

# Genetic and molecular analysis of Sanfilippo C syndrome. Generation of a neuronal model using human induced pluripotent stem (iPS) cells and therapeutic strategies

Issac Canals Monferrer



Aquesta tesi doctoral està subjecta a la llicència Reconeixement- NoComercial – SenseObraDerivada 3.0. Espanya de Creative Commons.

Esta tesis doctoral está sujeta a la licencia Reconocimiento - NoComercial – SinObraDerivada 3.0. España de Creative Commons.

This doctoral thesis is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0. Spain License.



**Genetic and molecular analysis of Sanfilippo C syndrome. Generation of a neuronal model using human induced pluripotent stem (iPS) cells and therapeutic strategies.**

Anàlisi genètic i molecular de la Síndrome de Sanfilippo C. Generació d'un model neuronal utilitzant cèl·lules pluripotent induïdes (iPS) humanes i estratègies terapèutiques

Memòria presentada per **Isaac Canals Montferrer** per optar al títol de

**Doctor per la Universitat de Barcelona**

Programa de Genètica

Departament de Genètica

Tesi realitzada sota la direcció de la **Dra. Lluïsa Vilageliu Arqués** i el **Dr. Daniel Grinberg Vaisman** al Departament de Genètica de la Facultat de Biologia de la Universitat de Barcelona

Dra. Lluïsa Vilageliu Arqués

Dr. Daniel Grinberg Vaisman

Isaac Canals Montferrer

Barcelona, 2014



*“Ever tried. Ever failed. No matter.*

*Try again. Fail again. Fail better.”*



## PRESENTATION

Sanfilippo C syndrome is a rare inherited monogenic disease that presents a pronounced neurodegenerative course since early stages of patient's life. It is caused by mutations in the *HGSNAT* gene, localized few years ago in the chromosome 8, which codes for the enzyme acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase, a lysosomal membrane protein. Its function is to transfer an acetyl group to glucosamine residues in the heparan sulfate chain that is being degraded. The abnormal action of this protein in patients promotes the accumulation of partially degraded heparan sulfate chains inside the lysosomes, causing their dysfunction. Because of this reason, Sanfilippo C syndrome is a lysosomal storage disease, particularly a mucopolysaccharidosis due to the nature of the stored substrate.

Heparan sulfate is a glycosaminoglycan (previously known as mucopolysaccharid) found in the extracellular matrix as part of proteoglycans. These molecules participate in many different and important cellular functions such as migration or adhesion. The impairment of their homeostasis leads to the dysfunction of many cellular processes.

This thesis makes an important contribution to the field of Sanfilippo C molecular study. Mutational analysis and the consequent mutation characterization have been performed in order to extend the knowledge of the disease. Moreover, several different therapeutic approaches have been tested as a first step in the achievement of a successful therapy for this devastating neurodegenerative disorder, for which no effective treatment exists. Finally, a neuronal model for Sanfilippo C syndrome have been generated using induced pluripotent stem cells, which may help understand the molecular processes that contribute to the development of the neural disease and will be a valuable tool for the pursuit of successful treatments.



# INDEX



## INTRODUCTION

---

<b>Lysosomal Storage Disorders</b>	<b>3</b>
<i>The lysosome</i>	3
<i>Lysosomal storage disorders</i>	4
<i>Mucopolysaccharidoses</i>	8
<b>Sanfilippo Syndrome</b>	<b>10</b>
<i>General aspects</i>	10
<i>Clinical aspects</i>	12
<i>Diagnosis</i>	14
<i>Incidence</i>	15
<i>Subtype A</i>	15
<i>Subtype B</i>	17
<i>Subtype C</i>	18
<i>Subtype D</i>	20
<b>Heparan Sulfate</b>	<b>22</b>
<i>Heparan sulfate and heparan sulfate proteoglycans</i>	22
<i>Heparan sulfate biosynthesis</i>	23
<i>Heparan sulfate degradation</i>	25
<b>Therapeutic Approaches</b>	<b>28</b>
<i>Enzyme replacement therapy</i>	28
<i>Substrate reduction therapy</i>	29
<i>Coenzyme Q<sub>10</sub></i>	32
<i>Pharmacological chaperones for enzyme-enhancement-therapy</i>	33
<i>Suppression of premature termination codons</i>	34
<i>Overexpression of transcription factor EB</i>	35
<i>Gene therapy</i>	36
<i>Stem cell therapy</i>	40

<b>RNA Splicing</b>	<b>43</b>
<i>The splicing process</i>	43
<i>The nonsense mediated decay quality control</i>	46
<i>Therapies to correct splicing mutations</i>	48
<b>RNA interference</b>	<b>50</b>
<i>History and mechanisms</i>	50
<i>Therapy based on this mechanism</i>	52
<i>RNA interference therapies for Sanfilippo syndrome</i>	53
<b>Induced Pluripotent Stem Cells</b>	<b>55</b>
<i>Reprogramming from fibroblasts to induced pluripotent stem cells</i>	55
<i>Use of human induced pluripotent stem cells in lysosomal storage disorders</i>	57
<i>Direct conversion of fibroblasts to neurons or neural stem cells</i>	58
<i>Cellular models of Sanfilippo syndrome</i>	59

## **OBJECTIVES**

---

<b>Objectives</b>	<b>65</b>
-------------------	-----------

## **RESULTS**

---

<b>Report of the thesis supervisors about the contribution of the PhD student to the articles in this doctoral thesis</b>	<b>69</b>
<b>Article 1</b>	<b>75</b>
<b>Article 2</b>	<b>87</b>

<b>Article 3</b>	<b>113</b>
<b>Article 4</b>	<b>127</b>
<b>Annex</b>	<b>173</b>

## **DISCUSSION**

---

<b>Discussion</b>	<b>177</b>
-------------------	------------

## **CONCLUSIONS**

---

<b>Conclusions</b>	<b>199</b>
--------------------	------------

## **REFERENCES**

---

<b>References</b>	<b>203</b>
-------------------	------------

## **RESUM**

---

<b>Resum</b>	<b>231</b>
--------------	------------

## **AGRAÏMENTS**

---

<b>Agraïments</b>	<b>235</b>
-------------------	------------

# ABBREVIATIONS

<b>AAV</b> – Adeno-associated virus	<b>IdoA</b> – Iduronic acid
<b>AON</b> – Anti-sense oligonucleotide	<b>IGFII</b> – Insulin-like growth factor II
<b>BBB</b> – Blood-brain barrier	<b>iNSC</b> – Induced neural stem cells
<b>BMT</b> – Bone marrow transplantation	<b>iPSC</b> – Induced pluripotent stem cell
<b>cDNA</b> – Complementary deoxyribonucleic acid	<b>ISE</b> – Intronic splicing enhancer
<b>CHX</b> – Cycloheximide	<b>ISS</b> – Intronic splicing silencer
<b>CoQ<sub>10</sub></b> – Coenzyme Q <sub>10</sub>	<b>LSD</b> – Lysosomal storage disorder
<b>CNS</b> – Central nervous system	<b>M6P</b> – Mannose-6-phosphate
<b>CS</b> – Chondroitin sulfate	<b>miRNA</b> – Micro ribonucleic acid
<b>DNA</b> – Deoxyribonucleic acid	<b>MPS</b> – Mucopolysaccharidoses
<b>DMMB</b> – 1,9-dimethylmethylene blue	<b>mRNA</b> – Messenger ribonucleic acid
<b>DS</b> – Dermatan sulfate	<b>MSC</b> – Mesenchymal stem cells
<b>dsRNA</b> – Double strand ribonucleic acid	<b>NMD</b> – nonsense-mediated decay
<b>EGFR</b> – Epidermal growth factor receptor	<b>NSC</b> – Neural stem cells
<b>EJC</b> – Exon junction complex	<b>PTC</b> – premature termination codon
<b>ER</b> – Endoplasmic reticulum	<b>RISC</b> – Ribonucleic acid-induced silencing complex
<b>ERT</b> – Enzyme replacement therapy	<b>RNAi</b> – Ribonucleic acid
<b>ESC</b> – Embryonic stem cells	<b>RNAi</b> – Ribonucleic acid interference
<b>ESE</b> – Exonic splicing enhancer	<b>shRNA</b> – Small hairpin ribonucleic acid
<b>ESS</b> – Exonic splicing silencer	<b>siRNA</b> – Small interfering ribonucleic acid
<b>GA</b> – Golgi apparatus	<b>acid</b>
<b>GAG</b> – Glycosaminoglycan	<b>SNM</b> – Spherical neural mass
<b>GlcNAc</b> – N-acetylglucosamine	<b>SNP</b> – Single nucleotide polymorphism
<b>GlcA</b> – glucuronic acid	<b>snRNA</b> – Small nuclear ribonucleic acid
<b>hnRNP</b> – Heterogeneous nuclear ribonucleoprotein	<b>SRT</b> – Substrate reduction therapy
<b>HS</b> – Heparan sulfate	<b>ss</b> – Splice site
<b>HSCT</b> – Hematopoietic stem cell	<b>TFEB</b> – Transcription factor EB
<b>HSPG</b> – Heparan sulfate proteoglycan	<b>UTR</b> – Untranslated region
	<b>WT</b> – Wild-type

# FIGURES

## INTRODUCTION

---

<b>Figure 1.</b>	Lysosomes	3
<b>Figure 2.</b>	Pathways with lysosomal involvement	4
<b>Figure 3.</b>	Organelles affected in LSDs	5
<b>Figure 4.</b>	Typical MPS cell vacuolization	9
<b>Figure 5.</b>	Sanfilippo syndrome lysosomes	10
<b>Figure 6.</b>	LSDs inside the neuron	12
<b>Figure 7.</b>	Sanfilippo patients	13
<b>Figure 8.</b>	<i>SGSH</i> gene	16
<b>Figure 9.</b>	<i>NAGLU</i> gene	17
<b>Figure 10.</b>	<i>HGSNAT</i> gene	19
<b>Figure 11.</b>	<i>HGSNAT</i> mutations grouped by type	19
<b>Figure 12.</b>	<i>GNS</i> gene	21
<b>Figure 13.</b>	HSPGs functions	22
<b>Figure 14.</b>	HS biosynthesis	23
<b>Figure 15.</b>	Endocytosis of HS prior to its degradation	25
<b>Figure 16.</b>	HS degradation	26
<b>Figure 17.</b>	SRT basis	29
<b>Figure 18.</b>	Mechanism of action of genistein	30
<b>Figure 19.</b>	Chaperone treatment possibilities	33
<b>Figure 20.</b>	Read-through mechanism	35
<b>Figure 21.</b>	M6P-mediated enzyme excretion	39
<b>Figure 22.</b>	Conserved splice sites	44
<b>Figure 23.</b>	Splicing reaction	45
<b>Figure 24.</b>	NMD mechanism models	47
<b>Figure 25.</b>	U1 snRNA structure	49
<b>Figure 26.</b>	siRNA structure	51
<b>Figure 27.</b>	RNAi mechanism	51
<b>Figure 28.</b>	iPSC basis and applications	55

<b>Figure 29.</b>	Disease-modelling using iPSC	57
<b>Figure 30.</b>	Neuronal-disease-modelling	58
<b>Figure 31.</b>	Neuronal storage	59

## **DISCUSSION**

---

<b>Figure 32.</b>	Percentage of alleles for each mutation in Spanish patients	179
<b>Figure 33.</b>	Origin and migration of the c.234+1G>A allele	180
<b>Figure 34.</b>	Intronic deletion	185
<b>Figure 35.</b>	Modified U1 snRNAs used	187
<b>Figure 36.</b>	Alternative ss near the natural one	188
<b>Figure 37.</b>	SNMs in suspension	191
<b>Figure 38.</b>	Neuronal cultures	192

## **TABLES**

### **INTRODUCTION**

---

<b>Table 1.</b>	Lysosomal Storage Disorders	7
<b>Table 2.</b>	ERT strategies for Sanfilippo syndrome in animal models	28
<b>Table 3.</b>	Gene therapy strategies for Sanfilippo syndrome	37
<b>Table 4.</b>	RNAi delivery systems	53

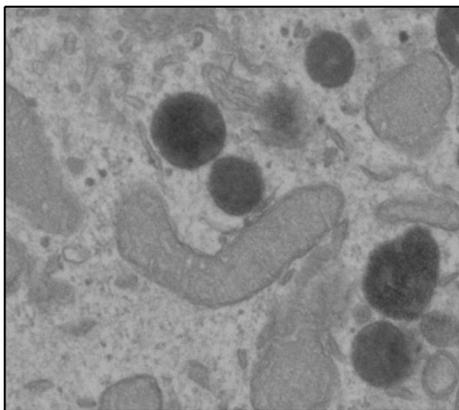
# **INTRODUCTION**



## Lysosomal Storage Disorders

### *The lysosome*

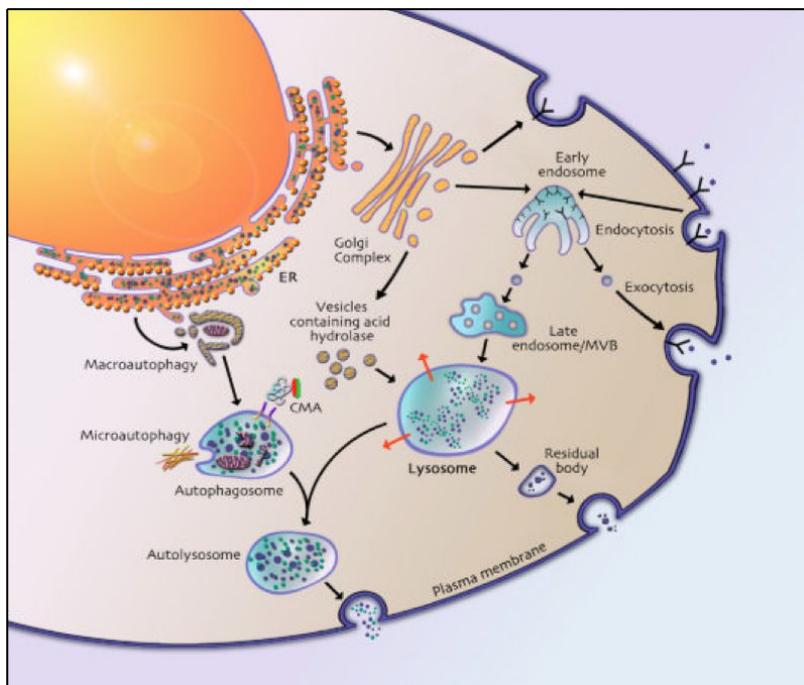
Lysosomes are subcellular organelles (figure 1) first described more than 60 years ago (De Duve et al., 1955). They are responsible for the turnover of cellular and extracellular constituents and organelles (De Duve and Wattiaux, 1966) and participate in some important cellular processes such as membrane repair, autophagy, cellular death via signal transduction, apoptosis, receptor recycling, neurotransmission regulation, skin pigmentation, bone biology and play an important role in cellular homeostasis and the endolysosomal system (Boustany, 2013; Saftig and Klumperman, 2009). Molecules degraded into lysosomes come either from outside the cell by phagocytosis, pinocytosis or receptor-mediated endocytosis, or from inside by microautophagy or macroautophagy (Ciechanover, 2005). In addition, the organelle is important for antigen presentation and phagocytosis, which are essential for the control of inflammation and autoimmunity. A scheme of some pathways in which lysosomes are involved can be seen in figure 2.



**Figure 1. Lysosomes.** Lysosomes display in a transmission electron microscopy image (dark vesicles).

Lysosomes present a single membrane and contain about 50-60 hydrolytic enzymes, integral membrane proteins and transporters. These hydrolases are active at the acidic pH (around 5) inside the organelle and most of them are soluble and localized in the lysosomal lumen. There are proton transporters located in the lysosomal membrane to maintain the acid pH at which the enzymes are functional. Lysosomal enzymes are synthesized in their inactive form at the endoplasmic reticulum (ER) and then glycosylated and transported to the Golgi apparatus (GA). Once in there, most of them acquire a

mannose-6-phosphate (M6P), which direct the protein to the lysosome via the M6P receptor. The low pH inside the lysosome allows the dissociation of the hydrolytic enzyme from the M6P receptor. Finally, the enzymes are matured and become functional. Some enzymes are not bound to the M6P receptor and are excreted outside the cell, where surrounding cells can take them. Alternative trafficking signals have been proposed for some lysosomal proteins (Lefrancois et al., 2003; Reczek et al., 2007). For instance, lysosomal membrane proteins are not directed via M6P. They present alternative signal peptides at the C-terminal end of the protein, which direct them to the lysosome.

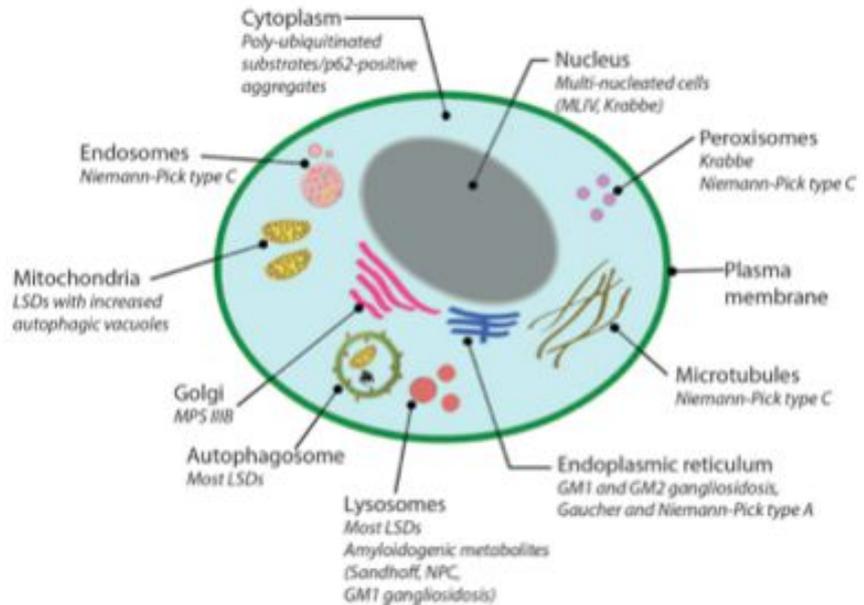


**Figure 2. Pathways with lysosomal involvement.** Different pathways in which an appropriate lysosomal function plays an important role such as autophagy, extracellular turnover or exocytosis (Schultz et al., 2011).

### *Lysosomal storage disorders*

Lysosomal storage disorders (LSDs) are a group of inherited metabolic diseases caused by mutations in genes encoding lysosomal proteins (from lumen and membrane), activator proteins, ER or GA modifying proteins that lead to an accumulation of partially degraded macromolecules inside the lysosomes. This accumulation promotes a cascade of events that affects not only the endosomal–autophagic–lysosomal system, but also other organelles (figure 3).

**Figure 3. Organelles affected in LSDs.** Examples of LSDs affecting different organelles (Platt et al., 2012).



More than 50 different LSDs have been reported until today (Table 1). All of them are monogenic disorders, and the majority present an autosomal recessive inheritance. Some present allele heterogeneity and some others present gene heterogeneity. It is thought that some of them remain to be described; since maybe one LSD exist for each lysosomal protein, but some of them could be lethal at early developmental stages (Futerman and van Meer, 2004). The first description of a LSD was that of Tay-Sachs disease in 1881 although the first link between an enzymatic deficiency and a LSD was not established until later, in 1963, for Pompe's disease (Hers, 1963).

Even though individually each LSD can be considered as a rare disease due to their low incidence in population ranging from 1:57000 (Gaucher disease) and 1:4200000 (sialidosis) live births, all together, LSDs present an incidence about 1:4000-8000 (Ortolano et al., 2014). LSDs can be classified either by the defective protein or by the storage material (reviewed in Futerman and van Meer, 2004), being that last criteria the most used. LSDs include sphingolipidoses, mucopolipidoses, glycoproteinoses, oligosaccharidoses or mucopolysaccharidoses (table 1). Clinical symptoms are variable among different LSDs from mild symptoms to severe disorders affecting a wide range of tissues. Despite the broad variability in the symptoms, some of them are common in most of the LSDs such as bone and central nervous system (CNS) affection, dysmorphism and visceromegaly.

Several therapies have been tested for different LSDs including selective treatments as enzyme replacement therapy (ERT), substrate reduction therapy (SRT), use of chemical chaperones and cell therapy using bone marrow transplantation (BMT), mesenchymal stem cells (MSC) or neural stem cells (NSC) (reviewed in Boustany, 2013; Hollak and Wijburg, 2014; Kim, 2014; Ortolano et al., 2014; Shayman, 2014). Despite all these attempts, until today no effective therapy has been developed for most of LSDs, especially for those affecting CNS, so only therapies aimed at improving patient's quality of life are available. To study and test new therapeutic approaches for these LSDs the development of neuronal models would be very useful.

Some animal models have been described or obtained for many LSDs (reviewed in Ellinwood et al., 2004). In the recent years, induced pluripotent stem cells allowed scientist to obtain cellular models for several LSDs (Higuchi et al., 2014; Huang et al., 2011; Lemonnier et al., 2011; Maetzel et al., 2014; Mazzulli et al., 2011; Panicker et al., 2012; Park et al., 2008; Tiscornia et al., 2013; Tolar et al., 2011; Trilck et al., 2013). This approach allows generating different cell types relevant for the disease. Moreover, these cells present exactly the same genome of the patients. Thus, cells that cannot be obtained directly from the patient can be generated by this technology. It also offers the advantage to test different therapeutic strategies in those cell types most affected by the disease.

**Table 1. Lysosomal Storage Disorders.** Defective enzyme and main storage material are indicated for each LSD (adapted from Futerman and van Meer, 2004).

Disease	Defective protein	Main storage materials
<b>Sphingolipidoses</b>		
Fabry	$\alpha$ -Galactosidase A	Globotriasylceramide and blood-group-B substances
Farber lipogranulomatosis	Ceramidase	Ceramide
Gaucher	$\beta$ -Glucosidase Saposin-C activator	Glucosylceramide Glucosylceramide
Globoid cell leukodystrophy (Krabbe)	Galactocerebrosidase $\beta$ -galactosidase	Galactosylceramide
Metachromatic leukodystrophy	Arylsulfatase A Saposin-B activator	Sulfated glycolipids Sulfated glycolipids and GM1 ganglioside
Niemann-Pick A and B	Sphingomyelinase	Sphingomyelin
Sphingolipid-activator deficiency	Sphingolipid activator	Glycolipids
GM1 gangliosidosis	$\beta$ -Galactosidase	GM1 ganglioside
GM2 gangliosidoses (Tay-Sachs)	$\beta$ -Hexosaminidase A	GM2 ganglioside and related glycolipids
GM2 gangliosidoses (Sandhoff)	$\beta$ -Hexosaminidase A and B	GM2 ganglioside and related glycolipids
GM2 gangliosidoses (GM2-activator deficiency)	GM2-activator protein	GM2 ganglioside and related glycolipids
<b>Mucopolysaccharidoses (MPS)</b>		
MPS I (Hurler, Scheie, Hurler/Scheie)	$\alpha$ -Iduronidase	Dermatan sulfate and heparan sulfate
MPS II (Hunter)	Iduronate-2-sulfatase	Dermatan sulfate and heparan sulfate
MPS IIIA (Sanfilippo)	Sulfamidase	Heparan sulfate
MPS IIIB (Sanfilippo)	N-Acetyl- $\alpha$ -glucosaminidase	Heparan sulfate
MPS IIIC (Sanfilippo)	Acetyl-CoA: $\alpha$ -glucosamide N-acetyltransferase	Heparan sulfate
MPS IIID (Sanfilippo)	N-Acetylglucosamine-6-sulfatase	Heparan sulfate
Morquio-A disease	N-Acetylgalactosamine-6-sulfate-sulfatase	Keratan sulfate and chondroitin-6-sulfate
Morquio-B disease	$\beta$ -Galactosidase	Keratan sulfate
MPS VI (Maroteaux-Lamy)	Arylsulfatase B	Dermatan sulfate
MPS VII (Sly)	$\beta$ -Glucuronidase	Heparan sulfate, dermatan sulfate, chondroitin-4- and -6-sulfates
MPS IX	Hyaluronidase	Hyaluronan
<b>Oligosaccharidoses and glycoproteinosis</b>		
Aspartylglucosaminuria	Aspartylglucosaminidase	Aspartylglucosamine
Fucosidosis	$\alpha$ -Fucosidase	Fucosides and glycolipids
$\alpha$ -Mannosidosis	$\alpha$ -Mannosidase	Mannose-containing oligosaccharides
$\beta$ -Mannosidosis	$\beta$ -Mannosidase	Man( $\beta$ 1 $\rightarrow$ 4)GlcNAc
Pompe (glycogen-storage-disease type II)	$\alpha$ -Glucosidase	Glycogen
Sialidosis	Sialidase	Sialyloligosaccharides and sialylglycopeptides
Schindler disease	$\alpha$ -N-Acetylgalactosaminidase	Glyco-conjugates containing $\alpha$ -N-acetylgalactosaminyl
<b>Lipidoses</b>		
Wolman disease and cholesterol-ester-storage disease	Acid lipase	Cholesterol esters and triglycerides
<b>Diseases caused by defects in integral membrane proteins</b>		
Cystinosis	Cystinosin	Cystine
Danon disease	LAMP2	Cytoplasmic debris and glycogen
Infantile sialic-acid-storage and Salla disease	Sialin	Sialic acid
Mucopolipidosis (ML) IV	Mucolipin-1	Lipids and acid mucopolysaccharides
Niemann-Pick C	NPC1 and 2	Cholesterol and sphingolipids
<b>Others</b>		
Galactosialidosis	Cathepsin A	Sialyloligosaccharides
I Cell and pseudo-Hurler polydystrophy (MLII and MLIII, respectively)	UDP-N-acetylglucosamine:lysosomal enzyme N-acetylglucosaminyl-1-phosphotransferase	Oligosaccharides, mucopolysaccharides and lipids
Multiple sulfatase deficiency	$\alpha$ -formylglycine-generating	Sulfatides
Neuronal ceroid lipofuscinosis (NCL)1 (Batten disease)	CLN1 (protein	Lipidated thioesters
NCL2 (Batten disease)	CLN2 (tripeptidyl amino peptidase-	Subunit c of the mitochondrial ATP synthase
NCL3 (Batten disease)	Arginine transporter	Subunit c of the mitochondrial ATP synthase
Pycnodysostosis	Cathepsin K	Bone proteins including collagen fibrils

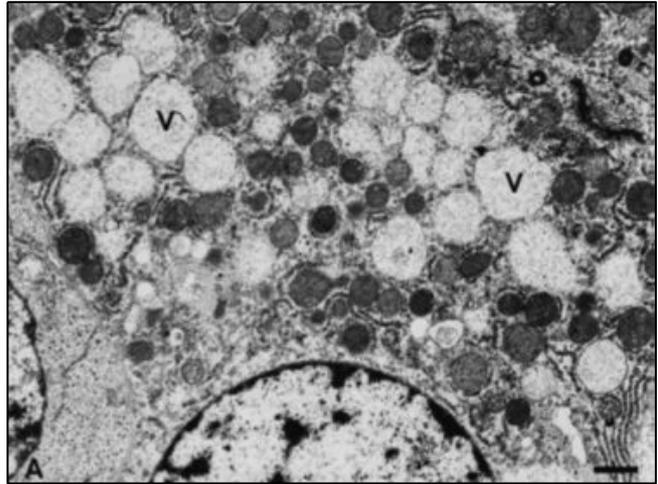
## *Mucopolysaccharidoses*

Mucopolysaccharidoses (MPS) are a LSD group characterized by the impairment in glycosaminoglycans (GAGs, previously known as mucopolysaccharides) degradation. Each MPS is due to mutations in genes coding for a lysosome enzyme involved in GAG catabolism. The enzyme deficiency causes the storage of partially degraded GAGs into lysosomes (and their increased excretion in urine) leading to progressive cellular damage that affects different tissues and organs (Neufeld and Muenzer, 2001).

MPS consist of a group of seven different diseases with an autosomal recessive inheritance pattern, except for MPS II, which presents X-linked recessive inheritance. Some reports differentiate among 11 different MPS since MPS III presents four different subtypes and MPS IV presents two subtypes. In each MPS, one or more GAGs are accumulated depending on the enzyme deficiency. Defective enzymes and the main storage products can be found in table 1. The incidence of MPS vary among populations from 1.9 to 4.5 for each 100.000 live births being MPS III, as a group, the most common (Applegarth et al., 2000; Baehner et al., 2005; Héron et al., 2011; Meikle et al., 1999; Nelson et al., 2003; Pinto et al., 2004; Poorthuis et al., 1999; Poupetová et al., 2010).

MPS present a large variety of symptoms, being organomegaly and coarse facial features the most frequent. Other organs and functions such as hearing, vision, cardiovascular system and joint mobility are usually affected. In severe MPS, the CNS function is also impaired. These clinical symptoms are chronic and course progressively during the patient's life. They are due to the accumulation of undegraded GAGs inside lysosomes that increase their size and number (figure 4) affecting other organelles and cellular functions.

**Figure 4. Typical MPS cell vacuolization.**  
Characteristic appearance of an MPS patient's cell (Neufeld and Muenzer, 2001).

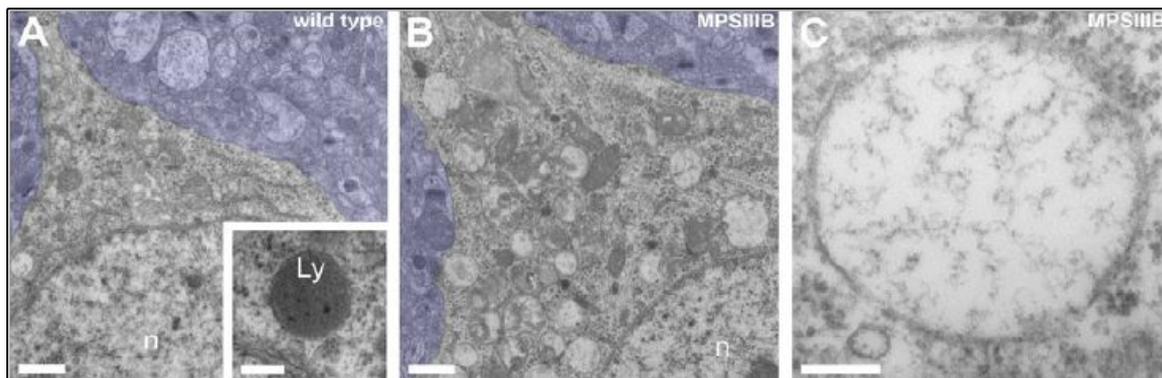


Early diagnosis is important for the treatment of MPS without neurological affection and available ERT (MPS I, II and VII). Usually diagnosis is delayed because of the wide clinical presentations. Typically they are diagnosed by the analysis of urinary GAGs that allows the differentiation between each MPS but not among the subtypes. For an accurate diagnosis, enzyme activity assays in fibroblasts, leukocytes, blood or plasma should be carried out (Neufeld and Muenzer, 2001). Many therapies have been tried for MPS with different results such as ERT, SRT, pharmacological chaperones, stop-codon read-through, gene therapy and cell therapy (reviewed in Giugliani et al., 2012).

## Sanfilippo Syndrome

### *General aspects*

Sanfilippo syndrome, also known as mucopolysaccharidoses III (MPS III), is a group of LSDs with an autosomal recessive inheritance pattern that was first described more than 50 years ago (Sanfilippo et al., 1963). Later MPS III was classified in four different subtypes (A: OMIM 252900, B: OMIM 252920, C: OMIM 252930, D: OMIM 252940), caused by mutations in four different genes, which code for the following enzymes: A, heparan N-sulfatase; B,  $\alpha$ -N-acetylglucosaminidase; C, acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase; and D, N-acetylglucosamine 6-sulfatase. These enzymes are involved in the degradation pathway of the glycosaminoglycan heparan sulfate (HS) (Neufeld and Muenzer, 2001). The deficient enzymes were identified during the 70's decade and in the last 20 years, so were the genes coding for these enzymes (Fan et al., 2006; Hřebíček et al., 2006; Robertson et al., 1988; Scott et al., 1995; Zhao et al., 1996). The lack of activity for one of these enzymes leads to the accumulation of partially degraded HS chains into the lysosomes. In patient's neurons and animal models, lysosomes increase in number and size due to this storage (Kurihara et al., 1996; Wilkinson et al., 2012) as showed in the cells of the mouse model of Sanfilippo subtype B (figure 5).



**Figure 5. Sanfilippo syndrome lysosomes.** A. WT mice lysosomes. B. MPS III B mice lysosomes. C. Typical lysosome from affected mice (Vitry et al., 2010).

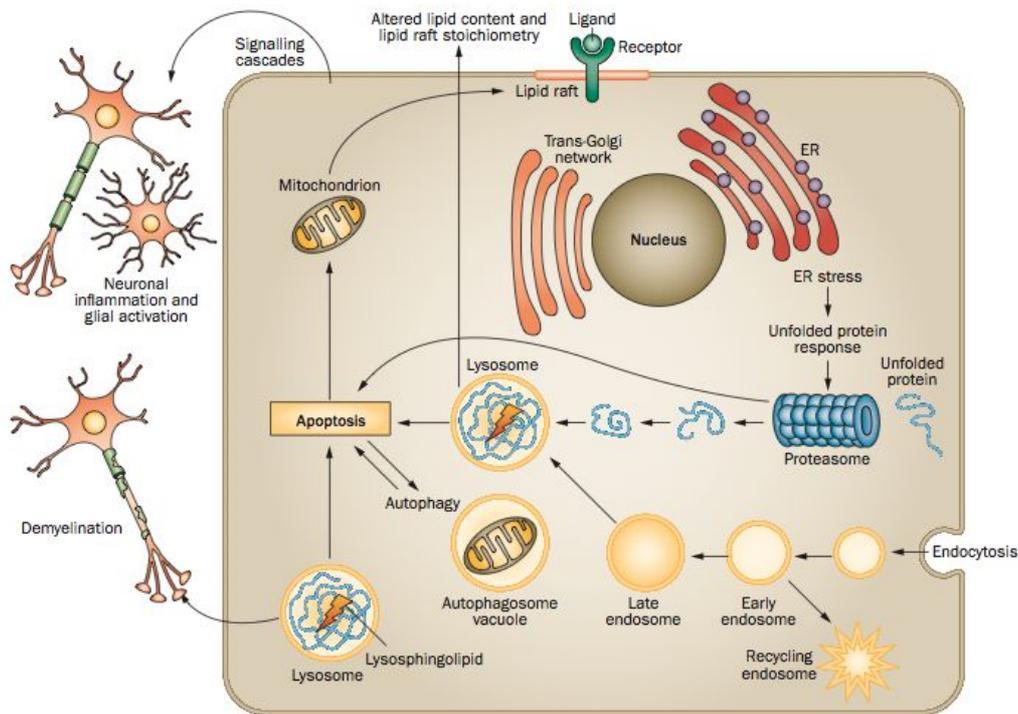
HS is one of the most common and important GAGs located in the extracellular matrix as a part of proteoglycans, which participate in many different cellular functions (Sarrazin et al., 2011). HS accumulation causes an alteration in the lysosomal environment

since it can bind to various hydrolases reducing their activity (Walkley, 2004). It causes the secondary accumulation of gangliosides and dermatan sulfate (DS) that may contribute to the central nervous system pathology (Lamanna et al., 2011).

