Generating isogenic NPC1 hiPSC lines for disease modeling using the CRISPR/Cas9 system

Joaquim Filipe Aresta Lopes Godinho[1,2,*]

Thesis to obtain the Masters of Science Degree in Biomedical Engineering – November 2018

Supervisors: Dr. Evguenia Bekman[1,2] and Dr. Tiago Fernandes[1]

[1]Stem Cell Engineering Research Group, Institute for Biosciences and Bioengineering, Instituto Superior Técnico, University of Lisbon, Portugal; [2]Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, Portugal;

*Email: filipe.godinho@tecnico.ulisboa.pt

Abstract: CRISPR/Cas9 is a powerful and versatile gene-editing tool. In this work we applied this tool to generate isogenic NPC1 hiPSC lines for disease modelling of Niemann-Pick type C disease. Our aim was to introduce a specific point mutation termed I1061T substitution, which occurs in exon 21 of the NPC1 gene. In order to do that the hiPSCs were transfected with a Cas9-GFP plasmid, an expression vector for the single-guide RNA (sgRNA) and a single-stranded oligonucleotide (ssODN) to serve as the homology-directed repair (HDR) template. The template included the I1061T substitution and two blocking mutations located on the protospacer adjacent motif (PAM), which is necessary for recognition by the CRISPR/Cas9 machinery. The blocking mutations were part of our strategy to increase the number of correctly-edited hiPSCs by reducing re-cutting by the Cas9. Out of the 62 clones that were successfully expanded only 3 showed mutations in the locus of interest, as determined by Sanger sequencing. Additionally, these 3 mutated clones did not incorporate the HDR template, but instead presented random mutations (indels), which is sign of DNA repair by non-homologous end joining (NHEJ). Of particular interest is a clone that presented a frameshift insertion which causes a premature stop of protein expression, as it might prove useful as an NPC1 knockout. Overall, the results obtained were not ideal, but we are confident that this work contributed to the advancement of future developments of NPC human disease modelling.

Keywords: CRISPR/Cas9; gene-editing; homology-directed repair; human induced pluripotent stem cells; NPC1; neurodegenerative disease modelling.

1. INTRODUCTION

1.1. Niemann-Pick disease type C

Niemann-Pick disease type C (NPC) is a rare lysosomal storage disease (LSD) that is mainly associated with an abnormal accumulation of cholesterol in late endosomes/lysosomes, leading to metabolic malfunctioning and cell death. [1–3]

The prevalence of NPC is only of approximately 1:150,000[2] and it is characterized by a fatal progressive neurodegeneration, while possessing a wide variety of clinical symptoms, including cerebellar ataxia, dystonia and dementia.[4]

While the mechanisms of the disease are still not fully comprehended, NPC is known to be an autosomal recessive disorder caused by mutations in the *NPC1* or *NPC2* genes. Around 95% of NPC-causing mutations are on the *NPC1* gene, which is located on chromosome 18 and contains 25 exons that code for a large transmembrane protein (NPC1), with a total of 1278 amino acids.[5–7] NPC1 is mostly found on lysosomal and endosomal membranes and although its specific function has not been fully defined it does play a critical role in intracellular cholesterol and lipid trafficking.[1,2]

One particularly frequent *NPC1* mutation, present in around 14% of disease alleles[8], is the

I1061T substitution, which is located in exon 21 and affects the C-terminal domain of NPC1, which appears to play an important role in cholesterol transport by interacting with the N-terminal domain.[2,7] The mutant NPC1 protein ends up being degraded, most likely due to protein misfolding, leading to a dysfunction of the metabolic pathways NPC1 is involved in.[9,10]

1.2. Stem cells and disease modeling

Stem cells are unspecialized cells with the ability to self-renew through cell division or to differentiate into cells from various tissues. Their self-renewal capacity allows them to multiply indefinitely, which is very useful for *in vitro* experimentation.[11,12] Most relevant to this work are pluripotent stem cells (PSCs), which can be differentiated into any cell type derived from the three embryonic germ layers (i.e. endoderm, ectoderm and mesoderm).[11,13]

