UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Modeling Neuropathological Type II Gaucher Disease using iPSCs, CRISPR/Cas9 and Transcription Factor-Driven Protocols

Mãdãlina Sãcultanu

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Isaac Canals e Gabriela Rodrigues

2020

Acknowledgments

I would like to express gratitude to Henrik Ahlenius for agreeing to take me into the Stem Cells, Aging and Neurodegeneration group and Isaac Canals, in particular, for allowing me to participate in this project, guiding me through every step and for all the support and helpful input throughout the process of making this dissertation. A special thanks to Oskar Zetterdahl for his exceptional work and all the members of the group as well for the discussions and brainstorming throughout these months. Additionally, I want to thank the constant presence of all my family and friends and all the incentive and interest they have shown, even though they weren't exactly sure where or what I was working with. To Filipe, thank you for always keeping me company and putting a smile on my face, even from far away. And finally, acknowledge all the wonderful work from both Lund University and the Faculty of Sciences of the University of Lisbon, arranging and making possible this exchange for my dissertation project, as it was truly an incredible opportunity, and to the coordinators and professors of the Master in Molecular Biology and Genetics , namely professor Gabriela Rodrigues for being my internal supervisor on this dissertation.

LUND UNIVERSITY

LUND STEM CELL CENTER

Abstract

Gaucher Disease (GD) is an autosomal recessive disorder caused by mutations in the *GBA1* gene. This gene encodes for the beta-glucocerebrosidase (GCase) protein, a lysosome housekeeping enzyme that degrades the glucocerebroside into glucose and ceramide. With an incidence of 1 in 40,000 to 1 in 60,000 births in the general population, and 1 in 800 births in the Ashkenazi Jewish population, it is the most common type of Lysosomal Storage Disease (LSD). *GBA1* mutations are associated with a higher risk of Parkinson's Disease and Lewy body dementia development, through mechanisms still unknown but likely to be related to compromised α-synuclein degradation by cells with impaired lysosomal function. During this project, we have initially characterized iPSCs from three patients with the most severe and neuropathic form of GD, type 2. We have further produced embryoid body formation and induced differentiation to the three germ layers endoderm, mesoderm and ectoderm; and performed RT-qPCR for pluripotency genes in order to demonstrate the pluripotent state of each cell line. Then, we have attempted to optimize CRISP/Cas9 genetic editing protocols and correct the single-base mutations in the *GBA1* gene responsible for the manifestations of GD in these patients, with the intent of obtaining isogenic controls for each cell line and, through the overexpression of transcription factors, we were able to induce the differentiation of the iPSCs to neurons and astrocytes. Finally, we have showed significantly reduced levels of GCase activity in the diseased cell lines, though enzymatic activity essays and altered lysosomal function through immunocytochemistry, when compared to *GBA1* wild-type controls. Overall this project has allowed to showcase the significant potential of using iPSCs as both a source of relevant cell types and their future application as *in vitro* models for the study of GD, particularly neuropathic GD, where the cell types affected are particularly challenging to obtain.

Key Words: (LSDs) Lysosomal Storage Diseases; (GD) Gaucher Disease; (GCase) beta-glucocerebrosidase; (iPSCs) induced pluripotent stem cells; CRISPR/Cas9

Resumo

A Doença de Gaucher (GD) é uma patologia com transmissão hereditária autossómica recessiva causada por mutações no gene *GBA1* que codifica para o enzima beta-glucocerebrosidase (GCase). Esta proteína é um enzima *housekeeping* nos lisossomas, responsável pela degradação de glucocerebrosidase em glicose e ceramida. Com uma incidência entre 1 em 40.000 e 1 em 60.000 nascimentos na população geral, e 1 em 800 nascimentos na população de descendência Judaica Ashkenazi, é a forma mais comum de Doença Lisossomal de Sobrecarga. Estão descritas mais de 495 mutações no gene *GBA1*, sendo as mais comuns mutações *missense* que levam à produção de proteínas *misfolded*. Em homozigotia ou em heterozigotia composta, certas mutações levam à manifestação de fenótipos mais severos com envolvimento neurológico, em que a atividade da GCase é quase nula, classificado como GD do tipo 2. Independentemente do nível de manifestação da doença, as mutações no gene *GBA1* estão associadas a um maior risco de desenvolvimento de Doença de Parkinson e Demência de Corpos de Lewy, através de mecanismos ainda desconhecidos, mas muito provavelmente relacionados com a degradação ineficiente de proteínas, nomeadamente α-sinucleína, em células com funcionamento lisossomal comprometido. Modelos animais têm sido gerados para o estudo de GD, e, embora bastante úteis para as manifestações somáticas desta patologia, o envolvimento neurológico tem sido muito limitado nestes modelos e as mutações induzidas em homozigotia levam a fenótipos agressivos em que os animais não sobrevivem muito tempo após o nascimento, pelo que a sua aplicação no contexto de GD do tipo 2 é bastante insuficiente. Neste sentido, as iPSCs produzidas através da reprogramação direta de células somáticas humanas obtidas de pacientes, a um estadio de pluripotência via o método desenvolvido por Yamanaka, *et al*, surgem como um possível modelo *in vitro*, possibilitando o estudo de diversas patologias. Através de métodos de edição genética como CRISPR/Cas9 e da sua otimização nos últimos anos, é possível a introdução, deleção ou correção de mutações nestes modelos celulares, de forma a replicar doenças hereditárias ou obter controlos isogénicos que permitem o seu estudo de forma fidedigna. Inicialmente neste projeto, caracterizámos células iPSCs provenientes de três pacientes com a forma mais severa e com manifestações neuropáticas de GD tipo 2. Através de RT-qPRC, da formação de corpos embrionários e cariotipagem, demonstrámos a estabilidade cromossómica nas amostras em estudo e o seu estado de pluripotência através da expressão de genes como *NANOG*, *SOX2* e *OCT4* (*POU5F1*), bem como a posterior diferenciação dos corpos embrionários nos três folhetos germinativos: endoderme, mesoderme e ectoderme. Seguidamente, trabalhámos na otimização de protocolos de edição genética com CRISPR/Cas9 com o intuito de corrigir as mutações *missense* de substituição de bases no gene *GBA1* das amostras em estudo, de forma a obter controlos isogénicos e, através da sobre-expressão de fatores de transcrição, induzimos a diferenciação das iPSCs em neurónios e astrócitos. Finalmente, demonstrámos níveis significativamente reduzidos de atividade enzimática da GCase nas linhas mutadas e alteração do funcionamento dos lisossomas, quando comparadas a controlos *wild type* e ao controlo isogénico e função lisossomal alterada, nas amostras de GD tipo 2. De uma forma geral, este estudo demonstrou o potencial significativo da utilização de células iPSCs como modelos *in vitro* de GD, particularmente da forma neuropática, em que a obtenção dos tipos celulares afetados é particularmente difícil e não existem, até à data desta dissertação, modelos animais que repliquem este fenótipo.

Palavras-chave: (LSDs) Doença Lisossomal de Sobrecarga; (GD) Doença de Gaucher; (GCase) beta-glucocerebrosidase; (iPSCs) Células Estamiais Pluripotentes induzidas; CRISPR/Cas9

Index

Figure Index

Table Index

Introduction

Lysosomes are acidic organelles (pH 4,5-5) enriched in hydrolytic enzymes and are an essential component of the endosomal-autophagic-lysosomal system. Other intervening structures in this system have been identified according to their function and roles: early endosomes, recycling endosomes and late endosomes or multivesicular bodies (MVB) (Hu et al., 2015) which can degrade their cargo by fusing with lysosomes, or release their cargo extracellularly by fusing to the cell membrane (Tancini et al., 2019). The lysosome is, therefore, typically considered a final destination of the endocytic pathway where proteolytic degradation takes place. Cells deficient in lysosomal hydrolytic enzymes or lysosomal membrane proteins accumulate excessive levels of undegraded substrates which sets off a cascade of events that affect other organelles such as mitochondria, endoplasmic reticulum, Golgi, peroxisomes, and overall cell function (Platt et al., 2012).

Lysosomal storage diseases (LSDs) are a heterogeneous group of inherited metabolic disorders that are characterized by the accumulation of undegraded substrates in the cells. This is caused by mutations in genes encoding for lysosomal enzymes or for transporters involved in the degradation or export of degraded molecules, which in turn lead to lysosomal dysfunction and to a broad spectrum of clinical manifestations depending on the specific substrate and site of accumulation (Sun et al., 2015). Gaucher Disease (GD, OMIM #230800, ORPHA355) is the most common type of LSDs, affecting 1 in 40,000 to 1 in 60,000 births in the general population, but reaching 1 in 800 births in the Ashkenazi Jewish population. GD presents an autosomal recessive inheritance pattern and it is caused by mutations in the *GBA1* gene which leads to a decrease in the activity of the lysosomal enzyme glucocerebrosidase (GCase, EC: 4.2.1.25) that hydrolyzes glucosylceramide (GlcCer) into ceramide and glucose. The consequence of this is the typical accumulation of GlcCer in the cells, particularly in macrophages. The pathophysiological mechanisms of neurological involvement in GD, however, remain poorly explained, with evidence for low GlcCer turnover in neurons and for high-level accumulation only when residual GCase activity is drastically decreased. The diagnosis of GD is typically performed through direct sequencing of *GBA1* and quantification of GCase enzymatic activity in leukocytes, mononuclear cells, or cultured fibroblasts. The enzyme activity in neuropathic GD cases is approximately 10-15% of the normal value and total absence of GCase is associated with early lethality (Weiss et al., 2015).

The *GBA1* gene is located on the long arm of chromosome 1 (1q21) and contains 11 exons. The existence of a highly homologous pseudogene (*GBAP1* - 96% sequence identity) just 16 kb downstream at the same locus, also accounts for many recombination events between the two. More than 495 mutations have been described in the *GBA1* gene (Do et al., 2019). Between the most common *GBA1* mutations, usually missense mutations leading to a misfolded protein, homozygotic c.1448T>C (L444P) mutation is associated with high risk for developing neurological impairment and patients carrying two null mutations as either homozygous or compound heterozygous usually present total absence of GCase activity and the most severe phenotypes(Stirnemann et al., 2017). Heterozygous or homozygous mutations in *GBA1* gene are also considered a risk factor for developing Parkinson's disease (PD) and Lewy Body disorders, with the prevalence of GD heterozygous mutations in caucasian PD ranging from 3-8%, and higher in the Ashkenazi Jewish population (15-31%) (Sidransky & Lopez, 2012). Based on large population studies, it is known that among *GBA1* mutation carriers, about 9.1% will develop PD, with some reports suggesting that the degree of penetrance rises to 30% at 80 years of age. Overall these patients also present an earlier age of onset, 3 to 6 years earlier in heterozygous carriers, and about 6 to 11 years earlier in patients with homozygous *GBA1* mutations, and the progression of disease has been characterized by a more pronounced cognitive deficit with a threefold increased risk of developing dementia (Riboldi & Di Fonzo, 2019). The cause for this link most likely involves the compromised α-

synuclein degradation by lysosomes with impaired GCase function. GlcCer has also been shown to stabilize soluble oligomers of α-synuclein, which then aggregate and form Lewy bodies in the nerve cells. Furthermore, α-synuclein polymers have an inhibitory effect on GCase, generating a vicious circle (figure 1) (Stirnemann et al., 2017).

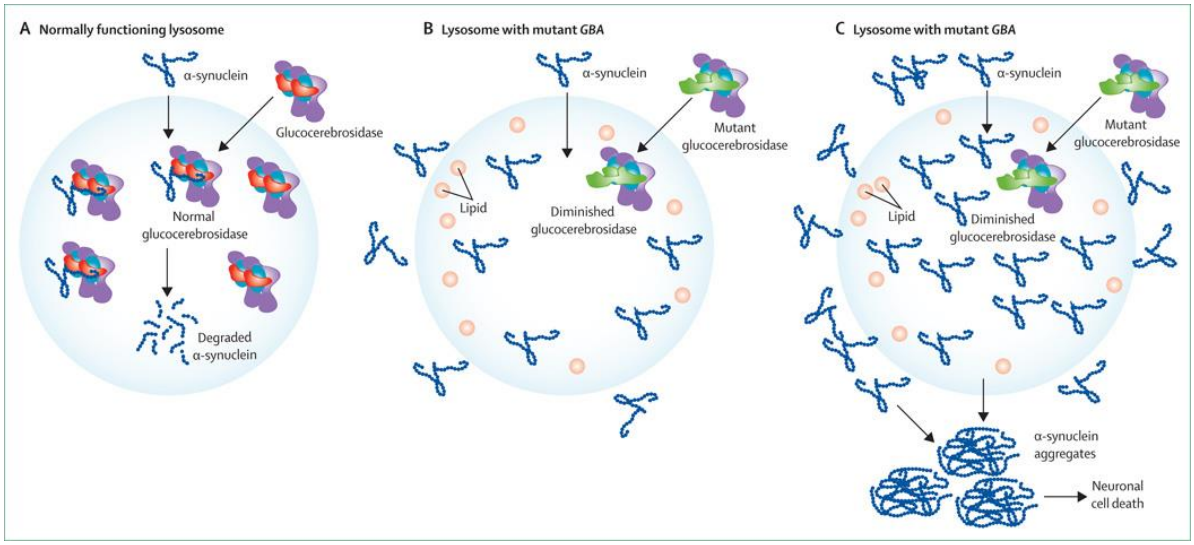


Figure 1 – Sidransky and Lopez, 2012 *Lancet Neurol*. Hypothesis on how mutant GCase increases α-synuclein aggregation: (A) In the normally functioning lysosome, WT GCase might interact with α-synuclein, facilitating α-synuclein degradation. (B) In most cases, when GCase is mutated, α-synuclein remains in the monomeric form and other processes are active in its degradation. (C) In some patients, GCase is mutated and the cell is unable to degrade α-synuclein. Lysosomal function is compromised and increased oligomeric forms of α-synuclein lead to neuronal cell death and the development of PD.