GAGs are important components of brain and peripheral nerves. The fact that these tissues present a limited capacity of regeneration, a high sensitivity to damage and a need of long cellular survival could explain the severe neural affection in Sanfilippo patients. It has been demonstrated in MPS IIIB mice model that not only the CNS is severely affected but also the peripheral nervous system (Fu et al., 2012). It is possible that the HS fragments released to the extracellular matrix interfere with many HS functions, favouring the disease development.

The lysosomal storage of GAGs in neurons can lead to brain atrophy (Palmucci et al., 2013). Since GAGs act as ligands for different factors, the abnormal storage can affect different signalling pathways. The injury in neurons activates microglia and the constant release of inflammatory mediators. The accumulation in storage vesicles has been detected also in microglial cells in a mouse model of Sanfilippo B disease (Li et al., 1999). These cells play an important role in the brain defence and may release different toxic products. Thus, affection of the glial cells together with the inflammation may contribute to neuronal degeneration in MPS III (Ohmi et al., 2003).

It has been established that the storage of gangliosides (secondarily accumulated in Sanfilippo disease) leads to a decreased uptake of calcium by the ER with the consequent increase of the cytosolic levels that promotes neuronal apoptosis, favouring neurodegeneration. The decreased levels of calcium in the ER and other alterations activate the unfolded protein response, which also leads to an increased apoptosis (reviewed in Boustany 2013) that is an important contributor to the severe neurodegeneration. Astrocyte dysfunction also plays a role in neuronal degeneration as proved in a mouse model of multiple sulfatase deficiency, another LSD (Di Malta et al., 2012), and for other neurological disorders (Scuderi et al., 2013). It has been reported that in MPS IIIA mice there is an increase in the number of autophagosomes due to an impairment in the autophagosome-lysosome function, which probably leads cells to death. Thus, LSDs could be considered autophagy disorders (Settembre et al., 2008). Some of the processes involved in Sanfilippo-depending neuronal toxicity are summarized in figure 6.



**Figure 6. LSDs inside the neuron.** Lysosomal storage and secondary mechanisms leading to cell death in neurons (Boustany, 2013)

### *Clinical aspects*

The four Sanfilippo subtypes present similar clinical symptoms, with phenotypic variation and probably with some very mild forms that are difficult to recognize (Neufeld and Muenzer, 2001). The most significant trait is the severe central nervous system degeneration, unique among MPS disorders, that leads to motor deterioration, learning difficulties, hyperactivity, aggressive behaviour, sleeping problems and pronounced mental retardation (Valstar et al., 2008). It is accompanied by mild somatic manifestations such as hirsutism, mild hepatosplenomegaly, joint stiffness, dysphagia, hypertrichosis, severe hypoacusia, speech loss, mild skeletal involvement, in some cases coarse facial features although not very pronounced and sometimes diarrhoea (Neufeld and Muenzer, 2001). Epileptic seizures (Muenzer, 1986) and retinitis pigmentosa (Berger-Plantinga et al., 2004; Ruijter et al., 2008) can also occur in old patients and an early onset of puberty has been described in some MPS III patients of different subtypes.

These mild somatic manifestations usually lead to a delayed diagnostic after the onset of symptoms. First symptoms appear in early life, among two and six years in apparently normal children, but can occur earlier or later and usually consist in developmental delay and behavioural problems, sometimes accompanied by speech loss (Valstar et al., 2008). Severe neurodegeneration occurs between six and ten years in most patients, together with a rapid deterioration in social skills and a profound mental retardation. The progressive degeneration leads to a severe cortical atrophy, progressive dementia, motor deterioration, sleep disturbances and severe behaviour problems with possible physical aggression. Death usually occurs at the second or third decade in severe patients while in mild patients, life may extend until the fifth or the sixth.

It is difficult to differentiate patients from different subtypes because the symptoms are similar for all of them and patients present a significant heterogeneity in the symptomatology (Neufeld and Muenzer, 2001). Even so, it has been found that subtype A is the most severe form of Sanfilippo syndrome, while subtype B and specially subtype C present a slower progress of the symptoms (van de Kamp et al., 1981). Many attenuated cases have been described for subtypes A, B and C with late onset (Coppa et al., 2013; Meyer et al., 2008; Moog et al., 2007; Ruijter et al., 2008; Valstar et al., 2011). At present, Sanfilippo patients are usually classified into severe, intermediate and attenuated cases depending on the severity of the symptoms (Valstar et al., 2010). Appearance of typical MPS patients can be appreciated in figure 7.



**Figure 7. Sanfilippo patients.** MPS IIIA patient (left image) and MPS IIIB patient (right image) both at seven years old (Neufeld and Muenzer, 2001).

## *Diagnosis*

Diagnosis usually begins with the measure of the increased GAGs in urine by different methods. The most frequent consists of a 1,9-dimethylmethylene blue (DMMB) test. DMMB binds GAGs and the complex can be quantified by spectrophotometry. This method was described more than 20 years ago but was improved later (Barbosa et al., 2003). Other methods include cetyltrimethylammonium bromide screening, toluidine blue spot test or alcian blue method. All these techniques allow the detection of the increase in urinary GAGs. Next step specifically detects the type of GAGs that are increased, either by a thin-layer chromatography (Humbel and Chamoles, 1972) or by electrophoretic separation (Stone, 1998). These methods are useful for the differential diagnosis of MPS III from other types of MPS, but still do not provide information about the MPS III subtype. In order to confirm diagnosis and establish the MPS III subtype, loss of specific enzyme activity in fibroblasts, leukocytes or plasma should be assessed with fluorogenic (4-methylumbelliferyl), spectrophotometric or radiolabeled substrates, and recently, new tandem mass spectrometric assays (Wolfe et al., 2012). Fluorogenic substrates are preferable over radiolabeled substrates because their sensitiveness and stability are higher (Kleijer et al., 1996).

Once the MPS III subtype is established, next step consists in the identification of mutations causing the disease in each patient. It is useful to analyze related carriers and to perform prenatal diagnosis. Since there are prevalent mutations in some populations for subtypes A, B and C (Bunge et al., 1997; Di Natale et al., 1998; Mangas et al., 2008; Montfort et al., 1998; Ruijter et al., 2008; Tanaka et al., 2002; Weber et al., 1998), the presence of these mutations can be tested first for patients belonging to these populations.

Prenatal diagnosis can be performed in families with an affected child or an affected relative. Amniotic fluid or chorionic villi are usually used and the previously described assays, such as enzyme activity assay or mutation analysis, allow the diagnosis (Hopwood, 2005).

Recently, a new technique has been proposed for newborn screening of MPS types I, II and III (Ruijter et al., 2012a). This new approach based on the detection of

disaccharides derived from HS and DS could make possible the earlier diagnosis of MPS III patients, which could be important for future therapeutic approaches to improve prognosis and delay the appearance of degeneration symptoms. Even so, further studies are needed to establish the potential of this new test.

### *Incidence*

Some studies in different populations indicated that Sanfilippo syndrome as a group, is the most prevalent MPS, with incidences between 0.29 and 1.89 for each 100000 live births (Applegarth et al., 2000; Baehner et al., 2005; Meikle et al., 1999; Nelson et al., 2003; Pinto et al., 2004; Poorthuis et al., 1999; Poupetová et al., 2010)(Meikle et al., 1999; Nelson et al., 2003; Pinto et al., 2004; Poorthuis et al., 1999). These incidences may underestimate the actual prevalence of different MPS III types because of the high heterogeneity in symptomatology and the difficulties in the correct diagnosis of mild forms.

According to published data, subtype A is more prevalent in Northern and Central European countries such as Germany, France, the Netherlands, United Kingdom, Sweden or the Czech Republic, and also other countries like Australia, while subtype B is more prevalent in Southern European countries, such as Portugal and Greece and other countries like Taiwan. In contrast, a recent study in Spain showed Sanfilippo syndrome type A as the most frequent subtype in Spanish Sanfilippo patients (62% of the cases), while type B (20%) and type C (18%) were less frequent (Delgado et al., 2013). Generally, subtype C and mainly subtype D are less common in all populations (Baehner et al., 2005; Emre et al., 2002; Héron et al., 2011; Lin et al., 2009; Malm and Månsson, 2010; Michelakakis et al., 1995; Pinto et al., 2004; Poorthuis et al., 1999; Poupetová et al., 2010).

### *Subtype A*

MPS IIIA or Sanfilippo syndrome type A is caused by mutations in the *SGSH* gene, coding for sulfamidase (also known as heparan sulfate sulfatase or N-sulfoglucosamine

sulfohydrolase, EC 3.10.1.1). This enzyme releases sulfate groups linked to the amino group of glucosamine. The gene is localized at 17q25.3 (Scott et al., 1995) with an approximated length of 11 Kb and contains eight exons (figure 8). It codes for a protein of 502 aminoacids with five possible glycosylation sites. A total of 139 mutations have been identified until now, including 106 missense/nonsense mutations, 17 small deletions, nine small insertions, three gross deletions, two splicing mutations, one small indel and one gross insertion/duplication (HGMD® Professional 2014.3). Prevalent mutations have been described for different populations such as p.R74C in Polish patients, accounting for 56% of mutant alleles (Bunge et al., 1997), p.R245H representing the 35% of German mutated alleles (Bunge et al., 1997), p.S66W as the 33% of Italian mutant alleles (Di Natale et al., 1998) or c.1091delC accounting for 45% of mutation alleles in Spanish patients (Montfort et al., 1998).



**Figure 8. *SGSH* gene.** *SGSH* gene structure, with 11.14 Kb of length and containing eight exons (www.ensembl.org).

Patients are divided in three phenotypic groups depending on the course of the disease: severe, intermediate and attenuated. Severe patients present the typical course of the disease, where intellectual and motor abilities are lost during teenage years. In contrast, intermediate patients show a slower progression of the disease and their lifespan is extended until adulthood. Finally, attenuated patients present speech and motor abilities until adulthood and could live until fifties (Valstar et al., 2010). Genotype-phenotype correlations have been established for Sanfilippo type A despite the different polymorphisms that can modify the enzyme activity making difficult these predictions.

A strong correlation has been established between homozygous or compound heterozygous for p.R245H, p.Q380R, p.S66W or c.1080delC mutations and a severe phenotype. In contrast, heterozygous patients for p.S298P and one of the mutations linked to the severe phenotype present intermediated course of the disease. Finally, patients homozygous for p.S298P mutation and presumably carriers of p.L12Q, p.P180L, and

p.T421R mutations showed an attenuated phenotype (Valstar et al., 2010). Other studies have also demonstrated that p.G122R, p.R206P, p.I322S and p.E369K mutations produce an attenuated phenotype (Gabrielli et al 2005; Yogalingam and Hopwood 2001).

Two different natural animal models were described for Sanfilippo A syndrome, a Dachshund dog presenting a 3 base-pair deletion resulting in the loss of a threonine in position 246 (Fischer et al., 1998) and a mouse with a point mutation resulting in the change p.D31N (Bhattacharyya et al., 2001; Bhaumik et al., 1999). Both animal models mimic the human phenotype, with a progressive neurodegeneration, loss of motor abilities and mild somatic symptoms such as hepatosplenomegaly, and the urinary excretion of GAGs (Bhattacharyya et al., 2001; Fischer et al., 1998; Hemsley and Hopwood, 2005). Impairment in the autophagy function has also been detected in the mouse model (Settembre et al., 2008).

### Subtype B

MPS IIIB or Sanfilippo syndrome type B is caused by mutations in the *NAGLU* gene, which encodes N-acetyl- $\alpha$ -glucosaminidase (EC 3.2.1.50), a lysosomal enzyme of 720 aminoacids with six possible glycosylation sites. The function of the enzyme is the hydrolysis of the linkage between N-acetylglucosamine (GlcNAc) and the uronic acid, the two saccharides that conform HS. The gene maps to 17q21.1 (Zhao et al., 1996), spans 8.3 Kb and contains six exons (Figure 9). Until now, 153 mutations have been identified including 104 missense/nonsense mutations, 23 small deletions, 13 small insertions, five splicing mutations, four gross deletions, three gross insertions/duplications and one small indel (HGMD® Professional 2014.3). Only one mutation, p.R565P, has been found as a common mutation in one population (Tanaka et al., 2002). Sanfilippo B presents a high heterogeneity of mutations in the rest of the affected populations.



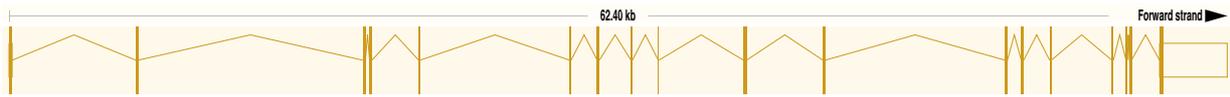
**Figure 9. *NAGLU* gene.** *NAGLU* gene structure, with 8.28 Kb of length and containing six exons (www.ensembl.org).

Genotype-phenotype correlations are difficult to establish because of the wide spectrum of phenotypes, the high number of described mutations with very low frequencies and the different polymorphisms that can modify the enzyme activity. Nevertheless p.F48L, p.G69S, p.S612G, and p.R643C mutations have been associated with a less severe phenotype in Sanfilippo B patients (Yogalingam and Hoopwood, 2001). A recent work correlates many missense mutations with an attenuated phenotype of Sanfilippo B syndrome, while nonsense mutations are correlated with a severe phenotype of the disease (Moog et al., 2007).

Two natural animal models were described for Sanfilippo B syndrome, a Schipperke dogs (Ellinwood et al., 2003) and an avian model, Emu (Aronovich et al., 2001). Furthermore, one mouse model has been generated (Li et al., 1999). Both natural animal models presented similarities with the human disease, such as motor deterioration, low enzyme activity and storage of GAGs in different tissues. The mutation in the avian model was found to be a 2 base-pair deletion, c.1098-1099delGG, in exon 6 (Aronovich et al., 2001). The mouse model was generated by the disruption of exon 6 of the mouse orthologous *Naglu* gene. These mice presented some symptomatology that mimics the human disorder, such as HS accumulation in different tissues, vacuolization in different cell types, secondary storage of gangliosides, hearing loss and a shorter lifespan (Li et al., 1999).

### *Subtype C*

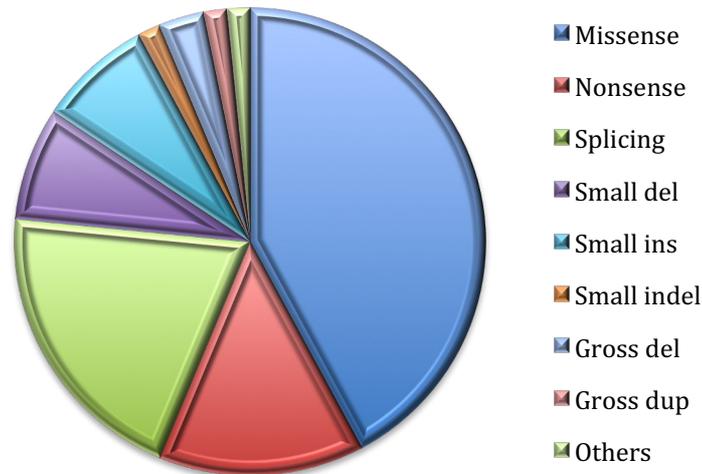
MPS IIIC or Sanfilippo syndrome type C is caused by mutations in the *HGSNAT* gene, which encodes the acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase (EC 2. 3.1.78), a lysosomal membrane protein. The *HGSNAT* gene was identified by two independent groups in 2006 (Fan et al., 2006; Hřebíček et al., 2006). The gene is located at chromosome 8p11.1, spans about 62.5 Kb, contains 18 exons and presents two possible initiation codons giving rise to proteins of 663 and 635 amino acids respectively (Figure 10). There was controversy about the real initiation codon. It was proposed that both were used at the same time (Durand et al., 2010), but a later publication suggested that only the second ATG codon worked as an initiation codon (Fan et al., 2011).



**Figure 10. *HGSNAT* gene.** *HGSNAT* gene structure, with 62.4 Kb of length and containing 18 exons (www.ensembl.org).

Until now, 64 mutations have been identified (figure 11), including 27 missense mutations, 13 splicing mutations, nine nonsense mutations, five small deletions, five small insertions, one small indel, two gross deletions, one gross duplication and one complex rearrangement (HGMD®Profesional 2014.3). Most of the missense mutations affect amino acid residues adjacent to or within a transmembrane domain, presumably interfering with the correct folding of the protein.

### ***HGSNAT* mutations**



**Figure 11. *HGSNAT* mutations grouped by type.** Missense, nonsense and splicing mutations represent more than 75% of total mutated alleles, while other type of mutations are found in a small proportion.

Sanfilippo C syndrome presents a very low frequency and many different mutations have been identified. Even so, different prevalent mutations have been reported such as p.R344C (accounting for 22%) and p.S518F (accounting for 29.3%) in the Dutch

population (Ruijter et al., 2008), c.852-1G>A that represent 30% of Italian alleles (Fedele et al., 2007) and c.525dupT that accounts for 62.5% of Portuguese mutated alleles (Coutinho et al., 2008).

Genotype-phenotype correlations have not been established, although p.G262R and p.S539C mutations are probably associated with an attenuated phenotype with late onset of symptomatology (Ruijter et al., 2008).

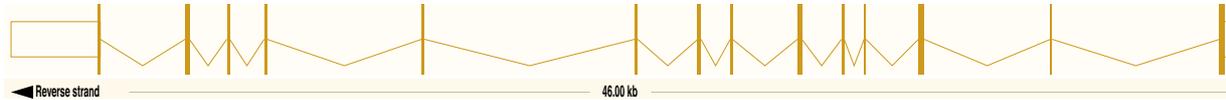
The enzyme encoded by the *HGSNAT* gene catalyses the acetylation of the terminal glucosamine residues of HS prior to its hydrolysis by  $\alpha$ -*N*-acetyl glucosaminidase (Klein et al., 1978). The protein contains 11 transmembrane domains and has a molecular weight of about 62 KDa. After the cleavage of the signal peptide in the ER, this enzyme is processed in two subunits into the lysosome by a cleavage in the first lysosomal loop, an  $\alpha$  chain of about 12-27 KDa containing two glycosylation sites and no transmembrane domains and a  $\beta$  chain of about 38-44 KDa containing three glycosylation sites and the 11 transmembrane domains (Fan et al., 2011). A ping-pong mechanism involving a big heterocomplex was proposed to explain the functional action of the enzyme (Durand et al., 2010) but it was demonstrated that the enzyme can work as a monomer even in the precursor form (Fan et al., 2011).

No animal model has been described for Sanfilippo C syndrome so far, which means that it is the only Sanfilippo syndrome subtype for which neither a natural nor an artificially generated animal model is available.

### *Subtype D*

MPS IIID or Sanfilippo syndrome type D is caused by mutations in the *GNS* gene, which codes for N-acetylglucosamine-6-sulfatase (EC 3.1.6.14), a lysosomal enzyme of 552 amino acids and 13 potential glycosylation sites (Robertson et al., 1988). This enzyme catalyses the removal of the sulfate in the N-acetylglucosamine residues. The gene is located at 12q14.3, is 46-Kb long and contains 14 exons (Figure 12). Until now, 23 mutations have been described, including seven missense/nonsense mutations, four small

deletions, four small insertions, three splicing mutations, two gross deletions, two complex rearrangements and one small indel (HGMD®Profesional 2014.3).



**Figure 12. *GNS* gene.** *GNS* gene structure, with 46 Kb of length and containing 14 exons ([www.ensembl.org](http://www.ensembl.org)).

Due to the extremely low frequency of this subtype and the low number of patients described, no genotype-phenotype correlations have been established until today.

Furthermore, and considering again the low number of patients for Sanfilippo type D, no prevalent mutations have been found.

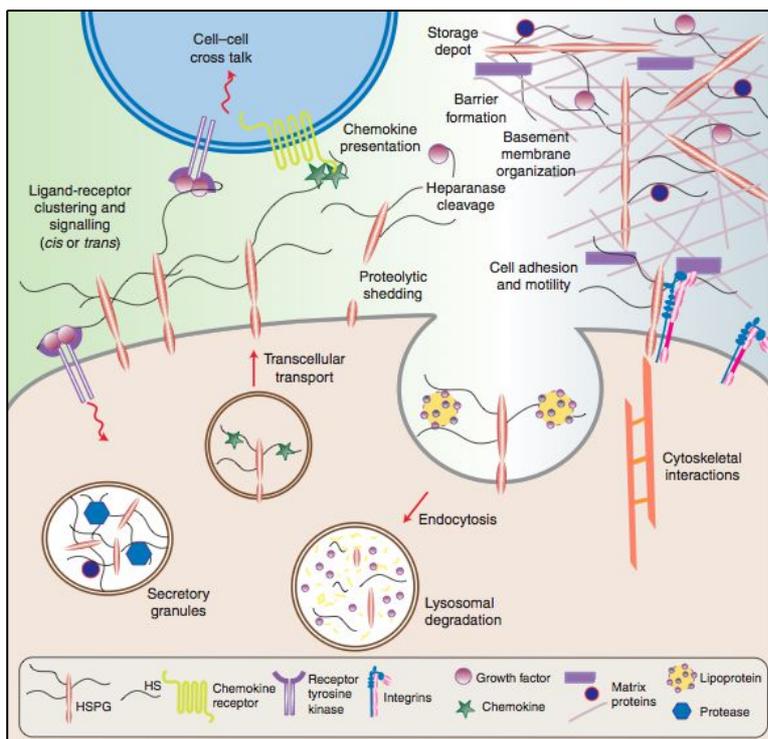
One natural animal model was described for Sanfilippo type D, a Nubian goat (Thompson et al., 1992), carrying a nonsense mutation. This model presented neurological and histological manifestations similar to those of human patients.

## Heparan Sulfate

### *Heparan sulfate and heparan sulfate proteoglycans*

HS is one of the most common and important GAGs located at the cell surface and in the extracellular matrix as a part of proteoglycans (Sarrazin et al., 2011). HS is composed of many disaccharide units, comprised of one type of uronic acid and one derivate of GlcNAc that can be sulfated (Esko and Selleck, 2002). The composition among different HS chains can differ, giving rise to different types of HS. The HS chains formed by these disaccharides repetitions are attached to a core protein being part of HS proteoglycans (HSPGs).

A great variety of HSPGs can be found in the cell surface and extracellular matrix based on the core protein and the type and number of HS chains linked to it (reviewed in Esko and Selleck, 2002). HSPGs participate in many different cellular functions and systems (Figure 13) such as cell migration, vesicle secretion system, endocytic system, cell adhesion and motility, membrane basement structure and recognition of different factors and molecules as receptors or co-receptors (reviewed in Sarrazin et al., 2011).



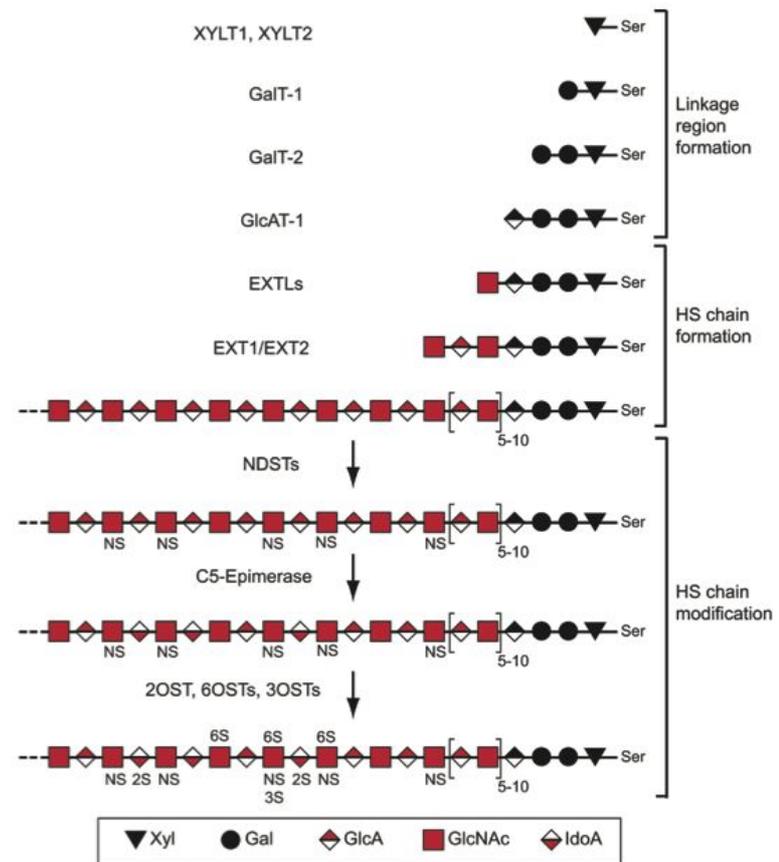
**Figure 13. HSPGs functions.** Different activities of HSPGs located in the extracellular matrix and the cell surface (Sarrazin et al., 2011)

## Heparan sulfate biosynthesis

The first step in the HSPGs synthesis is the formation of the different core proteins in the ER. The amount of these proteins, which could compete for the HS synthesis, could be the limiting factor in the synthetic pathway. HS synthesis *per se*, is widely accepted to take place in the GA, despite the first enzyme (xylosyltransferase) involved in this pathway is found in ER. All enzymes involved in HS synthesis are transmembrane proteins except for one sulfotransferase.

**Figure 14. HS biosynthesis.**

HS structure and three-steps biosynthesis scheme, including linkage region formation, HS chain elongation and HS chain modification (Kreuger and Kjellén, 2012).



HS synthetic pathway (figure 14) starts with the formation of the linkage region, a tetrasaccharide that binds the HS chain to the core protein. A xylose is transferred to a specific serine residue next to a glycine residue and flanked by acidic and hydrophobic residues in the core protein (Esko and Zhang, 1996; Zhang et al., 1995) by the action of one of the two xylosyltransferases (coded by *XYLT1* and *XYLT2*). Afterwards, two galactose residues are attached to the xylose by the sequential action of galactosyltransferases I and II (coded by *GALT1* and *GALT2*, respectively). Finally, the attachment of one glucuronic acid by a glucuronosyltransferase (GlcAT-1) completes the formation of the linkage

tetrasaccharide. This linkage region is common to HS, chondroitin sulfate (CS), DS and heparin. It has been proven that the xylose may be phosphorylated while the galactoses may be sulfated. These modifications could have an effect in the GAG to be synthesized. For instance the sulfation of the second galactose residue was reported to be associated with the CS synthesis (Ueno et al., 2001).

After the linkage region synthesis, the next step consists on the chain elongation. In the case of HS, elongation starts with the addition of one GlcNAc to the linkage region, step under control of the *EXTL* genes (*EXTL1*, *EXTL2* and *EXTL3*). While *EXTL1* is the only *EXTL* gene that is not ubiquitously expressed and whose function remains unclear, the *EXTL2* and *EXTL3* proteins have been proved to present GlcNAc-TI activity.

In the case of *EXTL2*, the enzyme has been reported to have also a low level of GlcNAc-TII activity, but only *in vitro* (Kitagawa et al., 1999). Inhibition of the *EXTL2* gene in fibroblasts from patients affected of MPSI and III using shRNAs resulted in a decrease in the HS synthesis and storage (Kaidonis et al., 2010).

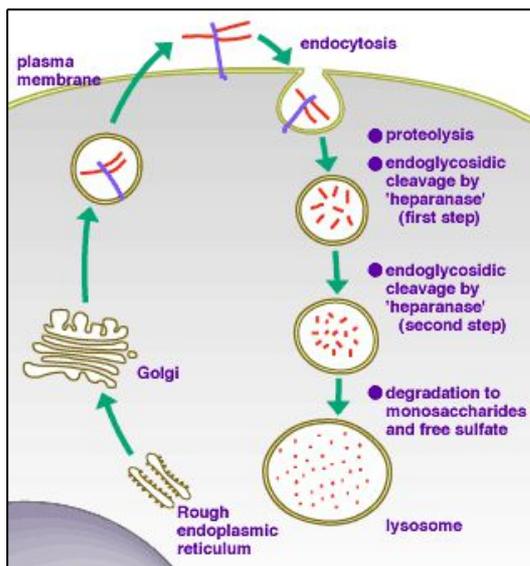
On the other hand, *EXTL3*, which also presents GlcNAc-TI activity, shows more ability to transfer the GlcNAc residue to the linkage region (Kim et al., 2001). *EXTL3* also presents *in vitro* GlcNAc-TII activity but it remains unclear if this activity also exists *in vivo*. When *EXTL3* is inhibited using specific siRNAs in 293 cells, longer HS chains are found, strongly suggesting that *EXTL3* participates in the initiation process. When *EXTL3* is inhibited, less linkage regions would have the GlcNAc residue necessary for the HS elongation, so the few HS chains available would grow longer (Busse et al., 2007). In contrast, its overexpression did not show any effect on HS chains, remaining unclear its GlcNAc-TII activity *in vivo*.

After this initiation step with participation of the *EXTL* genes, the elongation of the HS chain takes place by the action of the *EXT1-EXT2* complex, which adds alternative glucuronic acid (GlcA) and GlcNAc residues to the chain, forming polymers of different length. Inhibition of both genes using siRNAs resulted in shorter HS chains, while the overexpression of *EXT1* or both *EXT* genes resulted in longer HS chains and the overexpression of *EXT2* did not affect the HS chain length. It has been demonstrated that *EXT1* presents both activities *in vitro*, GlcA-TII and GlcNAc-TII, while the presence of *EXT2*

is essential *in vivo*, suggesting that EXT2 assists EXT1 in its translocation from ER to GA and plays a fundamental role for HS elongation even when it does not present any transferase activity (reviewed in Busse-Wicher et al., 2013). Mutations in *EXT1* and *EXT2* cause hereditary multiple exostoses (HME), a common (1:100000) autosomal dominant disorder affecting skeleton with a risk of malignant transformation. Mutations in the *EXT* genes result in reduced or absence of HS in the cartilage, producing a disorganization of chondrocytes (reviewed in Zak et al., 2002).

Finally, the last step in HS biosynthesis is the modification of the HS chain. These modifications take place while the chain is being synthesized and involve six different steps: deacetylation of the GlcNAc residues in some chain regions, sulfation of these residues, epimerization of many GlcA residues next to a modified GlcNAc residue to form iduronic acid (IdoA), sulfation of some IdoA and GlcA residues, 6-O-sulfatation and 3-O-sulfatation of some glucosamine residues in specific group contexts. Four different types of sulfotransferases and one epimerase are responsible for these modifications (reviewed in Esko and Selleck, 2002). All these modifications are important for the HS interactions and recognition of different factors and molecules.

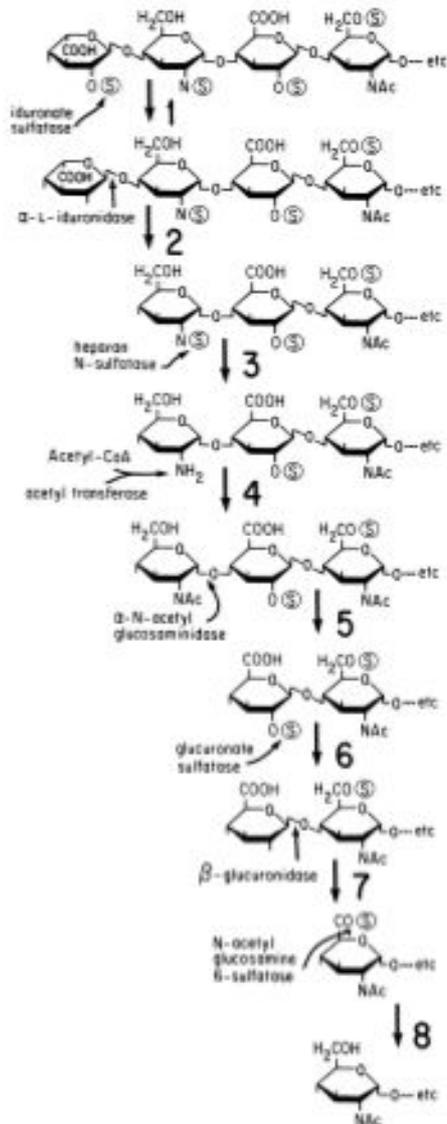
### *Heparan sulfate degradation*



**Figure 15. Endocytosis of HS prior to its degradation.** Internalizing of HSPGs from extracellular matrix to lysosomes to be degraded ([www.glycoforum.gr.jp](http://www.glycoforum.gr.jp)).

In the extracellular matrix, some endosulfatases and secreted heparanase could partially degrade HS chains giving rise to smaller and potentially active fragments (Gong et al., 2003). These fragments can be recycled with the same core protein, being internalized and brought under partial degradation and new rounds of biosynthesis and finally, exocytosed again to the cell membrane and the extracellular matrix (Fransson et al., 2004).

Final HS degradation takes place inside the lysosomes after internalization of HSPGs (figure 15) through the stepwise action of nine different enzymes (figure 16) (Neufeld and Muenzer, 2001). HS consists of GlcA and IdoA residues that can be sulfated or not, and glucosamine residues that can be acetylated or sulfated.



**Figure 16. HS degradation.** Stepwise degradation of HS with a scheme including all steps and all enzymes participating in the pathway (Neufeld and Muenzer, 2001).

The first enzyme of the pathway, heparanase, is an endoglucuronidase that cleaves HS chains into smaller fragments to facilitate the polymer degradation. Iduronate sulfatase is the enzyme that desulfates iduronic acid residues previous to their cleavage by the  $\alpha$ -L-iduronidase. Deficiency in the iduronate sulfatase causes MPS II while deficiency in  $\alpha$ -L-iduronidase causes MPS I, with consequent accumulation of HS and also DS, indicating their role also in the DS degradation pathway.

