FACULTAT
DE CIÈNCIES
I TECNOLOGIA
UVIC | UVIC·UCC

Final Degree Project

Generation of a cellular model, based on CaCO2 cells, to study tight junction's formation and dynamics. CRISPR/Cas mediated gene edition to fluorescently tag the occludin protein.

WRITTEN BY:
CARLOS RODRÍGUEZ PARDO

Degree in Biotechnology

Tutors: Joan Bertran Comulada i Carlo Manzo

Vic, June 2020

Acknowledgements

I'm extremely grateful to Joan Bertran Comulada, who has tutored and mentored me all along this project. He has taught me not only how to elaborate hypothesis and perform analysis, but also to think and act in a scientific manner, acquiring skills and knowledge that simply are not learn in a classroom. Moreover, his role in the redaction of this thesis allowed me to present our results in the best way it was possible, and at the same time improving my scientific writing abilities. For all this reasons, I wanted to acknowledge him for all the help he has brought to this work.

I'd also like to extend my gratitude to Carlo Manzo, for allowing me to participate in this project at the QuBI research group, on which I had the opportunity of collaborating in posterior applications of the tool we developed. I learned new analysis techniques and experienced not only the success of this project, but also the happiness of making use of the tool we created and the obtaining of results from it.

Acronyms:

OCLN	Occludin
TJ	Tight Junction
AJ	Adherent Junction
ZO	Zona Occludens
MAGUK	Membrane-Associated Guanylate Kinase
MDCK	Madin-Darby Canine Kidney
PEI	Polyethylenimine
UHA	Upstream Homology Arm
DHA	Downstream Homology Arm
TIRF	Total Internal Reflection Fluorescence
GFP	Green Fluorescent Protein
RFP	Red Fluorescent Protein
Clau	Claudin
DMEM	Dulbeccos Modifyed Eagle Medium
FBS	Fetal Bovine Serum
DMSO	Di-methyl sulfoxide
SPT	Single Particle Tracking

Summary

Title: Generation of a cellular model, based on CaCO2 cells, to study tight junction's formation and dynamics. CRISPR/Cas mediated gene edition to fluorescently tag the occludin protein.

Author: Carlos Rodriguez Pardo

Supervisor: Dr. Joan Bertran Comulada and Dr. Carlo Manzo

Date: June 2020

Keywords: *CRISPR-Cas9*, *fusion protein*, *occludin*, *CaCO2 cells*, *tight junctions*, *membrane protein dynamics*

Tight junctions are present in all epithelial tissues and are involved in several processes of signalling and barrier formation. Although the proteins that compose them are known, some of the exact mechanisms that drive them remain unclear. For this purpose, we have generated a CaCO2 cell line capable of expressing a fusion protein between occludin, a member of tight junction protein complex, and mEmerald, a fluorescent protein derived from GFP. This will allow the study of tight junctions in living cells.

Throughout this project, we generated two plasmids needed for carry out CaCO2 genome edition. We transfected the cells with one plasmid encoding custom Cas-9 endonuclease plus a specific single-guide RNA accompanied by a plasmid bearing the repair sequence, thus producing a double strand break in the cellular genome and re-writing it to generate a mEmerald-occludin fusion protein.

Furthermore, we verified the expression of the fusion protein by western blot analysis and successfully used the derived cells in single particle tracking studies.

Resum

Títol: Generació d'un model cel·lular basat en cèl·lules CaCO2 per estudiar la dinàmica i formació de les unions estretes. Edició genètica mediada per CRISPR_Cas per marcar fluorescentment la proteïna occludina.
Autor: Carlos Rodríguez Pardo
Tutors: Dr. Joan Bertran Comulada i Dr. Carlo Manzo
Data: Juny 2020
Paraules clau: *CRISPR-Cas9*, *proteïna de fusió*, *ocludina*, *cèl·lules CaCO2*, *unions estretes*, *dinàmica de proteïnes de membrana*

Les unions estretes estan presents en tots els teixits epitelials i estan involucrades en processos de senyalització i formació de barreres. Encara que les proteïnes que les componen són conegudes, alguns dels mecanismes que les controlen encara no estan clars. Per a aquest propòsit, hem generat una línia cel·lular de cèl·lules CaCO2 capaç d'expressar una proteïna de fusió entre l'ocludina, un component de les unions estretes, i mEmerald, una proteïna fluorescent derivada de GFP. Això pot permetre l'estudi de les unions estretes en cèl·lules vives.

Al llarg d'aquest projecte, hem generat dos plàsmids necessaris per dur a terme l'edició del genoma de les CaCO2. Vàrem transfectar les cèl·lules amb un plàsmid que codifica per l'endonucleasa Cas-9 amb un RNA guia i amb altre plàsmid portador de la seqüència de reparació. D'aquesta manera, vàrem produir una doble rotura de cadena a l'ADN i aquest es va reparar amb la seqüència que codifica per mEmerald-ocludina.

A més, vàrem confirmar l'edició per blotting dels extractes de proteïna dels diferents clons i varem obtenir resultats positius per a l'edició amb la tecnologia CRISPR. Per altra part, la proteïna de fusió s'ha provat en estudis de rastreig de partícules obtenint resultats similars als extrets amb altres tècniques de marcatge.

Contents

Figure index

1 Introduction

1.1 Tight junctions

Tight junctions are one of the cell-to-cell junctions involved in the process of tissue formation and regulation of epithelial and endothelial signalling [1]. They were discovered by M.G. Farquhar and G.E. Palade in 1963 using rat and guinea pig epithelia [2]. The major importance of tight junctions resides in the fact that they comprise the first barrier between the outer environment and the inner part of the organs that contain these complexes, including digestive tract lining, milk ducts, environment-exposed mucosa, inner lining of blood vessels and urine excretion ducts among others [3]. They are present in epithelial and endothelial cells where they contribute to the complete sealing of the intercellular space, allowing only the diffusion of small molecules. Examples of this function are the blood brain barrier and the enterocyte lining of the digestive tract, where there is a high specific selection of molecules that pass through them. Tight junctions are highly related to polarized cells; therefore, they are localized at specific sites at the cell. Specifically, tight junctions seal the apical localization in the epithelia, accomplishing a barrier function.

Tight junctions are considered multi-protein association complexes composed by transmembrane and peripheral membrane proteins. Mainly, they are composed by 3 protein families which interact among them to form membrane complexes (Figure 1). Furthermore, tight junction complexes can be more extensive than the main core, implicating interactions with signalling proteins involved in cellular processes as cell proliferation and cell-to-tell tightness regulation [4].

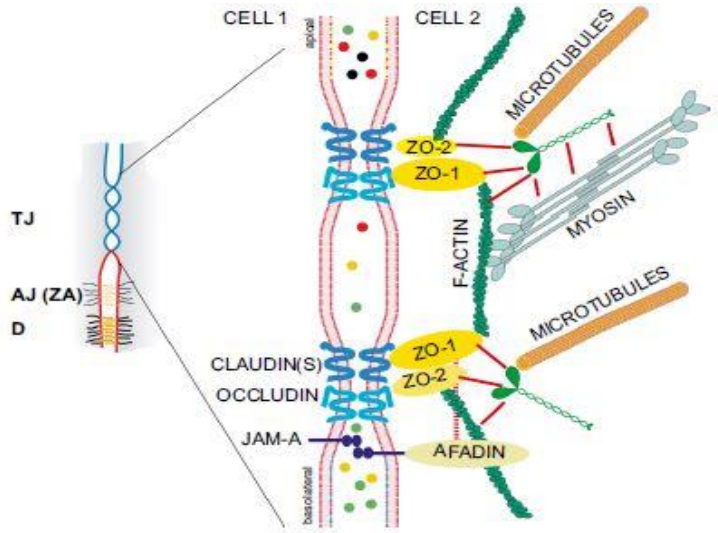


Figure 1; Schematic representation of how occludin and claudin interactions join cell membranes between two adjacent cells and how they associate with MAGUK proteins which, in last stage, anchor them to the actin cytoskeleton [5]

The main components of the tight junctions are claudins, MAGUK proteins (also known as zona occludens proteins) and occludin. Claudins are a family of more than twenty proteins [6], each of them accomplishes a different function thus they are expressed depending on the functions on the tissue [7]. Mainly, their structure is composed of four extracellular domains highly related to ionic permeability [8], therefore they are very important in determining tight junction selectivity. Proteins of the zona occludens family, also known as MAGUK (membrane-associated guanylate kinase) family, are highly involved in tight junction assembling and regulation. They are peripherally associated proteins, which have a role of scaffold for protein-protein interactions. On one hand, they serve as anchoring proteins for claudins and occludin, binding the tight junction complex to the actin cytoskeleton [9]. On the other hand, ZO proteins are highly related to g-coupled protein receptors and several transcription factors, hence the importance of TJ in signalling regulation at the endothelial and epithelial cells. Finally, occludin is another protein in the complex which physiological function has not been fully elucidated. There are different occludin splicing variants that yield different polypeptides susceptible to post-translational modifications [10]. The most common isoform is composed of two extracellular loops and two intracellular domains, linked with four transmembrane domain.

The process of tight junction assembly and tightening is highly regulated through phosphorylation and dephosphorylation of the proteins in the complex [11], ubiquination [12], G protein signalling coupled to VEGF receptors [13] and activation of signalling cascades that trigger transcription factor activation. All these processes are involved in the tight junctions response to different physiological stimulus including swelling, infection or digestive and vascular disorders. However, some of the mechanisms that drive them remain unknown.

It has been shown that tight junctions need adherent junctions in order to ensemble, since claudin family proteins recruit occludin to the plasma membrane [14]. Once the TJ complex is formed, adherent junctions(AJ) are no longer needed for its maintenance [14]. Tight junctions also bind and maintain a complex relation with the actin cytoskeleton and are involved in cell shape and polarity. This is due to zona occludens protein family, formed by ZO-1, ZO-2 and ZO-3, that act as scaffold proteins between TJ and actin cytoskeleton. It has been demonstrated that ZO-1 can form homodimers and heterodimers with the other ZO proteins [9] which consequently has a significant effect on TJ assembly. The zonula occludens proteins associate with occludin and claudin and use their proline-rich c-terminus to bind F-actin and actin-associated protein 4.1R [9]. Furthermore, it has been suggested that ZO proteins connect TJ with AJ forming heterodimers between them thanks to their PDZ, SH3 and guanylate kinase domains [14].