Embryonic stem cells (ESCs) are PSCs from the inner cell mass of the blastocyst.[13] However, their use for research is controversial due to the destruction of embryos involved in the process of collecting them.[14] In 2006, Yamanaka and his team discovered that it was possible to reverse engineer the differentiation process and generate induced pluripotent stem cells (iPSCs) by culturing somatic cells in the presence of a defined set of reprogramming factors.[15] This has already been applied to human somatic cells, namely fibroblasts, to produce human iPSCs (hiPSCs), opening up a new path for stem cell research.[16,17]

There are two main approaches to obtain disease-carrying hiPSCs. The first approach would be to use patient-specific hiPSCs derived from somatic cells of the patient with a certain genetic disease. Following the successful reprogramming, gene-editing tools can be used to correct the disease-causing mutations, generating healthy cells to be used as control or in cell-based therapies.[16,17] Alternatively, gene-editing tools can be used to introduce the disease-causing mutations into the genome of healthy cells.

Disease models are generated by differentiating the disease-carrying PSCs into relevant tissue types for phenotypical studies. Initially, this process was done using a 2D system to generate a tissue monolayer, however, the development of 3D culture systems lead to the generation of organoids. Organoids are 3D structures composed of several tissue types with the capacity to self-organize into a miniature organ, mimicking *in vitro* what happens *in vivo* to a much better extent than monolayers.

A protocol to generate cerebellar organoids from hiPSCs has been developed and successfully implemented in our lab. By applying this protocol on correctly edited NPC1 hiPSC lines, we would be able to create a robust and novel disease model for NPC that would allow a more in-depth study of the causes and mechanisms of cerebellar ataxia in the disease.

1.3. CRISPR/Cas9 system

In Bacteria and Archaea, clustered regularly interspaced short palindromic repeats (CRIPSR)/CRISPR-associated (Cas) systems constitute part of their adaptive immune response against invading genetic elements. CRISPR functions as a library that stores and posteriorly recognizes invading DNA, while the Cas9 nuclease cleaves these DNAs, effectively protecting the cells.[18,19]

The gene-targeting and DNA cleavage properties of CRIPSR/Cas system have been successfully re-engineered as a genome editing tool.[20] Currently, it is possible to introduce specific genomic DNA cuts with CRISPR in eukaryotic cells, by directing the Cas nuclease towards those specific locus using only short RNAs, which can be easily designed and synthesized to match the needs of the experiment.[21–24]

The most widely used CRISPR/Cas system for gene-editing is the type II CRISPR/Cas system, also known as CRISPR/Cas9 system, which has been adapted from *Streptococcus pyogenes* to introduce sequence-specific double-stranded breaks (DSBs) in genomic DNA.[23] Besides the Cas9 nuclease, which is responsible for the DNA cleavage, the system includes two noncoding CRIPSR RNAs (crRNAs), one of which possesses the protospacer sequences homologous to the target loci. The two crRNAs can be synthesized or expressed together as a single guide RNA (sgRNA) that binds to the Cas9 and directs it towards the targeted locus.[25]

Cas9 requires the presence of an 'NGG' protospacer-adjacent motif (PAM) in order to cleave the DNA, which has to be located on the non-complementary DNA strand and immediately downstream of the target site. The nuclease induces a DSB precisely 3 bp upstream of the PAM.[19]

Following DNA cleavage, the DNA repair mechanisms come into play, of which there are primarily two: non-homologous end joining (NHEJ) and homology-directed repair (HDR). The most frequent repair pathway is NHEJ, which does not require a large donor DNA template in order to re-attach the loose DNA ends, but results in random insertions or deletions, known as indels, which can be to generate gene knockouts.[26]

HDR allows specific modifications to be introduced into a genome without producing any undesired mutations, making it useful for precise gene-editing applications. In order to work, HDR requires a donor DNA template with sequence homologous to the region surrounding the DNA break, but also allows for small specific sequence changes to be introduced, which are incorporated into the locus during its repair. An HDR template can be synthesized as a plasmid DNA, a double-stranded DNA (dsDNA) fragment or a smaller single-stranded oligonucleotide (ssODN), and can be co-transfected with the other CRISPR/Cas9 components.[27,28]

2. RESEARCH METHODOLOGY

One of the main problems regarding the accuracy of the gene-editing with CRISPR/Cas9 is that, in most cases, the targeted locus can be detected more than once by the CRISPR/Cas9 complex. This means that even if a genome is correctly edited by HDR, which is already a rare event, it most likely will be re-cut by the Cas9 and re-edited by NHEJ, resulting in the introduction of undesired indels on top of the intended mutations.