The next step in the HS degradation consists in the desulfamation of the amino group of the glucosamine by the heparan N-sulfatase (sulfamidase). Deficiency in this enzyme causes MPS IIIA, with HS accumulation. This glucosamine has to be acetylated in its recently exposed amino group previous to its cleavage. This acetylation function corresponds to the acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase, the only lysosomal enzyme that is not a hydrolase, which takes the acetyl group from acetyl-CoA from cytoplasm and transfers it to the HS chain under degradation in the lysosomal lumen. Deficiency of this enzyme causes MPS IIIC due to the HS storage.

Once the glucosamine has been acetylated to obtain GlcNAc, it can be cleaved from the chain by the action of  $\alpha$ -N-acetylglucosaminidase. This is the enzyme deficient in MPS IIIB patients, who present HS accumulation. After that, GlcA residue has to be sulfated previous to its cleavage. The enzyme responsible for this sulfation is glucuronate 2-sulfatase, while the enzyme responsible for the cleavage of the sulfated residue is  $\beta$ -glucuronidase. Deficiency of the first enzyme is not yet associated with any known disease, while deficiency of the second one causes MPS VII, with consequent storage of HS and DS.

Finally, the last enzyme taking part in the HS degradation pathway is N-acetylglucosamine 6-sulfatase, which desulfates the non-deacetylated residues of glucosamine previous to their cleavage. This enzyme is deficient in MPS IIID patients, and causes HS storage. It has been suggested that all these enzymes function as a complex in the lysosomes (Freeman and Hopwood, 1992)

## Therapeutic Approaches

Non-effective therapy has already been developed for Sanfilippo syndrome patients, though different kinds of interesting approaches have been tested in cells and animal models of the disease, focused mainly in the treatment of the CNS involvement.

### *Enzyme replacement therapy*

Different approaches are available concerning this type of therapy, differing in the administration strategy (table 2). For a disease affecting the CNS such as MPS III, is important to consider the existence of the blood-brain barrier (BBB) that does not allow the pass of enzymes to the brain. Thus, intravenously administration is not as useful as for other LSDs without CNS affection, in which ERT is currently approved and in use (revised in Desnick and Schuchman, 2012). For Sanfilippo syndrome, intravenously administration was tested in subtype B murine model (Yu et al., 2000) and subtype D caprine model (Downs-Kelly et al., 2000) without positive results in the brain.

In contrast, intravenous administration to a mouse model of Sanfilippo type A from birth, when the BBB is not formed, showed a delay in the development of the disease, confirming the usefulness of the ERT when the enzyme can reach the brain (Gliddon and Hopwood, 2004). More recently, intravenously chemically modified enzymes at high doses have been showed as a good option for the treatment of MPS IIIA mice (Rozaklis et al., 2011).

**Table 2. ERT strategies for Sanfilippo syndrome in animal models.**

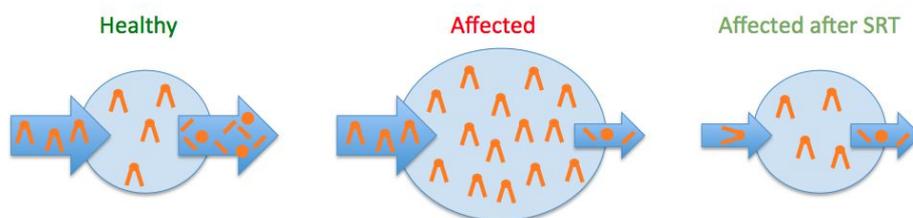
Enzyme	Administration	Subtype	Model
N-acetylglucosamine	Intravenous	B	Murine
N-acetylglucosamine 6-sulfatase	Intravenous	D	Caprine
Sulfamidase	Intravenous	A	Mouse
Chemically modified sulfamidase	Intravenous	A	Mouse
Sulfamidase	Intracerebral	A	Mouse
Sulfamidase	Intravenous + intracerebral	A	Mouse
Sulfamidase	Intracerebral	A	Dog
Sulfamidase + IGFII	Intraventricular	A	Mouse

Intrathecal, intracerebral, intracisternal and intracerebrospinal fluid injections of recombinant enzymes have been proved to be a considerable option for the treatment of MPS IIIA dogs (Crawley et al., 2011; Hemsley et al., 2009b; Jolly et al., 2011; Marshall et al., 2014) and MPS IIIA mice (Hemsley et al., 2008; Savas et al., 2004). Combined intravenous administration before adulthood and intracerebrospinal fluid administration during adulthood has been tested in MPS IIIA mice but this combination did not show better results than the use of only intracerebrospinal fluid administration (Hemsley et al., 2009a). It has been proved that this type of therapy is more efficient in early stages of the disease when compare to middle or late stages in mice (Hassiotis et al., 2014). Also the use of fusion protein containing NAGLU and insulin-like growth factor II (IGFII) has been tested in MPS IIIB mice to overcome the BBB. The fusion enzyme persists in the brain for around ten days after intraventricular administration, promoting a decrease in the brain levels of HS as well as in the liver levels to that of the control cases (Kan et al., 2014).

Direct brain administration for the treatment of neurological disorders presents some problems. It is an aggressive treatment that needs continued injections and the administration of an enzyme not naturally present in patient's brain may trigger an immune response.

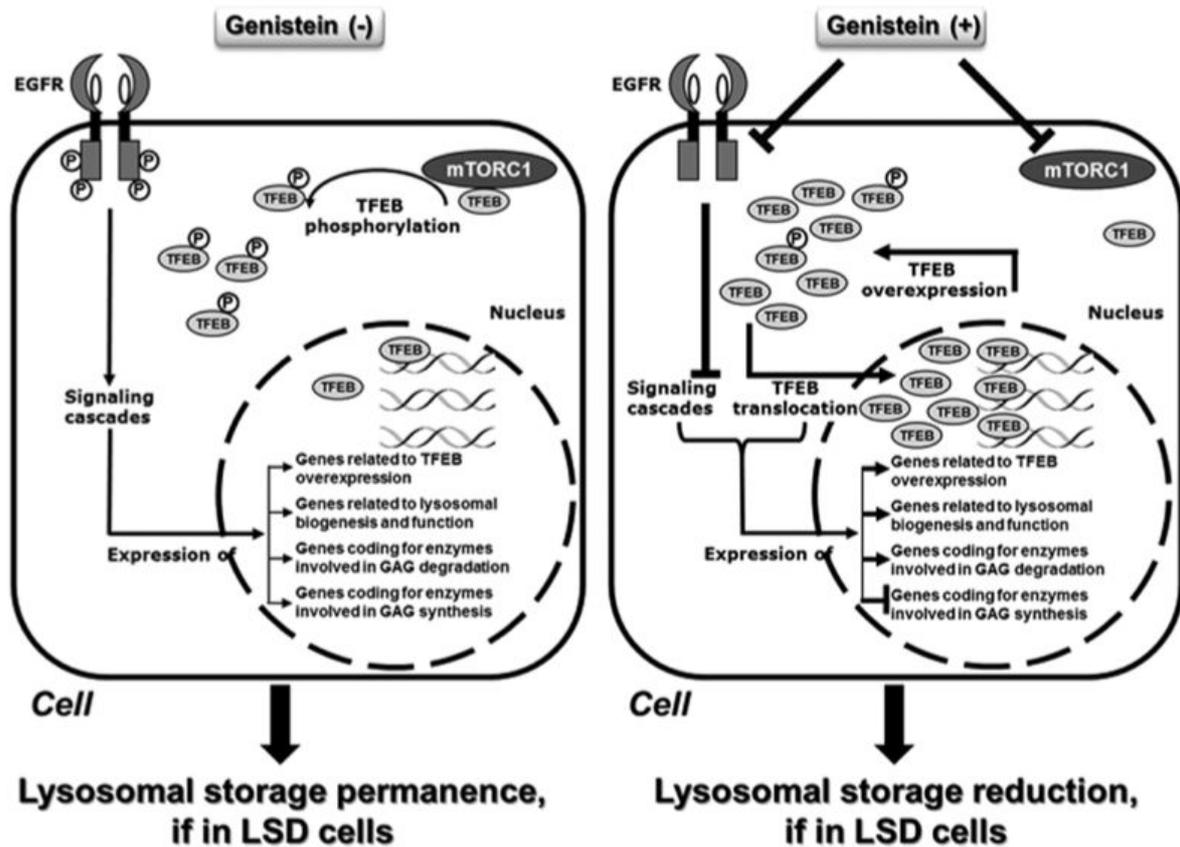
### *Substrate reduction therapy*

Taking into account the handicaps of ERT, SRT has been presented as a valid alternative approach. The objective of this therapy is to find molecular targets to decrease the production of the accumulated substrate and restore the balance between synthesis and degradation (figure 17). For this reason, it is important to remark the fact that the mutant enzyme has to maintain residual activity to achieve this restoration.



**Figure 17. SRT basis.** Restoration of the balance between synthesis and degradation is the main objective of SRT.

In the case of Sanfilippo disease, the goal of SRT is to decrease HS synthesis. In order to achieve this aim, several approaches have been developed using different types of molecules. All of them should be able to cross the BBB, since the main objective of any therapy for this disease is to reach the CNS. In the case of Sanfilippo syndrome, genistein and rhodamine B are the two small molecules more frequently studied and tested to check their SRT properties.



**Figure 18. Mechanisms of action of genistein.** Schematic representation of the putative mechanism of genistein to regulate the lysosomal metabolism-related genes and *TFEB*. (Moskot et al., 2014).

Genistein is a natural isoflavone that inhibits the kinase activity of epidermal growth factor receptor (EGFR), which is important for complete expression of genes encoding enzymes responsible for GAG production (Jakóbkiewicz-Banecka et al., 2009; Piotrowska et al., 2006). Genistein showed significant results in patients' fibroblasts (Piotrowska et al., 2006), peripheral tissues of Sanfilippo B mice (Malinowska et al., 2009)

and CNS of Sanfilippo B mice (Malinowska et al., 2010). Surprisingly, the reduction in the synthesis and storage of GAGs was also found in patients' cells carrying null mutations, indicating a combination of genistein mechanisms of action inside the cell (Banecka-Majkutewicz et al., 2012). A recent study established that genistein not only acts decreasing GAG synthesis, but also promotes transcription factor EB (TFEB) overexpression (figure 18), the lysosomal master regulator that promotes cellular clearance (Moskot et al., 2014).

Despite these promising results, low doses of genistein in human patients did not show significant amelioration in neuropathology although some improvement of the somatic symptoms have been described (Delgadillo et al., 2011; Ruijter et al., 2012b). The use of elevated doses of genistein has been proved to be safe in patients (Kim et al., 2013) but HS reduction results were inconclusive (Malinová et al., 2012).

Other flavonoids have also shown the ability to reduce the GAG synthesis in fibroblasts, independently of EGFR, and a combination of several of them resulted in a higher inhibition effect on the GAG synthesis (Kloska et al., 2011). Further studies with higher doses of genistein and other flavonoids should be carried out to establish the ability of this group of molecules to ameliorate CNS pathology in Sanfilippo patients.

On the other hand, rhodamine B is a colorant used in cosmetics that is thought to inhibit the GAG elongation step in the synthetic pathway (Kaji et al., 1991). Rhodamine B showed the ability to delay the somatic pathology in MPS IIIA mice (Roberts et al., 2006) and improved their behaviour and learning difficulties (Roberts et al., 2007). It is also thought that the effect of rhodamine B is stronger if there is some residual enzyme activity. Even though exposure to high doses of rhodamine B results in liver toxicity and reproductive detriment, none of these problems is detected after continued exposure at low doses, which have been reported to improve neurological defects in MPS IIIA mice (Roberts et al., 2010). Further studies should assess the ability of rhodamine B to ameliorate the neurological symptoms in animal models and patients without toxic effects.

Finally, some other molecules have been proposed as useful as SRT for Sanfilippo syndrome. One of them is miglustat, which was developed for the treatment of Gaucher disease type I patients (Cox et al., 2000). Some studies in Niemann-Pick disease type C

patients have shown a reduction in disease progression using miglustat, due to the decrease of the neurotoxic storage of G<sub>M2</sub> ganglioside (Patterson et al., 2007). Since this ganglioside is secondarily accumulated in Sanfilippo syndrome patients, miglustat could be an interesting therapeutic drug to reduce the neurodegeneration.

Another option is the use of specific RNAi directed to genes involved in the GAG synthesis. RNAi is a mechanism to selectively silence the expression of a particular gene by the specific degradation of the mRNA. Synthetic siRNAs and shRNAs have been widely used to downregulate the expression of a large number of genes. The use of shRNAs to downregulate *EXTL2* and *EXTL3* genes was found to reduce the GAG synthesis and storage in MPS IIIA fibroblasts (Kaidonis et al., 2010). The same effect was reported for the use of siRNAs to downregulate *XYLT1*, *XYLT2*, *GALTI* and *GALTI*, genes coding for enzymes responsible of the linkage region formation (Dziedzic et al., 2010).

### *Coenzyme Q<sub>10</sub>*

Coenzyme Q<sub>10</sub> (CoQ<sub>10</sub>) is a lipophilic molecule synthesized in all the cells and found in all the membranes of eukaryotic organisms. It works as an electron carrier in the mitochondrial respiratory chain, although it participates in many other processes (Turunen et al., 2004). It is known that lysosomal membrane contains similar amounts of CoQ<sub>10</sub> than mitochondrial membrane, and plays a role in the acidification of the lysosomal lumen (Gille and Nohl, 2000).

A recent study showed that CoQ<sub>10</sub> was low in Sanfilippo B patients when compared to healthy individuals (Delgadillo et al., 2011). In order to assess its potential therapeutic effect, CoQ<sub>10</sub> and also an antioxidant cocktail were administrated to MPS IIIA and B fibroblasts. The results about the decrease in GAG storage and enzyme activity enhancement were not conclusive, but they showed a slight increase in exocytosis probably due to the treatment (Matalonga et al., 2014). Further studies in cell and animal models are needed to establish whether or not CoQ<sub>10</sub> presents a therapeutic effect.

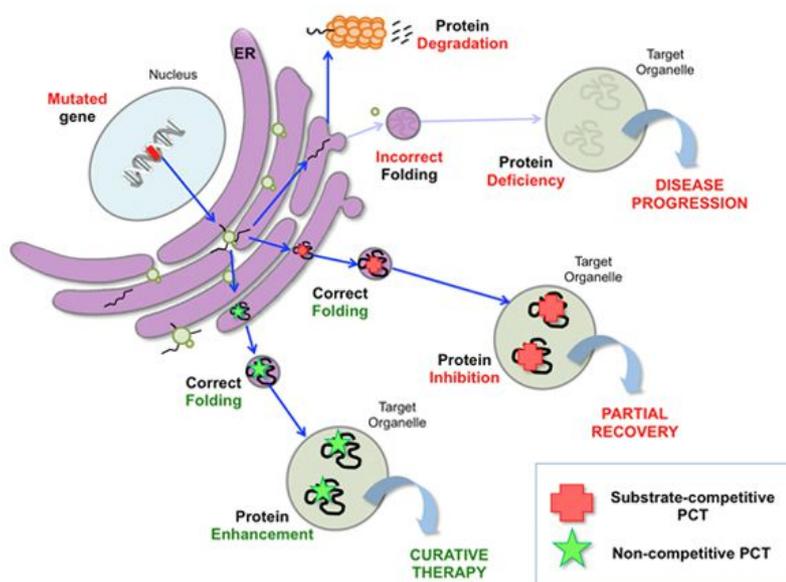
## Pharmacological chaperones for enzyme-enhancement-therapy

Many specific mutations, mainly missense, can give rise to misfolding proteins that can retain some residual activity but may be subjected to rapid intracellular degradation (reviewed in Suzuki, 2014).

Molecular chaperones are cellular proteins that act on the correct folding of other polypeptides (reviewed in Muchowski and Wacker, 2005). Pharmacological and chemical chaperones are small compounds that can be used, in a similar way, to avoid the misfolding of mutant proteins (figure 19). They are principally enzyme inhibitors, which at low concentrations interact specifically with the active site of proteins to restore the correct folding and to stabilize them (Bernier et al., 2004). In the case of LSDs, once in the lysosome, the enzyme substrate replaces the chaperone, completing enzyme activity restoration (Sawkar et al., 2002). Amino and iminosugars are the most common pharmacological chaperones used in enzyme-enhancement-therapy.

**Figure 19. Chaperone treatment possibilities.**

Action mechanisms of chaperones to restore correct folding of the lysosomal enzyme and its correct trafficking to the lysosome (www.minoryx.com).



Chaperones may represent an interesting approach for many LSDs. It is known that for these diseases, an enzyme activity above 10-20% is sufficient to preclude the development of clinical symptoms. Many chaperone approaches have been assayed at different levels for LSDs such as Fabry disease,  $G_{M1}$ -gangliosidosis, Morquio B disease, Pompe disease, Gaucher disease, Krabbe disease, Niemann-Pick A/B and C diseases, as

well as for other types of disorders such as retinitis pigmentosa, cystic fibrosis, Parkinson disease, Alzheimer disease or cancer (reviewed in Suzuki, 2014).

Several compounds were assayed as chaperones at the cellular level for Sanfilippo syndrome (Feldhammer et al., 2009b). One of them, glucosamine, which is a competitive inhibitor of the *HGSNAT* enzyme that is not expected to be highly toxic, showed to significantly increase enzyme activity in eight out of nine patients' fibroblasts, indicating its therapeutic potential. Further studies should be done in order to establish its efficacy and lack of toxicity in other cellular and animal models.

### *Suppression of premature termination codons*

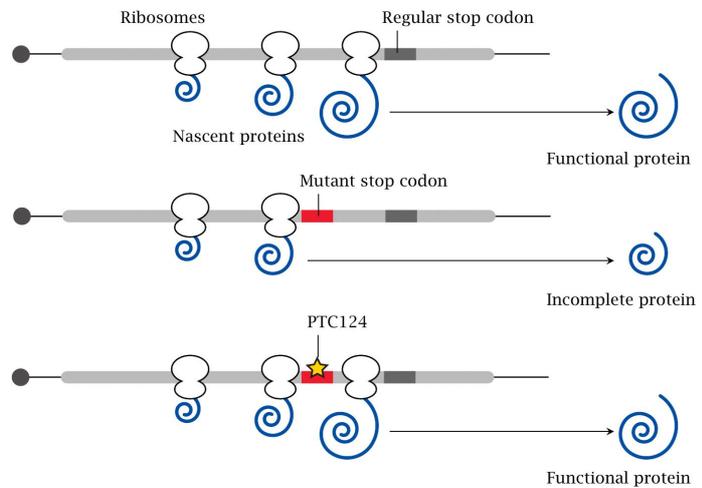
This therapy is based on the stop-codon read-through technology. It can be used in the cases where mutations introduce a premature termination codon (PTC) that gives rise to a truncated protein without proper enzymatic activity. Usually this PTC in the mRNA would activate the nonsense-mediated decay (NMD) response with a rapid degradation of the transcript. The use of small molecules that can read through the PTC would allow the cells to produce protein with a change in one amino acid but at least non-truncated. The read-through process occurs naturally in the cells at low frequencies (reviewed in Brooks et al., 2006).

Aminoglycosides are one type of molecules that have been tested for their read-through properties, especially gentamicin. Unfortunately, it promotes read-through also over well-positioned stop codons, resulting in unspecific effects. Moreover, long-term treatments using gentamicin can result in nephrotoxicity (Prayle et al., 2010).

Aminoglycosides derivatives have been developed in order to reduce their toxicity and improve their efficiency. Many different works have been reported using aminoglycosides or derivatives as a treatment for nonsense mutations in many cellular and animal models of different disorders (reviewed in Keeling et al., 2014). For example, NB54 aminoglycoside derivative was tested in MPS I fibroblasts (Nudelman et al., 2009) and mouse model (Wang et al., 2012) and showed improved properties when compared to

gentamicin. Gentamicin did not show a correction for MPS VI nonsense mutations (Bartolomeo et al., 2013).

**Figure 20. Read-through mechanism.** Example of read-through-mechanism of a PTC after PTC124 (Ataluren) treatment (watcut.uwaterloo.ca).



Other non-aminoglycoside antibiotics have been tested for this type of therapy such as PTC-124, also known as Ataluren (figure 20). Different works have been reported with different success (reviewed in Keeling et al., 2014). Ataluren has been tested in MPS VI (Bartolomeo et al., 2013) and MPS I (Peltz et al., 2013) with some degree of success. Nowadays, there is some controversy about the ability of Ataluren to restore enzyme activity.

Finally, the combination of PTC suppression together with NMD inhibition can result in an increase of the functional protein. One study in an MPS I mouse showed higher levels of enzyme activity in mice treated with an inhibitor of the NMD process together with a PTC suppression drug (Keeling et al., 2013).

Until now, no reports have been published for Sanfilippo syndrome regarding this type of therapy.

### *Overexpression of transcription factor EB*

TFEB has been proved to be a master regulator in the lysosome biogenesis and function. It is considered a target for any disorder that affects the lysosomal-autophagic

pathway since its overexpression promotes cellular clearance in cultured cells (Sardiello et al., 2009). It also reduces GAG levels and vacuolization in NSC from multiple sulfatase deficiency and MPS IIIA mouse models (Medina et al., 2011). Overexpression of TFEB in muscle cell cultures and Pompe mice enhance the lysosomal-autophagosomes fusion and result in an increase in the exocytosis of autolysosomes (Spampanato et al., 2013).

TFEB overexpression has been proved to be useful for other neurodegenerative disorders such as Parkinson disease (Dehay et al., 2010), Huntington disease (Tsunemi et al., 2012) and  $\alpha$ -1-anti-trypsin deficiency (Pastore et al., 2013).

How TFEB modulates cellular clearance should be further studied in order to determine the mechanism that activates the autolysosomal exocytosis. Different molecules that can promote TFEB dephosphorylation or inhibit its phosphorylation, enhancing its nuclear translocation, represent a novel putative therapeutic strategy for LSDs.

### *Gene therapy*

Gene therapy consists in the delivery of the correct copy of the gene (mainly the cDNA) to affected cells in order to recover the enzyme function. For Sanfilippo syndrome, many different viral vectors such as retroviruses, lentiviruses, adenoviruses and adeno-associated viruses (AAV) have been used to transduce cells (table 3). Also a gene therapy approach using a non-viral vector (pFAR4) administered via tail vein has been recently tested in MPS IIIA mice with significant results in the increase of the enzyme activity and the decrease of GAGs in different tissues and the reduction of the lysosomes in the brain (Quiviger et al., 2014). In these animals the liver became a sulfamidase secretor that promotes a reduction in GAG accumulation in others tissues (an scheme of the M6P transport of the lysosomal enzyme from a corrected cell to an affected cell is shown in figure 21)..

**Table 3. Gene therapy strategies for Sanfilippo syndrome.** AAV (adeno-associated virus), LV (lentivirus), IV (intravenous).

Model	Gene	Viral vector	Administration
MPS IIIA Fibroblasts	<i>SGSH</i>	Retrovirus	-
MPS IIIB Fibroblasts	<i>NAGLU</i>	LV	-
MPS IIIA mice	<i>SGSH</i>	LV	IV
MPS IIIA mice	<i>SGSH</i>	LV	Intraventricular
MPS IIIA mice	<i>SGSH</i>	Adenovirus 2	Intraventricular
MPS IIIB mice	<i>NAGLU</i>	AAV 2	IV + Intracinsternal
MPS IIIB dog	<i>NAGLU</i>	AAV 5	Intracerebral
MPS IIIB mice	<i>NAGLU</i>	AAV 5 + LV	Intracranial + IV
MPS IIIB mice	<i>NAGLU + SUMF1</i>	AAV 5	Intracerebral
MPS IIIA mice	<i>SGSH</i>	AAV 8	Intramuscular + IV
MPS IIIA mice	<i>SGSH</i> + signal peptide	AAV 8	IV
MPS IIIA mice	<i>SGSH</i>	AAV 9	Intracerebrospinal fluid
MPS IIIA dog	<i>SGSH</i>	AAV 9	Intracerebrospinal fluid

Retroviral constructs were reported to express the corresponding cDNA for MPS I, II, IIIA, VI and VII (reviewed in Neufeld and Muenzer, 2001). For MPS IIIA fibroblasts, continuous infections were necessary to obtain a high enzyme activity (Bielicki et al., 1996).

Lentiviral vectors are another option to deliver the correct copy of the gene into the cells. For Sanfilippo B fibroblasts, high amounts of the enzyme were produced using this strategy, but with a poor uptake of the secreted protein by untreated cells (Villani et al., 2002). Lentiviral gene therapy was also assessed in the mouse model of MPS IIIA by intravenous delivery (McIntyre et al., 2008). Somatic amelioration was detected in treated mice, with high enzyme production and reduction of the GAG storage, while in the CNS the storage was decreased but not significantly, suggesting a poor neural transduction. With intraventricular infusion of this lentiviral vector, a reduction in the GAG storage in the CNS was achieved in the mouse model of Sanfilippo A (McIntyre et al., 2010).

Since retroviruses and lentiviruses are integrative viruses and promote an immunogenic response, a canine adenovirus serotype 2, which induces a low level of innate response, was assayed in the mouse model of MPS IIIA (Lau et al., 2010b).

Intraventricular injection of this virus resulted in a long-term enzymatic expression and an improvement in the neuropathology.

Lately, the use of AAV, which are non-integrative, non-pathogenic in humans and with capacity to infect non-dividing cells providing long-term expression, have emerged as a good delivery tool for the treatment of neurological disorders. In the last few years several reports have been published concerning this AAV-mediated therapy using different AAV serotypes.

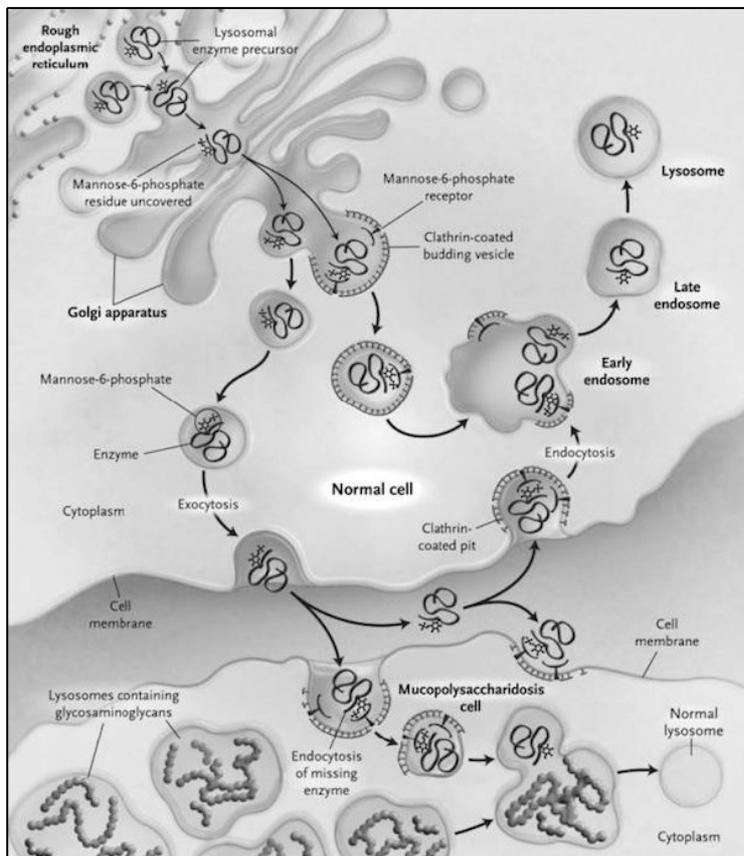
A combined administration via intravenous and intracisternal of an AAV2 carrying the *NAGLU* cDNA was shown to improve the lifespan and behaviour of the mouse model of MPS IIIB with a variable correction of the lysosomal storage pathology in the CNS and somatic tissues, and a restoration of the enzyme activity (Fu et al., 2007).

In MPS IIIB dogs, intracerebral administration of AAV serotype 5 plus immunosuppression showed to be a safe and an efficient method for gene therapy with an increase in the enzymatic activity and a decrease in the storage of GAGs (Ellinwood et al., 2011). A combination of intracranial AAV5 and intravenous lentiviral administration was assayed in MPS IIIB mice obtaining an improvement in the lifespan and motor activities (Heldermon et al., 2013). Intracerebral administration of AAV5 carrying the *SGSH* gene together with the *SUMF1* gene (coding for an essential and limiting factor for sulfatases) in the mouse model of MPS IIIA showed an increase in the SGSH activity in the brain, a decrease in the storage and inflammation and an improvement in the motor and cognitive function (Fraldi et al., 2007). After these results, a clinical trial phase I/II for MPS IIIA using AAV serotype 10 expressing the deficient SGSH enzyme and the SUMF1 enzyme was started. It recently finished showing no toxicity or lack of tolerance and a possible slight improvement in the patients' behaviour (NCT01474343, results at Tardieu et al., 2014).

AAV serotype 8 carrying the *SGSH* gene has been tested in a mouse model of Sanfilippo A, by intramuscular and intravenous administration (Ruzo et al., 2012a). Results showed no amelioration using intramuscular administration, while intravenous administration resulted in liver transduction and an improvement in the somatic tissues and some amelioration in the CNS symptoms, only in male mice. Another work, using AAV8 carrying *SGSH* together with a signal peptide to promote the secretion of the enzyme

and a BBB-binding domain to promote the cross of the BBB, resulted in a high enzyme activity in the brain together with an improvement in brain pathology and behaviour with minimal invasive treatment (Sorrentino et al., 2013).

Other works showed the ability of serotype 9 to transduce neural and some somatic cells and achieve gene correction, both in the mouse and in the canine models of MPS IIIA (Haurigot et al., 2013; Ruzo et al., 2012b). The mouse model, treated intravenously, showed an increase in the enzymatic activity together with a decrease in the GAG storage and neuroinflammation, expanding mice lifespan (Ruzo et al., 2012b). Intracerebrospinal fluid administration in the mouse and in the canine models of MPS IIIA resulted in an increase in the enzymatic activity in brain and serum, leading to whole body correction of GAG storage and lysosomal pathology, as well as a prolonged lifespan with low immune response (Haurigot et al., 2013). These results encourage the application of this approach in human patients.



**Figure 21. M6P-mediated enzyme excretion.** M6P-mediated transport of lysosomal enzymes from a corrected cell to an affected cell. This is an important process in the case of diseases affecting enzymes located at the lysosomal lumen that are directed to the lysosome via M6P signalling. For gene therapy and cell therapy, where not all affected cells are corrected due to the treatment, this process could play an important role (Kohan et al., 2011).

Finally, a recent study assayed the safety of AAV9 carrying the *NAGLU* gene in unaffected primates via intravenous injection (Murrey et al., 2014), since it seems to be the most efficient serotype for the treatment of neurological disorders. Persistent and high enzyme activity was detected in brain and serum with a low immune response. This result supports, again, the feasibility of this therapy for human patients.

### *Stem cell therapy*

In the last few years several stem cell applications have been described in the field of the treatment of neurobiological diseases in order to deliver the normal copy of the gene into the brain. Direct implantation in the brain of cells secreting the correct enzyme represents one of the options for this type of treatment (see figure 21). Several studies have tested these applications in many LSDs (reviewed in Ortolano et al., 2014) and specifically using NSC in MPS (reviewed in Giugliani et al., 2012). Taking into account that *HGSNAT* protein is not mannose-6-phosphatated, the secretion and uptake of this enzyme by deficient cells may not be successful in the case of Sanfilippo C syndrome.

Allogeneic BMT is used in the treatment of different LSDs with neurological affection, but its efficacy in the case of MPS III has been controversial. Intravenous administration of lentiviral-transduced bone marrow stem cells into the MPS IIIB mouse model reduces pathological manifestations in brain and GAG levels in liver and spleen (Zheng et al., 2004). However, recent works demonstrated the inefficacy of this treatment in the mouse model of MPS IIIA (Lau et al., 2010a) due to an insufficient production of enzyme by the donor cells or an inefficient uptake by the host cells (Lau et al., 2013), suggesting that it should not be used in affected children.

Hematopoietic stem cell (HSC) transplantation has been largely tested in many patients suffering from different LSDs. The aim of this type of therapy is to provide donor cells that can produce the enzyme deficient in patients' cells. In patients' brain, these cells can replace microglia and become enzyme-secreting donor cells (Krivit et al., 1995). Nevertheless, this process seems to be slow and not complete, and HSC transplantation has been considered an invalid therapy for neurological disorders with a rapid progress of

symptoms. Moreover, this process implies a high risk of morbidity and mortality, being appropriate only in some cases.

Nevertheless, some successful results have been obtained for HSC transplantation therapy in MPS I (Aldenhoven et al., 2008). It seems important to start the treatment early (at the age of 2-2.5 years), before the development of severe neurological symptoms (Ru et al., 2011). For other MPS such as type II and VI, results were less consistent, with an improvement in the somatic symptoms but no amelioration in the neurological and skeletal damage (Boelens et al., 2010). In the case of MPS IV, where the skeletal symptoms are the major feature, HSC transplantation does not show any effect.

In MPS III, no benefit has been described for this treatment but the reason remains unclear (Hoogerbrugge et al 1995; Peters and Steward 2003; Sivakumur and Wraith 1999; Vellodi et al 1992). Currently, this approach is no longer considered for the treatment of Sanfilippo syndrome. However, a recent work using HSC genetically modified with a lentivirus vector carrying the normal copy of the *SGSH* gene showed an improvement in the neurological pathology in MPS IIIA mice (Langford-Smith et al., 2012).

Administration of human umbilical cord blood cells to the MPS IIIB mouse model has been explored via intracerebral (Garbuzova-Davis et al., 2005) and intravenous delivery (Garbuzova-Davis et al., 2009) resulting in both cases in an amelioration of the neurological and somatic symptoms. However, it presents the inconvenience that the enzyme production declines with time. On the contrary, the transplantation of umbilical cord blood-derived stem cells in two type B patients before the disease onset did not prevent the neurological deterioration (Welling et al., 2014).

NSC can be isolated from human and murine brains, but recently, the development of iPSC technology (see below) has allowed the opportunity to easily obtain patient specific NSC. These cells have the ability to differentiate into neurons, astrocytes, and oligodendrocytes. They also have the property to migrate to distal sites, to differentiate and to integrate in the brain without disrupting its normal function. In conclusion, these cells present a unique opportunity to deliver the WT enzyme and also to restore neurological deficits.

Another option for therapy consists in the use of glial precursors cells derived from embryonic stem cells. It was assayed in the mouse model of MPS IIIA (Robinson et al., 2010) with cells genetically modified to overexpress the *SGSH* gene. These cells were previously differentiated to avoid the presence of stem cells in the implants. They were able to engraft in the brain, presented a long-term *SGSH* overexpression and survival and did not form teratomas, making it a promising therapeutic approach.