1.2 CaCO2 as a model

For these experiments, we will use CaCO2 cell lines. These cells come from a human colon adenocarcinoma, meaning it was originated at a secretary gland of the epithelium. This cell line was isolated and established in 1977 from a Caucasian male [15]. CaCO2 cells came from epithelial tissue and present tight junctions at confluence making them a good model to study cell to cell attachment. More recent studies revealed that, although they came from colon, their differentiation and behaviour are closer to enterocytes present at the jejunum, making them suitable for oral drug absorption predictions [16], [17], [18]. This cell line also presents the benefit that it is one of the few human cell lines that differentiate without needing external stimuli or other supplements [19]. Moreover, they also present cell polarity, differentiating into apical and basolateral physiognomy, and expresses some of the most common membrane transporters present at the enterocytes [20]. At microscopical level, they form a columnar monolayer with polygonal shape (Figure 2). Since they consist of a heterogeneous cell population [21], several minor differences can be observed, like different sizes and possibly different differentiation stages. Regardless of this differences, they still are a good model to study intestinal barrier functions like transport, subcellular localization of proteins and permeability studies.

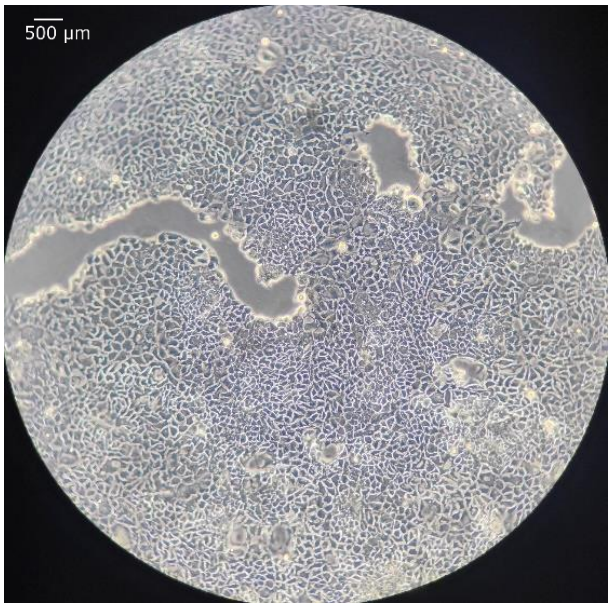


Figure 2; CaCO2 cells cultured approximately at 90% of confluence, observed at 200x in a phase contrast microscope. Here it is clearly visible the polygonal shape and heterogenicity from these cultures.

Although they are a non-homogeneous mix of cell phenotypes, the barrier they form lacks goblet cells, therefore, the jejunum model lacks a mucus layer and thus rends less meaningful the study of the transport of some drugs. On top of that, drug transport assays are very sensitive to culture conditions and require strict techniques and observation to obtain reliable results.

1.3 Occludin in tight junctions

In this work, we focused on the tight junction protein occludin, which role in tight junction formation and function is not fully understood. This protein is encoded by the *Ocln* gene, and the most common isoform consists of 522 amino acids. This isoform is considered the reference one due to its abundance compared to other isoforms. Occludin contains four transmembrane domains, one internal disulphide bond, three helix alpha motifs and several phosphorylation sites [22]. It has only been detected at tight junctions, in which it has interactions with zona occludens 1 and claudins [1], [5]. Moreover, occludin can carry several modifications like differential splicing and post-translational modifications. The former results in deletions and insertions in the mRNA sequence that produce different isoforms with altered functional domains. For example, an N-terminal extension that shields an SH3 domain which causes disfunction in the interaction with scaffold proteins has been reported [10]. Regarding post-translational modifications, they include phosphorylation which is involved in occludin dimer formation and other protein-protein interactions, and ubiquination, which is involved in the process of junction loss via proteasomal degradation [12].

Studies in occludin knock-out mice revealed postnatal growth retardation and male sterility. Histological studies did not show any evident change in TJ morphology but there was chronic inflammation of gastrointestinal tissues, calcification of the brain and thinning of compact bone among other minor differences in phenotype [23].

1.4 Previous approaches for the study of tight junctions

There are reports showing the expression of fluorescent occludin by tagging the N-terminus [24] or both ends of the protein [25]. These reports showed that the introduction of the tag does not affect occludin usual localization at tight junctions. Furthermore, it has been demonstrated that the C-terminus of the protein is involved in the localization process. In these reports, the authors used plasmids containing the cDNA, obtained from mRNA, to generate the fusion protein with the tags. Then, they transfected CaCO2 or MDCK cells in order to obtain stable transfectants. Although this is an easier approach to occludin labelling, it does not include possible alternative splicing at the transcript maturation stage, thus missing the possibility to include different isoforms of the protein. Moreover, integration and expression of fusion proteins, generated by cDNA cloning under the control of an exogenous promoter, often results in an overexpression of the protein, which may lead to artifacts. In addition, in stable cell lines one might end up with epigenetic silencing of the exogenous promoter. Our strategy consists on editing the genomic sequence, which consequently situates our fusion protein under the

regulation of the natural promoter and therefore increases resemblance with natural conditions.

Regarding the particular tag used, FLAG-tag [25], they present the same problems as before. This methodology consists of creating a fusion protein containing an epitope, which can be labelled later with antibodies *in vivo* or used for immunological techniques, such as co-immunoprecipitation. As it was said before, this approach does not consider the native forms of occludin, therefore the major part of the protein remains unlabelled. Moreover, the introduction of recombinant DNA encoding a fusion protein under the control of an exogenous promoter often results in unpredictable expression results.

Other studies created fusion proteins between other components of tight junctions like claudins and ZO proteins in order to study TJ assembly, dynamics and response to different stimulus. Some of these fusion proteins include ZO1-GFP [26], RFP-ZO-1, GFP-claudin1 and RFP-OCLN [27], and none of them reported negative influence of the tag as they presented normal growth and nominal values of membrane transport of substances and trans-epithelial electrical measurement. Moreover, the observation of the cells expressing these proteins showed normal distribution at the membrane in comparison to the ones expressing the native form of the same proteins labelled with fluorescent antibodies.

1.5 Genome edition with CRISPR-Cas9

To study the process of tight junction formation and the dynamics within this structure, we are planning to modify the genome of CaCO2 cells so that they produce fluorescently tagged variant forms of occludin under the control of its natural regulatory elements. The best way to achieve this goal nowadays is using the CRISPR-Cas9 system which allows to modify the cellular genome at will in an efficient manner [28].

CRISPR-Cas9 was discovered as an anti-viral defence system used by prokaryotic cells [29], which recognizes specific sequences in the genome and generates a double strand breaks, resulting in insertions and deletions that inactivate viruses. This system relies in an endonuclease associated with an RNA sequence composed by two active components; a short hairpin that associates with the nuclease, and a second part that plays the role of target recognition [30]. In nature, both parts are coded in the bacterial genome, and the guide-RNA is observed as tandem repeats on a list that can be extended with new viral sequences [31].

Based on its natural function, this mechanism has been used to generate organisms with a disrupted gene for functional studies [32], [33]. Furthermore, this technique can be applied on almost any kind of living being to be studied. In a more recent approach, this technique has been used to precisely edit regions on eukaryotic genomes by a process known as homology directed

repair, on which a template, from now on referred to as donor DNA, is used to rebuild the damaged DNA at the site of the double strand break [34]. This is a very low frequency event which can be enhanced by some methods as cell cycle synchronization and non-homologous end-joining inhibition, increasing the population of edited cells in a given experiment [35], [36].

2 Objectives

The availability of a cellular model system with properly regulated tagged-occludin will allow the monitoring of the protein in live cells, facilitating studies of tight junction formation, occludin dynamics and its interaction with other proteins

The aim of the present study was to obtain a CaCO2 derived cell line that expresses a tagged occludin under the control of the natural regulatory elements of the gene.

Specific aims:

1 Generate a plasmid encoding the Cas9 endonuclease associated with a specific guide RNA, to target the occludin gene at the 5' end of the coding sequence.

2 Generate a plasmid that will function as a donor/repair DNA containing the tag sequence flanked by homology arms to allow homologous recombination.

3 Obtain CaCO2 clonal cell lines, with targeted integration of the repair DNA, expressing the tagged-occludin protein.

3 Materials and methods

3.1 Materials

Oligonucleotides were ordered from Sigma and are listed in table 1(Appendix I). For cloning, PCR amplification was done with the high fidelity PrimeSTAR hot-start DNA polymerase (TAKARA). For routine PCR analysis, colourless NZYTaq master mix 2x (NZYtech) was used. For cell culture, medium, serum, glutamine and antibiotics were from Gibco (ThermoFisher). Salts, agarose and general reagents were purchased from Sigma.

Plasticware (plates, pipettes, pipette tips ...) were obtained from VWR, Eppendorf and Corning.

3.2 Plasmid construction

pLentiCRISPR V2-OCLN plasmid: The original plentiCrisprv2 vector was a gift from Feng Zhang (Addgene reference 52961). To generate specific target for OCLN gene two complementary oligos were selected to target a specific region on the OCLN gene, concretely right before the start codon on exon 2. Moreover, the target must be adjacent to a protospacer adjacent motif (PAM), so that Cas9 can align to the sequence to cut. The oligos were designed using the Benchling platform (https://www.benchling.com/crispr/), which shows the different PAM's present in the protein gene and provides scores to facilitate the selection of the most appropriate PAM. We uploaded sequences surrounding ATG at the beginning of exon two, and we chose the PAM at position 16938-16940 in the non-coding strand (NG_028291 RefSeqGene).