To prevent that and increase HDR efficiency, a strategy was implemented based on the framework defined by Kwart et al[29] named CORRECT, which is short for consecutive re-guide or re-Cas steps to erase CRISPR/Cas-blocked targets. The strategy consists on blocking the PAM, by introducing blocking mutations in the HDR template, so that the cut site is no longer recognized by the Cas9 after the template has been successfully incorporated.

In this particular case, the intended mutation lies within the coding region of *NPC1*, so the design of the HDR template had to be done in a way that it introduced only silent blocking-mutations that would also not disturb any splicing events.

Clonal isolation of hiPSCs co-transfected with Cas9-GFP plasmid, sgRNA and HDR template was done by fluorescence-activated cell sorting (FACS) and the clones were expanded in isolation to maintain isogenicity of the colonies until they were ready for analysis. The correct implementation of this methodology should yield isogenic NPC1 clones with the I1061T mutation.

3. MATERIALS AND METHODS

3.1. hiPSC culture

The cells used in this work were hiPSCs from two different cell lines: iPSC6.2 (Gibco®, Human

Episomal iPSC line) and F002.1A.13 (TCLab, Portugal). These cells were cultured on Matrigel™ (Corning)-coated plates with mTeSR™1 medium (StemCell Technologies). Cells were grown in a CO_2 incubator at 37° C with CO_2 levels of 5 %. Medium was changed daily.

3.2. Transfection methods

The transfection of hiPSCs was done using two methods. The first was the Neon™ electroporation system (Thermo Scientific). For this method $2x10^6$ cells per condition, treated with ROCKi for at least 1 hour prior, were collected with Accutase and resuspended in a buffer from the kit, where the DNA constructs are added. The cells are placed in a tube containing the electrolytic buffer via a plug-in pipette, which is where the electroporation occurs. The main cycle used consisted of 3 pulses of 1400 mV with a duration of 5 ms each, although other conditions were tested. The electroporated cells were replated in Matrigel-coated wells with mTeSR™1 supplemented with 10 µM ROCK inhibitor (ROCKi) and incubated at 37° C for 48 hours.

The second method was lipofection, for which we used the Lipofectamine® 3000 reagent (1.5 – 3 µL per well of a 12-well plate). This method consists on creating DNA-lipid complexes in Opti-MEM medium (Gibco®) that are carefully added directly to the wells containing the cells. After that the cells were incubated at 37° C or 32° C for 48 hours. The DNA constructs enter the cells after the lipid structure merges with the cell membrane.

3.3. Cell sorting

Cell sorting of GFP-positive cells was done by fluorescence-activated cell sorting (FACS) in a FACSAria III cell sorter (BD Biosciences). The cells were collected for sorting with Accutase, after being treated with 10 µM ROCKi for at least 1 hour prior, and resuspended in mTeSR™1 supplemented with 10 µM ROCKi. The GFP-positive cells were replated in Matrigel-coated wells in mTeSR™1 supplemented with 10 µM ROCKi and incubated at 37° C for 48 hours before any medium change was done.

3.4. Colony picking

The picking of isogenic colonies, treated with 10 µM ROCKi prior to the picking, was done by scraping them with a pipette and transferring them to separate Matrigel-coated wells of a 48-well plate in mTeSR™1 supplemented with 10 µM ROCKi. This was done with the aid of a JuLI™ smart fluorescent cell analyzer, which has a display LCD and a fluorescent light that facilitate the observation of colonies and the positioning of the pipette for picking. After picking the colonies were incubated at 37° C for 48 hours before changing the medium.

3.5. Bacteria culture

Solid culture of competent DH5α *Escherichia coli* (*E. coli*, Thermo Scientific) was done in LB agar at 37° C and liquid culture in LB medium at 37° C with 220 rpm shaking for 16h. Ampicillin (100 µg/mL, Sigma) was used for selection in both cultures.

3.6. Bacteria transformation

5 – 10 µL of plasmid DNA were added to 100 µL of the competent *E. coli* suspension (stored at -80° C in TB buffer with 7% DMSO) and incubated on ice for 30 mins. Then, the cells were heat shocked at 42° C for 45 secs on a dry bath followed by 2 mins. of incubation on ice. The cells were diluted in 900 µL of LB and incubated at 37° C with 220 rpm shaking for 45 – 60 mins. Afterward, the cells were centrifuged and resuspended in a small volume of LB to be spread across an LB agar dish with ampicillin, which was then incubated at 37° C overnight.