## RNA Splicing

### *The splicing process*

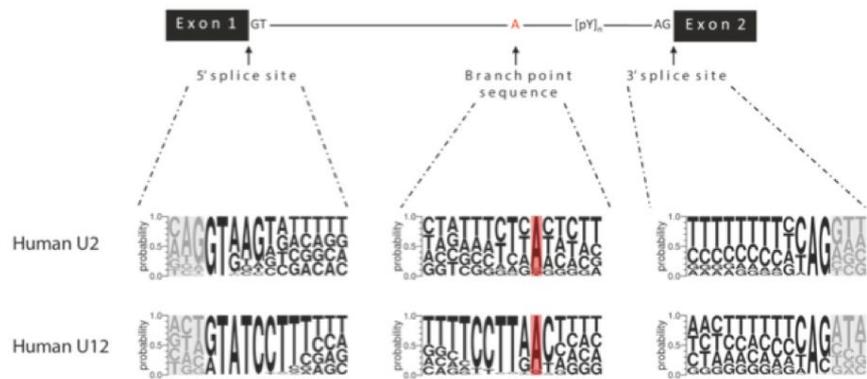
Splicing is an essential step in the expression of most of human genes, consisting in the excision of introns from precursor mRNAs and ligation of exons to generate the mature mRNA, prior to its translocation from the nucleus to the cytoplasm. This process is conducted by the spliceosome, which is a ribonucleoprotein complex formed by five snRNAs and more than 200 different proteins (reviewed in Wahl et al., 2009).

The spliceosome is able to recognise the different splice sequences that play a fundamental role, such as the 5' and the 3' splice sites (ss), the branch point and the polypyrimidine tract (the two latter located within 50 nucleotides upstream of the 3' ss). These sites present sequence variability along the human genome (Roca and Krainer, 2009).

Two different types of introns have been described and are excised by two different spliceosome complexes, U2 and U12, being the U2 type-introns the most frequent while the U12 represent a minor class of introns (less than one thousand U12 type-introns have been described in humans) (reviewed in Hertel, 2014).

The canonical U2-intron sequences in humans (figure 22), despite they are highly degenerated when compared to lower organisms, are AG/GTRAGT (R means purine and the slash corresponds to the exon-intron boundary) and YAG (Y means pyrimidine), for the 5' ss and the 3' ss respectively, while for the branch point it is YTNAN (N represents any nucleotide). The polypyrimidine tract is a sequence of 15 to 20 nucleotides rich in pyrimidines.

**Figure 22. Conserved splice sites.** Canonical sequences for splice components in human U2 and human U12 introns (Hertel, 2014).



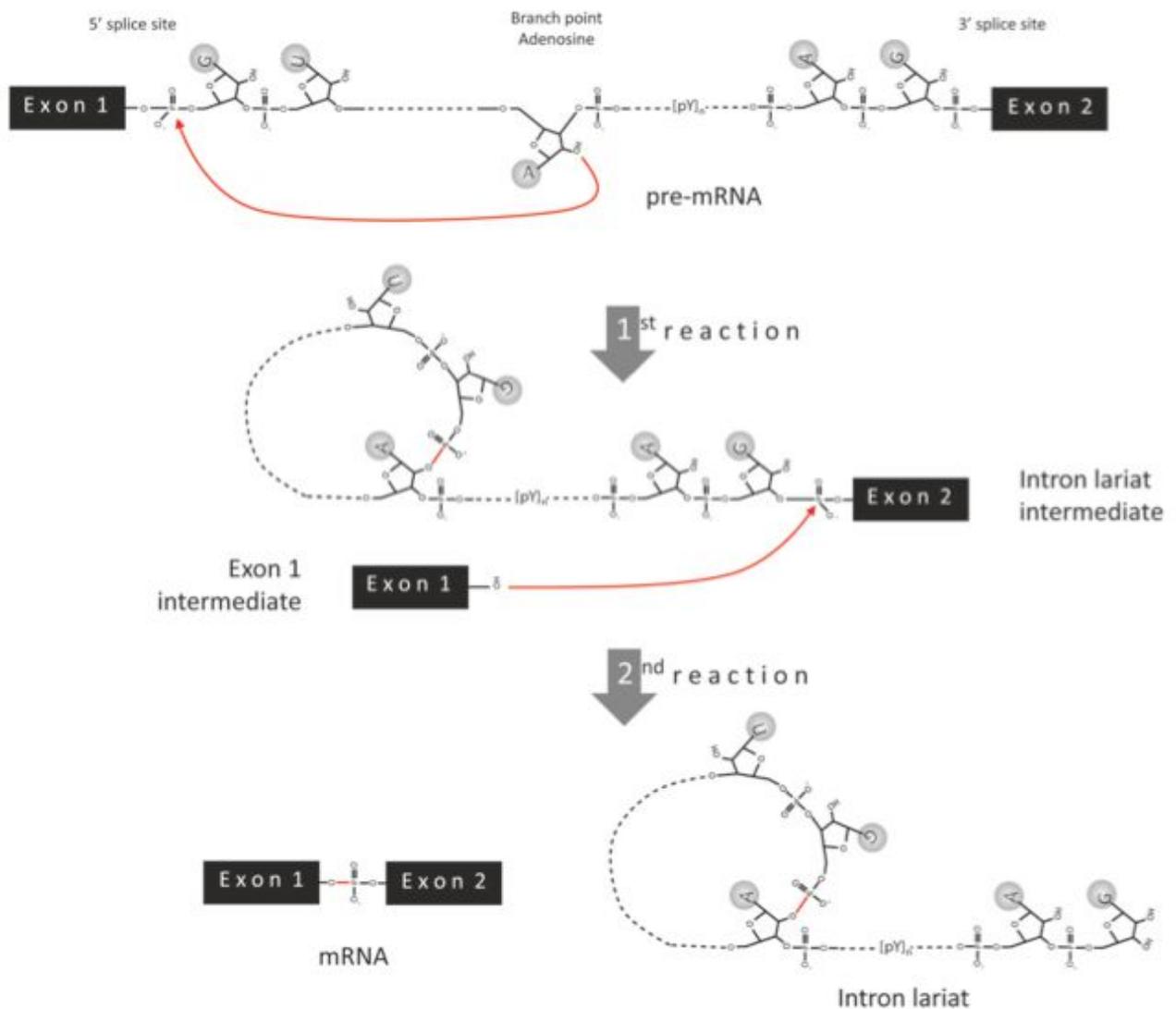
Other sequences play an important role in the intron definition, such as exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs), exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs). Members of the SR protein family usually recognize ESEs and ISEs while ESSs and ISSs are generally recognized by heterogeneous nuclear ribonucleoproteins (hnRNPs). These sequences participate in the splicing process and are important in the regulation of the different alternative splicing processes together with the secondary structure of the mRNA and RNA modifications. This is the reason why in some cases, exonic mutations can affect splicing process, if they are situated in ESEs or ESSs (Cartegni et al., 2002).

Thus, the splicing process is extremely complicated and regulated, and is crucial to obtain the correct gene expression that can generate more than one final protein product.

For the recognition of the 5' ss consensus motif, the U1 snRNA, which presents a partially complementary sequence to that motif, is essential. The U1 small nuclear ribonucleoprotein is composed of a 164-nt long U1 snRNA and several protein factors. The 5' region of the U1 snRNA is involved in the recognition of the 5' ss, being the C8 nucleotide the one that binds the first nucleotide (G) of the intron (reviewed in Roca et al., 2012).

For the recognition of the 3' ss, the branch point and the polypyrimidine tract, a complex set of different factors including SF1 and U2 snRNA play a fundamental role (reviewed in Hertel, 2014).

After the recognition of the sequences, the spliceosome formation is completed by the recruitment of many different components. Variation in their conformation to achieve the activated form of the complex is necessary to complete the process. The splicing of the intron *per se*, occurs in two steps; a first step where the branch point reacts with the 5'ss and a second step where this site reacts with the 3'ss (figure 23), resulting in the release of this intron and completing an essential process in gene expression (reviewed in Hertel, 2014).



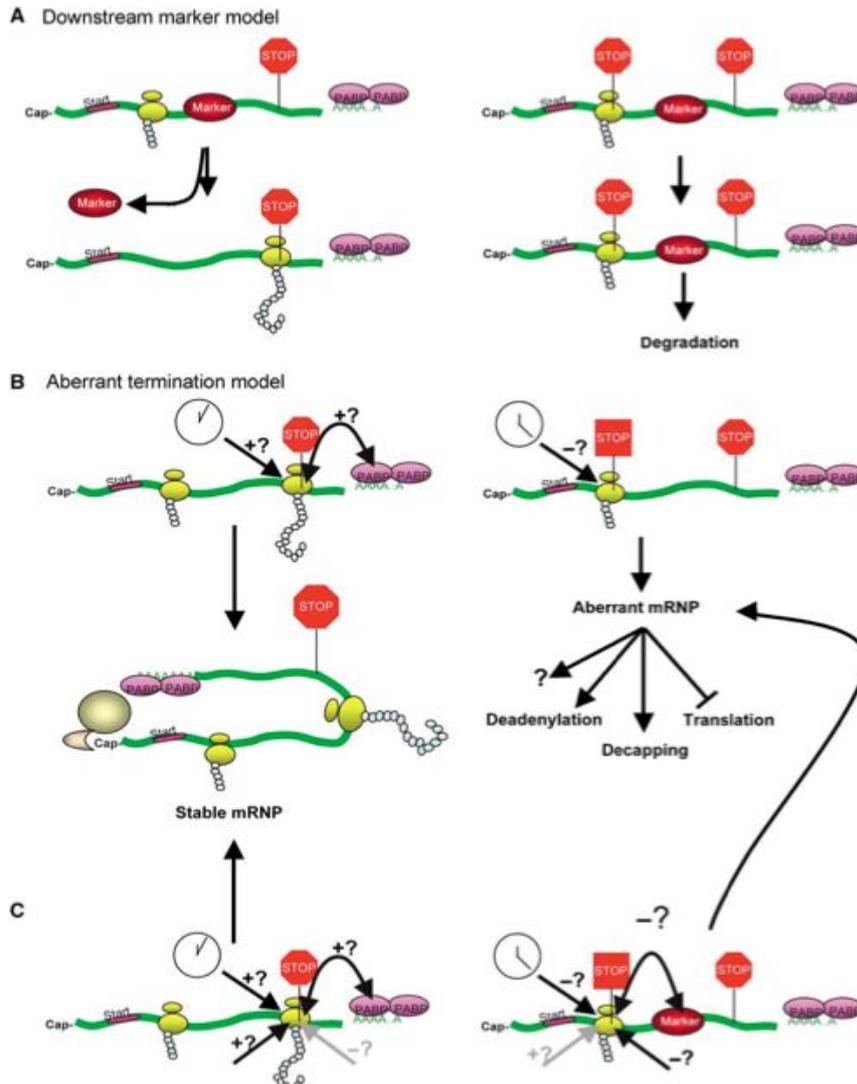
**Figure 23. Splicing reaction.** The two-steps splicing reaction for intron excision (Hertel, 2014).

### *The nonsense mediated decay quality control*

After splicing, the mature mRNA in the cytoplasm is submitted to different quality control processes to decide whether the mRNA is going to be translated or degraded. NMD is one of these quality control processes; a conserved mechanism responsible for the elimination of transcripts that would produce truncated proteins because of the presence of a PTC in their sequences. It has been estimated that around 30% of human disease alleles introduce a PTC changing a sense codon to a stop codon or introducing a nonsense codon indirectly (insertions, deletions and splicing mutations that vary the reading frame) (reviewed in Shyu et al., 2008).

The NMD mechanism is not yet fully understood, but two different models have been proposed: the downstream marker model and the aberrant termination model (figure 24). In the first model (figure 24A), it is supposed that a group of proteins called exon junction complex (EJC) is attached during the splicing process at approximately 20 nucleotides upstream of the exon-exon junctions, and they are used as the mark that triggers the NMD process (Le Hir et al., 2000). It seems that if the translational machinery detects an EJC after a stop codon, NMD occurs. Thus, PTCs situated more than 50-55 nucleotides before the last exon triggers NMD, while those situated in the last exon do not (Zhang et al., 1998).

The second model (figure 24B) is based on the difference in termination events when the transcript contains a PTC or not (Amrani et al., 2006). The basis of this model is the proximity between the stop codon and the 3' untranslated region (UTR). When this distance is longer because of the presence of a PTC, the remodelling steps required for the termination of the translation are not correct, either because the process occurs slower (Hilleren and Parker, 1999) or because the process is different at the biochemical level (Amrani et al., 2004).



**Figure 24. NMD mechanism models.** **A.** Downstream model with EJC markers. **B.** Aberrant termination model with PTC far from 3'UTR. **C.** Combination of both models (Shyu et al., 2008)

None of these models are applicable to all NMD cases and for both of them exceptions can be found. Recently, a combination of both models (figure 24C) has been proposed to explain NMD mechanism (Shyu et al., 2008), based on the fact that many features affect the translation termination and the decay of the transcripts.

In any case, still remains unclear how the transcript degradation takes place. In lower organisms, decapping of the mRNA or endocleavage have been described as markers to recruit degradation enzymes, while in mammals it seems that the rapid deadenylation of the transcript may be the marker for the recruitment of the catalytic enzymes (reviewed in Shyu et al., 2008).

### *Therapies to correct splicing mutations*

Different therapeutic approaches have been recently developed in order to correct splicing defects for many human disorders. The aim of these therapies is to obtain enough amounts of normal transcripts to produce levels of functional protein able to reverse the clinical phenotype of the patients.

One of these approaches consists in the overexpression of splicing factors such as SR family protein members or hnRNPs to modulate the transcript production (reviewed in Nissim-Rafinia and Kerem, 2005).

The use of small molecules to modulate the amount, the activation or the distribution of these splicing factors has also been assayed as a therapy to increase the correct splicing process (reviewed in Wang and Cooper, 2007).

In this regard, one common strategy consists in the use of anti-sense oligonucleotides (AON), a short RNA sequence complementary to that of the target mRNA. AONs can be used to restore normal splicing in the case of splicing mutations, or to alter the normal splicing in the case of mutations that introduce a PTC or that change the reading frame (reviewed in Aartsma-Rus and van Ommen, 2007). They have also been applied to block a cryptic splice site that appears because of a mutation or to block ESS or ISS in order to favour the inclusion of a constitutive exon.

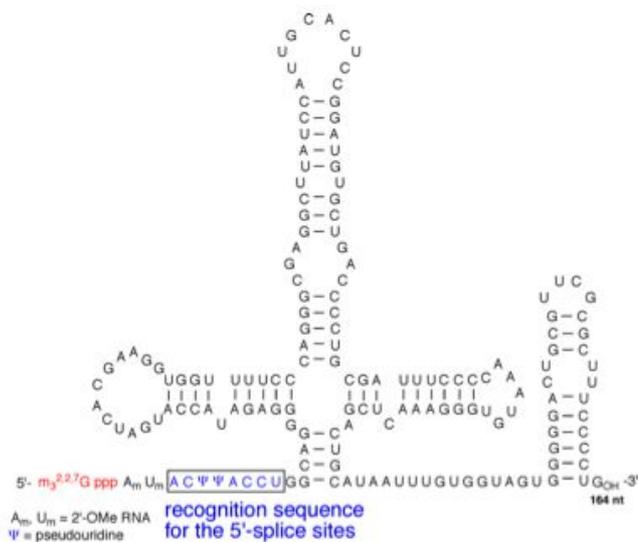
Finally, modified U1 snRNAs has been developed as another therapeutic approach to recover the normal splicing process by improving recognition of mutated 5'ss. This site can present mutations at different positions (both in the exon or in the intron), resulting in a decrease in the use of this site during the splicing process and resulting in the formation of aberrant transcripts.

Modified U1 snRNAs are designed to specifically present a higher complementary to the mutated site in order to improve the efficiency of its use by the splicing machinery as an attempt to correct the splicing defect (figure 25). They have been assayed as a possible therapy for several disorders and splicing mutations affecting different positions of the 5'

ss (Baralle et al., 2003; Fernandez Alanis et al., 2012; Glaus et al., 2011; Hartmann et al., 2010; Mattioli et al., 2014; Pinotti et al., 2008; Sánchez-Alcudia et al., 2011; Schmid et al., 2013; Susani et al., 2004; Tanner et al., 2009).

The success of these modified U1 snRNAs in the correction of splicing defects seems to depend not only in the complementarity of the sequence but also in the presence of proximate alternative sites, enhancers or silencers, the secondary structures and regulatory proteins (Roca et al., 2013).

Recently, an *in vivo* assay using modified U1 snRNAs has been performed in mice for the treatment of severe human factor VII deficiency (Balestra et al., 2014).



**Figure 25. U1 snRNA structure.** Splice site recognition sequence can be modified in order to increase complementarity to mutated sites (Ohkubo et al., 2013).

## RNA Interference

### *History and mechanisms*

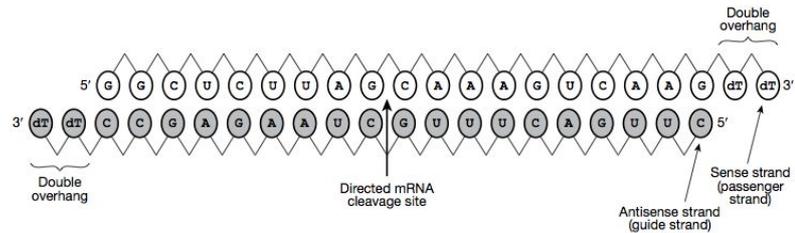
RNA interference (RNAi) is a post-transcriptional mechanism for gene silencing firstly described in *C. elegans* in the year 1998 by Andrew Fire and Craig Mello, who were awarded with the Nobel Prize in 2006 for their work (Fire et al., 1998). It consists in the silencing of mRNAs as a result of their hybridization with small antisense dsRNA molecules that inhibit their translation. Later, RNAi was described also in plants and in mammals (Hamilton and Baulcombe, 1999), opening the door for a revolution in the field of gene regulation.

A few years later, Elbashir and collaborators (Elbashir et al., 2001) showed that dsRNA are processed into 21 and 22 nucleotides small interfering RNA (known as siRNAs) by the ribonuclease III, and those siRNAs trigger the specific suppression of gene expression in many different organisms including humans. Those discoveries provided scientist with a new and powerful tool in the study of gene function and the development of new knockdown therapies.

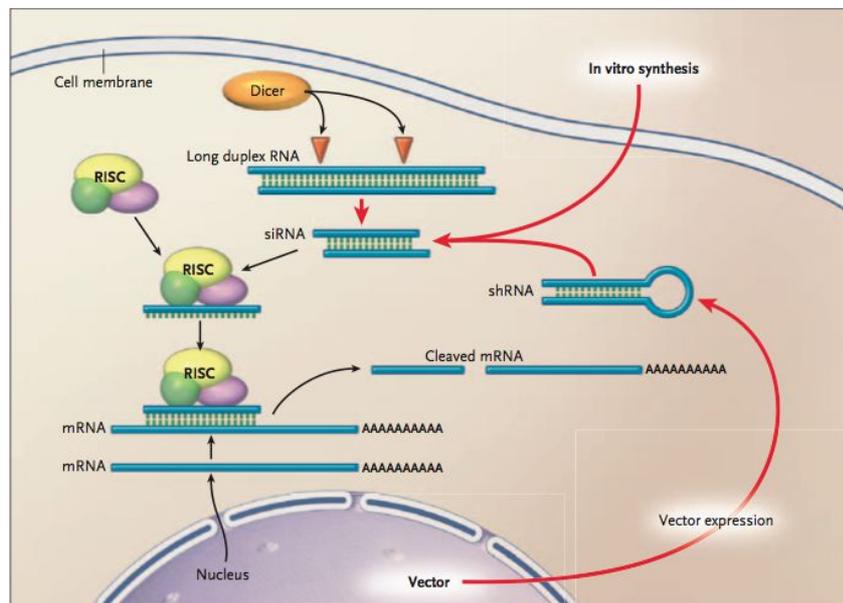
Three different types of small non-coding RNAs with the ability to induce RNAi have been described or developed. One of them, microRNAs (miRNAs), were firstly reported by three independent groups (Lee and Ambros, 2001; Lau et al., 2001; Lagos-Quintana et al., 2001) and are endogenous small RNAs with the function to regulate gene expression and consequently many cellular functions. They are transcribed in the nucleus as pri-miRNAs, which are processed into pre-miRNAs by the Drosha complex prior to their translocation to the cytoplasm. Once there, the Dicer complex processes the pre-miRNAs and generates 20 to 23 base-pair mature miRNAs (Bartel, 2004). These miRNAs are loaded in the RNA-induced silencing complex (RISC) and bind the target mRNAs by partial complementary to the 3' UTR (John et al., 2004). The fact that complete complementary is not required, allowed each miRNA to regulate hundreds of genes (Aagaard and Rossi, 2007).

**Figure 26. siRNA structure.**

Guide strand, passenger strand, overhanging nucleotides and the cleavage site (DeVincenzo, 2012).



The other two classes comprise siRNAs and short hairpin RNAs (shRNAs), both developed to exogenously modulate gene expression. On the one hand, siRNAs are short dsRNAs (21 base-pair) with two overhanging nucleotides and one guide strand (figure 26). This strand is directly incorporated to the RISC complex and leads it to the target mRNA. Two different responses can take place depending on the complementarity of the siRNA guide strand with the target mRNA sequence. If they are totally complementary, the mRNA is cleaved and the RISC with the guide strand binds another target mRNA (Elbashir et al., 2001). In the cases where the complementarity is partial, the target mRNA remains attached to the guide strand and it is not translated but neither cleaved (figure 27).

**Figure 27. RNAi mechanism.**

Dicer complex processes long RNAs or shRNA-coded products. In both cases, siRNAs guide strands are attached to RISC complex and recognize target mRNAs, which will be cleaved or repressed depending on the complementarity (Bernards, 2006).

On the other hand, shRNAs were developed to overcome the problem of the siRNAs short life inside the cells. They are introduced into the cells in a dsDNA vector that expresses the shRNA transcript, which is processed by the Drosha complex resulting in the pre-

shRNA. It is translocated from the nucleus to the cytoplasm, where Dicer processes it to obtain the mature shRNA as a siRNA that is incorporated to the RISC complex (figure 27) to function in the same manner than the siRNAs (Yu et al., 2002).

### *Therapies based on this mechanism*

Both siRNAs and shRNAs have been largely used for mRNA inhibition in several disorders and cell types due to their effectiveness in gene silencing. However, some barriers should be overcome to improve the results, such as the instability of naked siRNAs, the off-target effects, the activation of an immune response, and the delivery of both molecules to the target cells (reviewed in Deng et al., 2014).

To solve these problems, several issues have been addressed such as modifications in the chemical structure of the siRNAs (reviewed in Rana, 2007; reviewed in Zhou et al., 2014) or the encapsulation of siRNAs and shRNAs in many different types of particles to improve their stability and delivery (reviewed in Deng et al., 2014; reviewed in Joo et al., 2014; reviewed in Zhou et al., 2014).

In the recent years, several therapeutic approaches involving RNAi for many different disorders have been tested *in vitro* and *in vivo* (reviewed in Davidson and McCray, 2011; reviewed in Deng et al., 2014). For the *in vivo* assays, the target cell type and the class of molecule that is being used for gene silencing should be taking into account to decide the route of delivery. After some successful results, different clinical trials have been started using RNAi for some disorders (reviewed in Zhou et al., 2014). However, problems such as off-target effects, immune response or low efficacy commonly appeared in these trials, mainly when free siRNAs were administered. One crucial aspect in the near future is to evaluate the silencing effect of modified siRNAs conjugated with molecules or encapsulated in polymers as well as the efficacy of different delivery approaches such as viral systems (table 4).

**Table 4. RNAi delivery systems.** (adapted from Davidson and McCray, 2011)

Species/formulation	Packaging capacity	Applications and considerations
<b>Viral vector</b>		
Adenovirus	Up to 35 Kb, usually <10	dsDNA vector with large packaging capacity, transient expression, highly immunogenic
Adeno-associated virus	Around 4.5 Kb	ssDNA vector, small packaging capacity, mildly immunogenic, lasting expression in nondividing cells, capsid pseudotyping/engineering facilitates specific cell-targeting
Lentivirus	Up to 13.5 Kb	RNA vector, integration competent and incompetent forms available, less immunogenic, envelope pseudotyping facilitates cell targeting, clinical production more difficult
Herpes simplex virus	150 Kb	DNA vector, episomal, lasting expression, immunogenic
<b>Bacterial vector species</b>		
<i>E. Coli, S. Typhimurium</i>		Delivery of shRNAs or siRNAs to gut tissue
<b>Non-viral formulations</b>		
Nanoparticle		Self-assembling, may target specific receptors, require technical expertise to prepare
Stable nucleic acid lipid particle		Stable for systemic delivery, broad cell-type delivery
Aptamer		Targeting of specific receptors, requires sophisticated screening to develop
Cholesterol		Stable for systemic delivery, broad cell-type delivery

### *RNA interference therapies for Sanfilippo syndrome*

As mentioned above, in the synthetic pathway of GAGs, the linkage region formation is an essential step previous to the specific HS synthesis. Is in this point where *EXTL* genes (*EXTL1*, *EXTL2* and *EXTL3*) play a crucial role (Kreuger and Kjellén, 2012).

RNAi targeting these genes could be applied as a SRT for all Sanfilippo patients, since all of them present the inability to degrade HS. Specifically in the field of MPS, two different therapeutic approaches using RNAi have been previously tested, inhibiting different genes of the HS synthetic pathway (Dziedzic et al., 2010; Kaidonis et al., 2010).

In one of the studies, siRNAs were the molecule chosen to regulate the gene expression of the four genes implicated in the formation of the linkage region Results in MPS I and MPS IIIA fibroblasts showed an important decrease at the mRNA and protein levels for all the genes and a consequent significant decrease in the GAG synthesis after three days of treatment (Dziedzic et al., 2010).

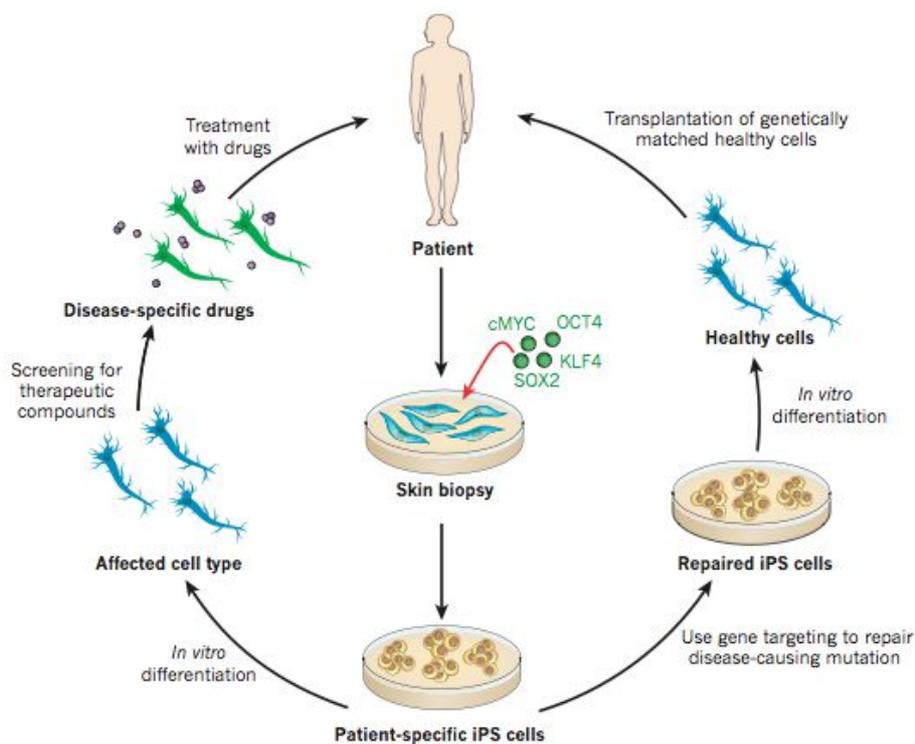
In the other case, the use of shRNAs inhibiting *EXTL2* and *EXTL3* genes was assayed in MPS IIIA fibroblasts. Results showed a reduction at the mRNA level as well as a decrease in the GAG synthesis and storage after three days of treatment. At longer times (seven days), only one shRNA was still capable to produce a decreased GAG synthesis although it was not significant (Kaidonis et al., 2010).

Both studies showed that it is possible to accomplish a reduction at the mRNA and protein level, GAG synthesis and storage at three days using RNAi, but failed in the attempt to extent these results in time. Further investigations have to be done in order to establish if RNAi could be a long-term therapy for Sanfilippo syndrome.

## Induced Pluripotent Stem Cells

### *Reprogramming from fibroblasts to induced pluripotent stem cells*

The possibility of reprogramming human adult cells to induced pluripotent stem cells (iPSC) was recently described by the group of Yamanaka (Takahashi et al., 2007), who was awarded with the Nobel Prize in 2012 due to the relevance and applicability in many biological fields of this new technology. This approach allows to obtain different human cell types minimizing ethical issues and becomes a powerful tool to model and treat a wide range of human disorders (figure 28).



**Figure 28. iPSC basis and applications.** The development of patient-specific iPSC represents an important tool in regenerative medicine and cell therapy as well as for disease modelling to further study the molecular basis of pathologies and to perform drug screenings (Robinton and Daley, 2012).

Adult somatic cells reprogramming into iPSC involves to go back to a pluripotent state. This induction was first accomplished in mice by introducing 4 reprogramming

factors, *OCT4*, *SOX2*, *KLF4* and *c-MYC* (also known as OSKM or Yamanaka set), with retroviruses in fibroblasts (Takahashi and Yamanaka, 2006). Later was also achieved in human fibroblasts using the same factors (Takahashi et al., 2007).

In the following years, new sets of reprogramming factors and different delivery systems were developed such as the set *OCT4*, *SOX2*, *NANOG* and *LIN28* and the use of lentiviruses or non-integrative viruses. Different types of cells, mainly fibroblasts and keratinocytes, can be used (reviewed in González et al., 2011). Still remains unclear the mechanisms through which the overexpression of these genes can lead to this induced pluripotent status.

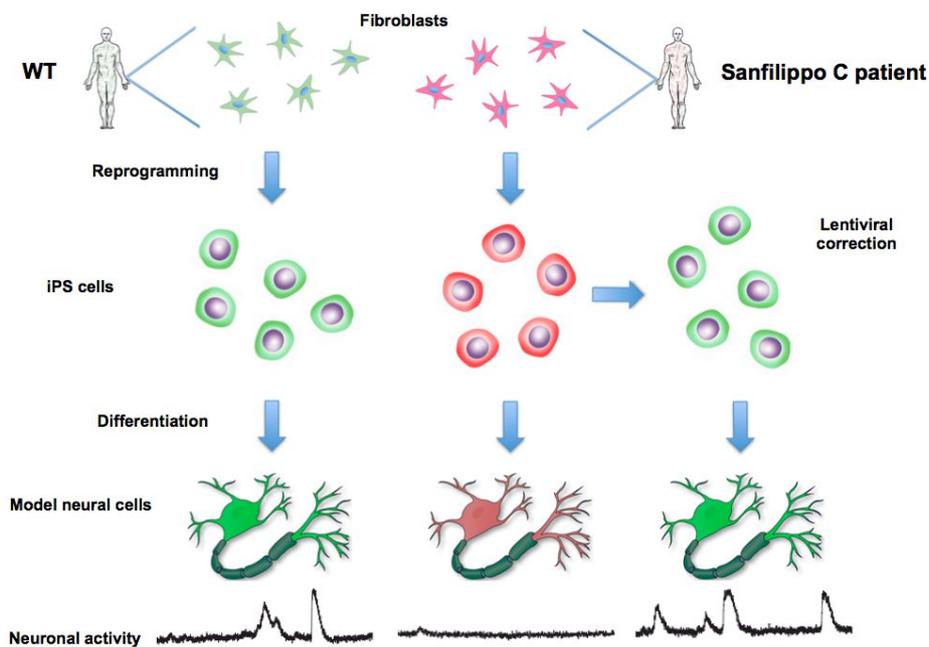
The human reprogramming process can last between one to three months depending on the cells and it can be considered an inefficient process since less than 0.01% of human fibroblasts become fully reprogrammed using retroviruses and the Yamanaka set of factors (González et al., 2011). When iPSC colonies emerge, each one from one reprogrammed cell (one clone), they have to be characterized performing the following set of validations (reviewed in Tiscornia et al., 2011):

- 1) The iPSC colonies morphology should be similar to that of embryonic stem cells (ESCs).
- 2) The iPSC have to be positive for alkaline phosphatase staining
- 3) Assessment of the presence and expression of the reprogramming genes in the cells
- 4) Positivity for the expression of different endogenous pluripotency genes
- 5) Demethylation status of promoters of important pluripotency genes
- 6) Karyotype stability of the iPSC
- 7) Ability of the iPSC to differentiate to the three germ layers both *in vitro* and *in vivo*

After iPSC validation, scientists are provided with a magnificent constant source of different cell types to model diseases in order to investigate their molecular and cellular basis, perform drug screenings into the affected cell type and develop cell therapies.

## Use of human induced pluripotent stem cells in lysosomal storage disorders

The differentiation of iPSC into neurons has been widely used in order to model central nervous system disorders causing neurodegeneration (reviewed in Marchetto et al., 2011). Some iPSC have been derived using fibroblasts from mouse models of different LSDs (Huang et al., 2012; Meng et al., 2010). Moreover, human cell models derived from iPSC of various LSDs have also been developed, such as Gaucher disease (Mazzulli et al., 2011; Panicker et al., 2012; Park et al., 2008; Tiscornia et al., 2013), Hurler syndrome (Tolar et al., 2011), Pompe disease (Higuchi et al., 2014; Huang et al., 2011), Sanfilippo B syndrome (Lemonnier et al., 2011) and Niemann-Pick type C1 (Trilck et al., 2013). These models can be used to achieve a better knowledge on the underlying molecular defects and cellular mechanisms that lead to the development of the diseases. They are also a good tool to perform drug screenings and test different therapeutic approaches in the affected cellular type (figure 29).