The two complementary oligos were aligned following the protocol published by Sanjana et al. [37] in order to generate target sgRNA which already contains overhangs, so we just have to prepare the plasmid for ligation. For this, we performed a heat-up and gradual cooling process which allows oligos to hybridise. Plasmid was cut open with BsmBI restriction enzyme and then treated with alkaline phosphatase. Then digested mix was run on a 1% agarose gel and the band corresponding to opened vector was cut off and purified using NZYGelpure kit(NZYTech). Next, 4 microlitres of the digested plasmid was mixed in an Eppendorf tube with one microlitre of the small guide annealed oligos for ligation with T4 DNA ligase. Ligated plasmids were transformed onto Top10 *E. coli* competent cells. This mixture was seeded onto LB plus ampicillin plates and cultured at 37 degrees until colonies were visible. First assessment of clones was performed with DNA extraction followed by XhoI digestion which was expected to give bands of 8, 1.6, 1.4, 1.2 and 0.2 kilobases Once a clone showed the expected band pattern, the plasmid was sequenced and produced at the mg scale using the Maxi Prep kit form Nzytech.

pBluescript-Repair plasmid: original PBluescript was obtained from Agilent. This plasmid was modified to generate a specific repair DNA sequence useful as template for homologous directed repair at the double strand break generated by Cas9. For its construction, we first amplified sequences corresponding to chromosome homology arms upstream and downstream of the predicted double-strand break. The upstream homology arm (UHA) was obtained by amplification from HTC116 cells genomic DNA. The forward primer contained a 5' extension with a SalI restriction site for easier cloning in the plasmid. The reverse primer contained 21 base pairs to anneal the genomic template and a 5' extension, of another 21 bases, to anneal the mEmerald coding sequence for later overlap extension. For the downstream homology arm (DHA), we designed a nested PCR based strategy to avoid contamination with other genome fragments that may be co-amplified with the primers initially designed. The nested primers also contained modifications. The forward primer contained a 28 base pair extension at the 5' end able to anneal the mEmerald PCR product. The reverse primer contained a BamHI restriction site at its 5' end for cloning purposes. Finally, we designed two primers to amplify the mEmerald sequence together with a linker using plasmid mEmerald-claudin7-C-12 (gift from Michael Davidson(Addgene reference 54041) as template. To ensemble the three fragments UHA+mEmerald-linker+DHA into one single molecule, the three units were joined by sequence overlap extension and finally cloned into the opened pBluescript vector by conventional cloning. Transformation was set up using Top10 competent cells and blue-white selection in S-gal ampicillin plates(Appendix I). White clones were grown on ampicillin LB medium, plasmid DNA was extracted and digested with XhoI for band pattern analysis. Clones with the expected band pattern were further analysed by sequencing. One clone was selected and used to prepare a large amount of plasmid using Maxi Prep kit from Nzytech.

All sequencing procedures were ordered to the company Eurofins Genomics, which returned the chromatogram files from our sequences. All files were loaded into ChromasPro software to obtain a consensus sequence from all the fragments, which will be aligned to the expected sequence.

3.3 Cell culture

CaCO2 cells were obtained from Dr. Raquel Martin (University of Barcelona, pharmacy faculty, ATCC reference HTB37) and cultured in DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin/glutamine solution provided as a 100X stock (ThermoFisher 10378016) (D10). Cells were cultured at 37 ^{o}C, 5% CO_2 in a humidified incubator. Passages were performed at a confluence superior to 70%. To this end, cells were washed three times with PBS (5 mL were usually used per each wash when using 10 cm diameter plates) and then trypsin was added and kept in the plate till the majority of the cells were detached from the surface (usually 1,5 mL of

trypsin solution were added per 10 cm plate and the incubation time was around 5 minutes). Cells were finally collected with culture media and seeded as needed for experiments.

To keep a stock of derived clones, cell freezing was performed by trypsinization, as described above and centrifuged at 1200 rpm for 5 minutes. Supernatant was removed and cells were resuspended in freezing medium (DMEM with 20% FBS and 10% DMSO) at 4ºC. Cells were frozen by gradually cooling down (1 ºC/minute) to -80 ºC. 3 vials containing 1 mL of cell suspension were obtained from each 10 cm diameter plates at confluency.

For transfections, cells were plated at 20% confluency and the next day they were transfected with a mixture of polyethilenimine (PEI) – DNA. PEI is a highly positively charged molecule that binds effectively to DNA, neutralizing negative charges in DNA and facilitating entry to the cell. We used 4 μL of PEI at 1mg/mL in water and pH 7.4 (linear PEI, MW 25000; Polysciences incorporation 9002-98-6) per 1 μg of DNA. Our transfection mixture contained 0.5 μg of plentiCrisprV2-OCLN, 5.5 μg of pBluescript-OCLN and 24 μL of PEI in 1 mL of serum free DMEM. Prior to mixing PEI with DNA, PEI was diluted in serum free medium and incubated for 5 minutes (usually 24 microliters of PEI in 1 mL of DMEM), then this was added to a tube containing the DNA, in a very small volume (5 to 50 μL). This final mixture was incubated for 15 minutes at room temperature and added to a 10 cm plate containing 30-40% confluent CaCO2 cells in 10 mL of D10.

For the selection of transfected cells, 24-48 hours after transfection we added puromycin (5μg/mL; ThermoFisher 10378016) for 72 hours. This resulted in an almost complete cell death for non-transfected control plates. At this time point, medium was replaced with D10 and cells were allowed to expand for two weeks and sorted according to their m-Emerald expression using the equipment FACSVantage at the CRG facilities in Barcelona. Sorting was done using single cells obtained by trypsinization and filtering to avoid clumps. DAPI staining (Molecular Probes) was used to exclude dead cells. All data were analysed with FlowJo software (Tree Star). Individual cells were placed in separated wells of 96 well plates and allowed to expand till filling up over 30% of the surface. Then cells were subsequently expanded in larger surfaces till enough cells were available to freeze two aliquots and get cell lysates to isolate genomic DNA and total proteins. Individual clones were also plated on glass bottom plates to get fluorescence images with Leica DMI 8 inverted microscope, using phase contrast and epifluorescence. Images were acquired with a microscope-coupled Photometrix prime 95B camera.

3.4 Genomic DNA isolation and PCR analysis

CaCO2 cells genomic DNA was obtained by scrapping cells grown at 50-80% confluency in PBS, pelleting them at 1600 rpm for 3 minutes and resuspending them in TE-Proteinase K buffer. The cell suspension was then incubated at 50 ºC for 1 hour and then DNA was precipitated by adding

Na-acetate and 1 volume of isopropanol (Appendix I). After shaking carefully but extensively, genomic DNA was evidenced as a jelly material in suspension which could be recovered by sticking it to the exterior of a pipette tip. DNA was transferred to a new tube, washed, briefly dried and resuspended in 10 mM Tris pH8(Appendix I). After quantification, 100 ng of genomic DNA were used in each PCR reaction to amplify the regions of interest by PCR using a DNA polymerase with high proof-reading activity. PCR products were then sequenced using Sanger's method. Primers used to characterize the edited genome are listed in table S1.

3.5 Western Blot analysis

To analyse tagged-occludin expression, cells were lysed in 1X reducing protein loading buffer (Appendix I), sonicated at 80% amplitude for 10 minutes at 10 seconds on/off intervals using the model FB120 Sonic Dismembrator (Fisher Scientific) and boiled for 15 minutes. The amount corresponding to approximately 10^5 cells (typically 15 μL) was loaded in 12.5% polyacrylamide gels, proteins were separated according to their molecular weight and transferred to PVDF membranes. Membranes were blocked in 5% non-fat dried milk, incubated overnight with the appropriate primary antibody, either anti-occludin (OC-3F10; ThermoFisher) or anti-GFP (Clonetech) (green fluorescent protein shares almost all structure with mEmerald), extensively washed and incubated with HRP conjugated secondary antibodies (DAKO). Membranes were then washed and developed by chemiluminescence using clarity western ECL substrate (BioRad; 1705061). Images were obtained with a Versadoc™ 4000 imaging system (BioRad) (Appendix I).

3.6 Single Particle Tracking(SPT) analysis

We performed fluorescence live-cell imaging experiments and analyse them to determine whether our fusion protein had potential for being used in single molecule applications, such as SPT [38]. To this end, we cultured cells and observed them at different culture times using a Leica DMI 8 inverted microscope with 100x objective coupled to a Photometrix prime 95B camera. Regarding the illumination geometry, we used total internal reflection fluorescence (TIRF). This geometry allows us to reduce background fluorescence and observe only a slice of about 90-150 nm above the glass slide of the culture plates, by changing the angle of incidence of a 488nm-wavelength CW laser. We recorded 500-frames-long videos of different field of views at different frame rates, ranging between 10 to 100 Hz.

The fluorescence live-cell videos obtained as described above were analyzed using the TrackMate plugin [39] of ImageJ analysis tool (https://imagej.net/Fiji). This allowed us to identify fluorescence spots corresponding to single molecules, determine their X and Y coordinates at each frame and reconnect them in time to generate trajectories. First, the detection parameters,

defining typical size and intensity, are given to the program and it localizes all particles compatible with those characteristics. Next, detected spots are linked in time using other parameters that connect them frame to frame. Finally, one CSV file is generated for each video and further analysed with MATLAB. We used different custom scripts in MATLAB that allow us to generate a matrix with trajectories information and joins all the matrices from the same experiment. For the quantification of the trajectories, we calculate the short time-lag diffusion constant D by a linear fit of the first points of the mean-squared displacement plot [38]. The diffusion coefficients of all the trajectories were displayed as a histogram depicting the distribution of the diffusion in a given experiment. The histogram also contains a vertical bar labelled as $D_{Boundary}$. Its value represents a statistical threshold for identifying a particle as mobile or immobile and is associated to the limited precision introduced by the imaging noise and by the spot localization procedure. This value was estimated from the analysis of movies obtained with fixed cells [38].

4 Results and discussion

Protein fusion generation has been proved to be a useful technique that allows to study proteins and their interactions, improve recombinant protein purification methods and generate more effective enzymes for industrial processes. In this study, we aimed to develop a new CaCO2-derived cell line expressing a fusion protein, tagged occludin, that allows the study of tight junctions and occludin in living cells and tested it with analytical methods.

In order to accomplish our goal, we generated two plasmids, one encoding Cas9 endonuclease and a sgRNA to direct the endonuclease to the occludin gene, and a second plasmid, the repair plasmid, bearing the sequences to be modified in the genome. We next transfected the plasmids into CaCO2 cells, selected genome-edited cells and characterized them.