Although the phenotype is very variable, three clinical forms of GD have been identified: type 1, the most common and with clinical manifestations limited to the hematopoietic system, skeletal system and visceral organs without neurological involvement, and types 2 and 3, typically characterized by neurological impairment. With the increasing research however, it became clearer that GD represents a phenotypic spectrum, varying between extrapyramidal syndrome in type 1 to hydrops fetalis in the most severe cases of type 2. Genotype-phenotype correlation is very challenging, which leads to believe that other genetic and environmental modifiers may play a role in the phenotypic manifestation of GD. Thus, the broader term "neuronopathic GD" describing types 2 and 3 has been proposed, which classically refer to acute and chronic presentations, respectively (Gupta et al., 2011). During this dissertation, however, the commonly known classification will be applied.

Type 2 GD (ORPHA77260) is the rarest and most severe form of GD with symptoms including seizures, splenomegaly, thrombocytopenia, growth retardation and lung lesions presenting either prenatally or during early infancy. Death typically occurs before the third year of life, with the mean survival age at 11.7 months. Fetal GD manifesting with hydrops fetalis is the rarest form of the disease, with death often occurring in utero or soon after birth (Stirnemann et al., 2017).

The treatment of somatic symptomatology of GD is currently done though enzyme replacement therapy (ERT) or substrate reduction therapy (SRT). ERT involves periodic infusions of wild type GCase recombinant protein into the bloodstream, however, due to its high molecular weight, it cannot cross the blood–brain barrier and is therefore not effective in treating neuronopathic symptoms. SRT pharmacologically inhibits glucosyl-ceramide synthase, decreasing the rate of synthesis of glucosylceramide. While these compounds are small enough to cross the blood–brain barrier, there has been no reported improvement in neurological symptoms in a clinical setting (Stirnemann et al., 2017). Other alternatives are being investigated such as molecular chaperones or gene therapy.

Several *in vitro* and *in vivo* models have been developed over the years for studying GD, however, for the neuropathic phenotype, this is more challenging than for the somatic manifestations of the disease. As tissue specimens from CNS is limited to post-mortem samples, which only represent the very advanced stages of the disease, several mouse models have been developed, starting with the administration of chemical inhibitors of GCase. However, this fails to represent the full physiological impact of the disease in the organism. Genetic mouse models have been developed as well, with very limited neurological impairment or with phenotypes so aggressive when in homozygoty, that the mouse pups did not survive long after birth. Nonetheless, these mouse models have been very useful for studying the non-neuropathic or somatic symptomatology of GD and, when combined with chemical inhibition of GCase have been applied to some studies regarding the progression of the neuropathic phenotypes (Santos & Tiscornia, 2017). The species to species variability in terms of genomic regulation and pathophysiology can limit the application of mouse models however, particularly for diseases with a high heterogenic manifestation such as neuropathic GD. Thus, induced pluripotent stem cells (iPSCs) have emerged as a very useful method for disease modelling in which human somatic cells can be reversed to a pluripotent state through direct reprogramming via ectopic expression of Oct4, Sox2, Klf4 and c-Myc transcription factors (Takahashi & Yamanaka, 2006), and used as a possible source of any type of cells needed for research or therapeutic application maintaining their original genetic background (Santos & Tiscornia, 2017).

In an effort to produce *in vitro* disease models, particularly to obtain isogenic controls, one genetic engineering technique has been wildly applied due to its high efficiency and enormous applicability in different organisms, RNA-guided clustered regularly interspaced short palindromic repeat/Cas9 (CRISPR/Cas9). This technique can be tailored to be highly specific by using a combination of single guide RNA (sgRNA) and a recognizable short protospacer-adjacent motif (PAM) to enhance nuclease Cas9 targeting. The double-strand break (DSB) resulting stimulates endogenous DNA damage repair pathways in the cell, homology-directed repair (HDR) or nonhomologous end-joining (NHEJ). The first one requires a homologous template for recombination and high-fidelity repair of DSBs while the second does not use a template and thus frequently produces insertions or deletions (indels). Exogenous single-stranded oligo DNA nucleotides (ssODNs) can be designed and introduced along with Cas9 and sgRNAs to promote a specific editing at a target locus. Contrastingly, a mutant of Cas9 (Cas9n) can be generated by an alanine substitution at the catalytic residues in the conserved nuclease domains of the enzyme, enabling a single cut at the time in each of the DNA strands, targeted by two separate sgRNA molecules, and a more effective recruitment of HDR. The on-target editing efficiency using two Cas9n enzymes oriented to the same locus and originating DSBs is comparable to the wild-type Cas9, but indels at predicted off-target sites are significantly reduced by applying the first method (Trevino & Zhang, 2014).

In this project we characterized iPSCs from three patients with type 2 GD and optimized CRISPR/Cas9 genetic editing to correct single-base mutations in the *GBA1* gene to obtain isogenic controls for each of the lines. Through overexpression of transcription factors, we induced the differentiation to neurons (iNs) and astrocytes (iAs) and were able to show the reduced enzymatic activity (EA) in the diseased cell lines, and the potential of using these cells as an *in vitro* model of neuropathic GD.

Methods

1. Fibroblasts cell lines

Fibroblasts from two compound heterozygous patients, G1 and G2, and one homozygous patient, G3 with mutations P415R/L444P (c.1361C>G/c.1448T>C), G325R/C342G (c.1090G>A/ c.1141T>G) and L444P/L444P (c.1448T>C), respectively, on *GBA1* were obtained from Coriell Institute. The fibroblasts were maintained in feeder-free conditions using MEF medium (DMEM, 10% FBS, 1% Glutamax) on T75 flasks (Thermo Fisher Scientific) in humidified air at 37°C with 5% CO_2. Cells were passaged at ~80% confluency using Trypsin (Thermo Fisher Scientific) and were split 1:3.

2. iPSCs generation

Generation of human induced Pluripotent Stem Cells (hiPSCs) through fibroblast reprogramming was performed by Lund Stem Cell Center ES/iPS Core Facility according to Stemgent® StemRNA™ 3rd Gen Reprogramming Kit for Reprogramming Adult and Neonatal Human Fibroblasts. The iPSCs were then maintained in feeder-free conditions using mTeSR1 medium (StemCell Technologies) on Matrigel-coated six-well plates (Corning) in humidified air at 37°C with 5% CO_2, with daily medium changes. When the culture reached ~80% confluency, cells were dissociated with Accutase (Thermo Fisher Scientific) and 2,5x10^5 cells were replated in mTeSR1 supplemented with 10µM ROCK inhibitor (Y-27632; StemCell Technologies) during the first 24h.

3. CRISPR/Cas9 plasmids generation

Semi-complementary oligonucleotides corresponding to the sgRNA target sequence (table: 1) where phosphorylated and annealed using 1 µl of each oligonucleotide (100 mM), 1 µl of 10x T4 Ligation Buffer, 0,5 µl T4 PNK and 6,5 µl ddH2O in a thermocycler at 37°C for 5min then 95°C for 5min followed by a temperature ramp down of 5°C/min to 25°C. The annealed oligonucleotides were then diluted 1:250. A digestion-ligation reaction as set up using 100 µg of the CRISPR-Cas9 vector pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid #62988), 2 µl of the diluted annealed oligonucleotide duplex, 2 µl 10X Fast Digest Buffer, 1 µl DTT (10 mM), 1 µl ATP (10 mM), 1 µl FastDigest BbsI, 0,5 µl T7 DNA ligase and ddH2O to a final volume of 20 µl. The mixture was incubated in a thermocycle at 37°C for 5 min followed by 23°C for 5 min for a total of 6 cycles. 11 µl of the ligation reaction was treated with 1,5 µl 10x PlasmidSafe Buffer, 1,5 µl ATP (10mM) and 1 µl PlasmidSafe Exonuclease and incubated at 37°C for 30min. 1-2 µl of plasmid was added to a vial of TOP10 E. coli cells thawed on ice. Bacteria were incubated on ice for 30 min before being heat shocked for 30 sec at 42°C and then incubated on ice for 2 min. 250 µl of room temperature S.O.C medium was added to the each vial before being incubated in a shaking rack for 1 h at 37°C. 100-250 µl of the transformation mix was spread on agar plates with appropriate selection antibiotic and incubated overnight at 37°C for no longer than 16 h. Plasmids were isolated using Qiagen Mini/Maxi Kits, in accordance with the manufacturer's instructions, quantified through NanoDrop 1000 Spectrophotometer (Thermo Scientific) and the kept at -20ºC.

Table 1: Oligonucleotides for sgRNA cloning

sgRNA2	**Sequence**	GCGGTGTGTCTCCCCTAGGG TGG
	Fw	CACCGGCGGTGTGTCTCCCTTAGGG
	Rev	AAACCCCTAAGGGAGACACACCGCC

sgRNA6	**Sequence**	CCTGTGTGGGCTCCAAGTTC TGG
	Fw	CACCGCCGGTGTGGGCTCCAAGTTC
	Rev	AAACGAACTTGGAGCCCACACCGGC

4. CRISPR/Cas9 genetic editing

On day -1, ~80% confluent hiPSCs were dissociated with Accutase and $6x10^6$ cells per well were plated on Matrigel-coated six-well plates with mTeSR1 supplemented with 10µM ROCK inhibitor during the first 24 h in humidified air at 37°C with 5% CO_2. One day later (Day 0), medium was replaced with fresh mTeSR-1 medium with 10 µM ROCK inhibitor. 6 µl LipoStem (Thermo Fisher Scientific) was added to 100 µL OptiMEM (Thermo Fisher Scientific) in one tube and different quantities of CRISPR-Cas9 and singe-stranded donor oligonucleotide (ssODN, table: 2), with a total amount of 2 µg (1 µg CRISPR-Cas9 + 1 µg ssODN; 1,5 µg CRISPR-Cas9 + 0,5 µg ssODN; 1 µg CRISPR-Cas9 + 0,5 µg ssODN1 + 0,5 µg ssODN2 ; 0,5 µg CRISPR-Cas9 + 0,75 µg ssODN1 + 0,75 µg ssODN2) were added to 100 µl OptiMEM in a second tube. The reactions were then mixed and incubated for 10 min at room temperature before being added to the cells. The two following days (day 1 and 2), medium was replaced with fresh mTeSR-1 with 10 µM ROCK inhibitor and Puromycin (1,25 µg/ml). On days 3 and 4, medium was replaced with fresh mTeSR-1 with 10 µM ROCK inhibitor. From day 5 medium was replaced with fresh mTeSR-1. When colonies started to grow, the cells were dissociated with Accutase and 50, 100 and $1x10^4$ cells per well were plated on Laminine-coated six-well plates with mTeSR1 supplemented with 10 µM ROCK inhibitor during the first 24 h in humidified air at 37°C with 5% CO_2. Medium was replaced with fresh mTeSR-1 with 10 µM ROCK inhibitor during the next 3 days, decreasing the ROCK inhibitor concentration to half on the 4th day, depending on cell confluency. When the cell colonies reached appropriate size, the cells were picked using a 100 µl pipette and plated on Matrigel-coated 24-well plates with mTeSR1 supplemented with 10 µM ROCK inhibitor and 1x PenStrep during the first 24 h in humidified air at 37°C with 5% CO_2. Because the picking of cell colonies was performed outside of the laminar flow chamber and to decrease the chance of bacterial contamination, 1x PenStrep was added to every mTeSR1 medium change during the first 3 days post-picking. Clones were then expanded and analyzed.

Table 2: Single-Stranded Donor Oligonucleotides

ssODN	**Sequence**
G325R sgRNA-2	T*G*GTACCTGGACTTTCTGGCTCCAGCCAAAGCgACCCTAGGGGAGACACACCGCC TGTTCCCCAACAC*C*A
C342G sgRNA-6	C*C*AGGAGCCTAGCCGCACACTCTGCTCCCaaatttagaGCCCACACAGGCCTCTGAGGC AAAGAGCATGGTGTTG*G*G

5. DNA extraction, PCR and Restriction Enzyme Digestion

Remaining cells from the regular cell split were collected in a pellet and DNA was extracted with DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. DNA concentration was verified through NanoDrop 1000 Spectrophotometer and samples were kept at -20ºC. Using specific primers (table 3), region of interest of the *GBA1* gene was amplified to check for correct editing into the genome, primarily by restriction enzyme digestion (table: 4) and then through direct sequencing. For restriction enzyme digestion, 10 µl of PCR product was mixed with 7 µl nuclease-free water, 2 µl FastDigest Buffer, and 1 µl FastDigest enzyme for each specific mutation. The mix was

incubated at 37°C for 45 min before being analyzed by Gel Electrophoresis with 1.5 - 2% Agarose in TAE buffer, supplemented with 1x GelRed Nucleic Acid Stain gels, at 90 W for 90 min, together with the untreated product and negative controls.

Table 3: *GBA1* primers

Target and Size	Sequence
***GBA1* Exon 8-9**	8 Fw: CACACACAAATTAGCTGGGT
1003 bp	9 Rev: CGACAAAGTTACGCACCCAA
***GBA1* Exon 8-10**	8 Fw: CACACACAAATTAGCTGGGT
1601 bp	10 Rev: TTCAGCCCACTTCCCAGACC

Table 4: Restriction enzymes

Product	Supplier	Cat. No
FastDigest Eco147I and 10x FastDigest Green Buffer	Thermo Fisher Scientific	FD0424
FastDigest XmaJI and 10x FastDigest Green Buffer	Thermo Fisher Scientific	FD1564

6. Zero blunt TOPO cloning

TOPO-cloning of the PCR product from the *GBA1* region of interest was performed in accordance with the manufacturer's instructions prior to being transformed into TOP10 competent cells as described above.

7. Sequencing PCR

Following PCR amplification as described above, 15 µl of diluted DNA (5 ng/µl for products between 300-1000 bp and 10 ng/µl for products > 1000 bp) and 2 µl of each primer (table: 3), was sent for sequencing using Mix2Seq Kit (Eurofins Genomics), according to the manufacturer's instructions. Sequencing data was analyzed using SnapGene (GSL Biotech).