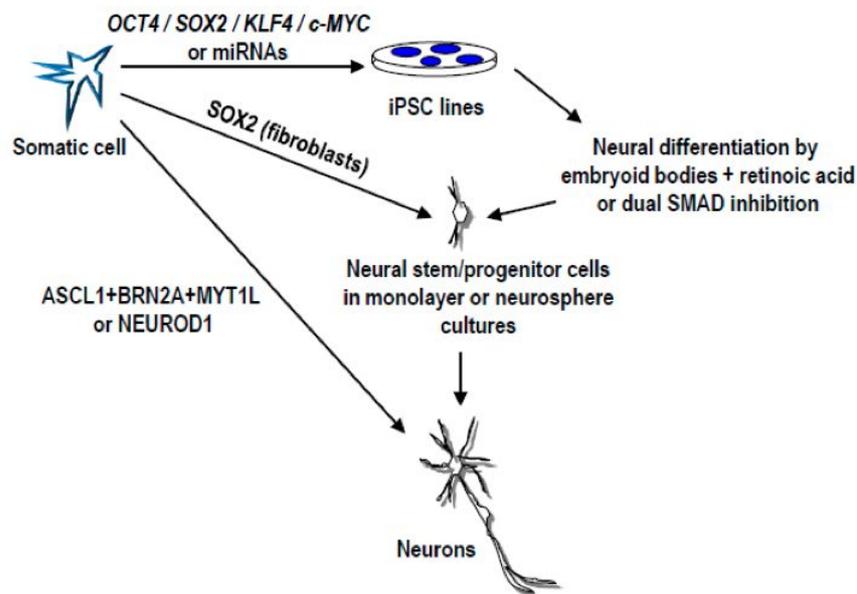


**Figure 29. Disease-modelling using iPSC.** Neuronal models derived from patient-specific iPSC will allow the study of the molecular basis of the disease in the neurons and how specific neuronal functions are affected due to the lysosomal dysfunction.

Moreover, iPSC could be used as a cellular therapy themselves for some diseases (Picanço-Castro et al., 2014), although it seems to be a long-term possibility that implies a better knowledge on the reprogramming mechanisms and its implications, the control of a complete differentiation, the response of the patient tissue cells and the adaptability of the new cells to the tissue environment.

### *Direct conversion of fibroblasts to neurons or neural stem cells*

More recently, direct conversion of fibroblasts to mature functional neurons (figure 30) has been developed using three factors (Vierbuchen et al., 2010) and later improved using only one of them, *ASCL1* (Chanda et al., 2014), avoiding the step of a pluripotent status cells, which can be considered potentially problematic in regenerative medicine. So, this technique provides an interesting tool for cell replacement therapy since pluripotent cells with the ability to develop teratomas seems not be appropriated for that purpose. Nevertheless, the low efficiency and the fact that a constant source of cells is not obtained, decrease the interest of its use to model diseases.



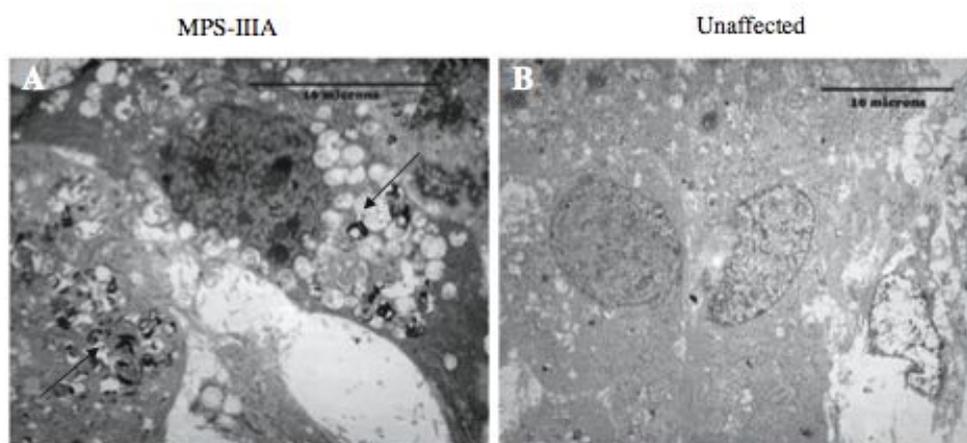
**Figure 30. Neuronal-disease-modelling.** Scheme showing different methods to obtain patient neurons from patient somatic cells, using reprogramming to iPSC, reprogramming to iNSC or direct conversion to neurons (Velasco et al., 2014).

Finally, direct reprogramming to multipotent induced neural stem cells (iNSC)(figure 23) with self-renewing ability, neural lineage restricted and no tumorigenic properties has been achieved using only one factor, SOX2 (Ring et al., 2012). This latter approach opens a new possibility for modelling and treatment of neurological diseases. Even so, the use of iPSC provides the advantage of obtaining cells that can potentially produce a wide range of cell types by differentiation in order to study diseases affecting different human tissues.

### *Cellular models of Sanfilippo syndrome*

Previously to the use of iPSC to model Sanfilippo syndrome, other systems were tested to develop cellular models that could allow the study of cellular and molecular mechanisms that promote the disease development.

The first attempt to establish a cellular model of Sanfilippo disease was performed in 2001 when caprine stem cells were treated to induce storage of gangliosides, which was previously reported in MPS IIID caprine brains. They demonstrate that not only the degradation pathway was affected but also the synthetic pathway of different gangliosides, suggesting an ER and GA affectation.



**Figure 31. Neuronal storage.** Typical MPS storage vesicles in neural cultures (Sutherland et al., 2008)

Next step was taken in 2008 when Sutherland and collaborators isolated neuronal cells from the cerebellum of newborn and adult MPS IIIA mice and were able to culture them. However, should be note that this work did not involve the use of pluripotent stem cells (Sutherland et al., 2008). These cultures apparently did not present any morphological difference between control and patients, and contained all different brain cell types in similar proportions, including neurons (less than 10% of the total), astrocytes and oligodendrocytes, with the ability to form culture neuronal networks. Typical MPS storage vacuoles appeared after some days (figure 31), and enzymatic activity of MPS IIIA cultures was always below 5% when compared to control cultures. Despite this, the enzymatic level slightly increased over time and was higher in newborn cultures (probably due to an increase in the protein amounts). The HS storage was also detected in all the cultures with an increase in the amounts over time, demonstrating the progressive nature of the accumulation. Taken together, all these results showed how the main features of the disease at the cellular level could be mimicked in a neuronal cultured model. Obviously, this technique could not be performed in humans since cells should be directly extracted from brains.

More recently, a cellular model of Sanfilippo disease was obtained using shRNAs to inhibit *NAGLU* gene (Sanfilippo type B) in HeLa cells (Roy et al., 2012). These HeLa cells accumulated storage vacuoles, over-express GM130 and showed an increased and disorganized Golgi complex. The authors showed how the downregulation of GM130 could rescue the normal phenotype and the upregulation in control HeLa cells reproduced the pathology. They suggested that the abnormal lysosomal formation is due to alterations in GA and that HS from the extracellular matrix modulate GM130 functions. This model presents the limitation that HeLa cells are not the cellular type affected in Sanfilippo patients. The development of neuronal models would represent a better option in order to study the molecular basis of this neurological disease.

Finally, the use of iPSC derived from patients' fibroblasts allowed the development of a Sanfilippo B neuronal model (Lemonnier et al., 2011). In this work, the authors remark the presence of storage vacuoles in the neurons and link the presence of these vacuoles to alterations in the Golgi complex and modified expression of Golgi matrix protein GM130, important for the structure and dynamics of GA. Human iPSC clones of one control and two patients using three and four factors were obtained in this work. In the case of patients,

NAGLU enzyme supplementation was needed to clear HS stored in fibroblasts that prevented the reprogramming process. They detected the storage lesions even in the iPSC clones and the expression profiles in NSC showed differences between control and patients in the Golgi related genes and in the genes coding for extracellular matrix components. The storage lesions did not affect the neural differentiation, but the neurons presented abundant storage vesicles LAMP1 and GM3 positive, confirming the typical secondary storage of gangliosides. Increase in the GM130 amounts indicating numerous abnormal organized Golgi complexes were also detected. In conclusion, this work remarks the Golgi affection in Sanfilippo patients' cells but remains unclear whether this is the cause or a consequence. Recently, the authors have described how cell polarisation and oriented migration are affected in this model and may contribute to the neurological phenotype of Sanfilippo patients (Bruyère et al., 2014).

To date, there are no animal or cellular model for Sanfilippo C disease. The development of a cellular model of Sanfilippo C would be of great interest since it is the only subtype due to mutations in a lysosomal membrane protein, which can lead to differences in the efficiency of some therapies despite the similarities in symptomatic manifestations and molecular basis between all subtypes.



# **OBJECTIVES**



The main objectives of this thesis are to perform a molecular analysis of Sanfilippo C syndrome, to test different therapeutic approaches for the disease and to develop a cellular neuronal model using the induced pluripotent stem cell technology due to the severe affection of the central nervous system in patients.

**Mutational analysis:**

- \* To perform a mutational analysis in Sanfilippo C patients
- \* To characterize the identified mutations at the mRNA and protein level and study the possible single origin of the common ones

**Therapeutic approaches:**

- \* To test the possible therapeutic effect of modified U1 snRNAs in the restoration of the normal splicing process for mutations affecting 5' splice site
- \* To evaluate the therapeutic effect as a substrate reduction therapy of different small interfering RNAs targeting *EXTL* genes to decrease the heparan sulfate synthesis

**Neuronal model:**

- \* To obtain induced pluripotent stem cells from fibroblasts of two different Sanfilippo C patients and a healthy control
- \* To differentiate the induced pluripotent stem cells to neurons and demonstrate that they recapitulate the main phenotypic features of the disease and to perform the analysis of neuronal activity and network connectivity to detect the brain affection



# RESULTS



## **REPORT OF THE THESIS SUPERVISORS ABOUT THE CONTRIBUTION OF THE PHD STUDENT TO THE ARTICLES IN THIS DOCTORAL THESIS**

**Thesis title:** “Genetic and molecular analysis of Sanfilippo C syndrome. Generation of a neuronal model using human induced pluripotent stem (iPS) cells and therapeutic strategies.”

**Author:** Isaac Canals Montferrer

**Supervisors:** Dr. Lluïsa Vilageliu Arqués and Dr. Daniel Grinberg Vaisman

### **Article 1**

**Title:** Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles.

**Authors:** Isaac Canals, Siham Chafai Elalaoui, Mercè Pineda, Verónica Delgadillo, Marina Szlago, Imane Cherkaoui Jaouad, Abdelaziz Sefiani, Amparo Chabás, Maria Josep Coll, Daniel Grinberg, Lluïsa Vilageliu.

**Publication:** Clinical Genetics. 2011; 80(4): 367-374

**Impact Factor (2011 JCR Science Edition):** 3.128

**Contribution of the student:** Isaac Canals was involved in the design of the study and performed the whole experimental work described in the article. Patients were diagnosed at Institut de Bioquímica Clínica (Spain), at the Hospital Sant Joan de Déu (Spain), at the Institut National d'Hygiène (Morocco) or at the Laboratorio Dr. N. A. Chamoles (Argentina). The student also participated in the discussion of the results and the elaboration of the first draft, as well as the final version of the manuscript.

## Article 2

**Title:** Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations.

**Authors:** Liliana Matos, Isaac Canals, Larbi Dridi, Yoo Choi, Maria João Prata, Peter Jordan, Lourdes R. Desviat, Belén Pérez, Alexey V. Pshezhetsky, Daniel Grinberg, Sandra Alves, Lluïsa Vilageliu.

**Publication:** Orphanet Journal of Rare Diseases. 2014 (In press)

**Impact Factor (2013 JCR Science Edition):** 3.958

**Contribution of the student:** Isaac Canals and Liliana Matos contributed equally to the work. Isaac Canals was involved in the conception and design of the study, performed the experimental work of the U1 snRNA part together with Liliana Matos, the analysis and interpretation of the data, and participated in the drafting and revising of the manuscript.

## Article 3

**Title:** *EXTL2* and *EXTL3* inhibition with siRNAs as a promising substrate reduction therapy for Sanfilippo C syndrome.

**Authors:** Isaac Canals, Mónica Cozar, Lluïsa Vilageliu, Daniel Grinberg.

**Publication:** submitted to Journal of Human Genetics

**Contribution of the student:** Isaac Canals participated in the design of the study, performed the experimental work with the collaboration of Mónica Cozar, contributed to the analysis of the data, elaborated the first draft and was involved in the revision of the final version of the manuscript.

## Article 4

**Title:** High-order connectivity alterations in Sanfilippo C patient-specific neural networks.

**Authors:** Isaac Canals, Jordi Soriano, Javier G. Orlandi, Roger Torrent, Yvonne Richaud-Patin, Senda Jiménez-Delgado, Simone Merlin, Antonia Follenzi, Antonella Consiglio, Lluïsa Vilageliu, Daniel Grinberg, Ángel Raya.

**Publication:** Manuscript in preparation.

**Contribution of the student:** Isaac Canals participated in the design of the study and performed the experimental work regarding the generation and validation of the iPSC, and the differentiation and characterization of the derived neurons. Other authors contributed to specific experimental parts of the work: Jordi Soriano and Javier G. Orlandi (neuronal activity experiments), Roger Torrent (in vivo differentiation immunoanalysis), Senda Jiménez-Delgado (RT-PCR and promotor methylation analysis), Antonella Consiglio (lentiviral production), Simone Merlin and Antonia Follenzi (cloning of the *HGSNAT* cDNA in the lentiviral vector). The student was also involved in the interpretation of the results as well as in the drafting and revising of the manuscript.

Barcelona, 6 November 2014

Supervisors' agreement,

Dr. Lluïsa Vilageliu Arqués

Dr. Daniel Grinberg Vaisman



# **Article 1**



# Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel *HGSNAT* mutations and characterization of the mutant alleles

Isaac Canals, Siham Chafai Elalaoui, Mercè Pineda, Verónica Delgadillo, Marina Szlago, Imane Cherkaoui Jaouad, Abdelaziz Sefiani, Amparo Chabás, Maria Josep Coll, Daniel Grinberg, Lluïsa Vilageliu.

## Summary

The Sanfilippo syndrome type C [mucopolysaccharidosis IIIC (MPS IIIC)] is caused by mutations in the *HGSNAT* gene, encoding an enzyme involved in heparan sulphate degradation. We report the first molecular study on several Spanish Sanfilippo syndrome type C patients. Seven Spanish patients, one Argentinean and three Moroccan patients were analysed. All mutant alleles were identified and comprised nine distinct mutant alleles, seven of which were novel, including four missense mutations (p.A54V, p.L113P, p.G424V and p.L445P) and three splicing mutations due to two point mutations (c.633+1G>A and c.1378-1G>A) and an intronic deletion (c.821-31\_821-13del). Furthermore, we found a new single nucleotide polymorphism (SNP) (c.564-98T>C). The two most frequent changes were the previously described c.372-2A>G and c.234+1G>A mutations. All five splicing mutations were experimentally confirmed by studies at the RNA level, and a minigene experiment was carried out in one case for which no fibroblasts were available. Expression assays allowed us to show the pathogenic effect of the four novel missense mutations and to confirm that the already known c.710C>A (p.P237Q) is a non-pathogenic SNP. Haplotype analyses suggested that the two mutations (c.234+1G>A and c.372-2A>G) that were present in more than one patient have a common origin, including one (c.234+1G>A) that was found in Spanish and Moroccan patients.



## Short Report

# Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel *HGSNAT* mutations and characterization of the mutant alleles

Canals I, Elalaoui SC, Pineda M, Delgadillo V, Szlago M, Jaouad IC, Sefiani A, Chabás A, Coll MJ, Grinberg D, Vilageliu L. Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel *HGSNAT* mutations and characterization of the mutant alleles. Clin Genet 2010. © John Wiley & Sons A/S, 2010

The Sanfilippo syndrome type C [mucopolysaccharidosis IIIC (MPS IIIC)] is caused by mutations in the *HGSNAT* gene, encoding an enzyme involved in heparan sulphate degradation. We report the first molecular study on several Spanish Sanfilippo syndrome type C patients. Seven Spanish patients, one Argentinean and three Moroccan patients were analysed. All mutant alleles were identified and comprised nine distinct mutant alleles, seven of which were novel, including four missense mutations (p.A54V, p.L113P, p.G424V and p.L445P) and three splicing mutations due to two point mutations (c.633+1G>A and c.1378-1G>A) and an intronic deletion (c.821-31\_821-13del). Furthermore, we found a new single nucleotide polymorphism (SNP) (c.564-98T>C). The two most frequent changes were the previously described c.372-2A>G and c.234+1G>A mutations. All five splicing mutations were experimentally confirmed by studies at the RNA level, and a minigene experiment was carried out in one case for which no fibroblasts were available. Expression assays allowed us to show the pathogenic effect of the four novel missense mutations and to confirm that the already known c.710C>A (p.P237Q) is a non-pathogenic SNP. Haplotype analyses suggested that the two mutations (c.234+1G>A and c.372-2A>G) that were present in more than one patient have a common origin, including one (c.234+1G>A) that was found in Spanish and Moroccan patients.

### Conflict of interest

All authors declare no conflicts of interest.

**I Canals<sup>a,b,c</sup>, SC Elalaoui<sup>d</sup>, M Pineda<sup>b,e</sup>, V Delgadillo<sup>b,e</sup>, M Szlago<sup>f</sup>, IC Jaouad<sup>d</sup>, A Sefiani<sup>d,g</sup>, A Chabás<sup>b,h</sup>, MJ Coll<sup>b,h</sup>, D Grinberg<sup>a,b,c</sup> and L Vilageliu<sup>a,b,c</sup>**

<sup>a</sup>Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain, <sup>b</sup>CIBER de Enfermedades Raras (CIBERER), Barcelona, Spain, <sup>c</sup>Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain, <sup>d</sup>Département de Génétique Médicale, Institut National d'Hygiène, Rabat, Maroc, <sup>e</sup>Servei de Neurologia, Hospital Sant Joan de Déu, Barcelona, Spain, <sup>f</sup>Laboratorio Dr. N.A. Chamoles, Buenos Aires, Argentina, <sup>g</sup>Centre de Génomique Humaine, Université Mohamed V Souissi, Rabat, Maroc, and <sup>h</sup>Institut de Bioquímica Clínica, Servei de Bioquímica i Genètica Molecular, Hospital Clínic, Barcelona, Spain

Key words: haplotype analysis – heterologous expression – *HGSNAT* gene – mucopolysaccharidosis IIIC – mutations

Corresponding author: Daniel Grinberg, Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain.  
 Tel: +34 93 4035716;  
 fax: +34 93 4034420;  
 e-mail: dgrinberg@ub.edu

Received 10 June 2010, revised and accepted for publication 27 July 2010

The Sanfilippo syndrome or mucopolysaccharidosis III (MPS III) is a lysosomal storage disorder with an autosomal recessive inheritance pattern. The Sanfilippo syndrome has four subtypes

(A: OMIM 252900, B: OMIM 252920, C: OMIM 252930, D: OMIM 252940) due to mutations in four different genes that result in the inability to degrade the glycosaminoglycan heparan sulphate.

**Canals et al.**

Clinically, the four subtypes are similar, with severe central nervous system (CNS) degeneration accompanied by mild somatic manifestations. Subtypes A and B are more prevalent, while subtype C and especially subtype D are less common (1, 2). MPS IIIC accounts for 12.4% of the total number of patients in the Spanish MPS III population (A. Chabás, M. J. Coll, unpublished data).

The *HGSNAT* gene, which encodes acetyl-CoA:α-glucosaminide *N*-acetyltransferase (EC 2.3.1.78), was identified as being responsible for MPS IIIC by two independent groups in 2006 (3, 4). The gene, located at chromosome 8p11.1, contains 18 exons, generates a 1908-bp cDNA, and codes for a 635-amino acid protein. The enzyme catalyses acetylation of the terminal glucosamine residues of heparan sulphate prior to its hydrolysis by α-*N*-acetyl glucosaminidase (5).

To date, about 50 mutations have been described in the *HGSNAT* gene (6). Here, we report the mutation analysis of seven Spanish, one Argentinean and three Moroccan patients, all unrelated. To date, only one Spanish Sanfilippo C patient has been described (4). In contrast, several Moroccan patients have been reported (4, 7). Interestingly, all the Moroccan patients and the only Spanish patient described bore the c.234+1G>A mutation.

**Materials and methods**

## Patients and control samples

The origin of the patients is described in Table 1. Biochemical diagnoses were performed at the Institut de Bioquímica Clínica, Barcelona, Spain (patients 1–7 and 12), at the Département de Génétique Médicale – Institut National d'Hygiène, Rabat, Morocco (patients 8.1, 8.2 and 9) and at the Laboratorio Dr Chamoles, Buenos Aires, Argentina (patient 10). Some of the samples from Moroccan control individuals were kindly provided by Dr P. Moral (University of Barcelona). The rest of the control samples were obtained from anonymous donors from Morocco and Spain.

## DNA extraction and mutation identification

Genomic DNA was extracted from patients' leukocytes or cultured skin fibroblasts using standard protocols. The exons and their flanking regions of the *HGSNAT* gene were PCR-amplified using previously described primers and conditions (3, 4), with minor modifications. Purification of PCR products and sequencing were performed as described (8). Gene nucleotides were numbered according to the sequence RefSeq NM\_152419.2,

with +1 as the A of the ATG start codon. The ATG codon represents +1 for amino acid numbering, according to the protein sequence NP\_689632.2.

## Splicing mutation analyses: NMD studies

RNA was isolated from patients and control skin fibroblast cultures as previously described (8). Sequences of the primers used to amplify cDNA fragments for splicing analysis are available on demand. For mutation c.234+1G>A, the control used to compare the relative intensity of the bands was a 419-bp fragment of the *GAPDH* gene. For the other mutations, the internal control used was either the other allele or another splicing product (not bearing any premature stop codon). For non-sense-mediated mRNA decay (NMD) studies, fibroblasts were cultured in the absence or presence of 1 μg/ml of cycloheximide (CHX) for 6 h.

## Splicing mutation analyses: minigene constructions

Human genomic DNA was amplified from normal and mutant (c.821-31\_821-13del) *HGSNAT* exon 9 and flanking regions (196 bp of the 5' intron and 162 bp of the 3' intron), using the forward primer 5'-CCAGGCTAGTCTCGAACTCC-3' and the reverse primer 5'-GGCACTGGGTCAAGTATGAGA-3'. The products were cloned into a previously developed minigene plasmid, pGLB1 (9).

The splicing assay was performed by transfecting 1 μg of each minigene plasmid with 5 μl of Lipofectamine™ 2000 Transfection Reagent (Invitrogen, Carlsbad, CA) into HeLa cells. Total RNA was isolated from cultured HeLa cells 24 h after transfection. To specifically amplify those transcripts expressed from the plasmid, a specific PCR was performed as described (9).

## Cloning of HGSNAT cDNA, site-directed mutagenesis and expression assays

The cDNA was obtained in two overlapping fragments of 348 (ordered from GeneScript Corporation, Piscataway, NJ, USA) and 1881 bp that were ligated and subcloned into a pUC19 vector. Mutagenesis and expression assays in COS-7 cells were performed as described (8). Sequences of the primers used are available on demand. Forty-eight hours after transfection, the cells were harvested and the acetyl-CoA:α-glucosaminide *N*-acetyltransferase activity was analysed using 4-methylumbelliferyl-β-D-glucosaminide as the substrate (Moscerdam Substrates, Oegstgeest, The Netherlands) in a Modulus™ Microplate Multimode

## Molecular analysis of Sanfilippo syndrome type C in Spain

Table 1. Genotypes and main clinical features

Patient	Sex	Origin	Genotype <sup>a</sup>	Age of onset	Age of diagnosis	Current age	Clinical signs	Consanguinity
SFC1	M	Spain	c.[234+1G>A]+[234+1G>A]	3 years 7 months	5 years	21 years	Motor deterioration, speech loss at 4 years, moderate bilateral hypoaacusia, epilepsy, coarse facial features, kyphoscoliosis, sleeping problems at 8 years, dysphagia to liquids, joint stiffness	Yes
SFC2	F	Spain	c.[372-2A>G]+[372-2A>G]	NA	9 years	15 years	Coarse facial features, hepatosplenomegaly, mental retardation, abnormal behaviour, skeletal involvement	No
SFC3	M	Morocco	c.[234+1G>A]+[234+1G>A]	3 years 6 months	7 years	13 years	Motor deterioration, speech loss at 7 years, bilateral hypoaacusia, epilepsy, coarse facial features, kyphoscoliosis, sleeping problems at 7 years, dysphagia to liquids, joint stiffness	Yes
SFC4	F	Spain	c.[1378-1G>A]+[1378-1G>A]	NA	NA	NA	NA	Unknown
SFC5	M	Spain	c.[372-2A>G]+[1271G>T]	4 years	7 years	14 years	Motor clumsiness and hyperactivity at 4 years, mental retardation, coarse facial features, hypoaacusia, dysphagia	No
SFC6	F	Spain	c.[633+1G>A]+[1334T>C]	3 years 7 months	4 years 7 months	16 years	Mental retardation, speech loss, obesity, coarse facial features, hirsutism	No
SFC7	F	Spain	c.[372-2A>G]+[372-2A>G]	NA	9 years	14 years	Severe mental retardation, coarse facial features	Unknown
SFC8.1 <sup>b</sup>	F	Morocco	c.[338T>C]+[338T>C]	3 years	8 years	9 years	Psychomotor development and language delay with limited speech, hyperactivity and aggressive behaviour at 3 years, coarse facial features, hypertrichosis, mild mental retardation, premature puberty at 7 years, delayed bone age in the wrists	Yes
SFC8.2 <sup>b</sup>	F	Morocco	c.[338T>C]+[338T>C]	4 years	7 years	8 years	Same features as her sister, premature puberty at 4 years	Yes
SFC9	F	Morocco	c.[234+1G>A]+[234+1G>A]	6 years	8 years	9 years	Delayed psychomotor development, mental retardation, hyperactivity, coarse facies, hypertrichosis, moderate mitral regurgitation, bilateral hypoaacusia, beaked vertebrae	Yes
SFC10	F	Argentina	c.[161C>T]+[821-31 821-13del]	3 years	12 years 4 months	13 years 7 months	Psychomotor development and language delay with speech loss at 3 years, motor clumsiness, hyperactivity, dysphagia, synophrys	No
SFC12	F	Spain	c.[372-2A>G]+[372-2A>G]	5 years	6 years 9 months	7 years	Motor clumsiness, learning difficulties, mild bilateral hypoaacusia, coarse facial features	No

NA, not available.

<sup>a</sup>Nucleotides were numbered according to cDNA sequence GenBank entry NM 152419.2, where +1 corresponds to the A of the ATG translation initiation codon.<sup>b</sup>Patients 8.1 and 8.2 were sisters.

**Canals et al.**

Reader (Turner Biosystems, Sunnyvale, CA), following the manufacturer's instructions.

**Results and discussion**

## Mutation analysis

Sanfilippo C is a rare disease and the gene responsible for it was only identified 4 years ago (3, 4). Thus, few studies on Sanfilippo C mutations have been published. This is the first report on Spanish patients, and it also includes three unrelated patients from Morocco and one from Argentina. We analysed samples from 11 patients and identified all the mutant alleles. In total, nine different mutations were found (Table 2), seven of which are novel. This figure is significant, given that only 50 mutations had been described prior to this study (6). Genotypes and main clinical features of the patients are shown in Table 1.

The c.372-2A>G mutation, the most prevalent in our series of patients, was previously described, only once, in a Portuguese patient (10). We found this mutation in four Spanish patients (one heterozygous and three homozygous) with a frequency of 50% (7/14 alleles) for the Spanish patients of this series.

Mutation c.234+1G>A, a prevalent mutation in Moroccan patients, was first described in one Spanish and two Moroccan patients, linked to change c.710C>A (p.P237Q), as a double mutation (4). This allelic association was confirmed in other studies, in Moroccan patients, but was not observed in patients from other origins (7, 11). We found this mutation in homozygosity in one Spanish and two Moroccan patients, also linked to p.P237Q.

Each of the seven novel mutations was present in only one patient (either heterozygous or homozygous). Four of them were missense mutations

(c.161C>T, c.338T>C, c.1271G>T and c.1334T>C) and three splicing mutations due to two point mutations (c.633+1G>A and c.1378-1G>A) and an intronic deletion (c.821-31\_821-13del). All of the new mutations were confirmed by restriction analysis. None of the new missense mutations were found in 100 control alleles.

## Polymorphism and haplotype analyses

We analysed 14 previously reported single nucleotide polymorphisms (SNPs) localized in the *HGSNAT* gene in patients bearing the same mutation (either c.234+1G>A or c.372-2A>G) to assess whether the haplotypes were consistent with a single origin for each of the mutations (Table 3). In addition, we analysed a novel polymorphism (c.564-98T>C, in intron 5) and the missense change c.710C>A (p.P237Q), which was controversially reported as a polymorphic variant (12) or as a pathogenic lesion associated with the c.234+1G>A mutation (4).

Analysis of samples from patients 1, 3 and 9, for mutation c.234+1G>A, showed that they shared the same haplotype, suggesting a single origin for this mutation in Spanish and Moroccan patients.

All the homozygous patients for mutation c.372-2A>G (patients 2, 7 and 12) presented a single haplotype associated with the mutation for all the markers. In the case of patient 5, one of the alleles of the only heterozygous SNP was consistent with the common haplotype (but phases could not be established). Again, these results suggest a single origin for this mutation in Spanish patients.

We identified a novel SNP in intron 5 (c.564-98T>C) in patients bearing the c.234+1G>A mutation together with the missense change p.P237Q. We analysed 114 Spanish control alleles and the c.564-98T>C change was not found.

Table 2. Mutations found in this study

Nucleotide change <sup>a</sup>	Effect on protein <sup>b</sup>	Location	Alleles	References
c.234+1G>A	p.D40VfsX19	Intron 2	6	(4)
c.372-2A>G	p.[L125_R128del, R124SfsX27]	Intron 3	7	(10)
c.633+1G>A	p.S188RfsX65	Intron 6	1	This report
c.1378-1G>A	p.[V460YfsX22, V460_Q488del]	Intron 13	2	This report
c.821-31_821-13del	p.G274GfsX14	Intron 8	1	This report
c.161C>T	p.A54V	Exon 2	1	This report
c.338T>C	p.L113P	Exon 3	2	This report
c.1271G>T	p.G424V <sup>c</sup>	Exon 13	1	This report
c.1334T>C	p.L445P	Exon 13	1	This report

<sup>a</sup>Nucleotides were numbered according to cDNA sequence GenBank entry NM\_152419.2, where +1 corresponds to the A of the ATG translation initiation codon.

<sup>b</sup>The ATG codon represents +1 for amino acid numbering, according to the protein sequence NP\_689632.2.

<sup>c</sup>Another change in the same residue, p.G424S (c.1270G>A), was previously described (12).

## Molecular analysis of Sanfilippo syndrome type C in Spain

Table 3. Haplotypes observed in patients bearing frequent mutations

Mutation c.234+1G>A																
	rs13256451 (C/A)	rs62517610 (C/T)	rs10958738 (C/T)	rs72647300 (A/G)	rs72647302 (G/T)	rs73675460 (C/T)	rs34962945 (-/C)	rs11348027 (C/-)	rs17603428 (G/A)	rs73569592 (A/C)	rs3208566 (C/G)	rs1126058 (T/C)	rs73675469 (G/A)	rs56304352 (G/T)	c.564-98 (T/C) <sup>a</sup>	P237Q (C/A)
SFC1	C	C	C	A	G	C	-	C	G	A	C	C	G	G	C	A
SFC3	C	C	C	A	G	C	-	C	G	A	C	C	G	G	C	A
SFC9	C	C	C	A	G	C	-	C	G	A	C	C	G	G	C	A
Mutation c.372-2A>G																
	rs13256451 (C/A)	rs62517610 (C/T)	rs10958738 (C/T)	rs72647300 (A/G)	rs72647302 (G/T)	rs73675460 (C/T)	rs34962945 (-/C)	rs11348027 (C/-)	rs17603428 (G/A)	rs73569592 (A/C)	rs3208566 (C/G)	rs1126058 (T/C)	rs73675469 (G/A)	rs56304352 (G/T)	c.564-98 (T/C) <sup>a</sup>	P237Q (C/A)
SFC2	C	C	T	A	G	C	-	C	G	A	C	C	G	G	T	C
SFC5 <sup>b</sup>	C	C	T	A	G	C	-	C	G/A	A	C	C	G	G	T	C
SFC7	C	C	T	A	G	C	-	C	G	A	C	C	G	G	T	C
SFC12	C	C	T	A	G	C	-	C	G	A	C	C	G	G	T	C

<sup>a</sup>Polymorphism not previously described.<sup>b</sup>Heterozygous patient (the rest are all homozygous).

Studies at the RNA level did not show any effect on the splicing process (data not shown).

## Splicing mutation analyses: NMD studies

We found four mutations that affect the conserved splice sites. These mutations were c.234+1G>A, c.372-2A>G, c.633+1G>A and c.1378-1G>A. RNA from patient fibroblasts was analysed to test how these mutations affect the splicing process. Transcripts were sequenced to confirm the results. As these mutations would probably generate mRNAs with premature termination codons (PTCs), which might be degraded by the NMD mechanism, fibroblasts were cultured in the absence or presence of CHX, an inhibitor of the NMD process.

Only one transcript was detected for mutation c.234+1G>A, which corresponded to the skipping of exon 2 (Fig. 1a). A similar amount of mRNA was observed in the presence of CHX. This result suggests that no NMD-dependent degradation occurred.

For mutation c.372-2A>G, two transcripts were detected (Fig. 1b). The lower band was the result of an alternative splicing in which exon 4 was skipped, creating a frameshift and, consequently, a premature stop codon. The upper band corresponds

to a different alternative spliced product, which used a new acceptor site, 12 nucleotides downstream from the normal one. As seen in the control samples, this site is not used when the wild-type acceptor site is present. The use of this alternative site would generate an in-frame deletion of four amino acids in the protein. In the presence of CHX, the ratio of the lower/upper band increased, which indicates that the smaller transcript, but not the larger one, is affected by the NMD mechanism.

The *HGSNAT* gene from patient 6 (compound heterozygote for c.633+1G>A and c.1334T>C) gave rise to two transcripts (Fig. 1c): the normal mRNA (bearing the missense mutation) and a second one, which resulted from the skipping of exon 6. The lower band showed an increase in intensity in the presence of CHX, indicating degradation by the NMD mechanism, as expected from the frameshift and PTC generated by the exon skipping. Note that there is an upper band, which was shown to correspond to heteroduplex DNA.