3.7. Agarose gel electrophoresis

Agarose gel electrophoresis was performed in custom-made apparatus (Acrílicos Fernando Gil Lda), in TAE (1x) buffer, at a voltage of 90 – 120 V (~6V/cm). After each run was over the gel was

analyzed by UV light in ChemiDoc™ XRS+ (Bio-Rad). The agarose 1 – 1.5% was dissolved in TAE (1x) by heat and 5 µL of Xpert Green DNA Stain (GRiSP) were used per 100 mL of agarose solution.

3.8. Plasmid DNA preparation and purification

The extraction and preparation of plasmid DNA from transformed *E. coli* was made using the NZYMiniprep kit from NZYTech. The purification of DNA samples, including DNA plasmids and PCR products, was performed by gel electrophoresis using the NZYGelPure kit (NZYTech).

3.9. sgRNA cloning and validation

The design of the sgRNA sequences was done using the CRISPOR web tool.[30] The selected sgRNAs were ordered from Sigma as a pair of primers that were annealed and ligated to the MLM3636 sgRNA expression vector (Addgene) previously digested with *BsmBI*. The plasmids were then cloned in competent *E. coli*. The validation of the sgRNAs was done using the Genomic Cleavage Detection Kit (GeneArt™) after the cells had been transfected with the Cas9-GFP and the sgRNA plasmids.

3.10. Mutation-block template design

The design of the mutation-block (MB) template was done using bioinformatics tools to ensure that the blocking mutations did not influence protein expression. Namely, the human splicing finder was used to analyze potential effects on splicing events.[31]

3.11. Analysis of expanded clones

Following the overnight cell lysis of the clones by buffered Proteinase K at 55 – 65° C, the genomic DNA (gDNA) was separated from the remaining lysate using phenol:chloroform:isoamyl and precipitated with isopropanol at -20° C overnight. The next day, the DNA was cleaned with EtOH, centrifuged at top-speed and resuspended in water. The region of interest of the gDNAs was amplified by PCR using the NZYLong polymerase (NZYTech) in a Bio-Rad T-100™ Thermocycler. The PCR products obtained were purified and sent to GATC Biotech for Sanger sequencing.

4. RESULTS AND DISCUSSION

4.1. sgRNA design and cloning

Using CRISPOR[30], we found 23 possible guide sequences of 20 bp within exon 21 of the *NPC1* gene. Besides finding the possible sgRNA designs it also scores their specificity and estimates their efficiency, off-target activity across the whole genome and other relevant information based on a selection of algorithms.

Out of the 23 possibilities, only 3 sgRNA designs were in sufficiently close proximity to the mutation site to result in a cut-to-mutation distance below 10 bp, which was necessary to introduce a homozygous mutation. These 3 sgRNAs were selected for testing (Table 1), as all three showed good specificity scores (>80), reasonable estimates of efficiency (35 – 55) and low off-target activity (<3 off-targets).

Table 1 - *Sequences of the suitable sgRNAs proposed by CRISPOR with the adjacent PAMs and cut-to-mutation distance associated with each.*

Name	Sequence (5' – 3')	PAM	Cut-to-mutation distance
sgRNA 3	CATGGTTTCGGTGACATTAC	TGG	8 bp
sgRNA 6	GACATTACTGGCTATAAGTC	GGG	3 bp
sgRNA 8	TGACATTACTGGCTATAAGT	CGG	2 bp

To prepare the cloning of the sgRNA, a pair of complementary primers was annealed together to form DNA duplexes corresponding to the sgRNA sequence and the MLM3636 plasmid, which is an sgRNA eukaryotic expression vector, was digested with *Esp3I* and purified. Per design, the overhangs from the sgRNA duplexes were complementary to those of the digested vector, which is what allowed the unidirectional ligation to happen. The ligated

sgRNA plasmids were each successfully transfected into competent *E. coli* and cloned. The correct inserts of the sgRNA plasmids were confirmed by Sanger sequencing.

4.2. sgRNA functional validation

The functional validation of the sgRNAs was done using the Genomic Cleavage Detection kit (GeneArt™). For this assay we transfected HEK-293 (ATCC® CRL-1573™) cells with the Cas9-GFP and each of the sgRNA plasmids.