8. Karyotyping

iPSCs at ~80% confluency from five six-well plates were incubated with KaryoMax Colcemid (Gibco) at 20 ng/ml in mTeSR1 for 45 min in humidified air at 37°C with 5% CO2. Cells were washed twice with PBS prior being dissociated with Accutase, centrifuged at 300 g for 5 min at RT, resuspended with PBS and centrifuged again. The supernatant was aspirated, and 2 ml of KaryoMax KCl 0,075 M (Gibco) previously warmed at 37ºC was added drop by drop while vortexing, and then incubated at 37°C for 10 min. 1 ml of Carnoy fixative (3:1 Methanol and Acetic Acid) previously cooled at -20°C was added drop by drop. The solution was then centrifuged at 1800 rpm for 10 min at RT and the supernatant was aspirated. Lastly, the pellet was resuspended with 2 mL of chilled Carnoy fixative drop by drop while vortexing, and additional 8 ml of Carnoy was added. Tubes were sealed and kept at -20°C before being sent for karyotyping analysis performed by Ambar Anàlisis Mèdiques (Barcelona).

9. Differentiation of Embryoid Bodies

On day -3, ~1x10^6, ~80% confluent hiPSCs were dissociated with Accutase, centrifuged at 300 g for 5min and resuspended in 12 ml of mTeSR1 supplemented with 10 µM ROCK inhibitor. 120 µl (~10^4 cells) of the solution was then plated into each well of a ninety-six-well plate with a V-bottom (Coster), centrifuged at 800 g for 10 min at RT and incubated for 72 h. On day 0, the embryoid bodies (EBs) were detached from the bottom of the wells and transferred to a 60 mm low attachment plate with mTeSR1 and incubated for 48h with medium change half-way. On day 1, EBs were transferred onto 4 Matrigel-coated chamber slides (Corning) (6 EBs per slide) and differentiation protocols were started as previously described in Pars et al 2018. For Endoderm differentiation the EBs were maintained in humidified air at 37°C with 5% CO2 for 28 days using EB medium (KO-DMEM, 10% Hyclone FBS, 1% GlutaMax, 1% NEAA, 0,1% β-mercaptoethanol) with medium changed every 2 days. For Mesoderm differentiation the EBs were maintained in humidified air at 37°C with 5% CO2 for 31 days using EB medium supplemented with 0,5 mM Ascorbic Acid, with medium changed every 2 days. For Ectoderm differentiation the EBs were maintained in humidified air at 37°C with 5% CO2 for 10 days on 60 mm low attachment plate in N2B27 medium (1:1 DMEM/F12 and Neurobasal medium, 0,5% N2 supplement, 1% B27 supplement, 1% GlutaMAX), supplemented with 10 ng/ml FGF2 with medium changed every 2 days. On day 10, EBs were transferred onto 3 Matrigel-coated chamber slides (6 EBs per slide) and maintained in humidified air at 37°C with 5% CO2 for 21 days in N2B27 medium, with medium changed every 2 days. At day 31, cells were fixed, stained and maintained in 1% potassium-phosphate-buffered saline with (KPBS) with 0,1% Sodium Azide (NaN_3) at 4ºC until further imaging.

10. Generation of induced Neurons

Neurogenin-2 (Ngn2) lentiviral overexpression was used for neuronal induction. The lentiviral vectors used were M2-rtTA (rtTA, reverse tetracycline-controlled transactivator, Addgene, Watertown, MA USA) and pTet-O-Ngn2-puro (Ngn2, Addgene), routinely produced in HEK 293T cells by Lund Stem Cell Center, Stem Cells, Aging and Neurodegeneration Group, as previously described (Zhang, Yingsha et al., 2013). On day -2, iPSCs at ~80% confluency were dissociated with Accutase and 3x10^5 cells were plated on Matrigel-coated 6-well plates with mTeSR1 supplemented with 10µM ROCK inhibitor. On day -1 medium was replaced with fresh mTeSR1, and 1 µl of each virus (rtTA and Ngn2) was added to each well. On day 0, medium was replaced with mTeSR-1 and 2,5 µg/ml doxycycline, which was added daily to the medium throughout the experiments. From day 1, supplemented BrainPhys medium (BrainPhys, 1% N2, 2% B27) was used and 72 h of puromycin (1,25 µg/ml) selection was performed. From day 5, the medium was supplemented with NT_3 (10 ng/ml) and BDNF (10 ng/ml). On day 7, the cells were dissociated with Accutase for 10 min, followed by mechanical dissociation with a 1000 µl pipette tip and incubated for another 5min, to ensure the disruption of bigger cell aggregates. Cells were then centrifuged at 300g for 5min at RT, supernatant was removed and pellets were used for RNA extraction or washed with PBS before centrifuging once more and maintained at -80°C for enzymatic activity assays.

11. Generation of induced Astrocytes

Transcription factor Sox9 (Sox9) and Nuclear factor 1 B-type (Nfib) lentiviral overexpression was used for astrocyte induction. The lentiviral vectors used were M2-rtTA (rtTA), tetO-Sox9-Puro (Sox9) and tetO-Nfib-Hygro (Nfib) routinely produced in HEK 293T cells by Lund Stem Cell Center, Stem Cells, Aging and Neurodegeneration Group, as previously described (Canals et al., 2018). On day -2 hiPSCs at ~80% confluency were dissociated with Accutase and 3x10^5 cells were plated on Matrigel-coated 6-well plates with mTeSR1 supplemented with 10 µM ROCK inhibitor. On day -1 medium was replaced

with fresh mTeSR1, and 1 µL of each virus (rtTA, Sox9 and Nfib) was added to each well. On day 0, medium was replaced with mTeSR-1 and 2,5 µg/ml Dox, which was added daily to the medium throughout the experiments. From day 1, supplemented Expansion medium (EM) (DMEM/F12, 10% FBS, 1% N2) was used and 72h of puromycin (1,25 µg/ml) and 120 h of hygromycin (200 µg/ml) selection were performed. From day 3, medium was gradually changed to FGF medium (Neurobasal, 2% B27, 1% NEEA, 1% GlutaMax, 1% FBS, 8 ng/ml FGF, 5 ng/ml CNTF, 10 ng/ml BMP4) starting from 75% of EM + 25% FGF on day 4, then 50% of each mediums on day 5 and 25% EM + 75% FGF on day 6. On day 7, the cells were then dissociated with Accutase over 10 minutes and centrifuged at 300 g for 5min at RT, the supernatant was removed and the pellets were used for RNA extraction or washed with PBS before centrifuging once more and maintained at -80°C for enzymatic activity assays.

12. Plating and maintenance of cocultures

For preparing cocultures, on day 7 of the previously described protocols, the dissociated iNs and iAs after centrifugation were resuspended in BrainPhys and FGF mediums, respectively, and were passed through 40 µm strainers to discard bigger cell aggregates. The cells were then manually counted and $7{,}4x10^4$ iN were plated with $2{,}6x10^4$ iAs on poly-d-lysine- (Sigma) and laminin-coated (Sigma) 13 mm coverslips. On day 8, medium was replaced with 1:1 BrainPhys and FGF medium. Cells were kept for 30–32 additional days, during which half of the medium would be changed every two days using 1:1 BrainPhys and Maturation medium (1:1 DMEM/F12 and Neurobasal, 1% N2, 1% Na Pyruvate, 10 µg/µl NAC, 10 µg/µl hbEGF, 10 ng/ml CNTF, 10 ng/ml BMP4, 500 µg/ml cAMP).

13. RNA extraction, cDNA synthesis and Real Time qPCR

All RNA extractions were performed according to the provided manufacturer's instructions on RNeasy Mini Kit (Qiagen). To prevent DNA contamination, RNA was treated with DNase I (Qiagen). The yield of RNA was determined with NanoDrop 1000 Spectrophotometer. 1 µg of RNA was then reverse-transcribed using qScript cDNA Synthesis Kit (QuantaBio) following the manufacturer's instructions. cDNA was diluted 1:6,75 with nuclease-free water and ~50 ng of cDNA were used for each well together with 1xTaqMan Universal PCR Master Mix and 1x TaqMan Gene Expression assay (table: 5). The assays were performed using iQ5 Multicolor Real-Time PCR Detection System (Bio Rad) during 45 cycles of amplification and the cycle threshold (Ct) value was used to quantify the expression of each gene. GAPDH expression was quantified in all experiments and used as a normalizer.

Table 5: TaqMan Assays

Assay	Supplier	Cat. No
GAPDH	Thermo Fisher Scientific	hs02786624_g1
NANOG	Thermo Fisher Scientific	hs02387400_g1
POU5F1	Thermo Fisher Scientific	hs00742896_s1
SOX2	Thermo Fisher Scientific	hs01053049_s1
MAP2	Thermo Fisher Scientific	hs00258900_m1
SYP	Thermo Fisher Scientific	hs0030053_m1
S100B	Thermo Fisher Scientific	hs00902901_m1
GFAP	Thermo Fisher Scientific	hs00909233_m1

14. GBA enzymatic activity

Frozen pellets of ~1,5x10^6 cells were resuspended in 90µL of extraction buffer (63,2% of 0,2M Na_2HPO_4 (Sigma), 36,8% of 0,1M Citric Acid (Sigma), 1% Sodium Taurocholate Hydrate (Sigma), 0,2% Triton X-100 (Acros Organics)) and were disrupted by sonication at 10s on/off intervals at 70 Amp on ice for a total of 2 min. Samples were then centrifuged at 20.000 g for 20 min at 4°C and supernatant was collected as a cell extract for protein quantification according to the manufacturer's instructions on CommasiePlus Bradford Assay (Thermo Fisher Scientific). For enzyme activity assay, 20 µL of cell extract was then added to 200 µL of GC assay buffer (63,2% 0,2 M Na2HPO4 , 36,8% 0,1 M Citric Acid, 15 mM 4-methylumbelliferylglucopyranoside (Sigma), 0,25% Sodium Taurocholate Hydrate, 0,15% Triton X-100, 0,1% BSA (Sigma)). The reaction was incubated at 37°C for 1 h before 800 µl of stop solution (500 ml: dH2O, 3,755 g Glycine (Sigma), 1,545 g NaOH (Riedel-de Haën), adjusted to pH 10,4) was added, and fluorescence was measured (Excitation: 365 nm and Emission: 460 nm) and the GCase activity was calculated, as one GCase activity unit (U) releases 1 nmol of 4-methylumbelliferone per hour, using the formula (Jie Lu et al., 2010) (Lee KO et al., 2005):

$$\left(\frac{\boldsymbol{X}\ ng\ 4MU\ (sample - blank)}{ml}\right) * \left(\frac{ml}{0{,}02\ mL\ sample}\right) * \left(\frac{mL\ sample}{\boldsymbol{Y}\ mg\ Pr}\right) * \left(\frac{1nmol\ 4MU}{176{,}17ng\ 4MU}\right) * \left(\frac{1}{1h}\right) =$$

$$0.283817 * (\boldsymbol{X}/\boldsymbol{Y})\ nmol\ /\ h\ \cdot\ mg\ protein$$

15. Immunocytochemistry

At the corresponding time-point, cells were washed two to three times with KPBS and fixed in 4% paraformaldehyde (PFA) for 20 min at RT, before being washed two to three times with KPBS and blocked in a blocking solution (KPBS containing 0.025% Triton X-100 (TKPBS) and 5% normal donkey serum (NDS; Millipore) for 60 min. The primary antibody was then added in blocking solution and incubated overnight at 4°C. Cells were washed twice for 5 min with 0.025% TKPBS and once for 5min with blocking solution and secondary antibodies and Phalloidin were added in blocking solution and incubated for 2 h at RT. Washes were performed once with 0.025% TKPBS and DAPI for nuclear staining for 10 min, and twice for 5 min with KPBS and rinsed with mili-Q water. Coverslips were mounted using PVA:DABCO.

Table 7: Primary Antibodies and Probes

Product	Concentration	Supplier	Cat. No
Rabbit Polyclonal anti-Tubulin β-3	1:1000	BioLegend	PRB-435P
Mouse Monoclonal Pax-6	1:100	Santa Cruz Technology	sc-32766
Mouse Monoclonal Anti-α-Fetoprotein	1:100	Sigma Aldrich	A0008
Mouse Monoclonal Anti-Actin, α-Smooth Muscle	1:500	Sigma Aldrich	A5228
Chicken polyclonal Anti-MAP2	1:1000	Abcam	ab5392
Goat Polyclonal anti- Human Cathepsin D	1:500	Bio-techne	AF1014

Rabbit polyclonal anti-Human CD63	1:500	Bio-techne	MAB50481
Alexa Fluor™ 647 Phalloidin	1:400	Thermo Fisher Scientific	A22287
Alexa Fluor™ 568 Phalloidin	1:200	Thermo Fisher Scientific	A12380

16. Image acquisition and analysis

Images were obtained using a Zeiss 780 confocal microscope and the Zen lite software. Processing of images was performed using Fiji Software (Schindelin et al., 2012).

17. Statistical analysis

Data are presented as the mean ± s.e.m of three or four independent experiments. Statistical analyses were performed using Prism 8 software and unpaired t-tests were applied. Significance was set at $p<0.05$ for all experiments.

References for all remaining regents can be found in Supplementary table 1.

Results

1. Characterization of iPSC lines obtained from GD patients' fibroblasts

In order to obtain iPS cell lines with *GBA1* mutations required for this study, fibroblasts from three type 2 GD patients were reprogramed by Lund Stem Cell Center ES/iPS Core Facility and maintained all throughout this project. G1 (GM08760) and G2 (GM02627) cell lines were derived from 11 and 3 months old, respectively, female compound heterozygous patients. G1 carried the P415R and L444P mutations, in which there is a G to C transversion in nucleotide 1361 of exon 9, resulting in an arginine being substituted for proline at codon 415 in one allele; and a T to C transition in nucleotide 1448 of exon 10, resulting in a proline being substituted for leucine at codon 444, in the second allele. G2 carries the G325R and C342G mutations within the exon 8 in which one allele has a 1090 G to A transition on nucleotide 1090, leading to a substitution of arginine for glycine at codon 325 and the second allele has a T to G transition on nucleotide 1141, resulting in a substitution of glycine for cysteine at codon 342. The G3 (GM08760) cell line is homozygous for the L444P mutation in exon 10. Both alleles present T to C transitions on nucleotide 1448, generating substitutions of proline for leucine at codon 444.