4.1 Plasmid construction

To generate pLentiCRISPRV2-OCLN, the original plasmid was cut with BsmBI, the relevant resulting fragment was purified and treated with alkaline phosphatase. Two complementary oligonucleotides, with the relevant sequence to direct Cas9 to the intended target site and appropriate 5' extensions, were aligned, phosphorylated and ligated to the vector. After transformation of competent cells, the generated plasmid was amplified and characterized by sequencing. Figure 3 shows the result of the analysis of the plasmid isolated from 4 transformed colonies. Data indicates that the process is very efficient since 100% of the colonies analysed yielded the expected restriction pattern after PstI digestion which clearly distinguishes the original vector from the desired molecule. One of the plasmids was selected and further analysed by sequencing using Sanger's method at the company Eurofins Genomics.

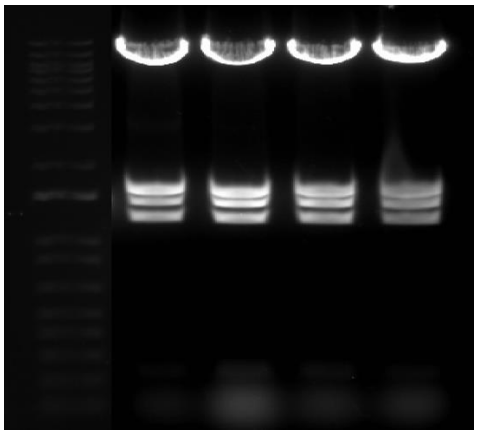


Figure 3; Agarose gel electrophoresis showing the PstI band pattern corresponding to 4 putative pLenti CRISPR V2-OCLN plasmids. The 4 clones gave the expected result, presenting bands at 1.6, 1.4 and 1.2 Kb, pattern that differs from that expected with the original vector.

The sequence obtained was aligned to the expected sequence using Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and confirmed to be correct (Figure 4). Once confirmed, a large-scale plasmid prep was prepared to continue with the project.

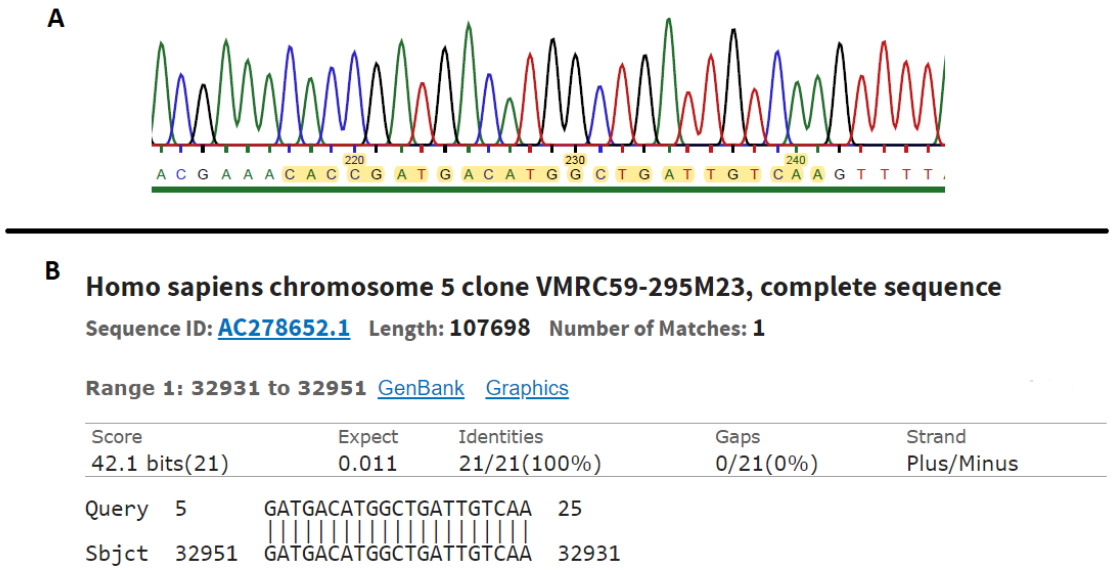


Figure 4; Chromatogram showing correct insertion of guide oligo(highlighted in yellow) in pLenti CRISPR-OCLN(A). Alignment of the choosed oligo against human genome, yielding 100% identity to the expected region in chromosome 5(B).

The pBluescript-OCLN-repair plasmid was generated by cloning the repair fragment into the properly linearized pBluescript vector. The repair fragment was produced by first amplifying homology arms and mEmerald coding sequence from HTC116 genomic DNA and plasmid mEmerald-claudin7-C-12, respectively. When necessary, PCR products were purified from an agarose gel to discard undesired amplicons using the Qiaquick™ gel purification kit (Qiagen). Primers had been designed with overlapping 5' extensions that permitted latter sequence overlap extension to create the final repair DNA. To facilitate cloning, the most external primers contained unique restriction sites (BamHI and SalI) that were used to create compatible ends with the vector. After restriction enzyme digestion, DNA purification and ligation, competent cells transformed, individual clones were amplified and, after plasmid purification, digestion with XhoI revealed successful cloning (Figure 5). One of the clones was chosen for sequencing and confirmed to be correct.

For sequencing, a single clone was chosen and a sample of its miniprep was sent to analyse by Sanger. Sequences were aligned to the expected sequence created *in silico*. A colony with the confirmed sequence was further amplified for large-scale plasmid preparation using a Maxiprep kit (NZYtech).

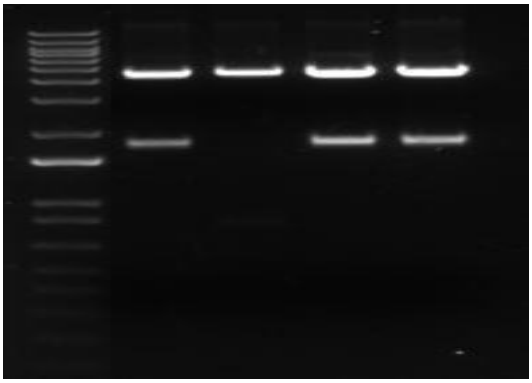


Figure 5; DNA gel electrophoresis showing band pattern of the four analysed clones. Upper band corresponds to 4.3Kb and lower band to 1.7Kb. It is clearly visible that clone in lane 3 is negative while the remaining ones at lanes 2, 4 and 5 are positive.

4.2 Cell edition

We initially set up the conditions for successful transfection and edition by testing different ratios of the two plasmids required for the process, plenticrisprv2-OCLN and pBluescript-OCLN-repair. Our standard use of 4 mg of PEI per mg of plasmid DNA in transfection mixtures was found to be appropriate in our experimental design, using CaCO2 cells, without further need for improvement (Figure 6).

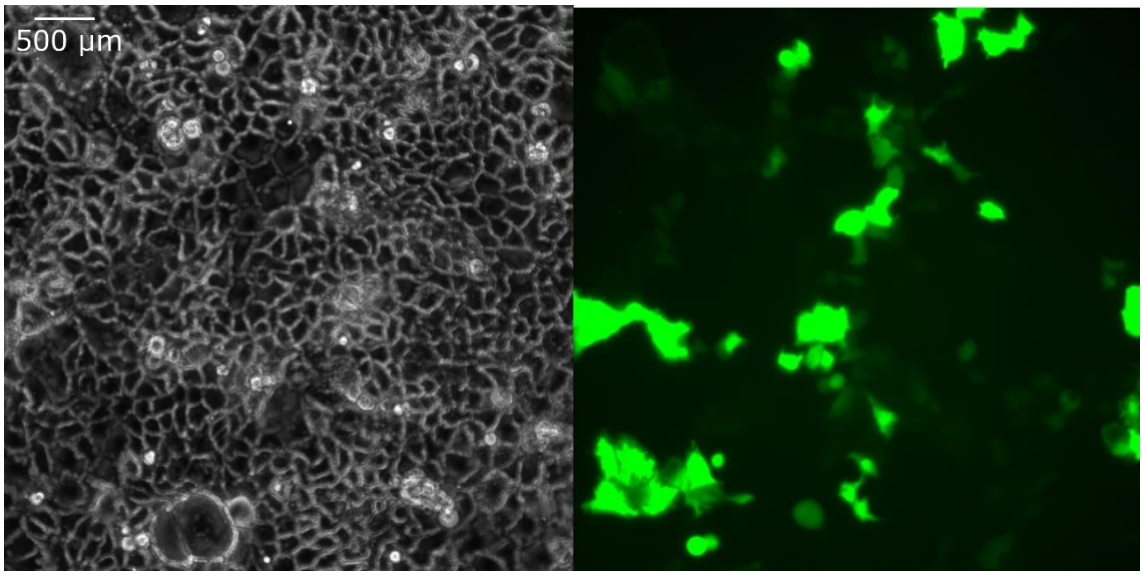


Figure 6; Image showing transfected cells expressing mEmerald protein in contrast to whole population. Phase contrast in left picture, fluorescent imaging in the right picture.

Since transfections were done with a mixture of the two plasmids, we reasoned that adding an excess of repair plasmid and selecting for the cells that received the DNA-*breaking* plasmid would maximize the recovery of successfully edited cells. At first, we decided to use mixtures of the two plasmids with molecular ratios of 1:10, 1:20 and 1:50, incubating the cells for 24 hours and replacing the medium with a puromycin containing medium (5-10 mg/L) to remove untransfected cells. The appropriate concentration of the drug was chosen based on a killing curve that we run (data not shown). Puromycin was only left in the medium for 3-4 days since we were not intending to select cells that integrated the DNA-*breaking* plasmid but just to keep in culture cells that had expressed such plasmid, mainly transiently. In our reasoning, these cells would have suffered a chromosome break just in exon 2 of the occludin gene and would have undergone DNA repair in the presence of a high amount of the repair DNA resulting in DNA repair through, to some extent, homologous recombination. Repair through non-homologous end joining, more frequent in these cells, would not yield cells expressing the fluorescent protein and should therefore be easy to discard. Two weeks after the transfection, when transient expression has mostly disappeared, we could observe that we had cells expressing a green fluorescent protein in the culture plates. The proportion of cells expressing such protein was around 10%. Importantly, many of these cells were showing a membrane location for the fluorescent protein, as we were expecting for the mEmerald-occludin fusion(Figure 7). At these stage, we decided to sort the green cells and plate individual cells in wells of 96 well plates to recover clones with homogeneous genome edition.