Genomic cleavage by the Cas9 is mostly repaired by NHEJ, which introduces indels in the region of the cut site. By denaturing and re-annealing the PCR fragments of that region, there will be mismatches in the re-annealed fragments that will be a measure of NHEJ and DNA cleavage, by association. The Detection Enzyme (GeneArt®) of the kit recognizes and cuts the mismatched DNA fragments allowing the analysis of cleavage efficiency by agarose gel electrophoresis, as shown in Figure 1.

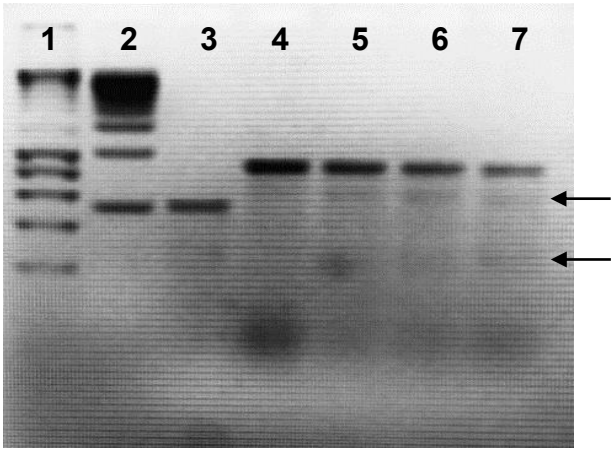

Figure 1 - Analysis by agarose gel electrophoresis of the digested PCR products of the cleavage detection assay. Lanes 1 and 2 are runs of the DNA Ladder III (NZYTech) and 500 bp (G-Biosciences), respectively. Lane 3 contains the run of the digestion control of the kit (GeneArt™). Lane 4 has the negative control with the PCR product from the gDNA of non-transfected cells Lanes 5 – 7 contain the digested PCR products of the gDNA of cells transfected with pCas9-GFP (all of them) and the sgRNA 3, sgRNA 6 and sgRNA 8 plasmids, respectively.

The control DNA from the kit shown in lane 3 should contain a stronger parent band with 546 bp and two weaker digested bands with 225 bp and 291 bp. However due to the poor quality of the image and the low size of the fragments it is extremely hard to detect them in this image. Lane 4 has the negative control with the PCR product from the gDNA of non-transfected cells, which shows only a strong band with approximately the expected size of 865 bp. By analyzing the DNA bands in lanes 5 – 7 it is possible to detect the strong parent band in all of them, with the expected length of 865 bp, but more importantly, the digested bands are also visible, despite being faint, and they have the expected sizes of approximately 585 bp and 280 bp (arrows). From this assay, it is valid to conclude that all the sgRNAs preformed aptly in guiding the Cas9 towards the region of interest.

4.3. Mutation-block template design

The mutation-block (MB) template plays a central role in this gene-editing experiment as it will serve as the DNA template for HDR. It is by modifying the sequence of this template and transfecting it into the cells that we attempt to manipulate the cells into using this template for HDR, resulting in the incorporation of the modifications made to the template into the edited genome. Thus, the potential MB templates for each sgRNA were designed to be a 100 bp single-stranded oligo deoxynucleotide (ssODN) with a sequence mostly homologous to the anti-sense strand of the wild type genome and centered on the cut site. The sequence changes made were the intended mutation (I1061T) and the blocking mutations on the PAM.

The selected template was designed to be used with sgRNA 8 because it was the only one that allowed both G's of the PAM to be mutated while having no effects on protein expression nor splicing events, as predicted by the Human Splicing Finder.[31]

4.4. Transfection of hiPSCs

The optimization of the transfection methods was prompted by the initial failure to obtain good transfection efficiencies with Neon™ electroporation. The optimization tests was performed on the iPSC6.2 cell line and led to the conclusion that using lipofection (with Lipofectamine® 3000) was the most efficient option, especially when using post-transfection incubation at 32° C (with efficiencies around 40%), instead of the standard 37° C (with efficiencies of around 15%). Several conditions were tested for the Neon™ electroporation system, but ultimately it proved to be inadequate to use in these cell lines as it caused too much cell death while presenting very poor transfection efficiencies (never above 4%).