1.1. iPSC lines are genotypically stable and express pluripotency genes

After reprogramming, two clones for each line were selected for karyotype evaluation: G1B1, G1B4; G2A4, G2B3; G3A1, G3C1. Since there is evidence for a relatively high incidence of aneuploidies and chromosomal aberrations in human iPS cell lines, even at early passages (Mayshar et al., 2010) (Laurent et al., 2011) it is very important to verify the genomic integrity of the cell lines and prevent the influence of chromosomal abnormalities in the differentiation capacity of the cells and the interpretation of biological findings. Except for clone G1B4, where no mitoses could be detected in the analysis, all karyotyping results were normal and one clone for each line could be selected for the experiments. G2A4 cell line was the main focus for this particular thesis, with a normal karyotype, matching the paternal fibroblast line: 46, XX (figure 2A).

Next, to prove the pluripotency of the cell line, RT-qPCRs were performed and expression of three pluripotency genes was analyzed. *NANOG*, together with octamer-binding transcription factor 4 (*OCT4*, also known as *POU5F1* (HGNC:HGNC:9221), and sex-determining region Y HMG-box 2 (*SOX2*) are genes responsible for maintaining cells in an undifferentiated state (Gawlik-Rzemieniewska & Bednarek, 2016). Expression of *NANOG* in particular, is associated to late-stage reprogramming(Schmidt & Plath, 2012). When compared to the relative expression in the fibroblast cell line, G2A4 iPSCs presented a pronounced upregulated expression of the pluripotency genes (figure 2B, 2C and 2D) which is an indication of successful reprogramming and maintenance of pluripotency state.

1.2. iPSC lines can differentiate into the three germ layers

For assessing the ability of the cells to differentiate towards the three germ layers, embryoid bodies were generated and differentiated using specific mediums containing exogenous growth factors and supplements as previously described. Fibroblast Growth Factors 2 (FGF2) was added to the differentiation medium for ectoderm, as it has been shown to have neural inductive activity by selectively expanding the neuroepithelial population, stabilizing the neural identity of the cells (Du & Zhang, 2004). Ascorbic acid, or vitamin C, was added to the differentiation medium for mesoderm, as it has been reported to drive stem cell differentiation toward myogenesis and osteogenesis, while at the same time inhibiting neurogenic differentiation(Rahman et al., 2017). After 31 days of differentiation, cells were fixed and stained for ectoderm markers βIII-Tubulin (βIIITub), a cytoskeletal protein that is

known as neuron-specific and often applied as a specific marker for neurons(Liu, Zhong, Apostolou, & Fang, 2013) and Pax6, a transcriptional determinant of the human neuroectoderm fate by differentially targeting pluripotent and neuroectoderm genes(Zhang, Xiaoqing et al., 2010). On the other hand, α-Smooth Muscle Actin (α-SMA) has been shown to mark the onset of cardiomyocyte differentiation during *in vivo* cardiogenesis(Clement et al., 2007) and thus, it has been used as a specific marker for mesoderm differentiation. And finally, α-fetoprotein (AFP), has been shown to be expressed in the yolk sac, hindgut/midgut endoderm, and the foregut hepatic diverticulum at 26 days postovulation in human early development(Mizejewski, 2004) therefore it has been used as a specific marker for endoderm. All stem cell-derived germ layers stained positive for the specific markers (figure 2I-2L).

Overall, our results show that the G2A4 iPSC have the capacity to differentiate into all three germ layers which proves their identity as pluripotent stem cell lines. This together with the genotypic assessment and expression of pluripotency genes validate their further application to the study in question.

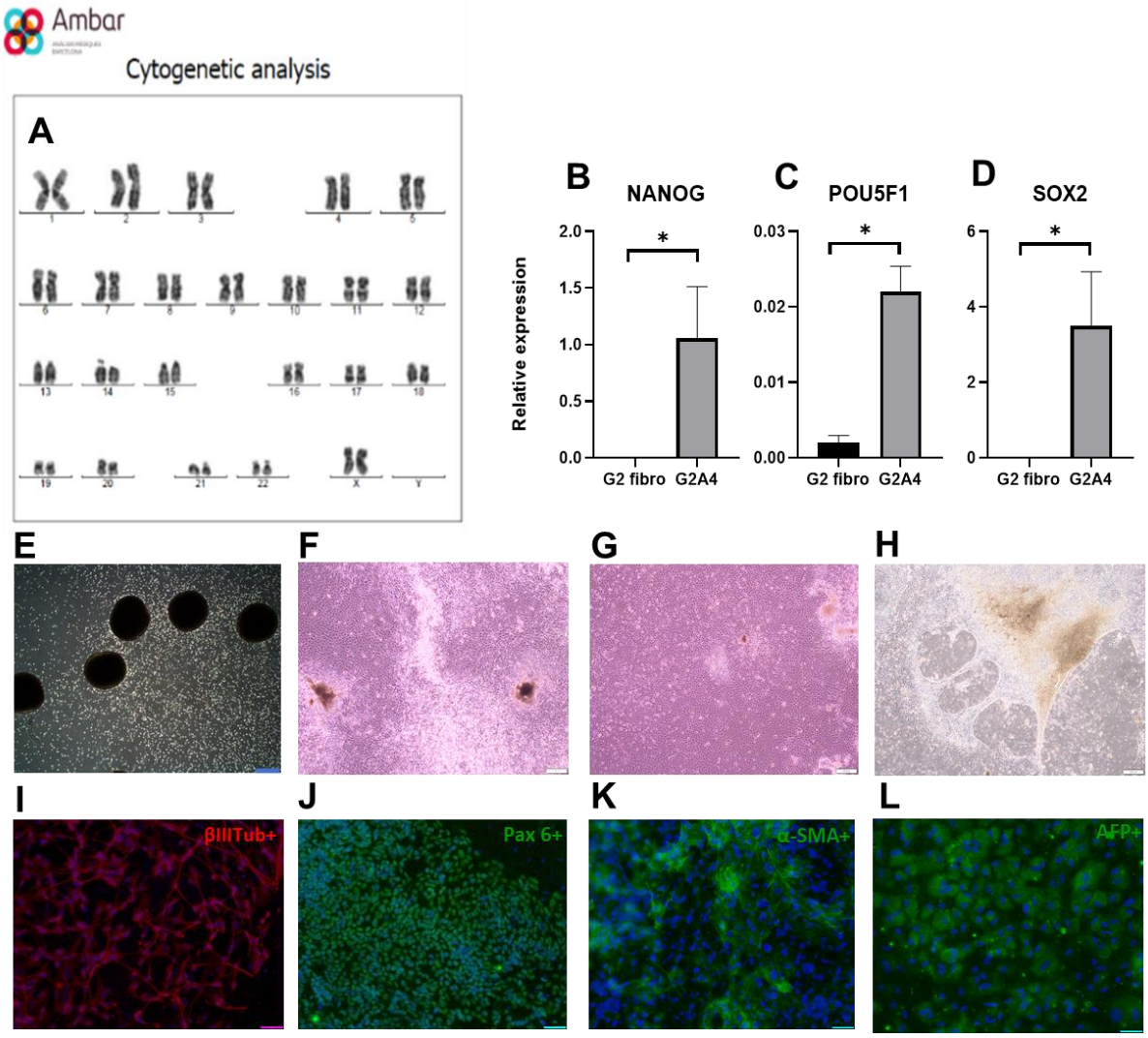


Figure 2: Characterization of G2A4 iPSCs

(A) Cytogenetic analysis of G2A4 showing 46 chromosomes and the sex chromosomes XX. (B) (C) (D) RT-qPCR gene expression analysis of NANOG, POU5F1 and SOX2, respectively. Gene expression is normalized to GAPDH expression. Data is shown as mean ± s.e.m. from three independent experiments with three technical replicates. **** p value < 0.0001, *** p value < 0.001, ** p value < 0.01, * p value < 0.05, ns: not significant, one-tailed unpaired t-test, comparing the iPSCs to their corresponding fibroblasts. (E) EBs in suspension after picking on day 4. (F) (G) (H) Images of EBs on day 18 of differentiation protocols: mesoderm, endoderm and ectoderm, respectively. Amp: 4X. Scale bars = 100 µm. Immunocytochemistry for germ layer specific markers after 31 days of differentiation: (I) (J) Ectoderm, BIII-tubulin (red), Pax6 (green); (K) Mesoderm, α-SMA (green); (L) endoderm, AFT (green). Amp: 20X. Scale bars = 50 µm.

2. Correction of disease-causative mutations with CRISPR/Cas9 genome-editing and generation of isogenic control iPSC lines

The quality of a disease model, particularly an *in vitro* one, depends largely on the ability to obtain appropriate controls. Ideally, phenotypes observed in mutant iPS derived cell-lines should be compared to control cell lines with the same genetic background and that differ only in the disease-related mutation. This is because the differences identified could result or be influenced by epigenetic markers, unmatched gender, age and ethnicity; differences in the protocols used for pluripotency induction or in passage number (Musunuru, 2013). By obtaining isogenic controls directly from the GD lines, we reduced the influence of external factors on biological observations and directly compared the findings in the corrected line for *GBA1* with the original diseased one.

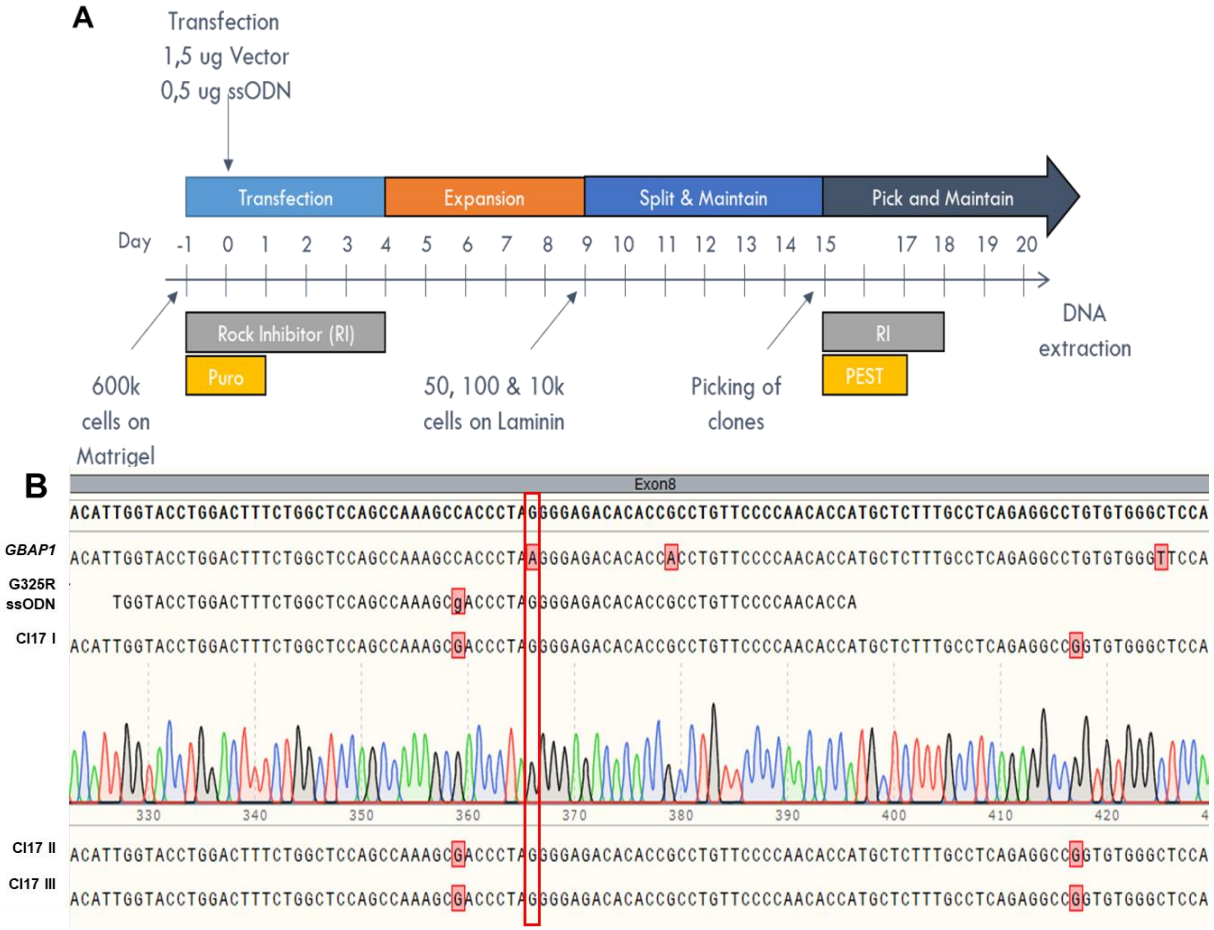
To obtain isogenic controls, we used CRISPR/Cas9 editing tool on the iPSC lines. In an initial stage, before the start of this dissertation project, sgRNAs efficiency was tested by cloning each individual sgRNA into pX459 CRISPR-Cas9 vector or pX462 CRISPR-Cas9n vector. The G2 line was then transfected with the individual pX459 vectors, each containing a different sgRNA; or with a combination of two pX462 vectors (expressing Cas9n), as each of the DNA strands need to be targeted for the genome editing to occur. Both vector variants carry a puromycin resistance gene which allow for selection of transfected cells. The efficiency of the different sgRNAs and vector combinations was assessed by their ability to induce small indels at the mutation site, thus the most efficient G325R pX459 sgRNA-2 and C342G pX459 sgRNA-6 were chosen to proceed with.

