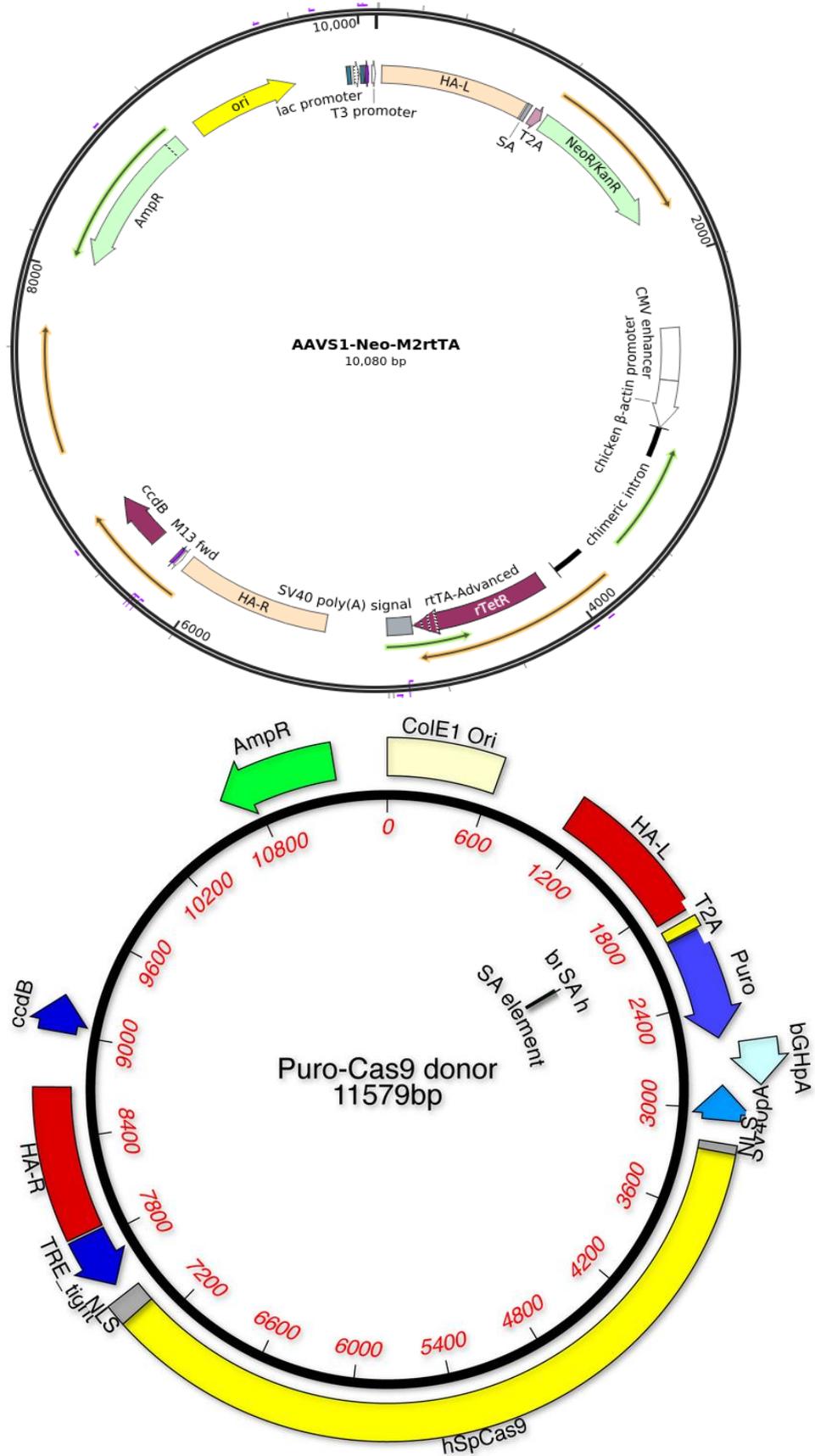
Several methods can be used for transfection due to variability in cell response, efficacy, and survival; we tested electroporation and lipofection in this project. Electroporation was tested first as it was the chosen method for previously established workflow. In our hands, it did not work correctly in A549 cells potentially due to lack of effective internal controls and thus needs further optimization. This has prompted us to try Lipofection prior to optimizing electroporation. We were able to observe colony formation in the lipofection experiment in the conditions where plasmids carrying the antibiotic resistance was added. This indicates that the transfection using lipofection may have succeeded. However further evidence and validation are needed to confirm this.

Looking towards future steps, we will perform cell isolation from the colonies that have emerged after the selection process. Cas9 expression will be induced with doxycycline and expression of cas9 transcripts and protein will be evaluated A549 iCRISPR cell line will be validated by DIG-labeled PCR genotyping.

The generation of the iCRISPR A549 cell line will be a remarkable success due to its potential capacities as a genome editing tool within the field of lung epithelial biology. This tool may provide several new implications in relevance to all chronic lung disease as it is not limited to IPF in any way. Within this project it will be considered a key mainstay in the evaluation of the TFs bioinformatically selected. The ability to identify disease causing transcription factor combinations may allow for finding new effective therapies for IPF.