The hiPSCs were co-transfected by lipofection with the sgRNA 8 plasmid, the Cas9-GFP plasmid and the MB template ssODN. The GFP-positive cells were then sorted by FACS and replated in a 6-well plate for expansion, at a density of approximately 6000 cell per well.

4.5. Analysis of hiPSC clones

The single cell clones were expanded until they had formed sizeable isogenic colonies, at which point the colonies were individually picked and transferred to separate wells for further expansion. Overall, the expansion process took about 20 – 25 days.

In regard to the iPSC6.2 cell line, out of the 73 colonies that were picked, 62 were successfully expanded and only 3 clones showed mutations in the targeted locus (named clones 29, 30 and 45), none of which were the intended mutation. In regard to the F002.1A.13 cell line, only 2 colonies were not expanded in time for its analysis to be included in this work, but have since been successfully expanded by a colleague from the lab.

The analysis of the clones was done by Sanger sequencing of the region of interest amplified by PCR from the genomic DNA extracted from each clone.

Clone 29 presented a one-nucleotide frameshift mutation caused by an insertion of a T between amino acid positions 1060 and 1061, very close to the intended mutation site. The frameshift results in the alteration of the open reading frame (ORF) and it leads to a premature stop of protein translation due to the appearance of an early stop codon in position 1064. The sequencing of this clone was repeated using a different primer and a new PCR product of the same region to confirm the insertion. In terms of effects on the cell, this mutation is likely to result in some or total loss of function of the NPC1 protein, which could prove useful for future studies of the disease.

The mutations of clones 30 and 45 were both substitutions. In clone 30, a G was substituted by a C very close to the intended mutation site, which results in an amino acid change from an Alanine to a Proline (1062). While in clone 45 a G was substituted by an A, which results in an amino acid change from a Serine to an Asparagine (1074). The Serine → Asparagine substitution of clone 45 will most likely not have a significant effect on protein function, as these two amino acids have similar properties and present a high exchangeability value.[32,33] On the other hand, the Alanine → Proline substitution in clone 30 will most likely affect protein conformation, since Proline is a rigid hinge amino acid that forces a tight angle shift of the protein chain, unlike Alanine.[32] The consequences of this in protein function will have to be determined experimentally, but it will likely be impacted by this mutation.

The chromatogram of each sequencing was examined to determine the zygosity of the mutations. The conclusion was that all mutations were heterozygous and that there was likely DNA contamination from other clones in all samples,

which can be confirmed by the double peaks that occur in the chromatograms at or after the mutation sites (marked in figures 2 – 4).

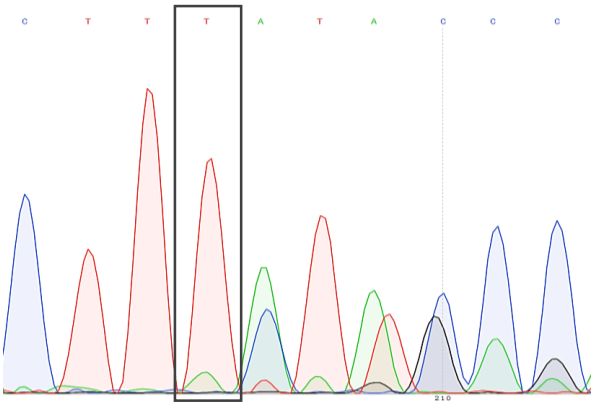

Figure 2 - Section of the chromatogram of the Sanger sequencing of clone 29.

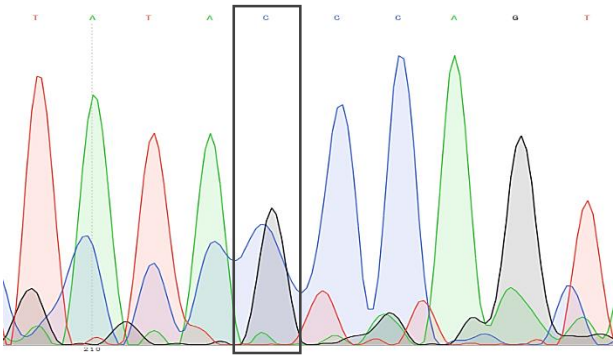

Figure 3 - Section of the chromatogram of the Sanger sequencing of clone 30.