To maximize the probability for correcting one of the mutations in one of the alleles, transfection experiments were performed side by side for both mutations and at the same time and in the same conditions (figure 3A). Both ssODN applied had between 75-85 nt with 30-35 nt homology arms on both sides and were complementary to the sgRNAs used, as this has been shown to result in a higher efficiency for introducing single-base substitutions, with the best results in balancing HR efficiency and the subsequent cell cytotoxicity (Okamoto et al., 2019). Phosphorothioate bonds were also introduced between 3 nucleotides at 5' and 3' ends of ssODN, in order to prevent their degradation by endogenous nucleases and thereby increasing the effective concentration of ssODN available during the DNA repair (Renaud et al., 2016). Additionally, both ssODN presented silent mutations within the PAM sequence (one for the G325R), or at the 5' neighboring base of PAM (three for C342G due to lack of silent mutation alternatives for codon corresponding to PAM) to prevent re-cutting by the CRISP/Cas9 machinery once the editing is performed.

All isolated and surviving clones that presented a normal iPSC morphology were analyzed through restriction enzyme digestion and further gel electrophoresis. For the G325R mutation, the XmaJI restriction enzyme would cut a mutated sequence once, originating two bands with 379 bp and 624 bp; and a corrected sequence twice, originating three bands with 79 bp, 300 bp and 624 bp. After this, promising clones were sent for direct sequencing. In total, 10 clones were sent for sequencing out of which three were successfully edited and integrated the change in the PAM sequence, resulting in an editing efficiency of 30%. Clone 17 (Cl17) presented both the editing corresponding to the change introduced to the PAM sequence and to the G325R mutation itself. This clone was then further verified through TOPO cloning of the 1003 bp sequence, in which 10 colonies were sent for sequencing and the same observations were made. When compared to the *GBAP1* sequence, there were no nucleotide matches, showing a clear difference between Cl17 sequence and *GBAP1* sequence, and confirming that the amplified and analyzed region was within the *GBA1* gene (figure 3B). All these results verified the successful editing on the G325R mutation in G2A4 Cl17.

Mutation C342G was targeted next for the G325R corrected Cl17. Surviving clones with normal iPSC morphology were selected as previously described using restriction enzyme digestion with Eco147I which only cuts the WT sequence resulting in two bands with 574 bp and 429 bp but not the mutant sequence. After this, promising clones were sent for direct sequencing. In total, nine clones were sent for sequencing out of which seven were successfully edited for the mutation, however, none integrated the change in the PAM sequence. Clone 17-7 (Cl17-7) was still selected for further evaluation through TOPO Cloning and sequencing of 10 colonies. The results showed corrections for both G325R and C342G missense mutations and the introduction of one of the silent mutations on the PAM sequence. However, there was no evidence for the introduced mutations in the PAM sequence for the C342G ssODN and some nucleotides matched the sequence of the *GBAP1* (figure 3C).

To clarify the editing for this clone, a larger region of 1601 bp, including exon 9 that naturally presents a clear difference between the *GBAP1* sequence and *GBA1*, corresponding to a large deletion in the pseudogene, was amplified and sequenced. While the editing seemed successful in the region corresponding to exon 8 (figure 3D), in the region of exon 9, Cl17 clearly shows the presence of the untargeted allele, which matches the WT *GBA1* sequence, and the presence of deletion or insertions in the targeted allele, resulting in the overlapping peaks in the sequence results of Cl17. When analyzing the sequence from Cl17-7 it was detected that a large sequence of exon 9 was missing for both alleles, and there was an overlap between this clone's and the *GBAP1* sequences (figure 3E), suggesting that an unexpected recombination might have happened and the pseudogene might have been used as a donor strand by the cell's HR instead of the ssODN supplied. Additional GCase enzymatic activity analysis showed no significant improvement for the edited clone when compared to a WT (supplementary figure 1), as it would have been expected for an isogenic control. For this reason, G2A4 Cl17-7 was discarded from all ongoing experiments.



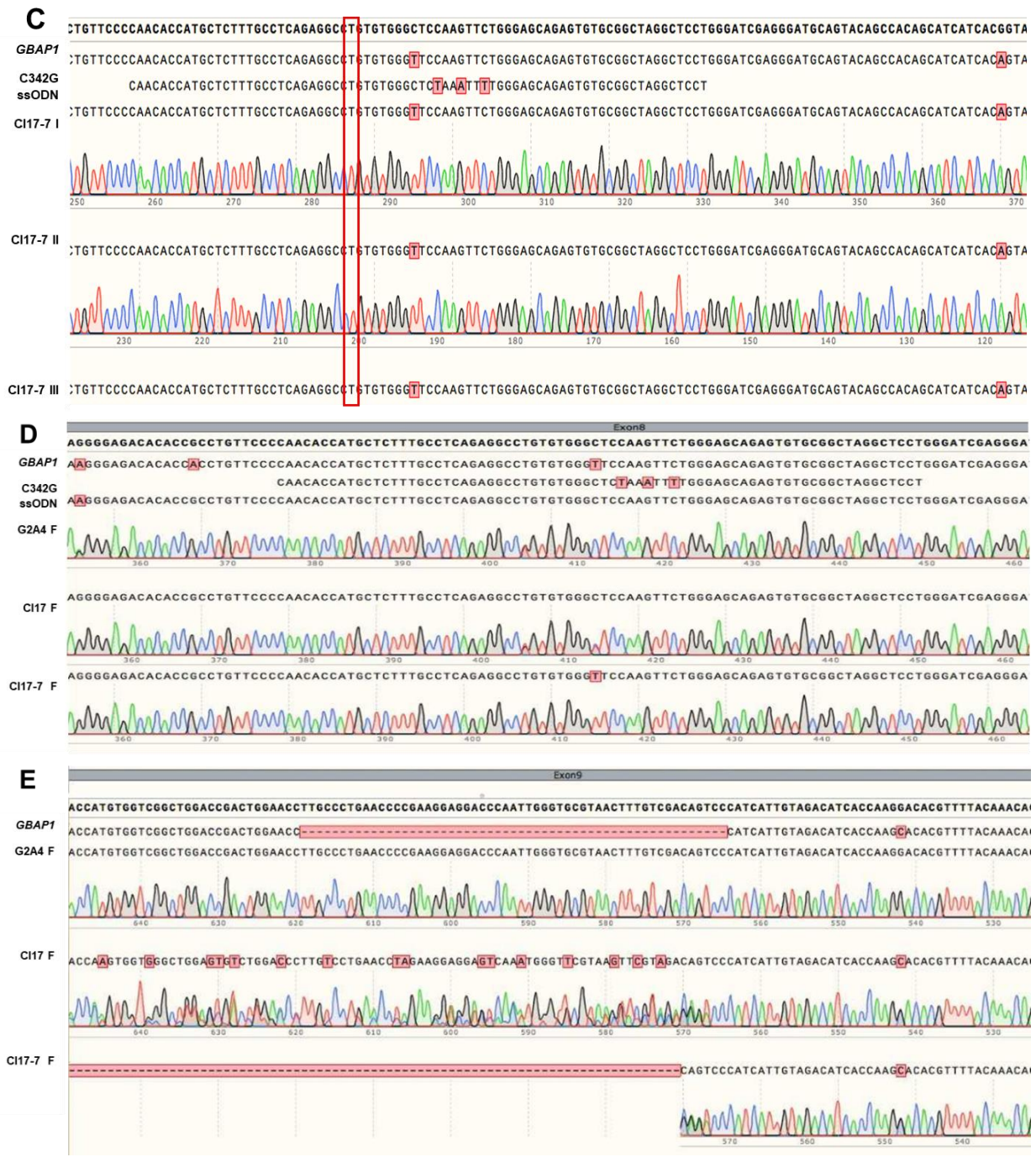


Figure 3: CRISPR/Cas9 genetic editing timeline and sequencing data of promising clones
(A) Timeline of genetic editing protocol including transfection, selection with puromycin, plating at low cell confluency, picking of clones and expansion. (B) Sequencing data of ~1003 bp amplified from Cl17 aligned with WT *GBA1*, G325R ssODN and *GBAP1* on exon 8 region to confirm that *GBA1* and not the pseudogene had been amplified. Highlighted in black is the location of G325R mutation and in red the silent mutation introduced within the PAM sequence. There is evidence for a successful editing as Cl17's sequence matches the ssODN. (C) Sequencing data of ~1003 bp amplified from potential isogenic control Cl17-7 aligned with WT *GBA1*, C342G ssODN and *GBAP1* on exon 8 region. Highlighted in red is the location of the C342G mutation and the silent mutations introduced to the ssODN. Also highlighted in red nucleotides coinciding with *GBAP1* sequence. (D) Sequencing data of ~1601 bp amplified from G2A4, Cl17 and potential isogenic control Cl17-7 aligned with WT *GBA1*, C342G ssODN and *GBAP1* on the exon 8 region. (E) Same sequencing data aligned on the exon 9 region with where a large deletion can be seen on Cl17-7 as well as several mismatches between Cl17 sequence and WT *GBA1*.

3. Generation and characterization of iNs and iAs from G2A4 and WT iPSCs lines

Given the very limited access to patient brain cells, the need for efficient and fast differentiation protocols from iPSCs became very important. Using lentiviral vectors to induce the overexpression of key transcription factors, Zhang et al, 2013 and Canals et al, 2018 succeeded in generating functional iNs and iAs, respectively, that can be used for studying the pathogenesis of neurological diseases, developing drug screening systems, or regenerative medicine (Zhang et al., 2013) (Canals et al., 2018).

Lentiviral transductions for G2A4 line and a WT iPSC control were performed to induce the overexpression of Ngn2 for iNs and Nfib and Sox9 for iAs, as described in methods (figure 4A, 4B and 4C). Exogenous Ngn2 probably induces neural differentiation by activating a transcription factor cascade, resulting in the production of excitatory neurons that have been shown to form robust synapses among themselves when cultured in the presence of astrocytes (Zhang et al., 2013). Astrocytes, one of the most abundant cell type in the CNS, play important roles in neurotransmitter trafficking and recycling, nutrient and ion metabolism, release of transmitters and growth factors, and protection against oxidative stress, thus supporting neuronal survival and function (Molofsky et al., 2012). Their development during early embryogenesis is thought to derive after a gliogenic switch, in which the division of neural progenitor cells decline, and glia (astrocytes and oligodendrocytes) start to be produced (Matuzelski et al., 2017). This process is bound to be tightly regulated and several pathways have been identified. Nuclear Factor I family members, of which Nfib is a part of, and Sox9 are intervenient in these (Wilczynska et al., 2009), and have been applied in several studies to successfully induce astrocytic differentiation (Canals et al., 2018) (Benetó et al., 2020) (Gao et al., 2019).

Differentiatiated cultures of iNs and iAs (figures 4D and 4E) were validated at day 7 post induction through RT-qPCRs for the expression of *NANOG* and *POU5F1* pluripotency markers and cell specific markers: Microtubule Associated Protein 2 (*MAP2*) and Synaptophysin (*SYP*) for iNs and Glial Fibrillary Acidic Protein (*GFAP*) and S100 calcium-binding protein B (S100B) for iAs.

Both *NANOG* and *POU5F1* expression were downregulated in G2A4 iNs and iAs when compared to the iPSC line (figures 4F and 4G), confirming the loss of the pluripotency state in these cells as expected after forcing differentiation towards neurons and astrocytes.

MAP2 encodes for a member of the cytoskeletal microtubule-associated proteins (MAPs) which participate in the binding, stabilization, and overall regulation of the microtubules mainly in neuronal dendrites (Jalava et al., 2007). *SYP* encodes for an integral transmembrane protein that in involved in many steps of endo- and exocytosis in the neuronal synaptic biogenesis (Leila Tarsa & Yukiko Goda, 2002). The expression of these genes was, therefore, used as markers for differentiated neurons capable of forming synapses. Both G2A4 and WT iNs overexpressed *MAP2* and *SYP* after 7 days, compared to the respective iPSCs (figure 4H and 4I).

GFAP is a vital intermediate filament III protein that is crucial for maintaining the thickened and elongated processes in activated astrocytes, as well as supporting neighboring neurons and the blood brain barrier in the human brain (Eng & Ghirnikar, 1994). S100B is a calcium-binding protein that is predominantly expressed in astrocytes in the CNS of mammals and evidence shows it can stimulate neural process outgrowth and astrocyte proliferation (Sakatani et al., 2008). Thus, the expression of these proteins are markers for differentiated and active astrocytes. *GFAP* and *S100B* expression was upregulated in G2A4 and WT iAs lines, compared to the corresponding iPSCs (figure 4J and 4K), confirming successful astrocytic differentiation.