## 7. ANNEXES





Annex 1. TALEN and Donor Plasmids.

Component	MW (g/mol)	Amount (g)	Concentration
KCl	74,56	0,186	5 mM
MgCl <sub>2</sub>	203,31	1,525	15 mM
NaCl	58,44	2,63	90 mM
Glucose	180,16	0,9	10 mM
Ca(NO <sub>3</sub> ) <sub>2</sub>	164,088	0,033	0,4 mM
Na <sub>2</sub> HPO <sub>4</sub>	268.07	3.57	0,0266 M
NaH <sub>2</sub> PO <sub>4</sub>	137.99	0,922	0,0134 M

*Annex 2. Electroporation Buffer*

## 8. BIBLIOGRAPHY

1. Barratt SL, Creamer A, Hayton C, Chaudhuri N. Idiopathic Pulmonary Fibrosis (IPF): An Overview. *J Clin Med* 2018; 7.
2. Ley B, Collard HR, King TE, Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 431-440.
3. Parimon T, Yao C, Stripp BR, Noble PW, Chen P. Alveolar Epithelial Type II Cells as Drivers of Lung Fibrosis in Idiopathic Pulmonary Fibrosis. *Int J Mol Sci* 2020; 21: 2269.
4. Corvol H, Flamein F, Epaud R, Clement A, Guillot L. Lung alveolar epithelium and interstitial lung disease. *Int J Biochem Cell Biol* 2009; 41: 1643-1651.
5. Barkauskas CE, Chung M-I, Fioret B, Gao X, Katsura H, Hogan BLM. Lung organoids: current uses and future promise. *Development* 2017; 144: 986-997.
6. Verheyden JM, Sun X. A transitional stem cell state in the lung. *Nature Cell Biology* 2020; 22: 1025-1026.
7. Glass DS, Grossfeld D, Renna HA, Agarwala P, Spiegler P, Kasselmann LJ, Glass AD, Deleon J, Reiss AB. Idiopathic pulmonary fibrosis: Molecular mechanisms and potential treatment approaches. *Respiratory Investigation* 2020; 58: 320-335.
8. Piersma B, Bank RA, Boersema M. Signaling in Fibrosis: TGF- $\beta$ , WNT, and YAP/TAZ Converge. *Frontiers in Medicine* 2015; 2.
9. Barry ER, Camargo FD. The Hippo superhighway: signaling crossroads converging on the Hippo/Yap pathway in stem cells and development. *Current Opinion in Cell Biology* 2013; 25: 247-253.
10. Hansen CG, Moroishi T, Guan K-L. YAP and TAZ: a nexus for Hippo signaling and beyond. *Trends in Cell Biology* 2015; 25: 499-513.
11. Brodowska K, Al-Moujahed A, Marmalidou A, Meyer Zu Horste M, Cichy J, Miller JW, Gragoudas E, Vavvas DG. The clinically used photosensitizer Verteporfin (VP) inhibits YAP-TEAD and human retinoblastoma cell growth in vitro without light activation. *Experimental Eye Research* 2014; 124: 67-73.
12. Zhu Z, González F, Huangfu D. The iCRISPR Platform for Rapid Genome Editing in Human Pluripotent Stem Cells. Elsevier; 2014. p. 215-250.
13. Song M, Kim Y-H, Kim J-S, Kim H. Genome Engineering in Human Cells. Elsevier; 2014. p. 93-118.
14. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 2013; 8: 2281-2308.

15. Dow LE, Fisher J, O'Rourke KP, Muley A, Kasthuber ER, Livshits G, Tschaharganeh DF, Socci ND, Lowe SW. Inducible in vivo genome editing with CRISPR-Cas9. *Nature Biotechnology* 2015; 33: 390-394.
16. Chen Y, Cao J, Xiong M, Andrew, Dong Y, Tao Y, Cindy, Du Z, Zhang S-C. Engineering Human Stem Cell Lines with Inducible Gene Knockout using CRISPR/Cas9. *Cell Stem Cell* 2015; 17: 233-244.
17. González F, Zhu Z, Shi Z-D, Lelli K, Verma N, Qing, Huangfu D. An iCRISPR Platform for Rapid, Multiplexable, and Inducible Genome Editing in Human Pluripotent Stem Cells. *Cell Stem Cell* 2014; 15: 215-226.
18. González F, Zhu Z, Shi ZD, Lelli K, Verma N, Li QV, Huangfu D. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* 2014; 15: 215-226.
19. DeKolver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, Sancak Y, Cui X, Steine EJ, Miller JC, Tam P, Bartsevich VV, Meng X, Rupniewski I, Gopalan SM, Sun HC, Pitz KJ, Rock JM, Zhang L, Davis GD, Rebar EJ, Cheeseman IM, Yamamoto KR, Sabatini DM, Jaenisch R, Gregory PD, Urnov FD. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res* 2010; 20: 1133-1142.
20. Life Sciences E. LDH Cytotoxicity WST Assay.