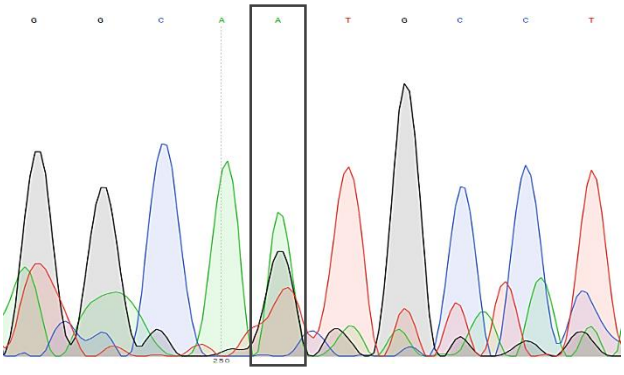

Figure 4 - Section of the chromatogram of the Sanger sequencing of clone 45.

5. CONCLUSIONS

During this work the design of the strategy for introducing the I1061T mutation in exon 21 of *NPC1* was made, as well as the optimization of the transfection conditions used in two different hiPSC lines. The implementation of the gene-editing strategy, which included the culture, transfection and sorting of the cells, culminated in the successful expansion of 62 isogenic clones, of which 3 contained mutations in the targeted locus.

Overall, this research work laid the groundwork for future studies of NPC disease through CRISPR/Cas9 gene-editing. Firstly, by having designed the sgRNAs and respective HDR templates to introduce the I1061T mutation and PAM-blocking mutations, and secondly, by having tested the efficiency of said sgRNAs.

The testing of Neon™ electroporation and lipofection conditions for transfection in the iPSC6.2 hiPSC line with the Cas9-GFP plasmid showed that Neon™ electroporation was not a viable option for transfection, while lipofection achieved good transfection efficiencies overall and showed the highest efficiencies when the cells were incubated at 32° C during the transfection period.

Because we were not able to generate any isogenic clone that had undergone a successful incorporation of the HDR template, we cannot comment on the effects that this would have had in reducing re-edits of the same locus by the CRISPR/Cas9 system. However, we are positive that the correct introduction of the PAM-blocking mutations would have been successful in preventing those events.

Also, due to the limited time frame attributed to the experimental work, the number of isogenic hiPSC colonies that successfully were picked and expanded was too low to allow any significant conclusions to be drawn in regard to the success of the gene-editing

and the accuracy of HDR. However, we recommend that other strategies be employed, if available, to increase the frequency of HDR, namely cell cycle manipulations, the use of culture components to inhibit NHEJ or using other HDR design strategies.

On the other hand, the number of mutant clones caused by NHEJ was still lower than expected (only 3 mutant clones out of the 73 expanded clones). However, we firmly believe this to be a result of the low purification of the transfected cells by FACS and not of the low efficiency of CRISPR/Cas9 editing or the NHEJ itself. From this experience, we conclude that it is essential that the FACS be performed with the highest purity to maximize the percentage of mutated clones among the sorted cells.

From the 3 *NPC1* mutants obtained from the iPSC6.2 cell line, clone 29 presents a promising mutation that most likely leads to the silencing of *NPC1* or at the very least a reduction in correct NPC1 expression in those cells. Also, due to the improved transfection conditions and high purity FACS used for the second round of gene-editing, we are confident that the two F002.1A.13 hiPSC clones that were already successfully expanded will have mutations in *NPC1*.

6. FUTURE WORK

We hope that future projects in search of reliable human disease models for NPC disease take advantage of the work and results presented here. Not only by using our designs for sgRNA and HDR templates, but also by taking into account the recommendations made based on the work done throughout this project.

It is still necessary to screen the mutant genomes for off-target mutations by sequencing specific locus of the gDNA or by full genome sequencing. Additionally, the pluripotency of these clones should be verified via a teratoma assay or by analyzing their expression of pluripotency markers. From there, a Western-blotting analysis should be performed on the contents of the hiPSCs expanded from all clones to study the changes in NPC1 expression.

Finally, these *NPC1* mutant hiPSCs can be differentiated, alongside the wt hiPSCs of the iPSC6.2 cell line for control, into cerebellar organoids, using the protocol developed in our lab, to analyze the differences in phenotype in an effort to shed new light on the mechanisms of NPC disease, particularly with respect to its symptoms of cerebellar ataxia.