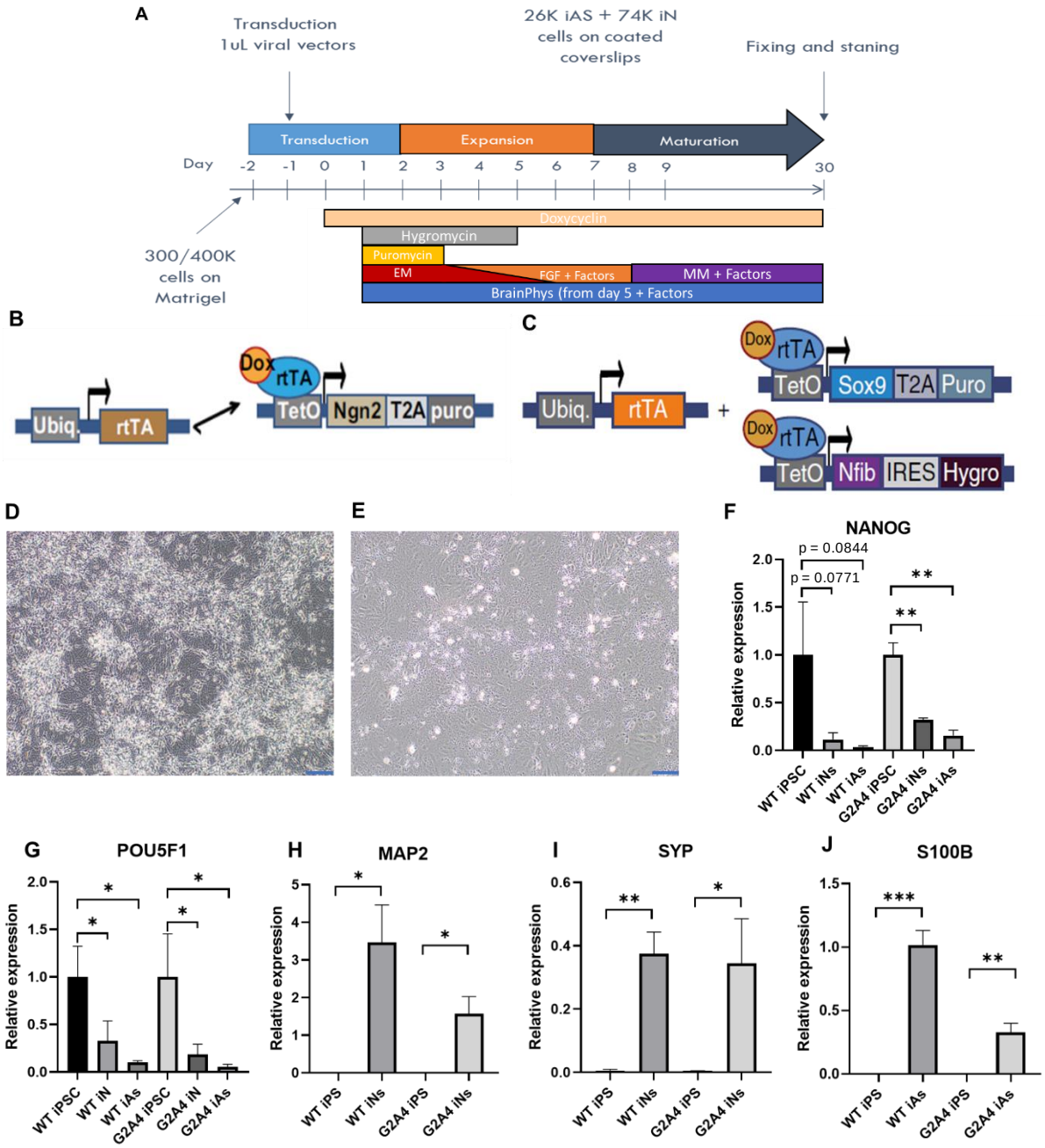


Figure 4: Differentiation timeline and lentiviral vectors and expression of iNs and iAs specific markers through RT-qPCR

(A) Timeline of differentiation protocols including transduction, induction with doxycycline, selection with puromycin and hygromycin and different medium changes. (B) Design of lentiviral vector for Ngn2-mediated conversion of iPSCs to iNs cells. Cells were transduced with viruses expressing rtTA and Ngn2/puromycin resistance. (C) Design of lentiviral vectors for Nfib and Sox9-mediated conversion of iPSCs to iAs cells. Cells were transduced with lentiviruses expressing rtTA Nfib/hygromycin resistance and Sox9/puromycin resistance. (D) (E) Images of experiment number 6 WT iNs and iAs, respectively, 5 days post induction. (F) (G) Relative expression of pluripotency genes *NANOG* and *POU5F1*, respectively, in WT and G2A4 iPSCs, iNs and iAs, normalized to GAPDH and iPSCs expression. (L) (M) Relative expression of neural-specific MAP2 and SYP in WT and G2A4 iPSCs and iNs. Gene expression is normalized to GAPDH expression. (N) (O) Relative expression of astrocyte-specific S100B and GFAP in WT and G2A4 iPSCs and iAs. Gene expression is normalized to GAPDH expression. Data are shown as the mean ± s.e.m. from three independent experiments with three technical replicates. **** p value < 0.0001, *** p value < 0.001, ** p value < 0.01, * p value < 0.05, ns: not significant, one-tailed independent t-test comparing the expression in iNs/iAs to the one in iPSC.

3.1. GCase Enzymatic Activity is reduced in GD iNs and iAs

As GD results from the deficiency or disfunction of GCase, caused by *GBA1* mutations, it is expected that a decrease in the activity of this lysosomal enzyme occurs in the diseased cells. The evaluation of GCase EA is, in fact, one of the diagnostic methods applied when GD is suspected, along with direct sequencing of *GBA1*(Weiss et al., 2015). As type 2 GD is typically accompanied by a severe phenotype and neuronal impairment, it was expected that G2A4 cell line presented a significant decrease in EA. Furthermore, it is hypothesized that the loss of GCase function may compromise lysosomal degradation of α-synuclein and lead to accumulation of oligomers resulting in neurotoxicity and the increased risk for developing PD (Stirnemann et al., 2017).

For calculating GCase activity, we collected pellets of iNs and iAs at day 7 postinduction and extracted GCase through sonication. The protein was then quantified through a colorimetric assay and the activity was assessed through the cleaving of a fluorogenically-modified substrate known as 4-methylumbelliferylglucopyranoside, which releases 4-methylumbelliferone (4-MU). A 4-MU standard curve was then used to calculate the enzymatic activity of each sample, in the form of nMol/(H*mg).

Overall, GCase activity was significantly decreased in the diseased line. G2A4 iNs presented a GCase activity of 13% relative to the WT control and 10% for iAs (figure 5A and 5B). Neuronal cultures harboring *GBA1* mutations have been shown to have lower levels of GCase activity and increased accumulation of GlcCer (Schöndorf et al., 2014) (Sun et al., 2015) (Aflaki et al., 2016). Recently, GD2 astrocyte cultures have also been shown to have decreased levels of GCase protein and reduction in GCase activity (<5% of WT control) (Aflaki et al., 2020).

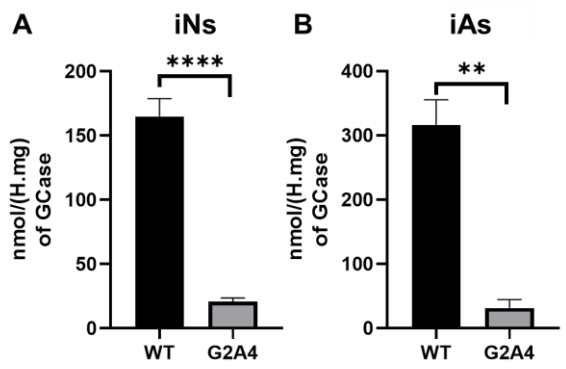
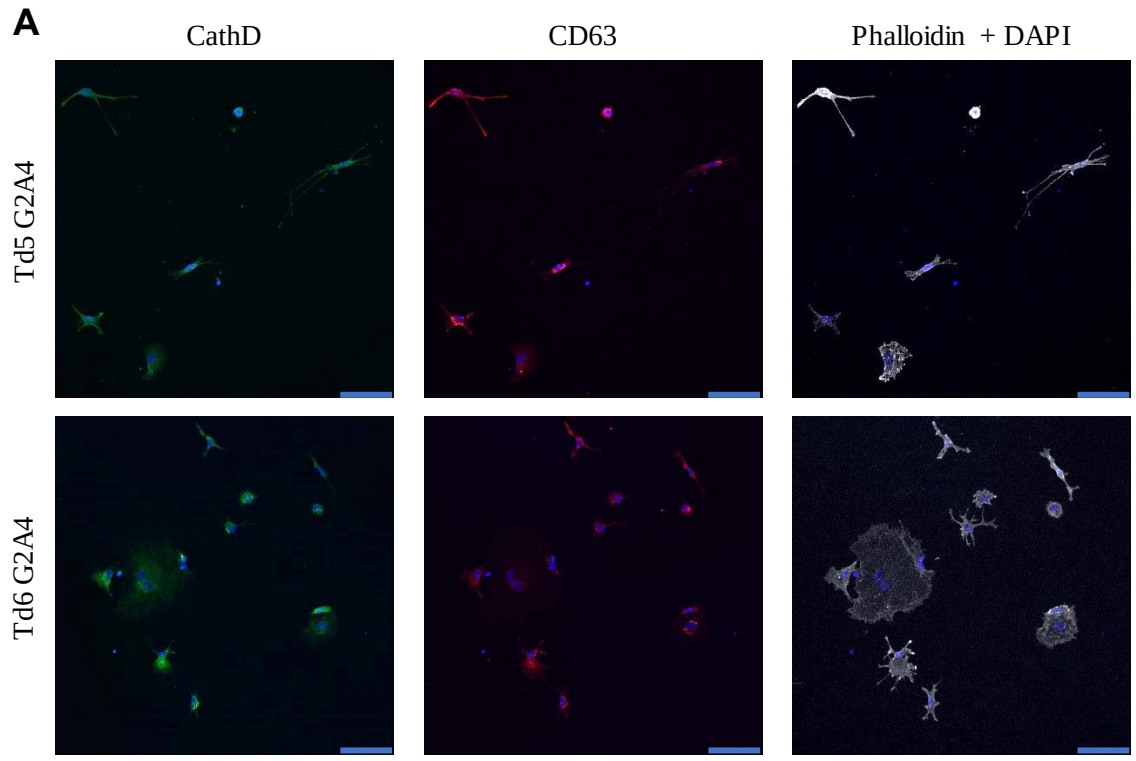


Figure 5: Enzymatic Activity in WT and G2A4 iNs and iAs
GCase activity is expressed in nMol/(H*mg) for WT and G2A4 (A) iNs and (B) iAs after 7 days of induction. Data are shown as the mean ± s.e.m. from three independent experiments, except for G2A4 iAs with data from two experiments. **** p value < 0.0001, *** p value < 0.001, ** p value < 0.01, * p value < 0.05, ns: not significant, two-tailed unpaired t-test comparing WT to G2A4.

3.2. GD iNs and iAS may have altered lysosomal function

Cocultures of iNs and iAs were plated at day 7 post induction and further differentiated for 23 days. After that, cells were stained for lysosomal markers Cathepsin D and CD63, as well as with a Phalloidin probe for F-Actine, allowing for the morphological distinction between iAs and iNs. Cathepsin D is a lysosomal protease, which is ubiquitously expressed and participates in protein turnover inside the cells and in the activation of apoptotic and necrotic pathways. In the CNS, cathepsin D is mainly released by activated microglia and has a potentially neurotoxic role. Disturbance of cathepsin balance inside the cell and extra-lysosomal localization of cathepsins has been observed in several neurodegenerative conditions, including neuropathic GD (Vitner et al., 2010) and has been shown in GD type 2 astrocyte cultures (Neurobiology of disease.1994). CD63 is a ubiquitously expressed protein highly enriched in the intraluminal vesicles (ILVs) of late endosomes, exosomes and multivesicular bodies (MVBs) and in the lysosomal membrane (Cashikar & Hanson, 2019). It participates not only in the endocytosis and intracellular transport from the plasma membrane but also in the transfer of lipids, proteins and possibly mRNA between neighboring cells (Pols & Klumperman, 2009). It is, thus, considered a marker for ILVs and MVBs (Tancini et al., 2019) and can be used to assess the overall state of intracellular trafficking.

Through immunocytochemistry, it was possible to quantify Cathepsin D and CD63 particles as well as their relative size for WT and G2A4-derived cocultures of iNs and iAs. Although non-significant, the results show a tendency of fewer Cathepsin D positive particles per cell and smaller relative sizes per particle (figure 6A and 6B) while no change was observed in CD63 positive particles in the G2A4 line compared to the WT, and no differences in the size of these particles was found either (figure 6D and 6E). These results indicate that the number degradative lysosomes might be decreased in GD neurons and astrocytes, while the endosomal pathway remains unaffected.



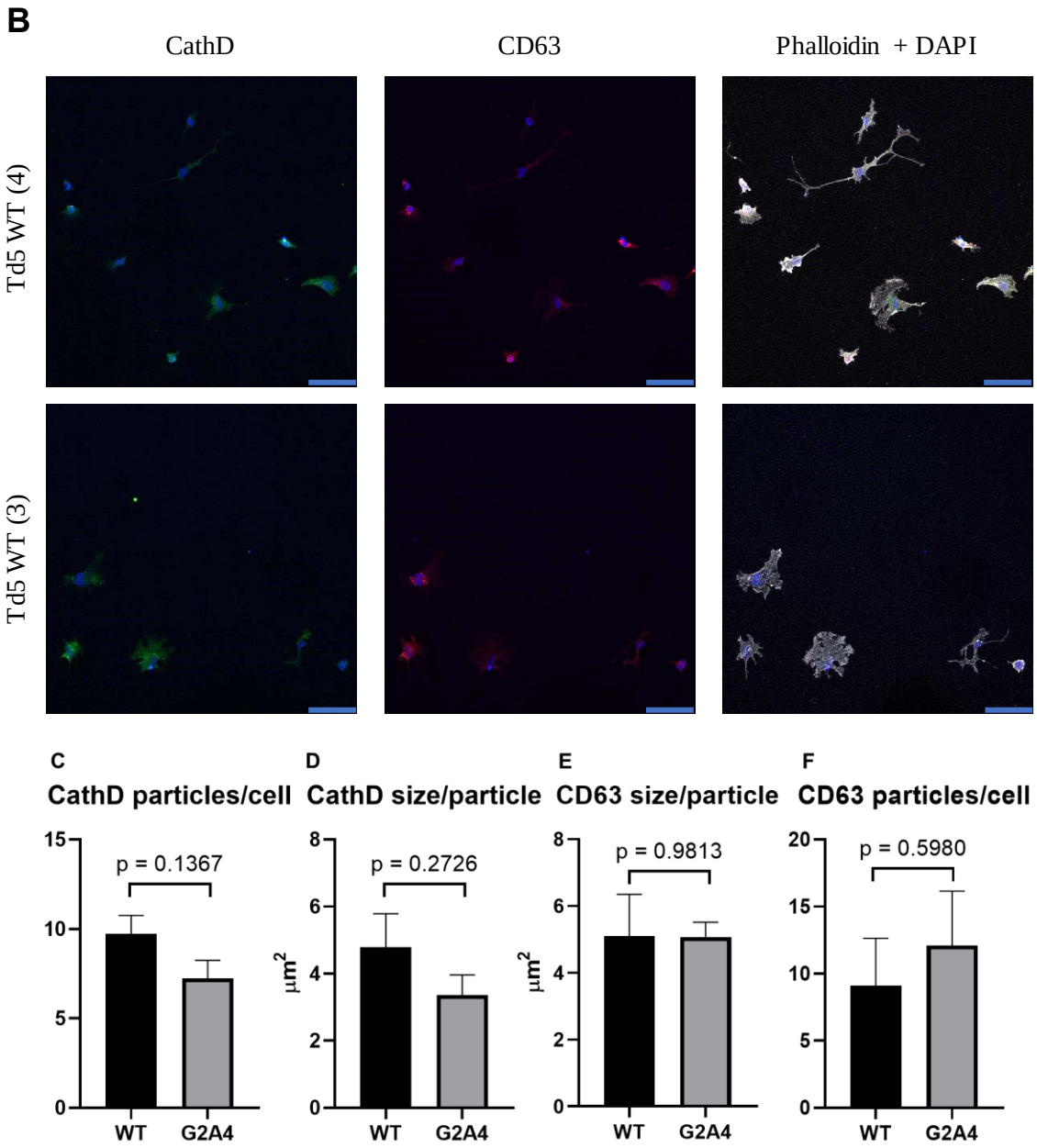


Figure 6: Imaging and quantification of Cathepsin D and CD63 particles in WT and G2A4 iN and iAs
As an example (A) G2A4 and (B) WT from experiments 5 and 6 (Td5 and Td6) and images from channels corresponding to Cathepsin D (CathD), CD63 and Phalloidin + DAPI, which were used for particle quantification, size analysis and cell counting. Amp: 20X. Scale bar = 50 µm. Particles were analyzed through Fiji win64 software on maximum intensity stacks. For CathD, threshold was set for RenyiEntropy and applied to all samples. For CD63 threshold was set at MaxEntropy and applied for all samples. (C) (D) (E) (F) Quantification results for WT and G2A4 iNs and iAs after ~30 days of differentiation. Data are shown as the mean ± s.e.m. from four independent experiments, two-tailed unpaired t-test, comparing number of particles and size between WT and G2A4.