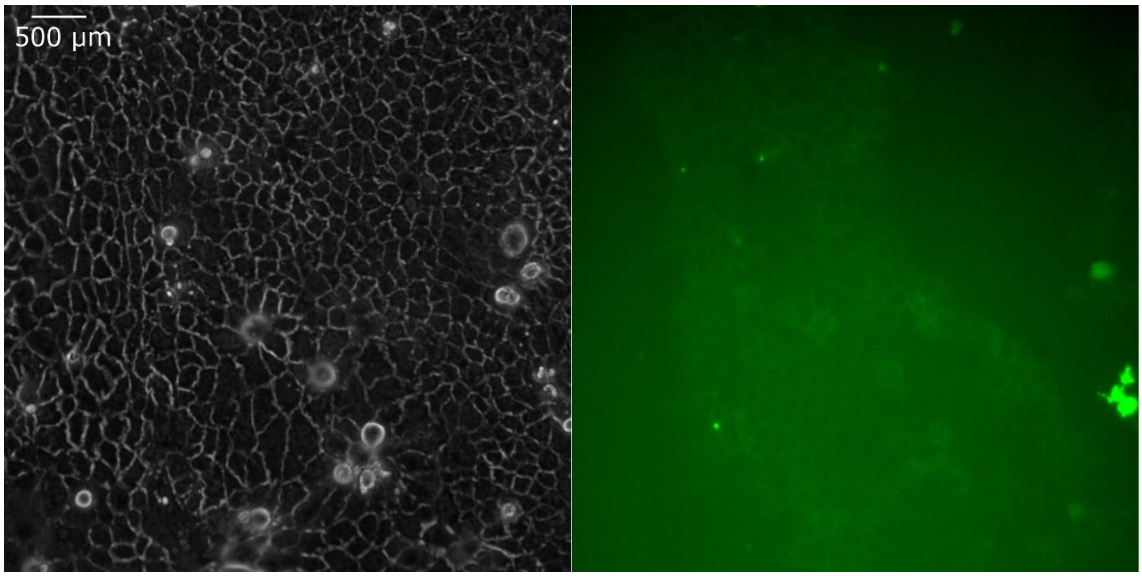


Figure 7; Picture showing clear differences among cell populations. Only the cells at the middle of the plane are edited and thus fluorescent(right image), while the ones in the sides are not modified and are only visible in phase contrast image(left)

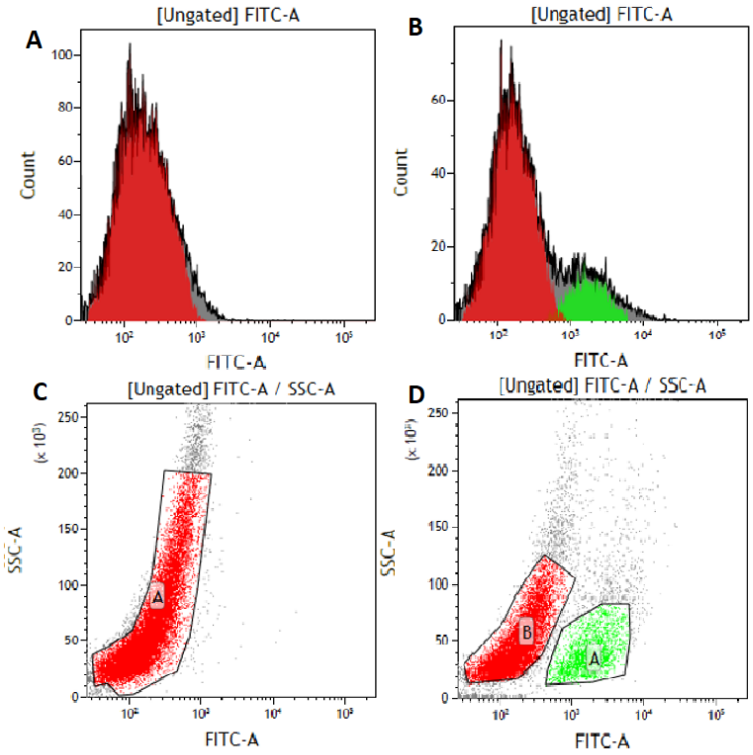
It must be mentioned that we were ready to modify our strategy by including inhibitors of NHEJ in the transfected cells, or modifying culture conditions, as means to increase the frequency of HDR [35], [36].

4.3 Clone selection

Edited cell pools were subcultured until we obtained three confluent 75 cm^2 flasks. At this point, one of them was frozen as stock and the remaining two followed the selection protocol. To recover single clones expressing the fusion protein, cells were trypsinized, filtered to eliminate clumps and run through a FACS-Vantage flow cytometer sorting instrument to recover cells expressing m-Emerald. DAPI was included in the buffer to discard dead cells. A preliminary cell cytometer analysis was performed before sorting in order to gain information about the heterogeneity in the cell pool and stablish gating strategy. We observed two differentiated groups; one composed by normal cells which resisted the puromycin treatment but did not maintain the mEmerald protein expression, and a second group which was positive for fluorescence (Figure 8). At this stage, we were happy to learn that around 13% of the cells were fluorescent prompting the selection by flow cytometry (Figure 8, E). Although these data are not directly related to transfection efficiency, they can be associated with edition efficiency in our setting, including the puromycin-mediated enrichment. Assuming that all the cells present in the pools 20 days after the transfection expressed the DNA-*breaking* construct and received the Repair-DNA, having over 10% of cells expressing the fusion protein suggests that the experimental model and design yield a pretty high efficiency of DNA repair by homologous recombination.

Individual cells expressing mEmerald were plated in wells of 96 well plates, using FACS-Vantage.

Regarding the gating strategy, we designed a protocol which first isolates cells negative to DAPI staining (Figure 9, A and C) and then separates them by FITC signal intensity, isolating the pools P4 and P5 from the positive cells (Figure 9, B and D). Figure 8 also demonstrates that our protocol distinguishes edited cells (top row) from an unedited population of CaCO2 cells (bottom row).



E

Gate	Number	%Total	%Gated
All	14.385	100,00	100,00
A	1.952	13,57	13,57
B	10.074	70,03	70,03

Figure 8; Flow cytometry analysis of a transfected cell pool. (A and C) Untransfected cells. (B and D) transfected cells. (A and B) Cell counts versus fluorescence intensity.(C and D) cell distribution according to size(y axis) and fluorescence(x axis). E shows cell counts in D.

Figure 9 shows FACS data associated to this particular cell isolation. Sorted cells were expanded for several weeks until above 30% of the culture surface was populated with cells. At this stage, cells were checked under the microscope and confirmed to express mEmerald (18 out of 18 clones that were successfully expanded expressed the mEmerald-OCLN fusion protein). Clones were expanded by successively increasing the culture plate surface, from 96 well plates to 24 well plates, then to 6 well plates and finally to 10 cm diameter plates. Once enough cells were available, 2 vials, with $5 \cdot 10^6$ cells each, were frozen and aliquots of cells were plated back to check for fluorescence distribution pattern, western blot analysis and genome edition assessment.

Unfortunately, a few of the clones failed to arrive at the last stage due to diminished growth abilities while being cultured alone, over-trypsinization or accidental aspiration of the smaller colonies. However, we improved the process and succeed at each clone expansion.

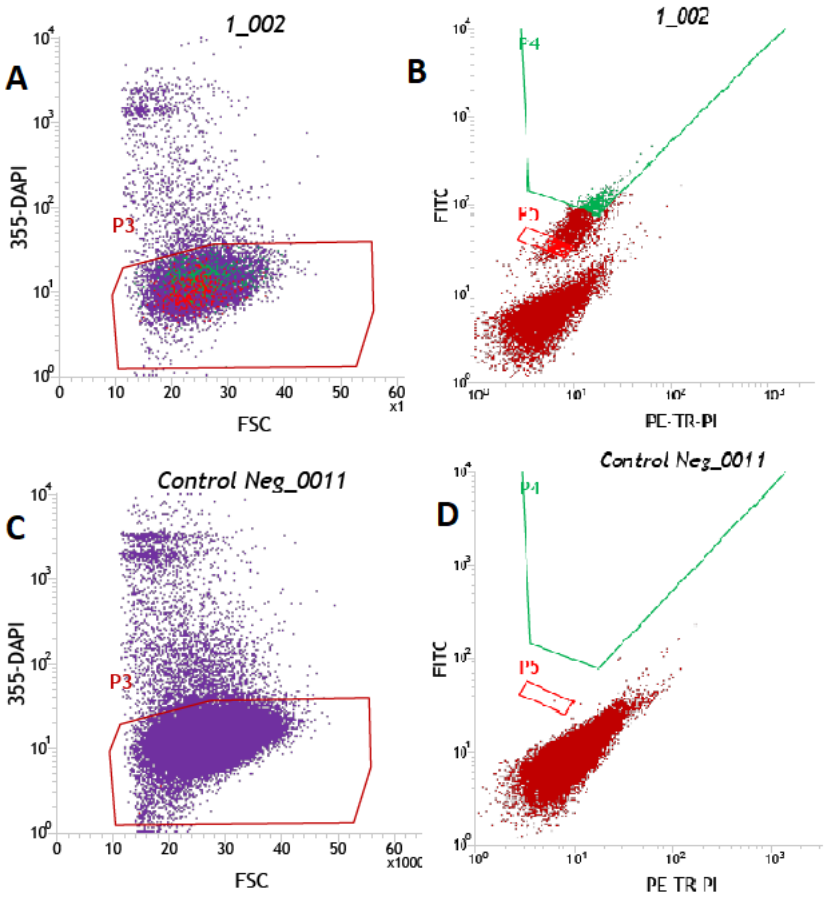


Figure 9; Gating strategy used to isolate edited cells from cultured pools. P3 separates cells negative to DAPI staining while P4 and P5 isolate fluorescence positive cells.

4.4 Clone characterization

To evaluate the relevant phenotype and genotype of the derived clones, cells were maintained in culture and analysed by fluorescence microscopy, western blot analysis and partial genome sequencing. To assess the fluorescence distribution, cells were plated in glass-bottom plates and when they reached more than 50% of confluence, to ensure tight junction presence, they were observed under 460nm light. All clones were analysed and reported positive, observing the expected fluorescence at the membrane (Figure 10). We can confirm that emission intensity was lower than in other fluorescent tags as we had to use high values of exposure time in order to acquire clear images. This can be due to expression restricted to natural OCLN promoter, resulting in lower concentration of the fluorescent molecule.