Finally, for mutation c.1378-1G>A, found in patient SFC4 in homozygosis, two transcripts were detected (Fig. 1d). The lower band corresponds to an alternative splicing in which exon 14 is skipped. As exon 14 contains 87 nucleotides, the reading frame is conserved and no premature stop codon appears. The protein coded by this transcript

Canals et al.

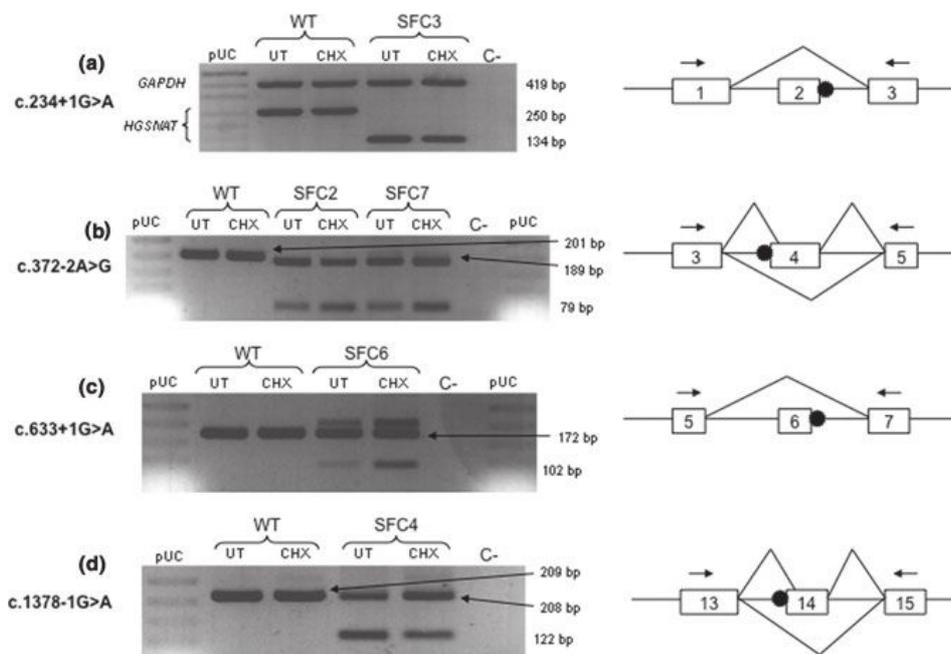


Fig. 1. Study of the effect of the splicing mutations at the RNA level. Reverse transcriptase polymerase chain reaction (RT-PCR) amplification of RNA extracted from some of the Sanfilippo C patients or wild-type (WT) fibroblasts, untreated (UT) or treated with cycloheximide (CHX) for 6 h, using specific primers for the *GAPDH* and *HGSNAT* genes (a) or only for *HGSNAT* (b–d) is shown. A scheme of part of the gene, indicating the position (black star) and the effect of each mutation, is shown on the right. In some cases, additional low mobility bands, corresponding to heteroduplex DNA, were observed. pUC: pUC mix marker 8 (Fermentas, Burlington, Ontario, Canada) was used as a molecular weight marker. Fragment sizes indicated on the right correspond to the bands shown by arrows.

would be 29 amino acids shorter. The upper band corresponds to an alternative splicing in which a new splicing acceptor site is used. This new splicing acceptor site was generated by the mutation. The normal acceptor site, AG, was changed to AA. However, the new A forms a new acceptor site with the G at the first position of the exon (Fig. 2). This G, previously exonic, is now intronic. Thus, it is lost in the mature mRNA, generating a frameshift and a premature stop codon. A change in the ratio of the two transcripts was observed in the presence of CHX. The upper one, which uses the new acceptor site, became more abundant, indicating degradation by the NMD mechanism when untreated.

Splicing mutation analyses: minigene constructions

Apart from the missense p.A54V, the sequence analysis of the patient SFC10 identified only a 19-bp deletion in intron 8. Given the lack of cells from this patient, a minigene assay was performed to assess whether the change c.821-31\_821-13del affects the splicing process.

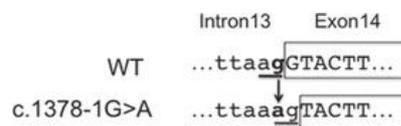


Fig. 2. Scheme of the intron 13 acceptor site in the wild-type (WT) allele and in the one bearing mutation c.1378-1G>A. Exonic sequences are boxed. The conserved AG dinucleotides are underlined. The change is indicated by an arrow and depicted in bold.

The *HGSNAT* exon 9 and intronic flanking regions, from the wild-type or the mutant allele, were cloned into the pGLB1 plasmid (Fig. 3a). Figure 3b shows the results of this minigene assay. The expected bands of 154 (1) and 301 bp (2) were obtained for the control plasmids. Different results were obtained when plasmids pGLB1-*HGSNAT*wt (3) and pGLB1-*HGSNAT*del (4) were assayed. A band of 332 bp corresponding to a mature transcript that includes the *HGSNAT* exon 9 was detected in pGLB1-*HGSNAT*wt (3). However, a 301-bp band was obtained from pGLB1-*HGSNAT*del, corresponding to the non-inclusion

## Molecular analysis of Sanfilippo syndrome type C in Spain

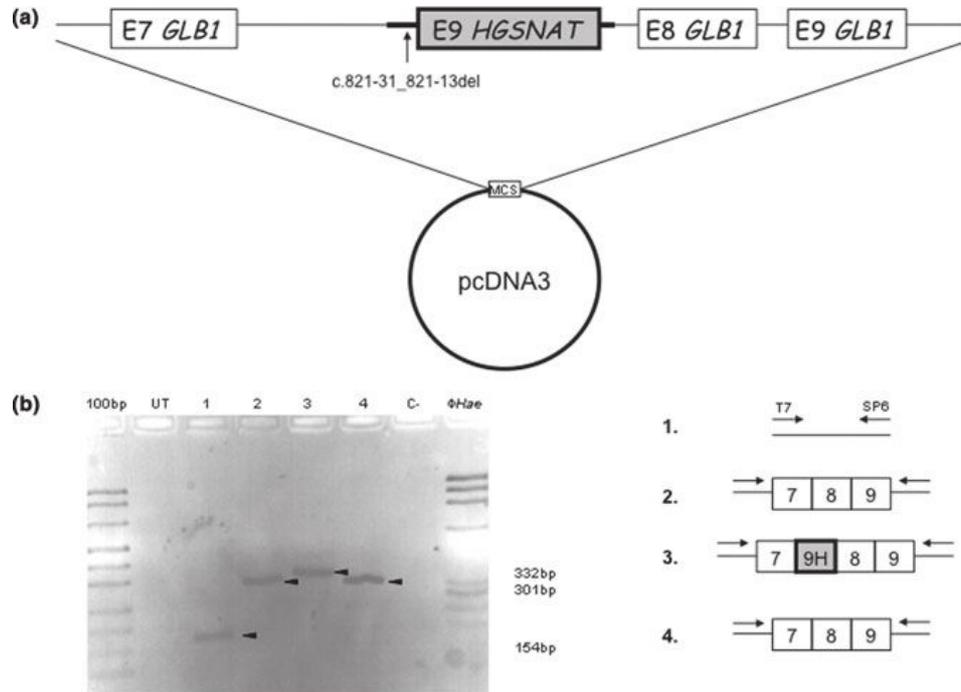


Fig. 3. Effect of the c.821-31\_821-13del mutation on splicing using a minigene assay. (a) Minigene construct in the pGLB1 vector. (b) The first three lanes show the results of reverse transcriptase polymerase chain reaction (RT-PCR) amplification using the T7 and SP6 primers in untransfected (UT) cells, cells transfected with the intact pcDNA3.1 plasmid (1) and cells transfected with the pGLB1 plasmid (2). Lane 3 shows the result of the RT-PCR amplification in cells transfected with the wild-type *HGSNAT* insert, whereas lane 4 shows the result for cells transfected with the *HGSNAT* deletion in intron 8. Lane C- is the negative control (no DNA) for the PCR amplification. Schemes of the amplifications corresponding to lanes 1–4 are depicted on the right.

of exon 9 (4). This shows that the mutation affects the splicing process and, thus, it can be considered a pathogenic mutation. The location and sequence of the deleted region are consistent with the hypothesis that the intron 8 branch point lies within the deletion.

## Expression assays

The functional effect of the missense mutations found in the Spanish, Moroccan and Argentinean patients described in this report (p.A54V, p.L113P, p.G424V and p.L445P), and that of two frequent previously described mutations (p.R344C and p.S518F), was evaluated by *in vitro* expression of the mutant cDNAs in COS-7 cells. The p.P237Q variant, described first as a mutation (4) and later as a polymorphism (12), was also analysed. Wild-type *HGSNAT* cDNA was also transfected in COS-7 cells as a positive control, and the empty pcDNA3.1 vector was used as a negative control. Western blots could not be performed because of inefficient binding of the only commercially available antibody.

The mean wild-type acetyl-CoA: $\alpha$ -glucosaminide *N*-acetyltransferase activity in COS-7 cells was  $95.5 \pm 25.7$  nmol/h/mg, whereas the average endogenous enzyme activity (COS-7 cells transfected with empty pcDNA3.1) was  $8.6 \pm 2.2$  nmol/h/mg. The residual enzymatic activities for each mutant protein, given as a percentage of that of the wild-type enzyme, are shown in Fig. 4. Alleles bearing the mutations showed practically no enzyme activity (ranging from 0% to 1.19%), showing the pathogenic effect of these changes.

Previous studies showed normal enzyme activity for the p.P237Q variant (12), raising questions regarding its initial description as a pathogenic mutation (4). We confirmed these results, observing 92.23% of the activity of the wild-type enzyme. Moreover, we found a novel rare polymorphism in intron 5 (c.564-98T>C), always linked to the p.P237Q variant. Studies at the RNA level did not show any effect on the splicing process for c.564-98T>C. The haplotype analyses of the patients with the c.234+1G>A mutation suggest a single origin for this change in Moroccan and Spanish patients.

## Canals et al.

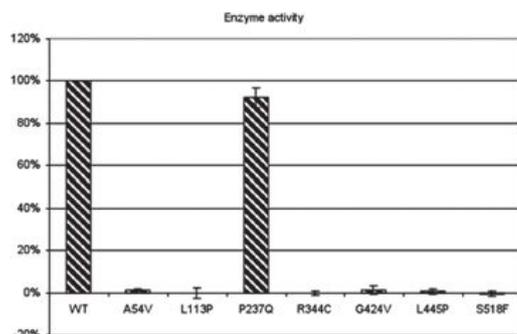


Fig. 4. Residual acetyl-CoA:α-glucosaminide N-acetyltransferase activity of the proteins from wild-type (WT) and mutant alleles expressed in COS-7 cells, given as the percentage of WT activity. The data are shown as the mean ± SD of three different experiments, each performed in triplicate. Error bars correspond to the standard deviation.

## Conclusion

We identified and characterized the molecular defects in 11 MPS IIIC patients, involving seven novel mutations and a new polymorphism. Expression studies confirmed the pathogenicity of the missense mutation and that p.P237Q is a polymorphism. The effect of splicing mutations was analysed at the RNA level on patients' fibroblast or using a minigene approach. Haplotype analyses suggested a single origin for mutation c.234+1G>A, found in Spanish and Moroccan patients, and for mutation c.372-2A>G, found in four Spanish patients.

## Acknowledgements

We thank the 'Federación de asociaciones de Mucopolisacaridos y síndromes relacionados' and, in particular, its president, Jordi Cruz, for their permanent support and encouragement. We also thank Drs Sierra and Parrilla (Jaén), del Castillo, Rodríguez Barrionuevo and Delgado (Málaga), and Eiris (Santiago de Compostela) for sending samples. The authors are also grateful to R. Rycroft for revising the English, to Dr L. Gort for technical help and to Dr P. Moral for providing control samples. The CIBER de Enfermedades Raras is an initiative of the ISCIII. This work was supported by grants from

the Spanish Ministerio de Educación y Ciencia (SAF2006-12276, SAF2009-11289) and the Generalitat de Catalunya (SGR2005-00848, 2009 SGR 971).

## References

- Neufeld EF, Muenzer J. The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The metabolic and molecular bases of inherited disease, 8th edn, Vol. 3. New York: McGraw-Hill, 2001: 3421–3452.
- Valstar MJ, Ruijter GJ, Van Diggelen OP, Poorthuis BJ, Wijburg FA. Sanfilippo syndrome: a mini-review. *J Inher Metab Dis* 2008; 31: 240–252.
- Fan X, Zhang H, Zhang S et al. Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). *Am J Hum Genet* 2006; 79: 738–744.
- Hrebicek M, Mrazova L, Seyrantepe V et al. Mutations in TMEM76\* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Am J Hum Genet* 2006; 79: 807–819.
- Klein U, Kresse H, Von Figura K. Sanfilippo syndrome type C: deficiency of acetyl-CoA:α-glucosaminide N-acetyltransferase in skin fibroblasts. *Proc Natl Acad Sci U S A* 1978; 75: 5185–5189.
- Feldhammer M, Durand S, Mrazova L et al. Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (*HGSNAT*) gene. *Hum Mutat* 2009; 30: 918–925.
- Ruijter GJ, Valstar MJ, Van De Kamp JM et al. Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. *Mol Genet Metab* 2008; 93: 104–111.
- Rodríguez-Pascau L, Gort L, Schuchman EH, Vilageliu L, Grinberg D, Chabas A. Identification and characterization of SMPD1 mutations causing Niemann-Pick types A and B in Spanish patients. *Hum Mutat* 2009; 30: 1117–1122.
- Díaz-Font A, Santamaria R, Cozar M et al. Clinical and mutational characterization of three patients with multiple sulfatase deficiency: report of a new splicing mutation. *Mol Genet Metab* 2005; 86: 206–211.
- Coutinho MF, Lacerda L, Prata MJ et al. Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the *HGSNAT* gene. *Clin Genet* 2008; 74: 194–195.
- Fedele AO, Filocamo M, Di Rocco M et al. Mutational analysis of the *HGSNAT* gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Mutation in brief* #959. Online. *Hum Mutat* 2007; 28: 523.
- Feldhammer M, Durand S, Pshzhetsky AV. Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C. *PLoS One* 2009; 4: e7434.

## **Article 2**



# Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations

Liliana Matos, Isaac Canals, Larbi Dridi, Yoo Choi, Maria João Prata, Peter Jordan, Lourdes R. Desviat, Belén Pérez, Alexey V. Pshezhetsky, Daniel Grinberg, Sandra Alves, Lluïsa Vilageliu.

## Summary

Mutations affecting RNA splicing represent more than 20% of the mutant alleles in Sanfilippo syndrome type C, a rare lysosomal storage disorder that causes severe neurodegeneration. Many of these mutations are localized in the conserved donor or acceptor splice sites, while few are found in the nearby nucleotides. In this study we tested several therapeutic approaches specifically designed for different splicing mutations depending on how the mutations affect mRNA processing. For three mutations that affect the donor site (c.234+1G>A, c.633+1G>A and c.1542+4dupA), different modified U1 snRNAs recognizing the mutated donor sites, have been developed in an attempt to rescue the normal splicing process. For another mutation that affects an acceptor splice site (c.372-2A>G) and gives rise to a protein lacking four amino acids, a competitive inhibitor of the HGSNAT protein, glucosamine, was tested as a pharmacological chaperone to correct the aberrant folding and to restore the normal trafficking of the protein to the lysosome. Partial correction of c.234+1G>A mutation was achieved with a modified U1 snRNA that completely matches the splice donor site suggesting that these molecules may have a therapeutic potential for some splicing mutations. Furthermore, the importance of the splice site sequence context is highlighted as a key factor in the success of this type of therapy. Additionally, glucosamine treatment resulted in an increase in the enzymatic activity, indicating a partial recovery of the correct folding. In conclusion, we have assayed two therapeutic strategies for different splicing mutations with promising results for the future applications.

Authors: Liliana Matos, Isaac Canals, Larbi Dridi, Yoo Choi, Maria J Prata, Peter Jordan, Lourdes R Desviat, Belén Pérez, Alexey V Pshezhetsky, Daniel Grinberg, Sandra Alves and Lluïsa Vilageliu

Title : Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations

Journal: Orphanet Journal of Rare Diseases

MS: 9291647101395415

Dear Mr Canals,

Peer review of your manuscript (above) is now complete and we are delighted to accept the manuscript for publication in Orphanet Journal of Rare Diseases.

Before publication, our production team needs to check the format of your manuscript, to ensure that it conforms to the standards of the journal. They will get in touch with you shortly to request any necessary changes or to confirm that none are needed.

If you have any problems or questions regarding your manuscript, please do get in touch.

Best wishes,

Ms Jonalyn Trampe  
on behalf of Dr Ségolène Aymé

e-mail: [ojrd@biomedcentral.com](mailto:ojrd@biomedcentral.com)

Web: <http://www.ojrd.com/>

# Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations

Liliana Matos<sup>1,2,†</sup>

Email: liliana.matos@insa.min-saude.pt

Isaac Canals<sup>3,4,5†</sup>

Email: icanals@ub.edu

Larbi Dridi<sup>6</sup>

Email: larbidridi.stejustine@gmail.com

Yoo Choi<sup>6</sup>

Email: yoochoi0730@gmail.com

Maria João Prata<sup>2,7</sup>

Email: mprata@ipatimup.pt

Peter Jordan<sup>8</sup>

Email: peter.jordan@insa.min-saude.pt

Lourdes R Desviat<sup>4,9</sup>

Email: lruiz@cbm.csic.es

Belén Pérez<sup>4,9</sup>

Email: bperez@cbm.uam.es

Alexey V Pshezhetsky<sup>6,10</sup>

Email: alexei.pchejetski@umontreal.ca

Daniel Grinberg<sup>3,4,5,\*</sup>

Email: dgrinberg@ub.edu

Sandra Alves<sup>1,†</sup>

Email: sandra.alves@insa.min-saude.pt

Lluïsa Vilageliu<sup>3,4,5,†</sup>

Email: lvilageliu@ub.edu

<sup>1</sup> Department of Human Genetics, Research and Development Unit, INSA, Porto, Portugal

<sup>2</sup> Department of Biology, Faculty of Sciences, Porto, Portugal

<sup>3</sup> Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

<sup>4</sup> CIBER de Enfermedades Raras (CIBERER), Barcelona, Spain

<sup>5</sup> Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain

<sup>6</sup> Department of Medical Genetics, Sainte-Justine University Hospital Centre, University of Montreal, Montreal, Canada

<sup>7</sup> IPATIMUP, Porto, Portugal

<sup>8</sup> Department of Human Genetics, Research and Development Unit, INSA, Lisbon, Portugal

<sup>9</sup> Centro de Diagnóstico de Enfermedades Moleculares, Centro de Biología Molecular Severo Ochoa, UAM-CSIC, Universidad Autónoma de Madrid, Madrid, Spain

<sup>10</sup> Department of Anatomy and Cell Biology, Faculty of Medicine, McGill University, Montreal, Canada

\* Corresponding author. Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

† Equal contributors.

## Abstract

### Background

Mutations affecting RNA splicing represent more than 20% of the mutant alleles in Sanfilippo syndrome type C, a rare lysosomal storage disorder that causes severe neurodegeneration. Many of these mutations are localized in the conserved donor or acceptor splice sites, while few are found in the nearby nucleotides.

### Methods

In this study we tested several therapeutic approaches specifically designed for different splicing mutations depending on how the mutations affect mRNA processing. For three mutations that affect the donor site (c.234 + 1G > A, c.633 + 1G > A and c.1542 + 4dupA), different modified U1 snRNAs recognizing the mutated donor sites, have been developed in an attempt to rescue the normal splicing process. For another mutation that affects an acceptor splice site (c.372-2A > G) and gives rise to a protein lacking four amino acids, a competitive inhibitor of the HGSNAT protein, glucosamine, was tested as a pharmacological chaperone to correct the aberrant folding and to restore the normal trafficking of the protein to the lysosome.

### Results

Partial correction of c.234 + 1G > A mutation was achieved with a modified U1 snRNA that completely matches the splice donor site suggesting that these molecules may have a therapeutic potential for some splicing mutations. Furthermore, the importance of the splice site sequence context is highlighted as a key factor in the success of this type of therapy. Additionally, glucosamine treatment resulted in an increase in the enzymatic activity, indicating a partial recovery of the correct folding.

### Conclusions

We have assayed two therapeutic strategies for different splicing mutations with promising results for the future applications.

## Keywords

Splicing mutations, Modified U1 snRNAs, Glucosamine, Sanfilippo C syndrome, Lysosomal storage disorder

## Background

Sanfilippo syndrome or Mucopolysaccharidosis III (MPS III), is a group of autosomal recessive lysosomal storage disorders caused by mutations in genes encoding enzymes responsible for heparan sulfate degradation [1]. There are four types of the disease, depending on the gene affected. They all present similar clinical symptoms, including severe central nervous system degeneration accompanied by mild somatic manifestations [2]. Sanfilippo syndrome type C (MPS IIIC) is caused by mutations in the *HGSNAT* gene. This gene codes for the acetyl-CoA:  $\alpha$ -glucosaminide N-acetyltransferase (EC 2.3.1.78), a protein localized in the lysosomal membrane which catalyses the acetylation of the terminal glucosamine residues of heparan sulphate prior to their hydrolysis by  $\alpha$ -N-acetyl glucosaminidase [3]. The *HGSNAT* gene, identified by two independent groups in 2006 [4,5], is located at the chromosome 8 (8p11.1) and contains 18 exons. The cDNA encodes a polypeptide of either 635 or 663 amino acids, since there is a controversy concerning the real initiation codon [6,7]. To date, 64 mutations have been reported, 16 of which (25%) involve splicing alterations: 13 are described as splicing mutations and three as small deletions and duplications that affect the splicing process (HGMD®Professional Spring 2014.1 Release).

Splicing is an essential step for the expression of most of human genes, in which the 5' and 3' splice sites (ss), the branch point sequence and the polypyrimidine tract (both the last two within 50 nucleotides upstream of the 3' ss) play a fundamental role. These sites present sequence variability throughout the human genome. The splicing process is conducted by the spliceosome, which is formed of five small nuclear RNAs (snRNAs) and more than 200 different proteins (reviewed in [8]). The U1 snRNA, which presents a partially complementary sequence to the 5' ss, is essential for the recognition of the 5' ss consensus motif (CAG/GTRAGT, exon/intron, R = purine). The U1 snRNP is composed of a 164-nt long U1 snRNA and several protein factors. The 5' region of the U1 snRNA is involved in the recognition of the 5' ss, with the C8 nucleotide being the one that binds the first nucleotide (G) of the intron (reviewed in [9]). Application of modified U1 snRNAs to improve recognition of mutated 5' ss represents a new strategy for recovering the normal splicing process. They have been assayed as a therapeutic approach for different diseases and splicing mutations affecting different positions of the 5' ss with variation in the efficacy of the treatment [10-19]. Recently, assays correcting an exogenous injected construct have been performed using modified U1 snRNAs in mice as a treatment for severe human factor VII deficiency [20].

In some cases, alternative splicing caused by specific mutations can give rise to misfolded proteins, which may be prone to rapid intracellular degradation (reviewed in [21]). Molecular chaperones are proteins that act on the correct folding of other polypeptides in cells. Pharmacological and chemical chaperones are small compounds that can be used, in a similar way, to avoid the misfolding of mutant proteins. They are principally potent enzyme inhibitors which interact specifically with their active sites to restore the correct folding and to increase stability [22]. In the case of lysosomal storage disorders, once in the lysosome, the enzyme substrate replaces the chaperone, thereby completing restoration of enzyme activity [23]. Aminosugars and iminosugars are the most common pharmacological chaperones used in enzyme enhancement therapy (EET) for lysosomal disorders. EET have been reported for

several of these diseases including Fabry disease, GM1-gangliosidosis, Morquio B disease, Pompe disease, Gaucher disease, Krabbe disease, Niemann-Pick A/B and C diseases; as well as for many other types of disorders such as retinitis pigmentosa, cystic fibrosis, Parkinson's disease, Alzheimer's disease and cancer (reviewed in Ref. [21]). In the case of Sanfilippo syndrome type C, glucosamine, a competitive HGSNAT inhibitor, has been shown to increase residual enzyme activity in cultured skin fibroblasts from patients affected with a number of missense mutations [24].

In this work we focus on *HGSNAT* mutations that affect the splicing process. Three previously described splicing mutations [5,25,26] were the object of the study with different modified U1 snRNAs to improve recognition of the donor ss and enhance the correct splicing process. The studies were performed using cells transfected with minigene constructs bearing the specific mutation as well as cultured patients' skin fibroblasts. Mutations include the most frequent change in Moroccan patients (c.234 + 1G > A), a mutation found in Spanish patients (c.633 + 1G > A) and a mutation found in an English patient (c.1542 + 4dupA). Furthermore, EET approach, was tested for a splicing mutation c.372-2A > G. This mutation is the most prevalent in Spanish and Portuguese patients [25,27] and affects the acceptor site at the end of the 4th intron of the *HGSNAT* gene [27], thereby altering the splicing process. The use of a downstream alternative cryptic ss generates an mRNA with an in-frame deletion that codes for a protein with the loss of four amino acids (p.[L125\_R128del]). Here, we show the effect of this mutation, which reduces enzyme activity, and the recovery of a part of that activity through treatment with glucosamine as a chaperone.

Our results show that, depending on the context of the mutated donor site, modified U1 snRNAs can be a promising therapeutic tool. The use of glucosamine as a chaperone improved enzyme activity, suggesting a therapeutic effect of this compound for Sanfilippo C patients.

## Methods

### Mutation analysis of the *HGSNAT* gene

This study included five MPSIIIC patients: three previously described, two Spanish and one Moroccan [25]; and two recently diagnosed, one French and one Portuguese, carrying mutations already reported by us [25,26] (Table 1). Studies were approved by the authors' Institutional Ethics Committee and conducted under the Declaration of Helsinki. Patients were encoded to protect their confidentiality. Genetic analysis was performed using control and patients' fibroblast cell lines as the source of RNA and genomic DNA whenever necessary. Total RNA was extracted using High Pure RNA Isolation Kit (Roche, Basel, Switzerland) and converted into cDNA using High-Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, CA) following the manufacturers' instructions. The RT-PCR amplifications were performed using the primers described in Additional file 1: Table S1 and information regarding specific conditions for each cDNA amplification and sequence analysis is available upon request.

**Table 1 Genotype and Origin of MPSIIIC affected patients**

Patient	Origin	Allele 1*	Allele 2*	Reference
SFCP	Portugal	c.234 + 1G > A	c.234 + 1G > A	This study
SFC3	Morocco	c.234 + 1G > A	c.234 + 1G > A	[25]
SFC6	Spain	c.633 + 1G > A	c.1334 T > C	[25]
SFC7	Spain	c.372-2A > G	c.372-2A > G	[25]
SFC13	France	c.1542 + 4dupA	c.1150C > T	This study

\* Mutation nomenclature is based on cDNA sequence (NM\_152419.2), with nucleotide +1 corresponding to the A of the ATG translation initiation codon.

## Minigene cloning and U1 snRNA expression constructs

For the *in vitro* splicing approaches, wild type (WT) and mutant minigenes were constructed for each mutation under study. A gene fragment including exon 2 and the flanking intronic regions was amplified from the DNA of the c.234 + 1G > A patient's fibroblast cells using the primers described in Additional file 1: Table S1, and cloned into the TOPO vector (Life Technologies). The insert was excised with EcoRI, purified using Wizard® SV Gel and the PCR clean-up system (Promega, Madison, WI), and subsequently cloned into the pSPL3 vector (Exon Trapping System, Life Technologies; kindly provided by Dr. B. Andresen) using the Rapid Ligation Kit (Roche Applied Science, Mannheim, Germany). The clone containing the desired mutant insert in the correct orientation was identified by restriction enzyme analysis and DNA sequencing. The exon 6 (c.633 + 1G > A) and exon 15 (c.1542 + 4dupA), together with their intronic flanking sequences, were also cloned in the pSPL3 vector. These mutated minigenes were ordered from GenScript (Piscataway, NJ). In all cases, the WT minigenes were obtained by site-directed mutagenesis using the QuikChange II XL site-directed Kit (Agilent Technologies, Santa Clara, CA) following the manufacturer's instructions and nucleotide changes were confirmed by sequencing analysis. The primers used are listed in Additional file 1: Table S1.

To express WT U1 snRNA (U1-WT), we used the vector pG3U1, which includes the sequence coding for human U1 ([28]; kindly provided by Dr. F. Pagani). The different U1 vectors adapted to the donor ss of *HGSNAT* exon 2, exon 6 and exon 15 were obtained by site-directed mutagenesis using the QuikChange II XL site-directed Kit (Agilent Technologies). For each construct (U1 suppressor 1 to 9) the PCR reaction was performed with the specific primers shown in Additional file 1: Table S1. The desired mutations were confirmed by sequence analysis.

## Cell culture and U1 transfection experiments

To perform the splicing assays, COS-7 cells and fibroblasts were grown in Dulbecco's modified Eagle's medium (Sigma-Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum (Life Technologies) and 1% PenStrep (Life Technologies) at 37 °C with 5% CO<sub>2</sub>. For co-transfection experiments, COS-7 cells at 90% of confluence were transfected in 6 or 12-well plates with either 1 or 2 µg of each WT or mutant minigene and 1 to 4 µg of each different mutation adapted U1 snRNA vector, using Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. When required, the amount of DNA was adjusted with the pSPL3 empty vector. For splicing analysis of the endogenous *HGSNAT* transcripts, healthy control and patients' fibroblasts at 90% of confluence were transfected in 6-well plates with 1, 2.5 or 3.5 µg of each modified U1 snRNA vector using either Lipofectamine 2000 or LTX (Life Technologies) according to the manufacturer's instructions. To estimate transfection

efficiency, healthy control and patients' cells were transfected with a control plasmid encoding either GFP or RFP and fluorescent cells were monitored by microscopy.

### **RT-PCR transcript analysis after transfection of modified U1 snRNAs**

Cells were harvested 24 or 48 h after transfection. Total RNA was extracted using High Pure RNA Isolation Kit (Roche Applied Science) and converted into cDNA using High-Capacity cDNA Reverse Transcription Kit (Life Technologies). The RT-PCR splicing analysis for minigene transfections was performed using the typical pSPL3 primers SD6 and SA2 (Additional file 1: Table S1). For the endogenous experiments, RT-PCRs were performed using the following primers: HGSNAT-Exon 2 F / HGSNAT-Exon 3R for mutation c.234 + 1G > A, HGSNAT-Exon 5 F / HGSNAT-Exon 13R for c.633 + 1G > A/c.1334 T > C and HGSNAT-Exon 12 F / HGSNAT-Exon 16R for the c.1542 + 4dupA/c.1150C > T mutation. In the case of the c.234 + 1G > A mutation, the forward primer was designed to anneal in the middle of exon 2 to amplify only the transcripts in which the correct splicing process was recovered. For the other mutations, the last nucleotide of one of the primers corresponded to the point mutation of the other allele (but with the WT nucleotide) to favour that only the cDNA from the splice-mutation allele was amplified. The RT-PCR products were sequenced to confirm their identity. All the primers used are listed in Additional file 1: Table S1 and information regarding the amplification conditions is available upon request.

### **Expression of recombinant human mutant and wild type *HGSNAT* in COS-7 cells**

COS-7 cells cultured in Eagle's minimal essential medium supplemented with 10% fetal bovine serum to 70% confluence were transfected with the pcTAP-HGSNAT plasmid and mutant pcDNA-HGSNAT-L125\_R128del plasmid as previously described [24]. 48 h after transfection the cells were harvested and analysed for N-acetyltransferase enzymatic activity or by Western blot.

### **Glucosamine-mediated refolding of mutant HGSNAT**

Twenty-four h after transfection with pcTAP-HGSNAT or mutant pcDNA-HGSNAT-L125\_R128del plasmids, COS-7 cells were grown for 72 h in Eagle's minimal essential medium supplemented with 10% fetal bovine serum containing 10 mM glucosamine in 6-well plates. Then the cells were kept for another 24 h in the normal culture medium without glucosamine. The cells were harvested, lysed by freeze-thaw in 750  $\mu$ l of water and assayed for HGSNAT activity. Three independent experiments (each with 2 cell plates) were performed on 3 different occasions.

Confluent primary cultures of skin fibroblasts of the MPS IIIC patient homozygous for the c.372-2A > G mutation and healthy control fibroblasts (n = 5) were grown in 6-well plates for 72 h in Dulbecco's minimal essential medium supplemented with 10% fetal bovine serum containing 10 mM glucosamine. Then the cells were kept for another 24 h in the medium without glucosamine, harvested, lysed in 500  $\mu$ l of water and assayed for HGSNAT activity. Three independent experiments (each with 2 cell plates) were performed on 3 different occasions.

## Enzyme assay

HGSNAT N-acetyltransferase enzymatic activity was measured using the fluorogenic substrate 4-methylumbelliferyl  $\beta$ -D-glucosaminide (4MU- $\beta$ -GlcN; Moscerdam, Oegstgeest, Netherlands) as previously described [24]. The protein concentration was measured according to the method of Bradford using a commercially available reagent (BioRad, Hercules, CA).

## Statistical analysis

Statistical analysis of the data has been performed by two-tailed unpaired t-test using the Prism GraphPad software.

## Western blotting

Cell lysates from 3 independent transfections (20  $\mu$ g of total protein each) were analysed by Western blot as previously described [24] using rabbit polyclonal antibodies raised against human HGSNAT Q52-N156 peptide (Sigma-Aldrich HPA029578, dilution 1:5000, incubation overnight at 4 °C). Detection was performed with an anti-rabbit IgG antibodies-HRP conjugate (ref. 7074S, Cell Signalling, Beverly, MA), and the enhanced chemiluminescence reagent (ref. 32106, Thermo Scientific, Waltham, MA).