Discussion

iPSCs have proved to be a common source of human samples, relatively simple to obtain and particularly useful for studying different phenotypes in neurodegenerative disorders and the role of each cell type in disease progression. By being able to adopt a wide array of different cell fates, iPSCs are able to provide samples that otherwise would be very difficult to obtain, and more importantly, that maintain the genetic background of the patients, thus being extremely useful for *in vitro* disease modelling. For this study we characterized and validated the application of three iPSCs lines from three patients with type 2 GD. We demonstrated the genomic stability of the cell lines, the expression of pluripotency markers and their ability to differentiate into the three germ layers *in vitro*. Although cell line G2A4 was the focus of this thesis, it is important to have different cell lines representative of the same pathology, particularly when there is such a high phenotypic heterogeny as in GD.

We then edited each cell line with CRISPR/Cas9 vectors and ssODN specific for each mutation. The selected clones were evaluated through TOPO cloning and direct sequencing until isogenic controls were found. This is a critical step for developing a reliable disease model, as an appropriate control is always required for analysing and comparing biological findings. An isogenic control should only differ from the original diseased line in terms of the mutation in the gene of interest. By editing this mutation, we would be able to have a WT form of the gene and by comparing it with the diseased cell line, we can identify the exact phenotype implications that derive from the particular mutation.

Overall, even though the percentage of cell survival was higher for the transfections with 1,5 µg of CRISPR/Cas9 vector and 0,5 µg of ssODN in all the experiments, the apparent editing efficiency for both mutations was higher when an equal amount of CRISPR/Cas9 vector and ssODN was added (1 µg each), ~12% for G325R mutation and ~18% for C342G mutations (data from one experimental replicate). As the sequencing of exon 9 from *GBA1* for both Cl17 and Cl17-7 showed clear signs of unwanted mutations and deletions resulting from the editing, the real efficiency was considered to be null. This could have resulted from several factors, the most probable being that following the CRISPR/Cas9 guided cut occurred, the cell utilized the pseudogene as a template for correcting the damaged DNA instead of the provided ssODN. As the mutations are in very close proximity of each other, a longer ssODN covering both mutations could be used to attempt a simultaneous editing.

Several unsuccessful transfection experiments were performed for two different clones of the iPS G3 cell line. This was attributed mainly to the fact that the original fibroblast line was already at a higher passage number, thus the reprograming might have not been as successful, which reflected on the cells not growing as well in culture. As cytogenetic analysis was normal and the differentiation to the three germ layers successful, further analysis of the expression of pluripotency genes could give more insight into the pluripotent state of the clones.

Although several studies for iPS-derived neurons have been performed in the context of GD, studies for astrocytes as well as for other brain cell types are still lacking. This is particularly essential as more evidence is growing regarding the role of this cell type in degrading pathological aggregates in several neuropathological disorders. We induced the differentiation of the GD iPSCs as well as a WT iPSC controls into neurons and astrocytes by lentiviral transduction and overexpression of Ngn2 and Nfib and Sox9 transcription factors, respectively. RT-qPCR at 7 days post-induction showed the successful differentiation through the expression of neuronal specific genes *SYP* and *MAP2* and astrocyte specific genes *S100β* and *GFAP*. As expected, both G1B1 and G2A4 iNs and iAs presented significantly decreased GCase activity when compared to WT and H1 iNs and iAs controls. This has been frequently

shown for neuronal cultures, as well as increased accumulation of GlcCer (Schöndorf et al., 2014), (Santos & Tiscornia, 2017). In astrocytes this is of particular interest as this cell type maintains crucial roles in phagocytosis and clearance of disease-specific protein aggregates, which has been described as a possible mechanism though which GD mutations might be a risk factor for later developing other neurodegenerative disorders, such as PD(Stirnemann et al., 2017). In a recent study, GD2 astrocyte cultures had decreased levels of GCase protein, GCase activity <5% of WT control and high accumulation of GlcCer and α-synuclein. In conformity with transcriptome data from human astrocytes and neurons that show that GCase is more expressed in astrocytes than neurons (Zhang, Ye et al., 2016) the GCase activity was doubled in the WT iAs when compared to the WT iNs, with these values decreasing dramatically in the diseased G2 line. This points to the importance of GD astrocytes and their role in the pathology of the disease and neurodegeneration and the need for further studies including this cell type. Even though for G2A4 it was not possible to rescue GCase activity in iNs and iAs, the results from the isogenic control for line G1B1, Cl2-5, which was studied in collaboration within the research group but not the focus of this dissertation, showed a rescue of GCase activity when compared to the unedited line (Supplementary figure 2A and 2B). Although not significant when compared to the WT controls, some evidence shows a correlation between a GCase activity around 50% and lack of symptomatology in heterozygous patients of GD (Santos & Tiscornia, 2017) which points to the possibility of still rescuing some of the pathological hallmarks of the disease.

iNs and iAs were also plated together in cocultures and further differentiated for 21 more days as this has been shown to increase iNs synaptic formation and provides a closer environment to the one *in vivo*. Through immunocytochemistry of the iNs and iAs co-cultures, it was possible to quantify the number and size of lysosomal components inside the cells. This was achieved though the staining for Cathepsin D and CD63. A tendency for a decline in Cathepsin D in the GD G2A4 line was observed, as it was expected according to the literature (Aflaki et al., 2020). It has been shown that the reduction in Cathepsin D activity alone is sufficient to accelerate the propagation of Lewy bodies by significantly declining lysosomal functions, while the ectopic expression of Cathepsin D was protective against α-synuclein aggregation and neurodegeneration in PD (Bae et al., 2015). The number of CD63 positive particles was unaffected in the diseased line, as well as the size of these particles. As this is a marker for late endosomes, exosomes and multivesicular bodies (MVBs), size and number increases could indicate substrate aggregation, which is expected in GD, or overall disruption of the normal flux of the endolysosomal pathway. Longer differentiation times might be needed for the cell cultures before significant alterations can be observed in lysosomal function and for the endosomal pathway to be affected. The quantification method applied could also have contributed to the lack of significance in these results. A higher amplification for the images acquired could have provided better particle resolution and analysis for the software used. Further characterization and investigation of these particles in GD cell cultures could help unveil the degree of disturbance of the endolysosomal system, possible communication between affected and healthy cells, as well as identify possible targets for drug delivery in the future. It would also be interesting to study the presence of α-synuclein and other protein aggregates in these co-cultures and if there are any significant contrasts between iNs and iAs results, as well as in co-cultures of WT iNs/GD iAs; GD iNs/WT iAs and the GD lines and with an isogenic control.

Overall, this study has contributed to the optimization of CRISPR/Cas9 gene editing method applied to single-base GD mutations in iPSCs. Through the application of transcription factor-based protocols, it was possible to show significantly low levels of GCase enzyme in GD iNs and iAs and use this as an indicator for successfully edited isogenic controls, and to characterize, to some extent, lysosomal function and alterations in iPSC-derived cells.

References

Aflaki, E., Borger, D. K., Moaven, N., Stubblefield, B. K., Rogers, S. A., Patnaik, S., et al. (2016). A new glucocerebrosidase chaperone reduces α-synuclein and glycolipid levels in iPSC-derived dopaminergic neurons from patients with gaucher disease and parkinsonism. *The Journal of Neuroscience, 36*(28), 7441-7452. doi:10.1523/JNEUROSCI.0636-16.2016

Aflaki E., Stubblefield B. K., McGlinchey R.P., McMahon B., Ory D. S. and Sidransky E. (2020). A characterization of Gaucher iPS-derived astrocytes: Potential implications for Parkinson's disease *Neurobiology of Disease*, (134), 104647. doi: 10.1016/j.nbd.2019.104647

Bae, E., Yang, N. Y., Lee, C., Kim, S., Lee, H., & Lee, S. (2015). Haploinsufficiency of cathepsin D leads to lysosomal dysfunction and promotes cell-to-cell transmission of α-synuclein aggregates. *Cell Death & Disease, 6*(10), e1901. doi:10.1038/cddis.2015.283

Benetó N., Cozar M., Castilla-Vallmanya L., Zetterdahl O. G., Sacultanu M., Segur-Bailach E., García-Morant M., Ribes A., Ahlenius H., Grinberg D., Vilageliu L and Canals I. (2020). Neuronal and Astrocytic Differentiation from Sanfilippo C Syndrome iPSCs for Disease Modeling and Drug Development. *Clinical medicine*, (9,3). 644. doi: 10.3390/jcm9030644

Canals, I., Ginisty, A., Quist, E., Timmerman, R., Fritze, J., Miskinyte, G., et al. (2018). Rapid and efficient induction of functional astrocytes from human pluripotent stem cells. *Nature Methods, 15*(9), 693-696. doi:10.1038/s41592-018-0103-2

Cashikar, A. G., & Hanson, P. I. (2019). A cell-based assay for CD63-containing extracellular vesicles. *PloS One, 14*(7), e0220007. doi:10.1371/journal.pone.0220007

Clement, S., Stouffs, M., Bettiol, E., Kampf, S., Krause, K. -., Chaponnier, C., et al. (2007). Expression and function of -smooth muscle actin during embryonic-stem-cell-derived cardiomyocyte differentiation. *Journal of Cell Science, 120*(2), 229-238. doi:10.1242/jcs.03340

Do, J., McKinney, C., Sharma, P., & Sidransky, E. (2019). Glucocerebrosidase and its relevance to parkinson disease. *Molecular Neurodegeneration, 14*(1), 36. doi:10.1186/s13024-019-0336-2

Du, Z., & Zhang, S. (2004). Neural differentiation from embryonic stem cells: Which way? *Stem Cells and Development, 13*(4), 372-381. doi:10.1089/1547328041797426

Eng, L. F., & Ghirnikar, R. S. (1994). GFAP and astrogliosis. *Brain Pathology (Zurich, Switzerland), 4*(3), 229-237. doi:10.1111/j.1750-3639.1994.tb00838.x

Gao, L., Zhang, Z., Lu, J., & Pei, G. (2019). Mitochondria are dynamically transferring between human neural cells and alexander disease-associated GFAP mutations impair the astrocytic transfer. *Frontiers in Cellular Neuroscience, 13*, 316. doi:10.3389/fncel.2019.00316

Gawlik-Rzemieniewska, N., & Bednarek, I. (2016). The role of NANOG transcriptional factor in the development of malignant phenotype of cancer cells. *Cancer Biology & Therapy, 17*(1), 1-10. doi:10.1080/15384047.2015.1121348

Gupta, N., Oppenheim, I. M., Kauvar, E. F., Tayebi, N., & Sidransky, E. (2011). Type 2 gaucher disease: Phenotypic variation and genotypic heterogeneity. *Blood Cells, Molecules and Diseases, 46*(1), 75-84. doi:10.1016/j.bcmd.2010.08.012

Hu, Y., Dammer, E. B., Ren, R., & Wang, G. (2015). The endosomal-lysosomal system: From acidification and cargo sorting to neurodegeneration. *Translational Neurodegeneration, 4*(1), 18. doi:10.1186/s40035-015-0041-1

Jalava, N. S., Lopez-Picon, F. R., Kukko-Lukjanov, T., & Holopainen, I. E. (2007). Changes in microtubule-associated protein-2 (MAP2) expression during development and after status epilepticus in the immature rat hippocampus. *International Journal of Developmental Neuroscience, 25*(2), 121-131. doi:10.1016/j.ijdevneu.2006.12.001

Laurent, L. C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., et al. (2011). *Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ES and iPS cells during reprogramming and time in culture* doi:10.1016/j.stem.2010.12.003

Lee KO, Luu N, Kaneski CR, Schiffmann R, Brady RO, Murray GJ. (2005). Improved intracellular delivery of glucocerebrosidase mediated by the HIV-1 TAT protein transduction domain. *Biochemical and Biophysical Research Communications,* 337(2), 701–707. https://doi.org/10.1016/j.bbrc.2005.05.207

Leila Tarsa, & Yukiko Goda. (2002). Synaptophysin regulates activity-dependent synapse formation in cultured hippocampal neurons. *Proceedings of the National Academy of Sciences - PNAS, 99*(2), 1012-1016. doi:10.1073/pnas.022575999