As a whole, the results achieved throughout this project drove forward the research efforts that are being made towards the generation of human disease models for NPC, putting us a few steps closer to reaching a more comprehensive understanding of the disease and, therefore, to the development of novel diagnostic methods and therapeutic strategies to mitigate or even cure this disease.

7. REFERENCES

1. Chang, T. Y. *et al.* Niemann-Pick type C disease and intracellular cholesterol trafficking. *J. Biol. Chem.* **280,** 20917–20920 (2005).
2. Vanier, M. & Millat, G. Niemann-Pick disease type C. *Clin. Genet.* **64,** 269–281 (2003).
3. Umeda, M. *et al.* Niemann-Pick C1 Disease Gene: Homology to Mediators of Cholesterol Homeostasis. *Science (80-.).* **277,** 228–231 (1997).
4. Patterson, M. Niemann-Pick Disease Type C. *GeneReviews by University of Washington, Seattle* (2000). Available at: https://www.ncbi.nlm.nih.gov/books/NBK1296/.
5. Kniffin, C. & Hamosh, A. NPC1 gene. *OMIM* (2003). Available at: http://omim.org/entry/607623.
6. NPC1 Human. *UniProtKB* (2000). Available at: https://www.uniprot.org/uniprot/O15118.
7. Li, X. *et al.* 3.3 Å structure of Niemann–Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol

transport. *Proc. Natl. Acad. Sci.* **114,** 201711716 (2017).
8. Millat, G. *et al.* Niemann-Pick C1 Disease: The I1061T Substitution Is a Frequent Mutant Allele in Patients of Western European Descent and Correlates with a Classic Juvenile Phenotype. *Am. J. Hum. Genet.* **65,** 1321–1329 (1999).
9. Gelsthorpe, M. E. *et al.* Niemann-Pick type C1 I1061T mutant encodes a functional protein that is selected for endoplasmic reticulum-associated degradation due to protein misfolding. *J. Biol. Chem.* **283,** 8229–8236 (2008).
10. Rauniyar, N. *et al.* Quantitative Proteomics of Human Fibroblasts with I1061T Mutation in Niemann–Pick C1 (NPC1) Protein Provides Insights into the Disease Pathogenesis. *Mol. Cell. Proteomics* **14,** 1734–1749 (2015).
11. NIH Stem Cell Basics. *Bethesda, MD: National Institutes of Health, U.S.* (2016). Available at: stemcells.nih.gov/info/basics.htm.
12. Zhu, Z. & Huangfu, D. Human pluripotent stem cells: an emerging model in developmental biology. *Development* **717,** 705–717 (2013).
13. Bongso, A. & Richards, M. History and perspective of stem cell research. *Best Pract. Res. Clin. Obstet. Gynaecol.* **18,** 827–842 (2004).
14. Hyun, I. The bioethics of stem cell research and therapy. *J. Clin. Invest.* **120,** 71–75 (2010).
15. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126,** 663–676 (2006).
16. Yu, J. *et al.* Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science (80-.).* **318,** 1917 LP-1920 (2007).
17. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–72 (2007).
18. Horvath, P. & Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science (80-.).* **327,** 167–170 (2010).
19. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci.* **109,** E2579–E2586 (2012).
20. Lander, E. S. The Heroes of CRISPR. *Cell* **164,** 18–28 (2016).
21. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science (80-.).* **339,** 823–826 (2013).
22. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2013,** 1–9 (2013).
23. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR / Cas Systems. *Science (80-.).* (2013). doi:10.1038/nbt1319
24. Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31,** 230–232 (2013).
25. Jinek, M. *et al.* A Programmable Dual-RNA – Guided. **337,** 816–822 (2012).
26. Lieber, M. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway. *Annu. Rev. Biochem.* 181–211 (2010).
27. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533,** 125–129 (2016).
28. Paix, A. *et al.* Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl. Acad. Sci.* (2017). doi:10.1073/pnas.1711979114
29. Kwart, D., Paquet, D., Teo, S. & Tessier-Lavigne, M. Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat. Protoc.* **12,** 329–334 (2017).
30. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17,** 1–12 (2016).
31. Desmet, F. O. *et al.* Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37,** 1–14 (2009).
32. Betts, M. J. & Russell, R. B. in Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data (Second Edition) **4,** 289–316 (2007).
33. Yampolsky, L. Y. & Stoltzfus, A. The exchangeability of amino acids in proteins. *Genetics* **170,** 1459–1472 (2005).