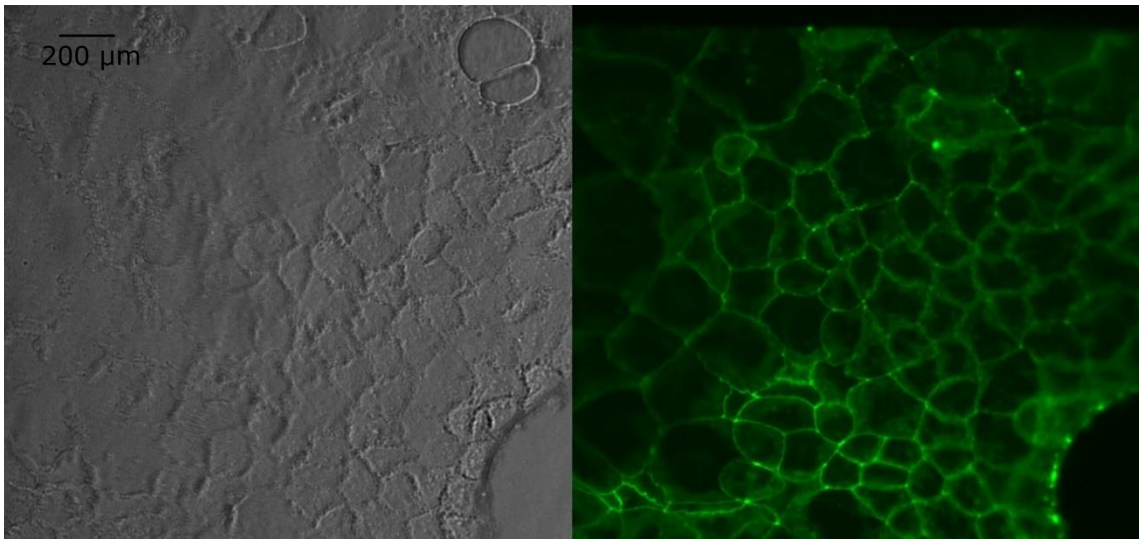


Figure 10; Fluorescence images from clone P4C8. On the left cells are observed in bright field. On the right only fluorescent cells can be observed. 100% of the cells in the clonal population present green fluorescence located at their membranes.

An additional piece of phenotypic information at the molecular level was provided by western blot analysis. Confluent cells were scrapped, pelleted and resuspended in 1x Laemmli sample buffer for western blot analysis(Appendix I). The occludin-specific antibody detected both, the wild type protein as a doublet at about 65 kD (corresponding to different occludin isoforms) and a new band of about 95 kD which is missing in non-transfected cells and matches the expected size of the fusion protein (Figure 11). This band is also detected with a GFP-specific antibody strongly suggesting that it corresponds to the fusion protein. Moreover, the bands corresponding to the natural forms of occludin reveals that not all the chromosomes were edited, leaving at least one copy of the OCLN gene unedited or repaired by NHEJ without the loss of open reading frame since the DNA is expected to be cut 16 nucleotides 5' of the starting ATG.

At this point, we have only analysed 8 of the clones and they all present the same band pattern with the antibodies used.

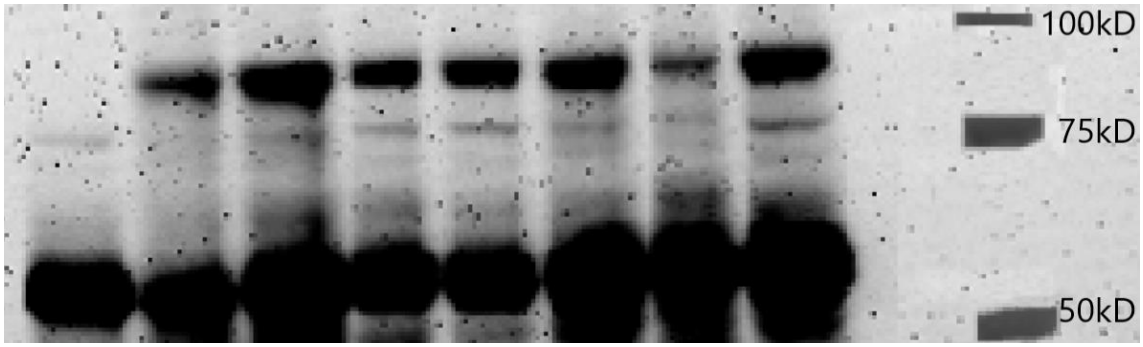


Figure 11; Western blot from control CaCO2 and clones P4C8, P5E8, P5E3, P5G10, P5E11, P4E4 and P5G5 respectively, followed by empty lane and molecular weight marker. It is clearly visible that all clones are positive in a band at around 92kD which control sample lacks, suggesting the presence of OCLN-mEmerald fusion protein.

Finally, we want to verify genome edition by sequencing the genome of the clones in a region of about 4 Kb which include the repair-DNA. To this end, we have isolated genomic DNA from a set of clones, and we are in the process of PCR-amplifying this 4KB genome region to proceed with the sequencing.

4.5 SPT applications

We obtained several videos corresponding to different experimental conditions of the cells and processed them with the ImageJ's plugin TrackMate [39]. Already while analysing and visually inspecting the tracks one by one, we notice the occurrence of two different diffusion patterns: (i) the first (Figure 12-A) corresponding to nearly immobilized particles, similar to the one observed in fixed cells; and (ii) the second (Figure 12-B) similar to the erratic and heterogeneous motion observed for several proteins in the cell membrane [38], alternating different levels of mobility.

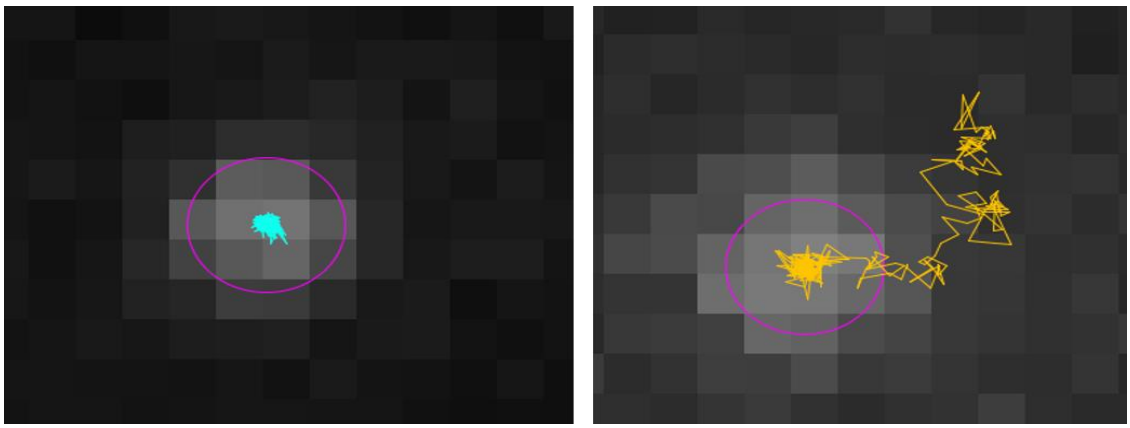


Figure 12; Tracks generated with TrackMate showing a clear difference between immobilized(left) and moving(right) particles

In order to quantify the observed mobility, we calculated the diffusion coefficient of each trajectory and represented their values in a histogram (Figure 13, blue histograms). The histogram shows a broad distribution of diffusion coefficients peaked at $\sim 0.1\ \mu m^2/s$, compatible with our visual observation of mobile and immobile particles, and consistent with values obtained by similar experiments for other membrane proteins. It is worth to mention that, to the best of our knowledge, this is the first measurement of the diffusion of occludin at the single molecule level.

As a further test, we compared the diffusion results of mEmerald-occludin cells to results obtained by another set of experiments in which we labeled occludin using an antibody [40] recognizing its extracellular loop linked to a quantum dots with same culture time. As figure 13 shows, the results obtained from our fusion protein match the observed in analyses from quantum dots. The slight differences that can be observed in the amount of immobilized particles (left side of histogram) could be attributed to a size or hindering effect. In fact, as

quantum dot-antibody-protein complexes have a size of approximately 20 nm and lie outside the cells, they may occasionally be hindered at the membrane-glass interface and thus appear as immobilized particles.

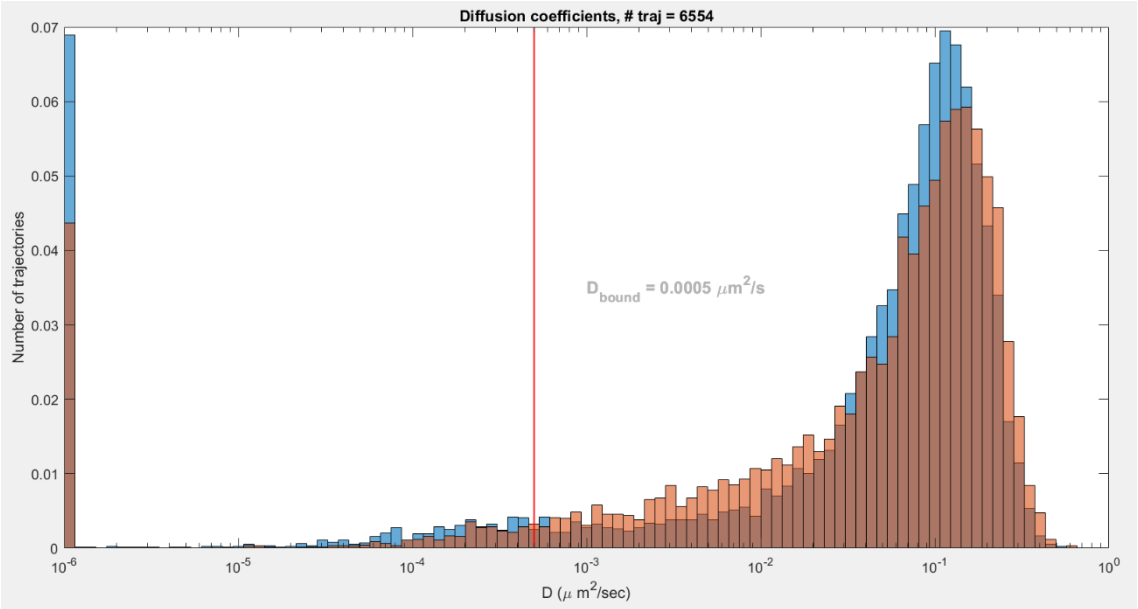


Figure 13; Histogram depicting the distribution of the diffusion of particles calculated with mean square displacement. Red histogram corresponds to quantum dot labelling of occludin, and blue histogram correspond to our cell line.