Liu, C., Zhong, Y., Apostolou, A., & Fang, S. (2013). Neural differentiation of human embryonic stem cells as an in vitro tool for the study of the expression patterns of the neuronal cytoskeleton during neurogenesis. *Biochemical and Biophysical Research Communications, 439*(1), 154-159. doi:10.1016/j.bbrc.2013.07.130

Lu J., Chiang J., Iyer R. R., Thompson E., Kaneski C. R., Xu D. S., et al. (2010). Decreased glucocerebrosidase activity in gaucher disease parallels quantitative enzyme loss due to abnormal interaction with TCP1 and c-cbl. *Proceedings of the National Academy of Sciences - PNAS,* 107(50), 21665-21670. doi:10.1073/pnas.1014376107

Matuzelski, E., Bunt, J., Harkins, D., Lim, J. W. C., Gronostajski, R. M., Richards, L. J., et al. (2017). Transcriptional regulation of nfix by NFIB drives astrocytic maturation within the developing spinal cord. *Developmental Biology, 432*(2), 286-297. doi:10.1016/j.ydbio.2017.10.019

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J., Yakir, B., Clark, A. T., et al. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell, 7*(4), 521-531. doi:10.1016/j.stem.2010.07.017

Mizejewski, G. J. (2004). *Biological roles of alpha-fetoprotein during pregnancy and perinatal development*. London, England: SAGE Publications. doi:10.1177/153537020422900602

Molofsky, A. V., Krenick, R., Ullian, E., Tsai, H. -., Deneen, B., Richardson, W. D., et al. (2012). Astrocytes and disease: A neurodevelopmental perspective. *Genes & Development, 26*(9), 891-907. doi:10.1101/gad.188326.112

Musunuru, K. (2013). Genome editing of human pluripotent stem cells to generate human cellular disease models. *Disease Models & Mechanisms, 6*(4), 896-904. doi:10.1242/dmm.012054

Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., & Mineno, J. (2019). Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. *Scientific Reports, 9*(1), 4811. doi:10.1038/s41598-019-41121-4

Pars S, Cristo F, Inácio JM, Rosas G, Carreira IM, Melo JB, Mendes P, Martins DS, de Almeida LP, Maio J, Anjos R, Belo JA. (2018). Generation and characterization of a human iPS cell line from a patient-related control to study disease mechanisms associated with DAND5 p.R152H alteration. *Stem Cell Research*, (29), 202-206.

Platt, F. M., Boland, B., & van der Spoel, Aarnoud C. (2012). Lysosomal storage disorders: The cellular impact of lysosomal dysfunction. *The Journal of Cell Biology, 199*(5), 723-734. doi:10.1083/jcb.201208152

Pols, M. S., & Klumperman, J. (2009). Trafficking and function of the tetraspanin CD63. *Experimental Cell Research, 315*(9), 1584-1592. doi:10.1016/j.yexcr.2008.09.020

Rahman, F., Bordignon, B., Culerrier, R., Peiretti, F., Spicuglia, S., Djabali, M., et al. (2017). Ascorbic acid drives the differentiation of mesoderm-derived embryonic stem cells. involvement of p38 MAPK/CREB and SVCT2 transporter. *Molecular Nutrition & Food Research, 61*(5), 1600506-n/a. doi:10.1002/mnfr.201600506

Renaud, J., Boix, C., Charpentier, M., De Cian, A., Cochennec, J., Duvernois-Berthet, E., et al. (2016). Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Reports (Cambridge), 14*(9), 2263-2272. doi:10.1016/j.celrep.2016.02.018

Riboldi, G. M., & Di Fonzo, A. B. (2019). GBA, gaucher disease, and parkinson's disease: From genetic to clinic to new therapeutic approaches. *Cells, 8*(4) doi:10.3390/cells8040364

Sakatani, S., Seto-Ohshima, A., Shinohara, Y., Yamamoto, Y., Yamamoto, H., Itohara, S., et al. (2008). Neural-activity-dependent release of S100B from astrocytes enhances kainate-induced gamma oscillations in vivo. *Journal of Neuroscience, 28*(43), 10928-10936. doi:10.1523/JNEUROSCI.3693-08.2008

Santos, D. M., & Tiscornia, G. (2017). Induced pluripotent stem cell modeling of gaucher's disease: What have we learned? *International Journal of Molecular Sciences, 18*(4), 888. doi:10.3390/ijms18040888

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods, 9*(7), 676. doi:10.1038/NMETH.2019

Schmidt, R., & Plath, K. (2012). The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *GenomeBiology.Com, 13*(10), 251. doi:10.1186/gb-2012-13-10-251

Schöndorf, D. C., Aureli, M., McAllister, F. E., Hindley, C. J., Mayer, F., Schmid, B., et al. (2014). iPSC-derived neurons from GBA1-associated parkinson's disease patients show autophagic defects and impaired calcium homeostasis. *Nature Communications, 5*(1), 4028. doi:10.1038/ncomms5028

Sidransky, E., & Lopez, G. (2012). The link between the GBA gene and parkinsonism. *Lancet Neurology, 11*(11), 986-998. doi:10.1016/s1474-4422(12)70190-4

Stirnemann, J., Belmatoug, N., Camou, F., Serratrice, C., Froissart, R., Caillaud, C., et al. (2017). A review of gaucher disease pathophysiology, clinical presentation and treatments. *International Journal of Molecular Sciences, 18*(2), 441. doi:10.3390/ijms18020441

Sun, Y., Florer, J., Mayhew, C. N., Jia, Z., Zhao, Z., Xu, K., et al. (2015). Properties of neurons derived from induced pluripotent stem cells of gaucher disease type 2 patient fibroblasts: Potential role in neuropathology. *PloS One, 10*(3), e0118771. doi:10.1371/journal.pone.0118771

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell, 126*(4), 663-676. doi:10.1016/j.cell.2006.07.024

Tancini, B., Buratta, S., Sagini, K., Costanzi, E., Delo, F., Urbanelli, L., et al. (2019). Insight into the role of extracellular vesicles in lysosomal storage disorders. *Genes, 10*(7), 510. doi:10.3390/genes10070510

Trevino, A. E., & Zhang, F. (2014). Genome editing using Cas9 nickases. *Methods in Enzymology, 546*, 161-174. doi:10.1016/B978-0-12-801185-0.00008-8

Vitner, E. B., Dekel, H., Zigdon, H., Shachar, T., Farfel-Becker, T., Eilam, R., et al. (2010). Altered expression and distribution of cathepsins in neuronopathic forms of gaucher disease and in other sphingolipidoses. *Human Molecular Genetics, 19*(18), 3583-3590. doi:10.1093/hmg/ddq273

Weiss, K., Gonzalez, A. N., Lopez, G., Pedoeim, L., Groden, C., & Sidransky, E. (2015). The clinical management of type 2 gaucher disease. *Molecular Genetics and Metabolism, 114*(2), 110-122. doi:10.1016/j.ymgme.2014.11.008

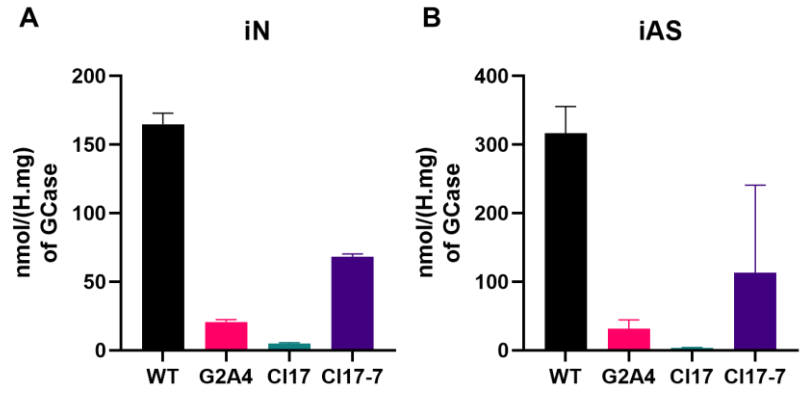
Wilczynska, K. M., Singh, S. K., Adams, B., Bryan, L., Rao, R. R., Valerie, K., et al. (2009). Nuclear factor I isoforms regulate gene expression during the differentiation of human neural progenitors to astrocytes. *Stem Cells (Dayton, Ohio), 27*(5), 1173-1181. doi:10.1002/stem.35

Zhang, X., Huang, C. T., Chen, J., Pankratz, M. T., Xi, J., Li, J., et al. (2010). Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell, 7*(1), 90-100. doi:10.1016/j.stem.2010.04.017

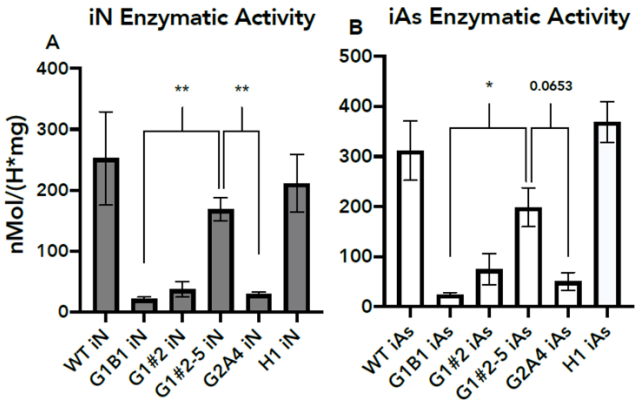
Zhang, Y., Sloan, S., Clarke, L., Caneda, C., Plaza, C., Blumenthal, P., et al. (2016). Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron (Cambridge, Mass.), 89*(1), 37-53. doi:10.1016/j.neuron.2015.11.013

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron (Cambridge, Mass.), 78*(5), 785-798. doi:10.1016/j.neuron.2013.05.029

Supplementary Data



Supplementary figure 1: GCase Enzymatic Activity results from discarded G2A4 edited clones Cl17 and Cl17-7. (A) EA in iNs and (B) iAs. No rescue in enzyme's activity was observed, as it would be expected in the isogenic control, leading to a further analysis of the sequencing data for these clones.



Supplementary figure 2: GCase Enzymatic Activity in G1B1 edited clones #2 and #2-5 iNs and iAs, compared to WT and H1 controls. Data are shown as the mean ± s.e.m. from three independent experiments. ** p value < 0.01, * p value < 0.05, two-tailed unpaired t-test comparing G1B1 & G2A4 iN and iAs to the isogenic control, 2-5.

Supplementary Table 1: Product Information

Product	Supplier	Cat. No
DMEM	Thermo Fisher Scientific	61965026
KO-DMEM	Thermo Fisher Scientific	10829-018
DMEM/F12	Thermo Fisher Scientific	31331-028
Essential 8	Thermo Fisher Scientific	A1517001
mTeSR-1	StemCell Technologies	05790
DPBS	Thermo Fisher Scientific	A1285601

Neurobasal	Thermo Fisher Scientific	21103049
BrainPhys	StemCell Technologies	05790
ROCK inhibitor Y-27632	StemCell Technologies	72308
PSC cryopreservation Kit	Thermo Fisher Scientific	A2644601
CryoStor CS10	Thermo Fisher Scientific	07930
FBS	Thermo Fisher Scientific	10082147
GlutaMAX	Thermo Fisher Scientific	35050061
NEAA	Thermo Fisher Scientific	11140050
N2 supplement	Thermo Fisher Scientific	17502048
B27 supplement	Thermo Fisher Scientific	12587010
Recombinant human FGF	PeproTech	100-18B
Recombinant human BMP4	PeproTech	120-05ET
hbEGF	Sigma Aldrich	E4643
5-Fluoro-2′-deoxyuridine	Sigma Aldrich	F0503
Laminin-521	BioLamina	LN521-02
hESC-qualified Matrigel	Thermo Fisher Scientific	08-774-552
Trypsin 0.25%	Thermo Fisher Scientific	25200056
TrypLE Select	Thermo Fisher Scientific	A1217702
StemPro Accutase Cell Dissociation Reagent	Thermo Fisher Scientific	A1110501
Stemgent® StemRNA™ 3rd Gen Reprogramming Kit for Reprogramming Adult and Neonatal Human Fibroblasts	ReproCell	00-0076
Lipofectamine RNAiMAX transfection reagent	Thermo Fisher Scientific	13778030
Lipofectamine Stem Transfection Reagent	Thermo Fisher Scientific	STEM00001
OptiMEM	Thermo Fisher Scientific	31985062
Puromycin	Thermo Fisher Scientific	A1113803
PenStrep	Thermo Fisher Scientific	15140122
Doxycycline	Sigma Aldrich	D9891
Hygromycin	Thermo Fisher Scientific	1068701
KryoMAX Colcemid	Thermo Fisher Scientific	15212012
KryoMAX KCl	Thermo Fisher Scientific	10575090
β-mercaptoethanol	Sigma Aldrich	M6250
L-Ascorbic Acid	Sigma Aldrich	A4544
T7 DNA ligase	New England BioLabs	M0318S/L

PlasmidSafe Exonuclease and 10x PlasmidSafe Buffer	Nordic BioLabs	E3101K
ATP	New England BioLabs	P0756S
TOP10 E. coli	Thermo Fisher Scientific	C404003
S.O.C. medium	Thermo Fisher Scientific	15544034
Phusion High-Fidelity Polymerase with HF-buffer	Thermo Fisher Scientific	F530L
Agarose	Sigma Aldrich	A9539
GelRed Nucleic Acid Stain	Biotium	41003
Na2HPO4	Sigma Aldrich	S3264
Citric Acid	Sigma Aldrich	C2404
Sodium Taurocholate	Sigma Aldrich	86339
Triton	Fisher Scientific	AC327371000
Glycine	Sigma Aldrich	G88898
NaOH	Honeywell	30620
4-methylumbelliferone	Sigma Aldrich	M1381
4-methylumbelliferylglucopyranoside	Sigma Aldrich	M3633
Bovine Serum Albumin	Sigma Aldrich	A2153
Qiagen Mini/Maxi Kit	Qiagen	27106/12963
Qiagen DNeasy Blood and Tissue Kit	Qiagen	69506
Qiagen RNeasy Mini Kit	Qiagen	74104
QuantaBio qScript cDNA Synthesis Kit	VWR International	101414-098