## Results

### Mutation analysis in MPS IIIC patients' fibroblasts

This study involved four splicing mutations found in five different patients: SFCP and SFC3 (both homozygous for c.234 + 1G > A), SFC6 (compound heterozygous for c.633 + 1G > A and the missense mutation c.1334 T > C; p.L445P), SFC13 (compound heterozygous for c.1542 + 4dupA and the nonsense mutation c.1150C > T; p.R384\*) and SFC7 (homozygous for c.372-2A > G) (see Table 1). For patients SFC3 (previously reported [25]) and SFCP (novel patient), the RT-PCR analysis of cDNA with primers in exon 1 and 3 revealed a fragment skipping exon 2 (252 bp) due to the presence of the c.234 + 1G > A mutation in homozygosity (Figure 1A). For patient SFC6 (previously described [25]) the cDNA analysis with primers in exons 5 and 7 showed two transcripts: one of a normal length of 172 bp arising from the missense mutation allele (c.1334 T > C; p.L445P); and a second one resulting from the skipping of exon 6 (102 bp), due to the splicing alteration c.633 + 1G > A in the other allele (Figure 1B). For a novel patient, SFC13, bearing the intron 15 mutation c.1542 + 4dupA in compound heterozygosity with the c.1150C > T nonsense change in exon 11, the RT-PCR analysis with primers in exons 11 and 16 revealed two aberrant transcripts. The first one showed the skipping of exon 15 (395 bp), and in the second, exon 15 was extended for the first 5 nucleotides of intron 15 (478 bp) due to the presence of a cryptic ss in the beginning of this intron (Figure 1C). For the fifth patient, SFC7 (previously reported [25]), who is homozygous for the c.372-2A > G mutation, the transcript analysis using primers in exons 3 and 5 revealed the presence of two transcripts: one of lower size (79 bp) in which exon 4 was skipped; and the other of higher size (189 bp) corresponding to a fragment with the deletion of the first 12 nucleotides of exon 4 due to the use of a cryptic 3' ss localized downstream of the constitutive one (Figure 1D), which gives rise to the p.[L125\_R128del] mutant protein.

---

Figure 1 Analysis of alternative *HGSNAT* transcripts from the mRNA of five MPS IIIC patients by RT-PCR. RT-PCR amplification of mutations c.234 + 1G > A (A), c.633 + 1G >

A (**B**), c.1542 + 4dupA (**C**), and c.372-2A > G (**D**), and of the WT allele, using the primers indicated in the schemes by arrows. The different splicing patterns are depicted in the schemes (right), in which the mutations are indicated by black circles. M: marker, C-: negative control, SFCP, -3, -6, -13, -7: patients. Boxes in the schemes correspond to exons.

### Development of splicing therapy approaches for *HGSNAT* gene mutations: minigene assays with modified U1 snRNA vectors

To reproduce the splicing defects in a cellular model that could be used to test U1 snRNA overexpression as a therapeutic strategy we constructed several mutant minigenes in the pSPL3 vector and expressed them in COS-7 cells. Each of the minigenes bore either one of the three mutations that affect the donor ss or the WT allele. The post-transfection cDNA analysis and sequencing revealed the skipping of exon 2 and exon 6 in the cases of the c.234 + 1G > A and c.633 + 1G > A mutant minigenes, respectively (Figure 2C,F). For the c.1542 + 4dupA minigene, the cDNA showed a band that corresponded to the skipping of exon 15 and another band that corresponded to the inclusion of this exon plus the first five nucleotides of intron 15 (Figure 2I). Meanwhile, the constructs with the WT sequences (relative to each mutation) revealed a single transcript for each of the three cases with the inclusion of exon 2, exon 6 or exon 15, respectively. These results show that the minigene-derived splicing patterns closely resemble the patterns observed in control and patients' cDNAs obtained from fibroblasts. Thus, the minigenes are reliable tools to test and optimize the overexpression of modified U1 snRNAs in our attempts to correct the splicing defect.

**Figure 2 Minigene therapeutic approaches with U1 adaptations to correct the pathogenic effect of different donor ss mutations on the *HGSNAT* gene.** (A), (D) and (G) Schematic illustrations of base pairing between the wild type U1 (U1-WT) and the 5' ss of wild type and mutant exons 2, 6 and 15 of the *HGSNAT* gene. The position of each mutation in the 5' ss is marked in grey and it is in italics. The different U1 snRNAs used for each mutated 5' ss of *HGSNAT* (designated as U1-sup, for suppressor) are also shown. The U1 sequence changes are illustrated in bold. (B), (C), (E), (F), (H) and (I) RT-PCR analysis of COS-7 cells in which splicing reporter minigenes for each mutation were co-transfected with the different modified U1 constructs as appropriate. Neither the U1-WT nor the various adapted U1s showed an effect on the splicing of exon 2, 6 or 15 wild type minigenes (B, E and H). In the case of the mutant minigenes, some of the U1 co-transfections [U1-sup2, 3 and 4 (C); U1-sup6 (F); U1-sup7, 8 and 9 (I)] changed the splicing pattern, not to the correct one, but mainly towards an alternative pattern using a "gt" 5'ss immediately downstream of the constitutive one. Importantly, the wild type minigene for exon 2 (Mini WT ex2) shows an extra lower band due to the amount of empty vector added to adjust the total quantity of DNA. Sequencing results for all RT-PCR products are illustrated by schematic drawings. M: molecular weight marker; NT: non-treated cells; C-: negative control.

To evaluate the efficiency of different U1s for correcting missplicing of the *HGSNAT* exons 2, 6 and 15, several constructs, with different degrees of complementarity to each mutated donor ss, were generated (Figure 2A,D,G). For each specific case, the WT and mutant splicing reporter constructs were co-transfected with the different U1 snRNAs into COS-7 cells for 24 and 48 h. RT-PCR analysis was then performed and showed that U1-WT and the different U1 snRNA modifications do not interfere with the splicing pattern of any WT minigene (Figure 2B,E,H).

In the case of the mutant minigenes, different splicing patterns were observed after overexpression of the different U1 snRNAs. In the case of the mutant c.234 + 1G > A

minigene, the cDNA analysis revealed no change in splicing after co-expression with the U1-WT and U1-sup1 (+1 T) constructs. With the remaining modified U1s [U1-sup2 (-1G + 1T); U1-sup3 (-1G + 4A); U1-sup4 (-1G + 1T + 4A)], a band of the size of 376 bp, expected for the normal splicing, was observed. However, sequence analysis showed that the fragment included exon 2 and the first four base pairs of intron 2 (ATAT), due to the use of an alternative donor site with the canonical “gt” in positions +5 and +6. A very faint band corresponding to exon 2 skipping was still observed when U1-sup2 and U1-sup3 were overexpressed (Figure 2C).

For the mutant c.633 + 1G > A minigene, the splicing pattern was not altered after overexpression of U1-WT and U1-sup5 (+1 T), while an apparently normal band was detected with the U1 matching all the nucleotides of the mutated donor ss [U1-sup6 (-1A-2G-3A + 1 T + 7A + 8 T)]. Sequencing of this band demonstrated aberrant splicing which includes, apart from exon 6, the first four nucleotides (ATAA) of intron 6. Again, this was due to the presence of an alternative donor site with the canonical “gt” in positions +5 and +6. A band of low molecular weight, which corresponded to the skipping of exon 6 was also detected upon U1-sup6 treatment (Figure 2F).

In the c.1542 + 4dupA mutant minigene, the co-transfection of the totally complementary U1 [U1-sup9 (-1A-2A + 5 T + 6C + 7A)] induced the appearance of a single band of the predicted normal size. Sequence analysis revealed an abnormal fragment including exon 15 and the first five base pairs (GTAAA) of intron 15. As in the two previous cases, this was due to the presence of a “gt” in positions +6 and +7 (the duplication of an A in this mutant allele moved the “gt” from the original +5 and +6 positions). This band was also observed in the untreated mutant minigene. The overexpression of U1-sup7 (-1A-2A) produced an increase in intensity of the band that corresponded to the use of the alternative site; while no effect of U1-sup8 (+5 T + 6C + 7A) or U1-WT was detected (Figure 2I).

### **Development of splicing therapy approaches for *HGSNAT* gene mutations: treatment of patients' cells with modified U1 snRNA vectors**

Despite the data of minigene assays showing that the three ss defects were not corrected by the expression of the different modified U1 vectors, we explored the feasibility of this approach to correct the endogenously misspliced *HGSNAT* transcripts through transfection of the same U1 variants in patients' derived cell lines. The spliced transcripts obtained after transfection of each fibroblast cell line were analysed by RT-PCR (Figure 3A-F).

---

**Figure 3 Analysis by RT-PCR of the endogenous splicing pattern of control and patients SFCP, SFC3, SFC6 and SFC13 derived fibroblasts after transfection with different U1 isoforms.** (A), (C) and (E) The constitutive splicing of exons 2, 6 and 15 of the *HGSNAT* gene was not altered in control fibroblasts after overexpression of U1-WT or any of the modified U1 constructs. (B) In the patients SFCP and SFC3, bearing the homozygous mutation c.234 + 1G > A, only the fully adapted U1 (2.5 µg of U1-sup4) resulted in partial correction of exon 2 skipping. The same result was obtained with 1 or 3.5 µg of the U1-sup4 construct (data not shown). (D) and (F) For patients SFC6 and SFC13, bearing genotypes c.633 + 1G > A / c.1334 T > C and c.1542 + 4dupA / c.1150C > T, respectively, the transfection of 2.5 or 3.5 µg of U1-WT or any generated U1 suppressor vector did not produced any change in the endogenous aberrant splicing pattern. Sequencing results for all RT-PCR products are illustrated by schematic drawings. M: molecular weight marker; NT: non-treated cells; C-: negative control; RFP: red fluorescent protein.

---

Interestingly, for patients SFCP and SFC3, homozygous for the c.234 + 1G > A mutation, treatment with different quantities of fully adapted U1 [U1-sup4 (-1G + 1T + 4A)] resulted in a partial recovery from the splicing defect. Sequence analysis revealed two different sequences: one with the normal splicing; and the other which included the first four base pairs of intron 2 (Figure 3B and Additional file 2: Figure S1), as detected in the minigene approaches in COS-7 cells (Figure 2C). To estimate the levels of correct splicing, the obtained PCR product was cloned in a pUC19 vector and amounts of around 50% were detected for the correct sequence (10 out of 22 clones with the correct spliced fragment). Due to this partial correction obtained, an experiment to measure the enzymatic activity in patient's cells after U1-sup4 transfection was carried out. However, no improvement in enzyme activity was observed. The values obtained for both patients, without treatment and using 2.5 µg U1sup4, were negligible and far below the normal range of controls (1.95-13.4 nmol/17 h/mg). The expression of the remaining U1 modifications had no detectable effect on exon 2 splicing.

In the case of the patients SFC6 (c.633 + 1G > A/c.1334 T > C) and SFC13 (c.1542 + 4dupA/c.1150C > T), the RT-PCR analysis showed no effect of any modified U1 snRNA on the endogenous splicing process (Figure 3D,F). Similarly, WT *HGSNAT* gene splicing was not affected by any of the U1 isoforms (Figure 3A,C,E).

To rule out the possibility of the negative results being due to poor fibroblast transfection efficiency, fluorescence expression GFP/RFP vectors were used; efficient acquisition of the vectors was observed (data not shown). The uptake of the U1 snRNA was confirmed by PCR (Additional file 3: Figure S2). Moreover, the electroporation technique was also tried but no improvement in transfection efficiency was achieved (data not shown).

### **Analysis of the protein encoded by the mutant *HGSNAT* cDNA with 12 nucleotides deleted**

In order to test whether mutant cDNA with a 12-nucleotide in-frame deletion (encoding p.[L125\_R128del]) produces a stable protein and, if so, whether it is enzymatically active, we transfected cultured COS-7 cells with a corresponding pcDNA-HGSNAT-L125\_R128del plasmid and measured N-acetyltransferase activity in the cell homogenate. Cells transfected with an empty plasmid or those transfected with the pcTAP-HGSNAT plasmid described previously [6] encoding WT human HGSNAT were used as positive and negative controls, respectively. Our data (Figure 4A) showed a small but significant increase of N-acetyltransferase activity in the cells transfected with the mutant *HGSNAT* compared to that in the cells transfected with the empty plasmid. The activity in the cells transfected with WT pcTAP-HGSNAT increased about 10-fold (Figure 4A).

---

**Figure 4 Enzymatic activity of mutant HGSNAT-L125\_R128del protein can be partially restored by the pharmacological chaperone, glucosamine.** (A) N-acetyltransferase activity of COS-7 cells transfected with the pCDNA-HGSNAT-L125\_R128del plasmid is significantly increased compared to that of cells transfected with the empty pCDNA plasmid (mock). The data show means (±S.D.) of individual measurements. Three transfections (each in duplicate) were performed on separate occasions. \*\* and \*\*\*: statistically different from mock-transfected cells ( $p < 0.01$  and  $p < 0.001$ , respectively) according to unpaired t-test. (B) COS-7 cells transfected with the pCDNA-HGSNAT-L125\_R128del plasmid produce 160 kDa dimmers and 78 kDa monomers of HGSNAT precursor protein but show drastically reduced amounts of 44 kDa and 25 kDa mature HGSNAT chains produced by intra-lysosomal enzymatic cleavage. Panel shows representative data of 3 independent transfections. (C) N-acetyltransferase activity of COS-7 cells transfected with the pCDNA-HGSNAT-L125\_R128del plasmid and of cultured

primary fibroblasts of the patient homozygous for the c.372-2A > G mutation is significantly increased after treating the cells in culture with 10 mM glucosamine for 72 h (+GA). The data show means ( $\pm$ S.D.) of individual measurements. Three independent experiments measurements were performed each of them with 2 cell plates. \* and \*\*: statistically different from untreated cells ( $p < 0.05$  and  $p < 0.01$ , respectively) by unpaired t-test.

We further analysed cell homogenates by Western blot and detected a cross-reacting band in the cells transfected with both WT and mutant plasmids (Figure 4B). The amount of 160 kDa dimer of the HGSNAT precursor was comparable in both homogenates, but the amount of 44 kDa and 25 kDa enzymatically cleaved mature HGSNAT chains was reduced in the cells transfected with the pcDNA-HGSNAT-L125\_R128del plasmid. Since the enzymatic cleavage of the HGSNAT precursor into the 2-chain form occurs in the lysosome [6], this result is consistent with misfolding of the majority of the HGSNAT-L125\_R128del mutant and its retention in the endoplasmic reticulum (ER), as previously demonstrated for most HGSNAT mutants with amino acid substitutions [24].

### **Partial recovery of the enzyme activity of p.[L125\_R128del] by glucosamine as a pharmacological chaperone**

Our previous data demonstrated that the treatment of cultured cells of MPS IIIC patients bearing missense mutations with the competitive HGSNAT inhibitor ( $K_i = 0.28$  mM) glucosamine, significantly increased the level of the residual N-acetyltransferase activity [24]. In order to determine whether glucosamine can also restore the folding and activity of the p.[L125\_R128del] mutant protein, we treated COS-7 cells transfected with the pcDNA-HGSNAT-L125\_R128del plasmid and cultured primary fibroblasts of the patient homozygous for the c.372-2A > G mutation with 10 mM glucosamine. Our data (Figure 4C) show that glucosamine significantly increases residual N-acetyltransferase activity in both COS-7 cells transfected with the pcDNA-HGSNAT-L125\_R128del plasmid and in the patient's fibroblasts. This suggests that the active conformation of the mutant HGSNAT can be stabilized by glucosamine, resulting in part of the mutant enzyme pool being properly processed and targeted at the lysosomes.

## **Discussion**

In this study we describe two therapeutic approaches specific for several splicing mutations in the *HGSNAT* gene that lead to defects in mRNA processing in five Sanfilippo C patients, each carrying at least one splicing mutation. To our knowledge, this is the first attempt to treat Sanfilippo syndrome type C splicing mutations with modified U1 snRNAs. The chaperone treatment, the second therapeutic strategy examined in the present work, was previously tested in vitro for several missense Sanfilippo C syndrome mutations with promising results [24] and it is applied here for the treatment of a mutant protein lacking four amino acids.

For patients carrying mutations in the donor ss (SFCP, SFC3, SFC6 and SFC13), modified U1 snRNAs have been tested as a therapeutic tool to recover the normal splicing process. This approach has previously been tested for different mutations in several disorders showing different efficiencies at rescuing the normal transcripts [10-19]. A total of 9 different U1s have been developed, as well as the U1-WT. None of them affected the normal splicing process in WT minigenes or healthy control fibroblasts when overexpressed.

Few assays to correct a +1 5' ss mutation have been reported. In some of those, the efficiency of modified U1 snRNAs for splicing mutations in canonical positions +1 and +2 has been shown to be inexistent [10,13]. However, Hartmann *et al.* [19] showed a partial correction of a +1 mutation in a case in which some degree of normal splicing was conserved in the mutant allele and the alternative donor site with a "gt" dinucleotide in positions +5 and +6 presented a low score according to different predictors. In this report, we present three cases with splicing mutations at position +1, two homozygous patients with the c.234 + 1G > A mutation (SFCP and SFC3) and one heterozygous patient with the c.633 + 1G > A mutation (SFC6). When we analysed the treatment efficiency of different modified U1 snRNAs in minigene constructions carrying these mutations, we were not able to detect restoration of the normal splicing process. Instead, we detected alternative splicing patterns due to the use of highly conserved "gt" nucleotides at positions +5 and +6 as a donor site, giving rise to a transcript that includes 4 intronic nucleotides. This result was obtained using three of the four modified U1 snRNAs for the c.234 + 1G > A mutation and the total complementary U1 snRNA for the c.633 + 1G > A mutation. Due to the presence of these "gt" dinucleotides at position +5 and +6 in many introns, it is important to sequence putative rescued transcripts to check whether this alternative site was used.

Despite these results, and taking into account that minigenes only included partial intronic sequences which could lack some splicing regulatory sites and that they were assayed in non-human cells, modified U1 snRNAs were tested on patients' fibroblasts. Partial rescue (almost 50%) of the normal splicing for the c.234 + 1G > A mutation was observed in patients SFCP and SFC3 overexpressing the U1-sup4. In the case of the c.633 + 1G > A mutation (patient SFC6), no rescue was observed after the modified U1 snRNAs overexpression. Analysis of the ss sequences using the Human Splicing Finder predictor [29], indicated that the alternative site of patient SFC6 had a high score (68.17/100), while that of patients SFCP and SFC3 had a null score. This could explain the difference in the rescue results between these two "+1" mutations: the mutant allele of patient SFC6 would efficiently use the alternative site and, thus, would not be rescued.

The molecular reason why the modified U1 promotes, in general, a new splicing process using the +5 + 6 "gt" donor site when its sequence perfectly matches the mutated +1 + 2 site remains unclear. One explanation could be the involvement of U6 snRNA complementarity, in addition to that of U1, in the choice of alternative ss in proximity to the normal one, which has been previously described [30,31].

The rescue of the splicing mutation shown here is one of the few positive results for a "+1" mutation using a modified U1 snRNA and the first one performed on an allele that did not produce any of the correctly spliced mRNAs when untreated. It is important to note that a few natural U2-type 5'ss that present "au" instead of the normal "gu" at the first two nucleotides of the intron have been described [32,33]. These introns present "ac" instead of "ag" at the last two nucleotides. Thus, one could think that mutated "au" sites could promote an alternative splicing using an "ac" acceptor site. However, for this +1 mutation the "au-ac" alternative splicing type was not observed. On the contrary, we found that the "au" 5'ss correctly paired with the normal "ag" and not with a possible cryptic "ac" 3'ss, restoring the normal splicing of intron 2.

Further studies are needed to improve the efficacy and to test the toxicity and side effects of U1 overexpression in order to confirm the feasibility of the use of this modified U1 snRNAs in therapy. Overexpression of U1 snRNA vectors introduced with adeno-associated virus into

mice liver has been shown to be toxic at high viral doses but safe at low doses, suggesting the viability of this treatment when low amounts of U1 snRNA viral particles are injected [20].

In the case of the c.1542 + 4dupA mutation (patient SFC13), it is important to note that the canonical “gt” site is not affected by the mutation. However, in this case there is also another “gt” dinucleotide in positions +6 and +7 (one nucleotide downstream due to the duplication). The minigene analysis showed that the splicing that uses the alternative donor site occurs, even without treatment, together with that causing exon 15 skipping. This alternative splicing is slightly enhanced with U1-sup7 and greatly enhanced with U1-sup9. Many different ss predictors were used in order to estimate the score for each donor site in this case (Additional file 4: Table S2). Clearly, in the absence of mutation, the scores of both sites are similar; but when the mutation is present, the normal site presents a lower score while the alternative site increases its score. This is consistent with the fact that in the presence of the mutation the alternative “gt” site is the only one used (see Figure 2I). As it was previously discussed, it remains unclear why the modified U1 snRNAs, designed to perfectly match the mutated site, enhance the splicing using the alternative site. When the modified U1 snRNAs were tested on patient’s fibroblasts, no rescue was detected: either for the normal mRNA or for the one including the five intronic nucleotides. The latter was detected without any treatment, indicating that it takes place in normal conditions in the patient’s cells. In any case, the alternative transcript would not restore the enzyme activity since it does not keep the frame, so it could not have any therapeutic use. Many different mutations in the +3 to +6 positions have been partially or totally rescued using U1 snRNAs [10-13,15,17,18], while one +5 mutation was not corrected [15]. Our current results point to the importance of the presence of an additional “gt” dinucleotide in the region which, depending on the sequence context, may be used as an alternative donor site. In previous reports where mutations have been partially or totally corrected and the “gt” dinucleotide was present at positions +5 and +6 [15,19], the alternative site presented a low score in accordance with different predictors, in contrast to the cases described here.

Finally, we attempted to rescue the c.372-2G > A mutation (present in SFC7), using molecular chaperone. This mutation gives rise to two different splicing processes in fibroblasts: one causing exon 4 skipping; the other using an alternative acceptor site, 12 nucleotides downstream of the normal site and producing a transcript with an in-frame deletion. We showed that this alternative splicing produces a protein lacking 4 amino acids p.[L125\_R128del] that has some residual activity but the majority of the protein does not reach the lysosome, remaining in the ER due to its misfolding.

In order to restore the correct protein folding, cells transfected with the plasmid carrying the mutant protein and the patient’s fibroblasts were treated with glucosamine chaperone, which resulted in a significant increase in the residual HGSNAT activity. This indicates the feasibility of this therapeutic approach for patients carrying this splicing mutation. Different chaperones assays for lysosomal disorders have been performed before and some have been tested in humans, showing their safety and potential as a therapeutic tool reviewed in Ref. [21].

## Conclusions

In conclusion, the results of two therapeutic approaches for different splicing mutations varied depending on the nature of the mutation. For the treatment of donor ss mutations, U1 snRNAs could represent a feasible option. This would depend on the presence of alternative donor sites close to the normal site that could interfere with the correction process and, thus, with the

success of the therapy. This is important since many introns present a “gt” in positions +5 and +6. In the present study, we have shown that it is possible to partially recover the normal splicing process for +1 mutations, which was reported only once before. Additionally, a chaperone treatment using glucosamine for a mutant protein with a loss of 4 amino acids, caused by an acceptor ss mutation, has been shown to result in a significant increase in the enzyme activity. These promising results encourage further research into the therapeutic use of U1 snRNAs and chaperones to treat Sanfilippo syndrome type C patients.

## Abbreviations

Bp, Base pair; EET, Enzyme enhancement therapy; MPS III, Mucopolysaccharidoses type III; MPS IIIC, Mucopolysaccharidoses type IIIC or Sanfilippo syndrome type C; snRNAs, Small nuclear ribonucleic acids; ss, Splice site; WT, Wild type.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

LM and IC were involved in the conception and design of the study, performed the experimental work of the U1 snRNA part, the analysis and interpretation of the data, and participated in the drafting and revising of the manuscript. LD and YC were involved in the conception, design and experimental work, as well as in the interpretation of the data and the drafting of the manuscript regarding to the glucosamine part. SA, DG, LV and AVP supervised all the research contributing critically to the design of the work and data interpretation as well as in the revision of the manuscript. MJP, PJ, LRD and BP are collaborators with experience in the field and supervised the research. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to thank Dr. Lúcia Lacerda from Centro de Genética Médica Dr. Jacinto Magalhães - Centro Hospitalar do Porto, Portugal for providing the SFCP patient fibroblasts sample and Helena Ribeiro from the same institution for the technical support. We thank also Dr. Mónica Sousa and Dr. Elsa Logarinho research groups from Instituto de Biologia Molecular e Celular (IBMC), Porto for the collaboration in the electroporation studies. We would like equally to thank Xavier Roca from the School of Biological Sciences, Nanyang Technological University, Singapore, for advice on the U1s constructs and to the Institut de Bioquímica Clínica, Barcelona, for their collaboration. The authors are also grateful for the support of the *Centro de Investigación Biomédica en Red de Enfermedades Raras* (CIBERER), which is an initiative of the ISCIII. This study was partially funded by a grant from the Spanish Ministry of Science and Innovation (SAF2011-25431) and from the Catalan Government (2009SGR971). We are also grateful for the permanent support, including financial aid, from 'patient-support' associations, such as Jonah's Just Begun-Foundation to Cure Sanfilippo Inc. (USA), Association Sanfilippo Sud (France), Fundación Stop Sanfilippo (Spain), Asociación MPS España (Spain). LM was supported by a grant (SFRH/BD/64592/2009) from the Fundação para a Ciência e Tecnologia (FCT), Portugal, IC by a grant from the University of Barcelona (APIF), Spain and AVP by an operating grant MOP111068 from Canadian Institutes of Health Research.

## References

1. Neufeld EF, Muenzer J: **The Mucopolysaccharidoses**. In *The metabolic and molecular bases of inherited disease*, Volume 3. 8th edition. Edited by Scriver CR, Beaudet AL, Sly WS, Valle D. New York: McGraw-Hill; 2001:3421–3452.
2. Valstar MJ, Ruijter GJG, van Diggelen OP, Poorthuis BJ, Wijburg FA: **Sanfilippo syndrome: a mini-review**. *J Inherit Metab Dis* 2008, **31**:240–252.
3. Klein U, Kresse H, von Figura K: **Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts**. *Proc Natl Acad Sci U S A* 1978, **75**:5185–5189.
4. Fan X, Zhang H, Zhang S, Bagshaw RD, Tropak MB, *et al*: **Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C)**. *Am J Hum Genet* 2006, **79**:738–744.
5. Hřebíček M, Mrázová L, Seyrantepé V: **Mutations in TMEM76 Cause Mucopolysaccharidosis IIIC (Sanfilippo C Syndrome)**. *Am J Hum Genet* 2006, **79**:807–819.
6. Durand S, Feldhammer M, Bonneil E, Thibault P, Pshezhetsky AV: **Analysis of the biogenesis of heparan sulfate acetyl-CoA:alpha-glucosaminide N-acetyltransferase provides insights into the mechanism underlying its complete deficiency in mucopolysaccharidosis IIIC**. *J Biol Chem* 2010, **285**:31233–31242.
7. Fan X, Tkachyova I, Sinha A, Rigat B, Mahuran D: **Characterization of the biosynthesis, processing and kinetic mechanism of action of the enzyme deficient in mucopolysaccharidosis IIIC**. *PLoS One* 2011, **6**:e24951.
8. Wahl MC, Will CL, Lührmann R: **The spliceosome: design principles of a dynamic RNP machine**. *Cell* 2009, **136**:701–718.
9. Roca X, Akerman M, Gaus H, Berdeja A, Bennett CF, *et al*: **Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides**. *Genes Dev* 2012, **26**:1098–1109.
10. Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, *et al*: **An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects**. *Hum Mol Genet* 2012, **21**:2389–2398.
11. Pinotti M, Rizzotto L, Balestra D, Lewandowska MA, Cavallari N, *et al*: **U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency**. *Blood* 2008, **111**:2681–2684.
12. Sánchez-Alcudia R, Pérez B, Pérez-Cerdá C, Ugarte M, Desviat LR: **Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia**. *Mol Genet Metab* 2011, **102**:134–138.
13. Schmid F, Hiller T, Korner G, Glaus E, Berger W, *et al*: **A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs**. *Hum Gene Ther* 2013, **24**:97–104.

14. Tanner G, Glaus E, Barthelmes D, Ader M, Fleischhauer J, *et al*: **Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA.** *Hum Mutat* 2009, **30**:255–263.
15. Susani L, Pangrazio A, Sobacchi C, Taranta A, Mortier G, *et al*: **TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA.** *Hum Mutat* 2004, **24**:225–235.
16. Mattioli C, Pianigiani G, De Rocco D, Bianco AMR, Cappelli E, *et al*: **Unusual splice site mutations disrupt FANCA exon 8 definition.** *Biochim Biophys Acta* 1842, **2014**:1052–1058.
17. Glaus E, Schmid F, Da Costa R, Berger W, Neidhardt J: **Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells.** *Mol Ther* 2011, **19**:936–941.
18. Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, *et al*: **Identification of a mutation that perturbs NF1 gene splicing using genomic DNA samples and a minigene assay.** *J Med Genet* 2003, **40**:220–222.
19. Hartmann L, Neveling K, Borkens S, Schneider H, Freund M, *et al*: **Correct mRNA processing at a mutant TT splice donor in FANCC ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs.** *Am J Hum Genet* 2010, **87**:480–493.
20. Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, *et al*: **An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice.** *J Thromb Haemost* 2014, **12**:177–185.
21. Suzuki Y: **Emerging novel concept of chaperone therapies for protein misfolding diseases.** *Proc Japan Acad Ser B* 2014, **90**:145–162.
22. Bernier V, Lagacé M, Bichet DG, Bouvier M: **Pharmacological chaperones: potential treatment for conformational diseases.** *Trends Endocrinol Metab* 2004, **15**:222–228.
23. Sawkar AR, Cheng W-C, Beutler E, Wong C-H, Balch WE, *et al*: **Chemical chaperones increase the cellular activity of N370S beta -glucosidase: a therapeutic strategy for Gaucher disease.** *Proc Natl Acad Sci U S A* 2002, **99**:15428–15433.
24. Feldhammer M, Durand S, Pshezhetsky AV: **Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C.** *PLoS One* 2009, **4**:e7434.
25. Canals I, Elalaoui SC, Pineda M, Delgadillo V, Szlago M, *et al*: **Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles.** *Clin Genet* 2011, **80**:367–374.
26. Feldhammer M, Durand S, Mrázová L, Boucher R-M, Laframboise R, *et al*: **Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene.** *Hum Mutat* 2009, **30**:918–925.

27. Coutinho MF, Lacerda L, Prata MJ, Ribeiro H, Lopes L, *et al*: **Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene.** *Clin Genet* 2008, **74**:194–195.
28. Lund E, Dahlberg JE: **True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation.** *J Biol Chem* 1984, **259**:2013–2021.
29. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, *et al*: **Human splicing finder: an online bioinformatics tool to predict splicing signals.** *Nucleic Acids Res* 2009, **37**:e67.
30. Brackenridge S, Wilkie AO, Screaton GR: **Efficient use of a 'dead-end' GA 5' splice site in the human fibroblast growth factor receptor genes.** *EMBO J* 2003, **22**:1620–31.
31. Hwang DY, Cohen JB: **U1 snRNA promotes the selection of nearby 5' splice sites by U6 snRNA in mammalian cells.** *Genes Dev* 1996, **10**:338–50.
32. Wu Q, Krainer AR: **AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes.** *Mol Cell Biol* 1999, **19**(5):3225–36.
33. Kubota T, Roca X, Kimura T, Kokunai Y, Nishino I, *et al*: **A mutation in a rare type of intron in a sodium-channel gene results in aberrant splicing and causes myotonia.** *Hum Mutat* 2011, **32**(7):773–82.

## Additional files

### Additional\_file\_1 as PDF

Additional file 1 **Table S1**. Oligonucleotide sequences for PCR amplifications. List of the oligonucleotides used in this work. (PDF 15 kb).

### Additional\_file\_2 as PDF

Additional file 2 **Figure S1**. Electropherograms of the bands obtained in RT-PCR after transfection of the U1-sup4 in control and c.234 + 1G > A patient's fibroblasts. (A) Wild type (control) sequence showing the identity of the normal fragment with exon 2 and 3. (B) Normal and aberrant sequence of the rescue band obtained with U1-sup4 transfection in patient's fibroblasts. (PDF 64 kb).

### Additional\_file\_3 as PDF

Additional file 3 **Figure S2**. PCR amplification to test the uptake of a modified U1 snRNA vector by patient fibroblasts. Agarose gel electrophoresis shows bands corresponding to the U1 vector after fibroblast transfection (by 2.5 µg and 3.5 µg of vector as indicated) and lower molecular weight bands which correspond to the *HGSNAT* exon 4 amplification, as a control. M: molecular weight marker; NT: non-treated cells; C-: negative control. (PDF 63 kb).

### Additional\_file\_4 as PDF

Additional file 4 **Table S2**. Intron 15 donor ss scores using different predictors. Comparison of the ss scores for the normal site and the alternative site either in the presence or in the absence of the mutation. (PDF 17 kb).