We can conclude that the fusion protein we generated can be used for single molecule applications such as SPT. The diffusion properties we retrieved from these experiments are comparable to those obtained with antibody and quantum dot labeling. Although mEmerald photophysics makes it more prone to bleaching, it stills performs rather well for tracking experiments. In addition, this kind of labelling is less invasive as the protein can move freely, nearly as in natural conditions; and it also allows to label a large part of the protein population. Modified occludin appears to function properly as no anomalous effects are observed in the cells and the cellular localization from fluorescence observation matches the expectations, thus offering a promising tool for future applications such as super-resolution fluorescence imaging.

6 Conclusions

We managed to accomplish our objectives first by generation of a new plasmid (pLentiCrisprV2-OCLN) that cuts the human genome in exon 2 of the occludin gene. Moreover, we generated another plasmid(pBluescript-OCLN-repair) containing mEmerald coding sequence flanked by occludin genomic sequences, which can be used as a donor template to repair a DNA double strand break in exon 2 of the human occludin gene by homologous recombination.

Transfection of both plasmids into CaCO2 cells allows the recovery of clonal populations producing a mEmerald-tagged occludin under the natural occludin promoter and regulatory elements. All the edited and isolated clones present the mEmerald-occludin protein located at the plasma membrane. This modification opens new perspectives of observing tight junction dynamics in live cells.

Moreover, our cell line proved to be useful in SPT applications since the experiments we conducted proved to yield results similar to other labelling methods.

7 Bibliography

[1] K. Shin, V. C. Fogg, and B. Margolis, “Tight Junctions and Cell Polarity,” *Annu. Rev. Cell Dev. Biol.*, vol. 22, no. 1, pp. 207–235, 2006, doi: 10.1146/annurev.cellbio.22.010305.104219.

[2] M. G. FARQUHAR and G. E. PALADE, “Junctional complexes in various epithelia.,” *J. Cell Biol.*, vol. 17, pp. 375–412, 1963, doi: 10.1083/jcb.17.2.375.

[3] B. M. Gumbiner, “Breaking through the tight junction barrier,” *J. Cell Biol.*, vol. 123, no. 6 II, pp. 1631–1633, 1993, doi: 10.1083/jcb.123.6.1631.

[4] L. González-Mariscal, R. Tapia, and D. Chamorro, “Crosstalk of tight junction components with signaling pathways,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1778, no. 3, pp. 729–756, 2008, doi: 10.1016/j.bbamem.2007.08.018.

[5] S. Citi, “The mechanobiology of tight junctions,” *Biophys. Rev.*, vol. 11, no. 5, pp. 783–793, 2019, doi: 10.1007/s12551-019-00582-7.

[6] S. Tsukita, H. Tanaka, and A. Tamura, “The Claudins: From Tight Junctions to Biological Systems,” *Trends Biochem. Sci.*, vol. 44, no. 2, pp. 141–152, 2019, doi: 10.1016/j.tibs.2018.09.008.

[7] M. K. Findley and M. Koval, “Regulation and roles for claudin-family tight junction proteins,” *IUBMB Life*, vol. 61, no. 4, pp. 431–437, 2009, doi: 10.1002/iub.175.

[8] D. Günzel, “Claudins: vital partners in transcellular and paracellular transport coupling,” *Pflugers Arch. Eur. J. Physiol.*, vol. 469, no. 1, pp. 35–44, 2017, doi: 10.1007/s00424-016-1909-3.

[9] A. Hartsock and W. J. Nelson, “Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1778, no. 3, pp. 660–669, 2008, doi: 10.1016/j.bbamem.2007.07.012.

[10] P. M. Cummins, “Occludin: One Protein, Many Forms,” *Mol. Cell. Biol.*, vol. 32, no. 2, pp. 242–250, 2012, doi: 10.1128/mcb.06029-11.

[11] O’connor, “乳鼠心肌提取 HHS Public Access,” *Physiol. Behav.*, vol. 176, no. 1, pp. 139–148, 2016, doi: 10.1016/j.physbeh.2017.03.040.

[12] T. Murakami, E. A. Felinski, and D. A. Antonetti, “Occludin phosphorylation and ubiquitination regulate tight junction trafficking and vascular endothelial growth factor-induced permeability,” *J. Biol. Chem.*, vol. 284, no. 31, pp. 21036–21046, 2009, doi: 10.1074/jbc.M109.016766.

[13] X. Liu *et al.*, “Occludin S490 Phosphorylation Regulates Vascular Endothelial Growth Factor–Induced Retinal Neovascularization,” *Am. J. Pathol.*, vol. 186, no. 9, pp. 2486–2499, 2016, doi: 10.1016/j.ajpath.2016.04.018.

[14] H. K. Campbell, J. L. Maiers, and K. A. DeMali, “Interplay between tight junctions & adherens junctions,” *Exp. Cell Res.*, vol. 358, no. 1, pp. 39–44, 2017, doi: 10.1016/j.yexcr.2017.03.061.

[15] J. Fogh, J. M. Fogh, and T. Orfeo, “One hundred and twenty seven cultured human tumor cell lines producing tumors in nude mice,” *J. Natl. Cancer Inst.*, vol. 59, no. 1, pp. 221–226, 1977, doi: 10.1093/jnci/59.1.221.

[16] P. Shah, V. Jogani, T. Bagchi, and A. Misra, “Role of Caco-2 cell monolayers in prediction of intestinal drug absorption,” *Biotechnol. Prog.*, vol. 22, no. 1, pp. 186–198, 2006, doi: 10.1021/bp050208u.

[17] Y. Yang, Y. Zhao, A. Yu, D. Sun, and L. X. Yu, “Oral drug absorption: Evaluation and prediction,” *Dev. Solid Oral Dos. Forms Pharm. Theory Pract. Second Ed.*, pp. 331–354, 2017, doi: 10.1016/B978-0-12-802447-8.00012-1.

[18] I. de Angelis and L. Turco, “Caco-2 cells as a model for intestinal absorption,” *Curr. Protoc. Toxicol.*, no. SUPPL.47, pp. 1–15, 2011, doi: 10.1002/0471140856.tx2006s47.

[19] S. M. Moyes, J. F. Morris, and K. E. Carr, “Culture conditions and treatments affect Caco-2 characteristics and particle uptake,” *Int. J. Pharm.*, vol. 387, no. 1–2, pp. 7–18, 2010, doi: 10.1016/j.ijpharm.2009.11.027.

[20] T. Han *et al.*, “Organic cation transporter 1 (OCT1/mOct1) is localized in the apical membrane of Caco-2 cell monolayers and enterocytes,” *Mol. Pharmacol.*, vol. 84, no. 2, pp. 182–189, 2013, doi: 10.1124/mol.112.084517.

[21] E. Walter and T. Kissel, “Heterogeneity in the human intestinal cell line Caco-2 leads to differences in transepithelial transport,” *Eur. J. Pharm. Sci.*, vol. 3, no. 4, pp. 215–230, 1995, doi: 10.1016/0928-0987(95)00010-B.

[22] A. Y. Andreeva, E. Krause, E. C. Müller, I. E. Blasig, and D. I. Utepbergenov, “Protein Kinase C Regulates the Phosphorylation and Cellular Localization of Occludin,” *J. Biol. Chem.*, vol. 276, no. 42, pp. 38480–38486, 2001, doi: 10.1074/jbc.M104923200.

[23] M. Saitou *et al.*, “Complex phenotype of mice lacking occludin, a component of tight junction strands,” *Mol. Biol. Cell*, vol. 11, no. 12, pp. 4131–4142, 2000, doi: 10.1091/mbc.11.12.4131.

[24] S. M. Stamatovic, R. F. Keep, M. M. Wang, I. Jankovic, and A. V. Andjelkovic, “Caveolae-mediated internalization of occludin and claudin-5 during CCL2-induced tight junction remodeling in brain endothelial cells,” *J. Biol. Chem.*, vol. 284, no. 28, pp. 19053–19056, 2009, doi: 10.1074/jbc.M109.000521.

[25] Y. H. Chen, C. Merzdorf, D. L. Paul, and D. A. Goodenough, “COOH terminus of occludin is required for tight junction barrier function in early Xenopus embryos,” *J. Cell Biol.*, vol. 138, no. 4, pp. 891–899, 1997, doi: 10.1083/jcb.138.4.891.

[26] F. K. Riesen, B. Rothen-Rutishauser, and H. Wunderli-Allenspach, “A ZO1-GFP fusion protein to study the dynamics of tight junctions in living cells,” *Histochem. Cell Biol.*, vol. 117, no. 4, pp. 307–315, 2002, doi: 10.1007/s00418-002-0398-y.

[27] V. S. Subramanian, J. S. Marchant, D. Ye, T. Y. Ma, and H. M. Said, “Tight junction targeting and intracellular trafficking of occludin in polarized epithelial cells,” *Am. J. Physiol. - Cell Physiol.*, vol. 293, no. 5, pp. 1717–1726, 2007, doi: 10.1152/ajpcell.00309.2007.

[28] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang, “Genome engineering using the CRISPR-Cas9 system,” *Nat. Protoc.*, vol. 8, no. 11, pp. 2281–2308, 2013, doi: 10.1038/nprot.2013.143.

[29] F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, and E. Soria, “Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements,” *J. Mol. Evol.*, vol. 60, no. 2, pp. 174–182, 2005, doi: 10.1007/s00239-004-0046-3.

[30] H. Nishimasu *et al.*, “Crystal structure of Cas9 in complex with guide RNA and target DNA,” *Cell*, vol. 156, no. 5, pp. 935–949, 2014, doi: 10.1016/j.cell.2014.02.001.

[31] Stan J. J. Brouns,1* Matthijs M. Jore,1* Magnus Lundgren,1 Edze R. Westra,1 Rik J. H. Slijkhuis,1 Ambrosius P. L. Snijders,2 Mark J. Dickman,2 Kira S. Makarova,3 Eugene V. Koonin,3 John van der Oost., “Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes,” no. August 2008, pp. 960–965, 2013.

[32] and J. J. L. Alfred Doyle1, Michael P. McGarry1, Nancy A. Lee2, “The Construction of Transgenic and Gene Knockout/Knockin Mouse Models of Human Disease,” vol. 21, no. 2, p. 3, 2013, doi: 10.1007/s11248-011-9537-3.The.

[33] H. Bouabe and K. Okkenhaug, “Europe PMC Funders Group Gene Targeting in Mice : a Review,” *J. Innate Immun.*, vol. 1064, no. September, pp. 315–336, 201AD, doi: 10.1007/978-1-62703-601-6.

[34] S. J. Gratz *et al.*, “Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila,” *Genetics*, vol. 196, no. 4, pp. 961–971, 2014, doi: 10.1534/genetics.113.160713.

[35] S. Lin, B. T. Staahl, R. K. Alla, and J. A. Doudna, “Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery,” *Elife*, vol. 3, p. e04766, 2014, doi: 10.7554/eLife.04766.

[36] Q. Guo *et al.*, “‘Cold shock’ increases the frequency of homology directed repair gene editing in induced pluripotent stem cells,” *Sci. Rep.*, vol. 8, no. 1, pp. 1–11, 2018, doi: 10.1038/s41598-018-20358-5.

[37] N. E. Sanjana, O. Shalem, and F. Zhang, “Sanjana, Shalem et Zhang,” *Nat. Med.*, vol. 11, no. 8, pp. 783–784, 2014, doi: 10.1038/nmeth.3047.Improved.

[38] C. Manzo and M. F. Garcia-Parajo, “A review of progress in single particle tracking: From methods to biophysical insights,” *Reports Prog. Phys.*, vol. 78, no. 12, p. 124601, 2015, doi: 10.1088/0034-4885/78/12/124601.

[39] J. Y. Tinevez *et al.*, “TrackMate: An open and extensible platform for single-particle tracking,” *Methods*, vol. 115, pp. 80–90, 2017, doi: 10.1016/j.ymeth.2016.09.016.

[40] Y. Shimizu *et al.*, “Characterization of monoclonal antibodies recognizing each extracellular loop domain of occludin,” *J. Biochem.*, vol. 166, no. 4, pp. 297–308, Oct. 2019, doi: 10.1093/jb/mvz037.

Appendix

Primer ID	Primer description	Sequence(5'->3')
HAdOCL_NS_FW	Forward nested for DHA	CGAGGAAGCGGCGGATCTGGGTCCGGATC ATCCAGGCCTCTTGAAAGTCCACCTCCTTACAGGC
HAdOCL_NS_RV	Reverse nested for DHA	AATAATGGATCCCCAAAACCCCTCAACACTGC
HAdOCL_FW	Unspecific forward for DHA	CCAGGCCTCTTGAAAGTCCA
HAdOCL_RV	Unspecific reverse for DHA	TTCACATGACCAAGGAGCCC
HAuOCL_FW	Reverse for UHA	AATAATGTCGACAACTTACACCTCCCTTGCCG
HAuOCL_RV	Forward for UHA	ACCGGTAGCGCTAGCGGATCTTGGTCAGCTGATCTTCAGGATGAG
mEm_FW	Forward for mEmerald	CTCATCCTGAAGATCAGCTGACCAAGATCCGCTAGCGCTACCGGT
mEM_RV	Reverse for mEmerald	GAGGCCTGGATGATCCGGACCCAGATCCGC CGCTTCCTCGAGATCTGAGTCCGGACTTGT
sgOCL_1	Small guide oligo plus strand	CACCGATGACATGGCTGATTGTCAA
sgOCL_2	Small guide oligo minus strand	AAACTTGACAATCAGCCATGTCATC

Table 1; primers used in the construction of pLentiCRISPR-OCLN and pBluescript-OCLN

Protocols:

- **Protein electrophoresis**

For protein electrophoresis we used SDS-PAGE procedure. First cell pellets scraped from sample plates were resuspended in protein loading buffer 1x and sonicated 10 minutes with 10 second on/off interval at 80% amplitude, followed by boiling at 95ºC for 15 minutes and used directly or kept frozen at -20ºC.

For each electrophoresis we prepared 10ml of running gel containing 12% acrylamide, 375 mM Tris pH 8.8 and 0.1% SDS polymerized with 100 microlitres of APS (10%) and 10 microlitres of TEMED. Stacking gel was prepared as a 5 ml solution composed by 4% acrylamide, 125 mM Tris pH 6.8 and 0.1% SDS, polymerized with 50 microlitres of APS (10%) and 5 microlitres of TEMED.

Each gel was loaded with equal amounts of protein from 14 samples plus molecular weight marker and run at 35 mA in Tris-Glycine buffer until electrophoretic front (bromophenol blue) reached the bottom of the glass. At this point, proteins should be separated in a gradient of size from bigger proteins at the top of the gel to the smaller ones at the bottom.

- **Western blot**

After SDS-PAGE gel is separated from the glass mould, it is placed in a tray containing transfer buffer and a sponge covered by a gel-size piece of whattman paper. On top of the gel, a PVDF membrane is placed followed by another whattman paper piece, making sure that no air bubbles remains between the gel and the membrane. Then the paper is covered by another sponge forming a layer sandwich. This composition is placed in the blotting cuvette filled with transfer buffer and containing a plastic cooler block. It is essential that the membrane side must be face the positive pole of the cuvette as proteins will move from negative to positive.

Transfer is performed at 400 mA for 60 minutes and then the blotting sandwich is unmounted, and the membrane recovered. It is important at this stage to mark the side of the membrane which contains transferred proteins. Then the membrane is blocked for one hour with blocking buffer followed by overnight incubation at 4ºC with the appropriate solution of each antibody in blocking buffer.

After incubation, the membrane is washed with TBS-T for a total of 6 washes of 5 minutes each. Then membrane is incubated with the secondary antibody diluted in blocking buffer for 1 hour at room temperature. For last, membrane is washed as before and developed using clarity western ECL substrate (BioRad; 1705061), which can be detected and imaged using Versadoc4000 (Bio-Rad).

- **DNA electrophoresis**

Gels were prepared as a solution of 0.5gr of agarose in 50 millilitres of TAE buffer(1% agarose gels) and microwaved until complete dissolution of agarose. Once temperature decreases to about 70ºC, 5 microlitres of 10000x SybrSafe(Invitrogen) is added to the mixture and transferred to the taped mould mounted with the 10-well comb.

Once solidified, the tape and comb are retired and the gel placed in the electrophoresis cuvette, previously filled with TAE buffer. Each gel fits 9 samples and a molecular weight

marker. Gels are run at 100 volts for 40 to 60 minutes and results are revealed with ultraviolet light using Versadoc4000 imaging system form Bio-Rad.

- **Plasmid DNA isolation**

Plasmid DNA was extracted using Miniprep kit(Qiagen). A total of 1.5 ml from an overnight culture are centrifuged at 14000rpm for 1 minute. Next supernatant is discarded, and pellet is resuspended it 200 microlitres of buffer P1. Once homogenized, 200 microlitres of buffer P2 are added and contents are mixed by inversion followed by the addition and homogenization of 300 microlitres of buffer N3. Next mixture is centrifuged at 14000 rpm for 10 minutes, and supernatant is transferred to a new Eppendorf tube.

Afterward, one volume of isopropanol (corresponds to the volume of supernatant recovered) is added to the supernatant to precipitate DNA, followed by centrifugation at 14000 rpm for 10 minutes. Then supernatant is discarded, and pellet washed with 1 ml of ethanol 70% followed by centrifugation at 14000 rpm for 5 minutes. Finally, supernatant is discarded, and pellet is allowed to air-dry for 5 to 10 minutes, followed by adding of 50 microlitres of milliQ-H_2O to dissolve DNA pellet.

- **CaCO2 DNA isolation**

Cell pellets scrapped from sample plates were resuspended in 500 microlitres of TE-Proteinase K buffer and incubated at 55ºC for 60 minutes. Next, 55 microlitres of a solution of sodium acetate 3M is added and mixed by inversion, followed by 600 microlitres of isopropanol homogenized gently until the appearance of a jelly-like material is evident. This material, composed mostly by genomic DNA, is picked using a clean P200 pippete tip and transferred to a new Eppendorf tube. Afterwards the material is washed with 500 microlitres of ethanol 70% followed by centrifugation at 14000 rpm for 5 minutes. Finally, supernatant is discarded, and pellet allowed to air-dry for 5 to 10 minutes. Once ethanol has evaporated, pellet is dissolved using 100 to 300 microlitres of 10 mM Tris pH8, depending on the size of the pellet(for bigger pellets use values closer to 300 and for smaller pellets smaller volumes of water).

- **Competent cell transformation**

Plasmid transformations were carried on E. coli competent cells. In one Eppendorf tube it is added the required volume of plasmid plus 27 microlitres of polyethylene glycol 800 10% in water, 10 microlitres of KCM(KCl 5mM, $CaCl_2$ 15mM, $MgCl_2$ 50mM) and topped up to 100 microlitres with water. To this mixture we added 100 microlitres of competent cells and homogenized briefly by pipetting, followed by 30-minute incubation at 4ºC. Once incubation is finished, the mixture is heated at 37ºC for 2 minutes, mixed with 800 microlitres of LB plus ampicillin and incubated at 37ºC for another 20 minutes. For last, 100 microliters of the cells are seeded at each plate, being these plates the appropriate ones for the selection method.

Buffers:

- Laemmli buffer: 2% SDS, 5% B-mercaptoethanol, 10% glycerol, 0.002% bromophenol blue and 0.0625mM Tris-HCl
- Tris-Glycine: 3.03g/l Trizma-base, 14.2g/l glycine and 1% SDS.
- Transfer buffer: 3.03g/l Trizma-base, 14.2g/l glycine in 20% methanol in water solution.
- TBS-T: 150mM NaCl, 10mM Trizma base and 0.1% Tween20.
- Blocking buffer: 150mM NaCl, 10mM Trizma base, 0.1% Tween20 and 5% non-fatty powder milk.
- TE-Proteinase K buffer: 10mM Tris, 1mM EDTA, 0.5% SDS and 1ug/ml proteinase K.
- TAE buffer: 40mM Tris, 20 mM acetate and 1mM EDTA
- Bacteria freezing media: LB broth with 20% glycerol
- Cell freezing media: DMEM supplemented with 20% FBS and 10% DMS