Figure 1

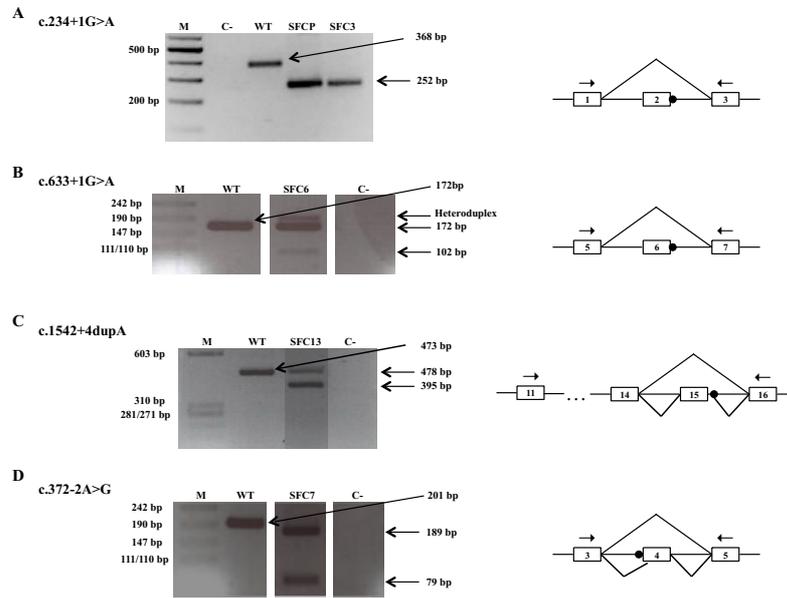


Figure 2

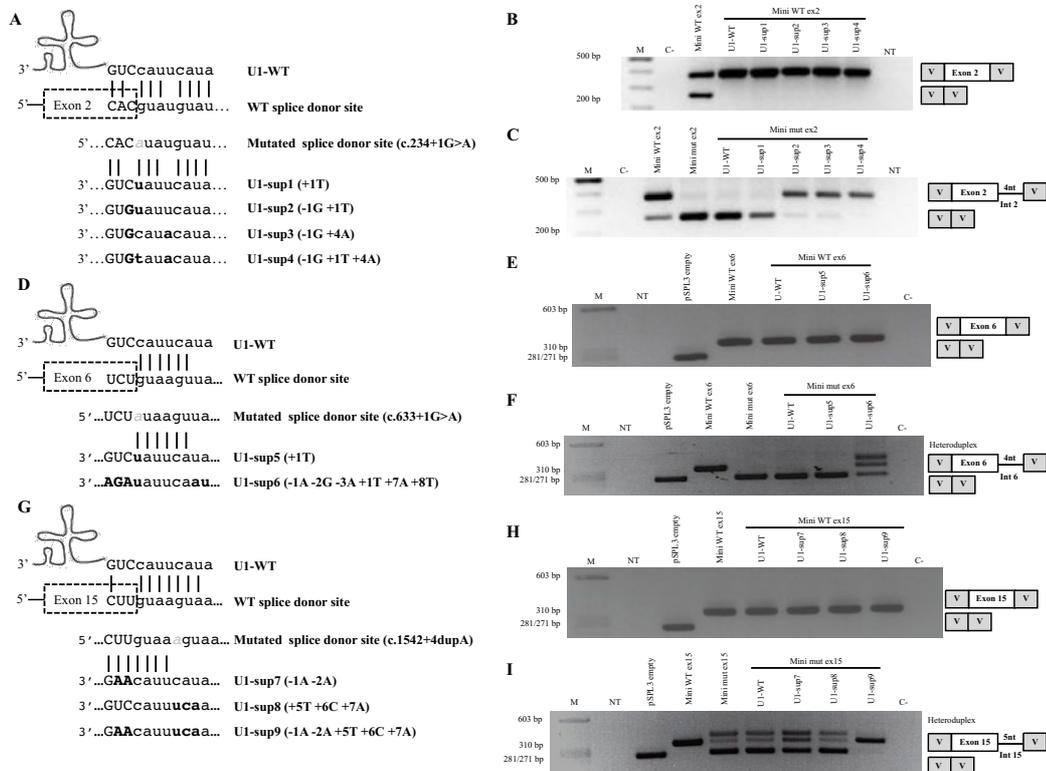


Figure 3

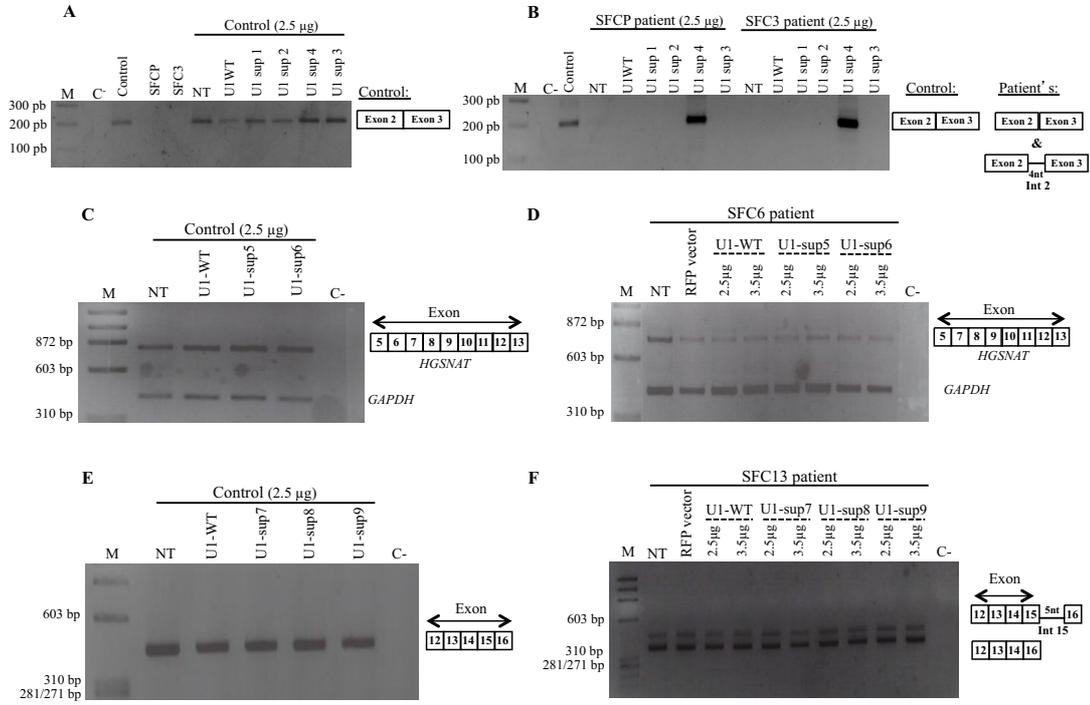
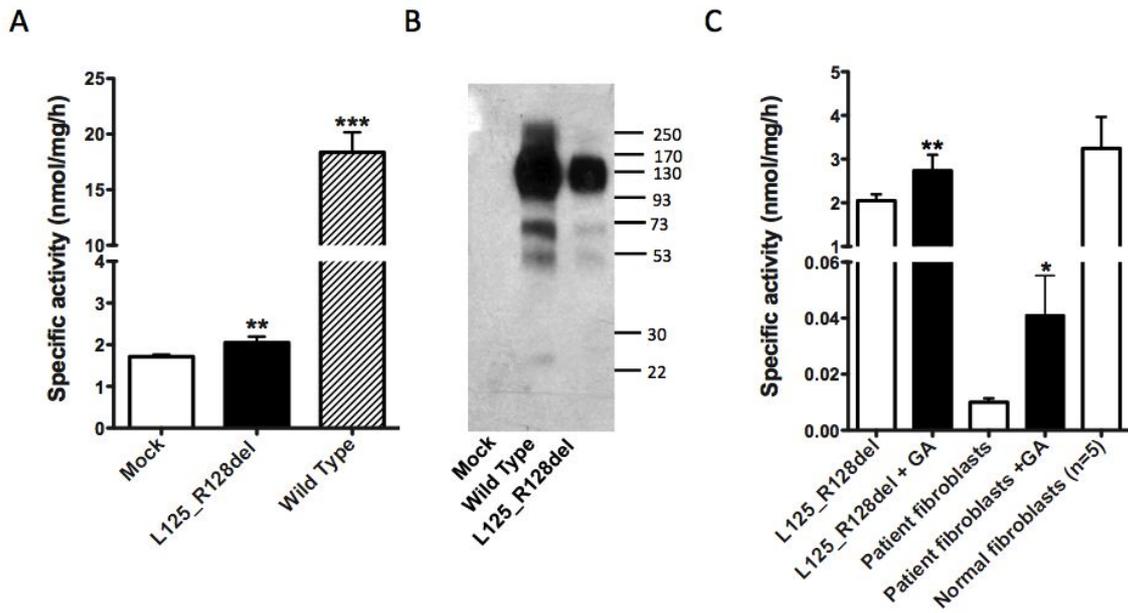


Figure 4



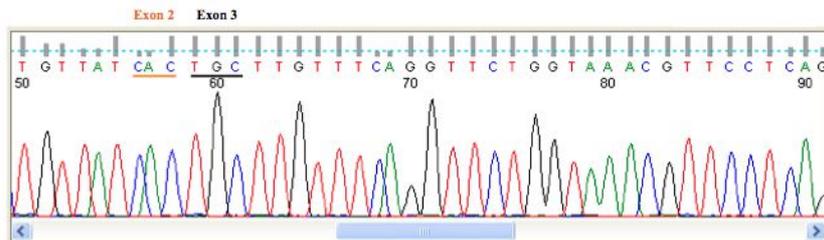
Additional file 1

Oligonucleotide	Sequence (5' to 3')
<b>RT-PCR fragments*</b>	
HGSNAT-Exon 1F	GCAGCGGGCAGGCAAGGGGG
HGSNAT-Exon 2F	ACATGCAGAGCTGAAGATGGA
HGSNAT-Exon 3R	GATAGATCCGTGCTGGGTG
HGSNAT-Exon 4F	AGCTGAACGACACCTTGGAA
HGSNAT-Exon 4R	CAATGATGACGACAGACAA
HGSNAT-Exon 6F	TGCACTCCCTTATTTGGCTTBC
HGSNAT-Exon 6R	ATCACCATCGAGAGGGTCTG
c.1542-4dupA F	ACCTGTGCCCTGAACATTGTG
c.1542-4dupA R	CAGAAACCTTCGTCAGAGCA
HGSNAT-Exon 5F	GCTTCTGAGGGCTTGTGTG
HGSNAT-Exon 12F	AGSAGCTGGCTTCTCTTC
HGSNAT-Exon 13R	TGATCGTCTCCGACGACGA
HGSNAT-Exon 16R	ACCTTCGTCAGAGCAACAG
<b>DNA fragments*</b>	
HGSNAT-Exon 4 F	TTATTCTGCCTCCATGATATTAGC
HGSNAT-Exon 4 R	GTACAGAAAAGCGCTAGGACTGC
<b>Cloning fragments*</b>	
Intron 1 F	GCAAAGGAGACCTGTGTGTG
Intron 3 R	TCATCCCTGAGAACTGGCTTT
<b>Plasmid Vectors</b>	
SD6 F	TCTGAGTCACCTGGACAACC
SA2 R	ATCTGAGTGGTATTTGTGAGC
pGEM3	ATCGAAATTAATACGACTCA
U1-QR	CTGGAAAACGACCTTGGT
<b>Site-directed mutagenesis</b>	
<b>pSPL3 WT minigenes</b>	
Exon 2 F	GGAAATCTGAATGCTGTATCACATATGTATCAGTTTCACACTCAG
Exon 2 R	CTGAGTGTGAACGTACATATGTGATAACAGCATTGAGATTCC
Exon 6 F	CTGATCGCCTCATCAATCTGTAAAGTATGAGATGATAGTG
Exon 6 R	CACATGTCATCTCAAACTTACAGAAATTGATGAGGCGATCAG
Exon 15 F	GGTGTGTATCTTTGTAAGTAAGCAGCATTCTCTCGC
Exon 15 R	GCGAGGAATGCTGCTTACTTACAGAAATACAAACAGC
<b>U1siRNA vectors</b>	
U1 suppressor 1 F	GATCTCACTTATCTGGCAGGGGAGATAC
U1 suppressor 1 R	GTATCTCCCTGCCAGATAAGTATGAGATC
U1 suppressor 2 F	GATCTCACTTATGTTGGCAGGGGAGATAC
U1 suppressor 2 R	GTATCTCCCTGCCACATAAGTATGAGATC
U1 suppressor 3 F	CAAGATGTCATACATACGTTGGCAGGGGAGATAC
U1 suppressor 3 R	GTATCTCCCTGCCACGTAAGTATGAGATGTTG
U1 suppressor 4 F	CAAGATGTCATACATATGTTGGCAGGGGAG
U1 suppressor 4 R	CTCCCTGCCACATAATGATGAGATCTTGG
U1 suppressor 5 F	GATCTCACTTATCTGGCAGGGGAGATAC
U1 suppressor 5 R	GTATCTCCCTGCCAGATAAGTATGAGATC
U1 suppressor 6 F	GAGGGCCCAAGATCTCAACTTATAGSACGAGGGGAGATACCATG
U1 suppressor 6 R	CATGTTATCTCCCTGCCCTGCTATAAGTATGAGATCTTGGGCCTC
U1 suppressor 7 F	AGCCCAAGATCTCACTTACAAAGCAGGGGAGATACC
U1 suppressor 7 R	GGTATCTCCCTGCCCTTGAAGTATGAGATCTTGGGCCT
U1 suppressor 8 F	GCAGAGCCCAAGATCTCAACTTACCTGGCAGGGGAGATA
U1 suppressor 8 R	TATCTCCCTGCCAGGTAAGTTGAGATCTTGGGCCTTGGC
U1 suppressor 9 F	GTGTCCGCGCAGAGGGCCCAAGATCTCAAGTTTACAAGGCAAGGGGAGATAC
U1 suppressor 9 R	GTATCTCCCTGCCCTTGTAAAGTTGAGATCTTGGGCCTTGGCCGAGAG

\* Primers were designed according to the sequence described in the ENSEMBL database (www.ensembl.org; ENSG00000165102). F – Forward, R – Reverse

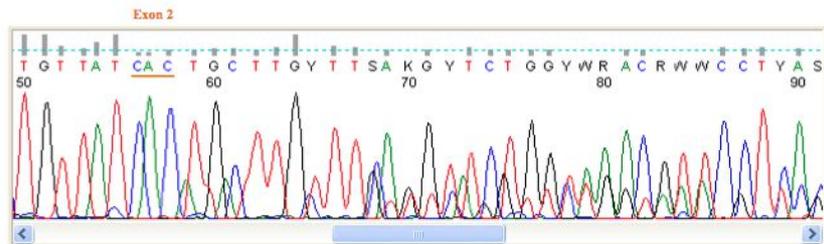
Additional file 2

A) Wild-type sequence



Normal exon 2 – exon 3 sequence ...TATCACTGCTTGTTCAGGTTCTGGTAAACGTTCTCTCAG...

B) c.234+1G>A sequence after U1 sup 4 treatment

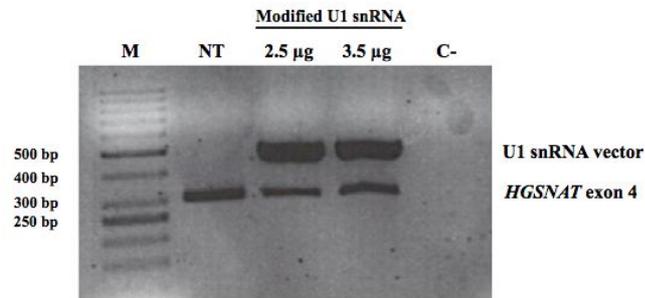


Two sequences:

Normal exon 2 – exon 3 sequence ...TATCACTGCTTGTTCAGGTTCTGGTAAACGTTCTCTCAG...

Aberrant exon 2 – exon 3 sequence...TATCACA**t**ATGCTTGTTCAGGTTCTGGTAAACGTTCTCTCAG...

## Additional file 3



## Additional file 4

Predictor	Normal	Alternative	Normal mutated	Alternative mutated
	CTT/gtaagt	taa/gtaagc	CTT/gtaaag	aaa/gtaagc
<b>Splice site score calculation (Max = 12.6)<sup>1</sup></b>	6.9	6.2	2	7.6
<b>Analyzer splice tool (Max = 100)<sup>2</sup></b>	78.94	78.17	61.92	81.99
<b>Splice site prediction by neuronal network (Max = 1) [29]</b>	0.95	0	0	0.97
<b>MaxEntScan (Max = 12) [30]</b>	8	5.66	-3	7.31
<b>Human splicing finder (Max = 100) [31]</b>	84.38	85.27	70.28	86.29

1. [http://rulai.cshl.edu/new\\_alt\\_exon\\_db2/HTML/score.html](http://rulai.cshl.edu/new_alt_exon_db2/HTML/score.html)

2. <http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm>

## **Article 3**



# ***EXTL2* and *EXTL3* inhibition with siRNAs as a promising substrate reduction therapy for Sanfilippo C syndrome.**

Isaac Canals, Mónica Cozar, Lluïsa Vilageliu, Daniel Grinberg.

## **Summary**

Sanfilippo syndrome is a rare lysosomal storage disorder caused by an impaired degradation of heparan sulfate (HS), presenting a severe and progressive neurodegeneration and for which no effective treatment exists. Substrate reduction therapy (SRT) may be a useful option for the treatment of neurological disorders such as Sanfilippo syndrome and different approaches have been tested until today. In this work we attempt the use of different siRNAs targeting *EXTL2* and *EXTL3* genes, important for HS synthesis, as a SRT in Sanfilippo C patients' fibroblasts to decrease the glycosaminoglycan (GAG) storage inside the lysosomes. The results show a high inhibition of the *EXTL* genes mRNAs (around 90%) and a consequent decrease in GAG synthesis after 3 days (30-60%) and a reduction in the GAG storage after 14 days (up to 24%). The use of siRNAs inhibiting HS synthesis genes is a promising therapeutic option for Sanfilippo C syndrome.



# ***EXTL2* and *EXTL3* inhibition with siRNAs as a promising substrate reduction therapy for Sanfilippo C syndrome.**

Isaac Canals<sup>1,2,3</sup>, Mónica Cozar<sup>1,2,3</sup>, Lluïsa Vilageliu<sup>1,2,3\*</sup>, Daniel Grinberg<sup>1,2,3\*</sup>.

<sup>1</sup> Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

<sup>2</sup> Centro de Investigación Biomédica En Red de Enfermedades Raras (CIBERER), Spain

<sup>3</sup> Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain

\* Co-last authors.

## **Abstract**

Sanfilippo syndrome is a rare lysosomal storage disorder caused by an impaired degradation of heparan sulfate (HS). It presents severe and progressive neurodegeneration and currently there is no effective treatment. Substrate reduction therapy (SRT) may be a useful option for neurological disorders of this kind, and several approaches have been tested to date. Here we use different siRNAs targeting *EXTL2* and *EXTL3* genes, which are important for HS synthesis, as SRT in Sanfilippo C patients' fibroblasts in order to decrease glycosaminoglycan (GAG) storage inside the lysosomes. The results show a high inhibition of the *EXTL* gene mRNAs (around 90%), a decrease in GAG synthesis after three days (30-60%) and a decrease in GAG storage after 14 days (up to 24%). The use of siRNAs inhibiting HS synthesis genes is a promising therapeutic option for Sanfilippo C syndrome.

## **Key words**

*EXTL* genes / heparan sulfate / Sanfilippo C syndrome / short interference RNA / substrate reduction therapy

## Introduction

Sanfilippo C syndrome, or mucopolysaccharidosis (MPS) IIIC, is an autosomal recessive lysosomal storage disorder caused by mutations in the *HGSNAT* gene. This gene codes for an enzyme responsible for heparan sulfate (HS) degradation, and its dysfunction leads to the storage of partially degraded molecules inside the lysosomes.<sup>1</sup> HS is one of the most common glycosaminoglycans (GAGs) located in the extracellular matrix as a part of proteoglycans, which participate in many different cellular functions.<sup>2</sup> In the synthetic pathway of HS, there is an essential step in which the *EXTL* genes play a crucial role (reviewed in Ref. 3). In particular, *EXTL2* and *EXTL3* are the main proteins involved in the HS chain elongation. Small interference RNAs (siRNAs), discovered in the late 1990s,<sup>4</sup> were shown to inhibit mammalian genes,<sup>5</sup> and may be applicable in substrate reduction therapy (SRT) for Sanfilippo C patients. For other types of MPS, two different approaches using siRNAs or shRNAs have been previously described which inhibit different genes in the HS synthetic pathway.<sup>6,7</sup>

In this study we used four different siRNAs, two of which inhibit *EXTL2* and the other two inhibit *EXTL3*. These siRNAs were transfected into fibroblasts of two patients in an attempt to reduce HS synthesis. All four siRNAs obtained a notable reduction in the mRNA levels of the corresponding gene and decreased the rates of GAG synthesis and storage.

## Materials and Methods

Fibroblasts from two patients (SFC6 and SFC7) were used in this study. The patients and the culture conditions were described in a previous work.<sup>8</sup> Fibroblasts were transfected with four different siRNAs using Lipofectamine 2000 as the transfection agent. Quantities of 5, 2.5 and 1.25  $\mu$ l were used for 6-well, 12-well and 24-well plates respectively. Two Silencer® Select siRNAs (Ambion) for *EXTL2* (si4899 and si4900), two for *EXTL3* (si4901 and si4902) and one negative control (siC-) were used at a final concentration of 10  $\eta$ M. siRNA sequences are available on demand. Quantitative real-time

PCR was performed using TaqMan® Gene Expression Assays for *EXTL2*, *EXTL3*, *HPRT* and *SDHA* genes, the last two as endogenous controls. PCRs were carried out in a LightCycler® 480 System (Roche). The efficiency of GAG synthesis was quantified testing the cell incorporation of <sup>35</sup>S sodium sulfate from the medium as previously described.<sup>6</sup> The sulfated GAG storage was quantified using the Rheumera® Proteoglycan Detection Kit (Astarte Biologics) following the manufacturer's instructions. GAG extraction was performed as previously described.<sup>9</sup>

## Results

### Inhibition of *EXTL2* and *EXTL3* mRNA with specific siRNAs

All *EXTL* siRNAs were tested in fibroblasts from both patients and high inhibition levels were observed from day 3 to day 14 after transfection. All treated cells had expression levels of around 10% compared to cells treated with negative control siRNAs (Figure 1).

### Decrease in GAG synthesis

After three days of transfection, fibroblasts showed a decrease in the incorporation of <sup>35</sup>S sulfate of about 30% to 60% (depending on the siRNA used) compared to control samples (Figure 2). Both patients showed similar results for each siRNA, indicating a more efficient inhibition of GAG synthesis for siRNAs designed to target the *EXTL2* gene (more than 50%) than in those targeting the *EXTL3* gene (less than 50%).

### Decrease in GAG storage

Sulfated GAG storage was quantified in patients' fibroblasts after treatment (Figure 3). After three days, SFC6 fibroblasts showed  $1.346 \pm 0.085$   $\mu\text{g GAGs}/\mu\text{g DNA}$ , while treated fibroblasts presented between 1.075 and 1.235  $\mu\text{g GAGs}/\mu\text{g DNA}$ . SFC7 fibroblasts had  $0.953 \pm 0.049$   $\mu\text{g GAGs}/\mu\text{g DNA}$  without treatment, and between 0.812 and 1.168  $\mu\text{g GAGs}/\mu\text{g DNA}$  after treatment (Figure 3A). After seven days, SFC6 fibroblasts showed  $1.824 \pm 0.136$   $\mu\text{g GAGs}/\mu\text{g DNA}$ , while treated fibroblasts presented between 1.538 and

1.701  $\mu\text{g GAGs}/\mu\text{g DNA}$ . SFC7 fibroblasts had  $1.521\pm 0.081$   $\mu\text{g GAGs}/\mu\text{g DNA}$  without treatment and between 1.354 and 1.677  $\mu\text{g GAGs}/\mu\text{g DNA}$  after treatment (Figure 3B). After 14 days, SFC6 fibroblasts showed  $3.259\pm 0.085$   $\mu\text{g GAGs}/\mu\text{g DNA}$ , while treated fibroblasts presented between 2.618 and 2.945  $\mu\text{g GAGs}/\mu\text{g DNA}$ . SFC7 fibroblasts had  $2.569\pm 0.049$   $\mu\text{g GAGs}/\mu\text{g DNA}$  without treatment and between 1.973 and 2.66  $\mu\text{g GAGs}/\mu\text{g DNA}$  after treatment (Figure 3C). In general, patients' storage decreased at each time point, reaching a maximum reduction of 24% for siRNAs si4901 at 14 days in the fibroblasts of patient SFC7.

## Discussion

At present there is no effective treatment for Sanfilippo syndrome. In this report, different siRNAs were tested as SRT to inhibit key genes for HS synthesis in Sanfilippo C patients. A similar strategy has been applied previously, using shRNAs to inhibit *EXTL* genes<sup>7</sup> and siRNAs to inhibit other genes participating in GAG synthesis.<sup>6</sup>

The siRNAs used in this study showed high mRNA inhibition capacity (around 90%) for as long as 14 days at low concentrations (10 nM). Higher concentrations did not improve the result (data not shown), indicating the high efficiency of these siRNAs. These inhibition rates were higher than those previously obtained for the same genes with shRNAs.<sup>7</sup>

The results for GAG synthesis were interesting. After three days of transfection, all siRNAs achieved notable decreases in the synthetic pathway in both patients. Moreover, the reduction may have been even higher, considering that HS synthesis was inhibited but all types of GAG synthesis were quantified. Compared with previous results,<sup>6,7</sup> the higher mRNA inhibition levels obtained here achieved greater GAG synthesis inhibition. Another important point to note is that *EXTL2* inhibition seemed to be more efficient in decreasing GAG synthesis both in our study and in the study by Kaidonis et al.,<sup>7</sup> suggesting that *EXTL2* may be a better target candidate. It should be noted that at 7 and 14 days after transfection GAG synthesis was decreased in untreated cells, probably due to the fact that

they reached confluence, which may have promoted a decrease in the synthetic pathway of extracellular matrix components such as GAGs. It would be interesting to be able to work with more suitable cell types – neural cells, for example - which are highly relevant to this disease, and have lower growth rates that would allow long-term studies.

Patients' fibroblast cultures showed a pronounced increase in GAG storage with time, since these cells are not able to degrade HS. A slight decrease in storage has been detected after treatment with some of the siRNAs. The best results were observed at 14 days, although a trend towards a reduction was detected from three days of treatment onwards. We stress, again, that we decreased HS storage but quantified all GAG amounts, which may have led to an underestimation of the reduction.

Taken together, our results indicate that these siRNAs promote a reduction in mRNA levels of target genes, a notable reduction in the GAG synthetic pathway after 3 days, and a slower accumulation rate in patients' cells over two weeks. Despite this decrease, treated patients' fibroblasts still presented higher GAG levels than those from wild-type individuals. This suggests that RNAi-based therapies, although promising, should be used as a complementary therapy to obtain synergic effects with other treatments in order to accelerate the rate of HS degradation and/or excretion out of the cell. The search for novel shRNAs, stable and highly inhibitory for *EXTL2* (the best target, according to our results and those of other authors) is a necessary next step on the way to achieving a successful SRT for Sanfilippo syndrome.

## **Conflict of interest statement**

The authors declare no conflict of interest.

## **Acknowledgements**

We thank L.Gort for advice in setting up some techniques. The authors are also grateful for the support of the *Centro de Investigación Biomédica en Red de Enfermedades*

*Raras* (CIBERER), which is an initiative of the ISCIII. This study was partially funded by a grant from the Spanish Ministry of Science and Innovation (SAF2011-25431) and from the Catalan Government (2009SGR971 and 2014SGR932). We are also grateful for the permanent support, including financial aid, from 'patient-support' associations, such as Jonah's Just Begun-Foundation to Cure Sanfilippo Inc. (USA), Association Sanfilippo Sud (France), Fundación Stop Sanfilippo (Spain), Asociación MPS España (Spain). IC was supported by a grant from the University of Barcelona (APIF), Spain.

## References

- 1 Neufeld EF, Muenzer J. The Mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *Themetabolic and molecular bases of inherited disease*, 8th edn, Vol. 3. New York: McGraw-Hill, 2001: 3421–3452.
- 2 Sarrazin S, Lamanna WC, Esko JD. Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol* 2011; **3**: 1–33.
- 3 Kreuger J, Kjellén L. Heparan sulfate biosynthesis: regulation and variability. *J Histochem Cytochem* 2012; **60**: 898–907.
- 4 Fire A, Xu S, Montgomery M, Kostas S, Driver S, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; **391**: 806–811.
- 5 Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001; **411**: 494–8.
- 6 Dziejcz D, Wegrzyn G, Jakóbkiewicz-Banecka J. Impairment of glycosaminoglycan synthesis in mucopolysaccharidosis type IIIA cells by using siRNA: a potential therapeutic approach for Sanfilippo disease. *Eur J Hum Genet* 2010; **18**: 200–5.
- 7 Kaidonis X, Liaw WC, Roberts AD, Ly M, Anson D, Byers S. Gene silencing of EXTL2 and EXTL3 as a substrate deprivation therapy for heparan sulphate storing mucopolysaccharidoses. *Eur J Hum Genet* 2010; **18**: 194–9.
- 8 Canals I, Elalaoui SC, Pineda M *et al.* Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles. *Clin Genet* 2011; **80**: 367–74.
- 9 Barbosa I, Garcia S, Barbier-Chassefière V, Caruelle J-P, Martelly I, Papy-García D. Improved and simple micro assay for sulfated glycosaminoglycans quantification in biological extracts and its use in skin and muscle tissue studies. *Glycobiology* 2003; **13**: 647–53.

## Titles and legends to figures

**Figure 1** *EXTL2* and *EXTL3* gene silencing. SFC6 and SFC7 fibroblasts were transfected with all four siRNAs and a negative control siRNA, and after 3, 7, 10 and 14 days, gene expression was analysed by real-time PCR. Results are expressed as the percentage of gene expression compared to cells transfected with the negative control siRNA and using two different reference genes (*HPRT* and *SDHA*).

**Figure 2** Inhibition of GAG synthesis. SFC6 and SFC7 fibroblasts were transfected with all four siRNAs and a negative control siRNA and after 3 days incorporation of <sup>35</sup>S sodium sulfate was analysed. Results for both patients are the mean ± standard error of three experiments performed in quadruplicate and are expressed as disintegrations per minute per µg of DNA. Differences between *EXTL* siRNAs with respect to negative control siRNA were evaluated using the non-parametric Mann-Whitney U test, and statistical significance was set at  $p < 0.05$  (\*),  $p < 0.01$  (\*\*) or  $p < 0.001$  (\*\*\*)).

**Figure 3** Decrease in GAG storage. SFC6 and SFC7 fibroblasts were transfected with all four siRNAs and a negative control siRNA and after 3, 7 and 14 days, GAG storage was analysed. Results for both patients are the mean ± standard error of three experiments performed in duplicate and are expressed in µg GAGs / µg DNA at 3 days (**a**), 7 days (**b**) and 14 days after transfection (**c**). Differences between *EXTL* siRNAs respect to negative control siRNA were evaluated using the non-parametric Mann-Whitney U test, and statistical significance was set at  $p < 0.05$  (\*) or  $p < 0.01$  (\*\*).

Figure 1

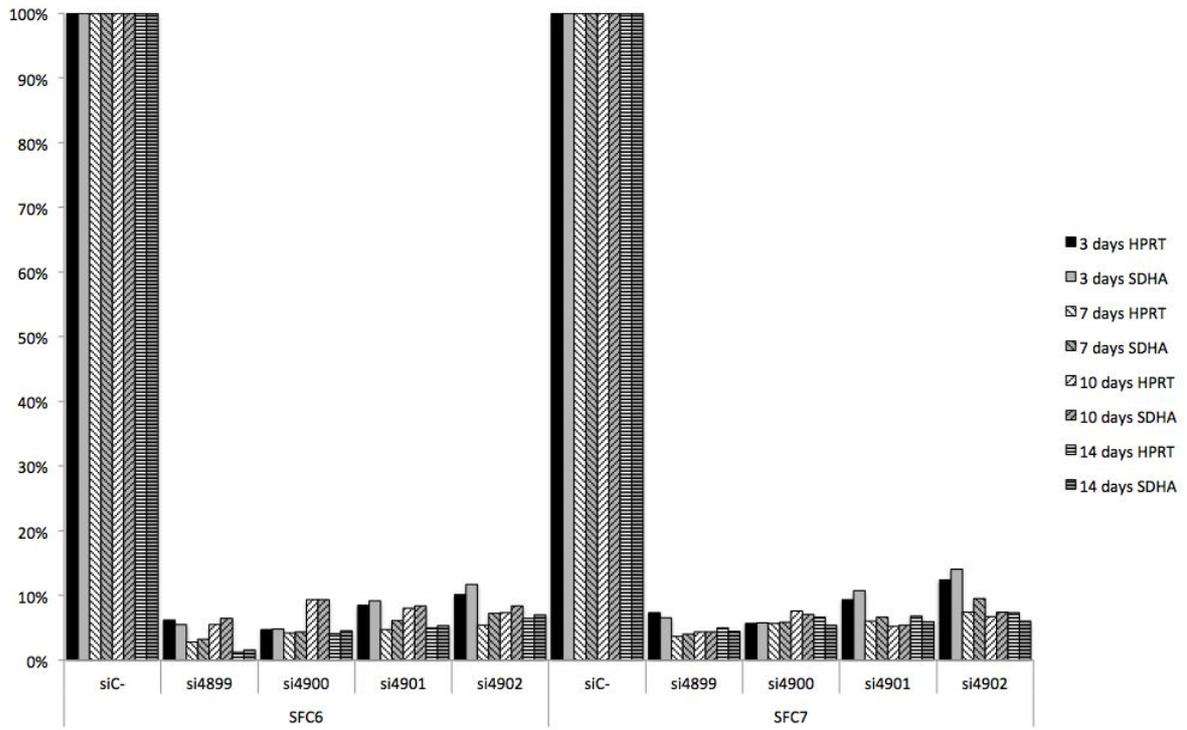


Figure 2

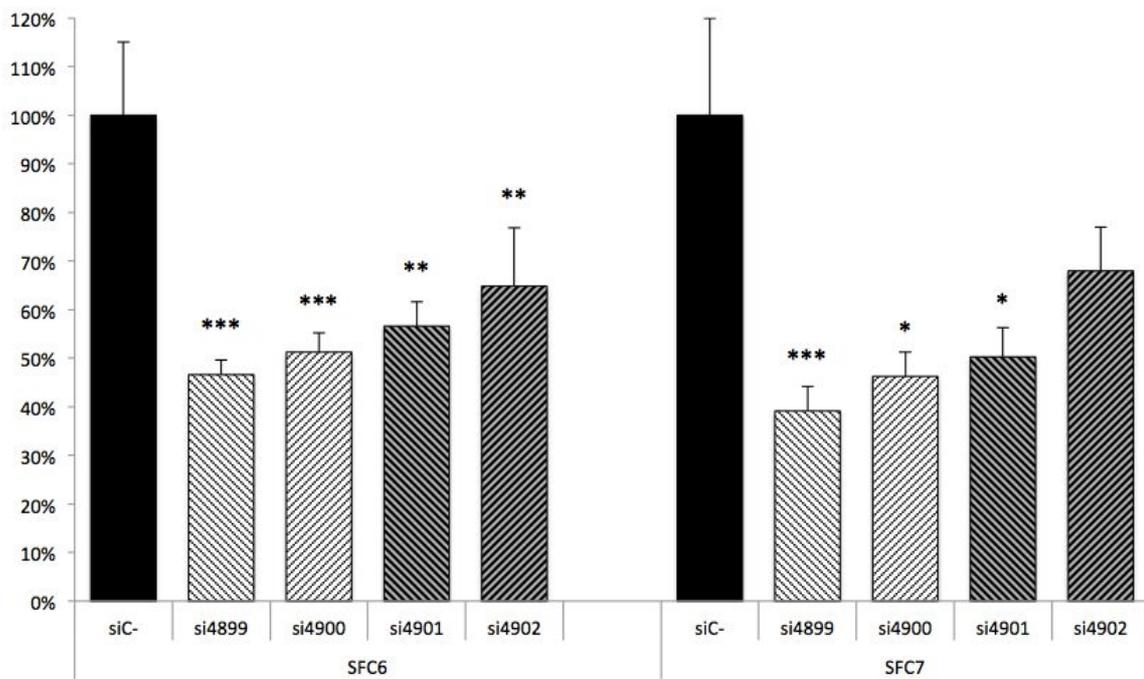
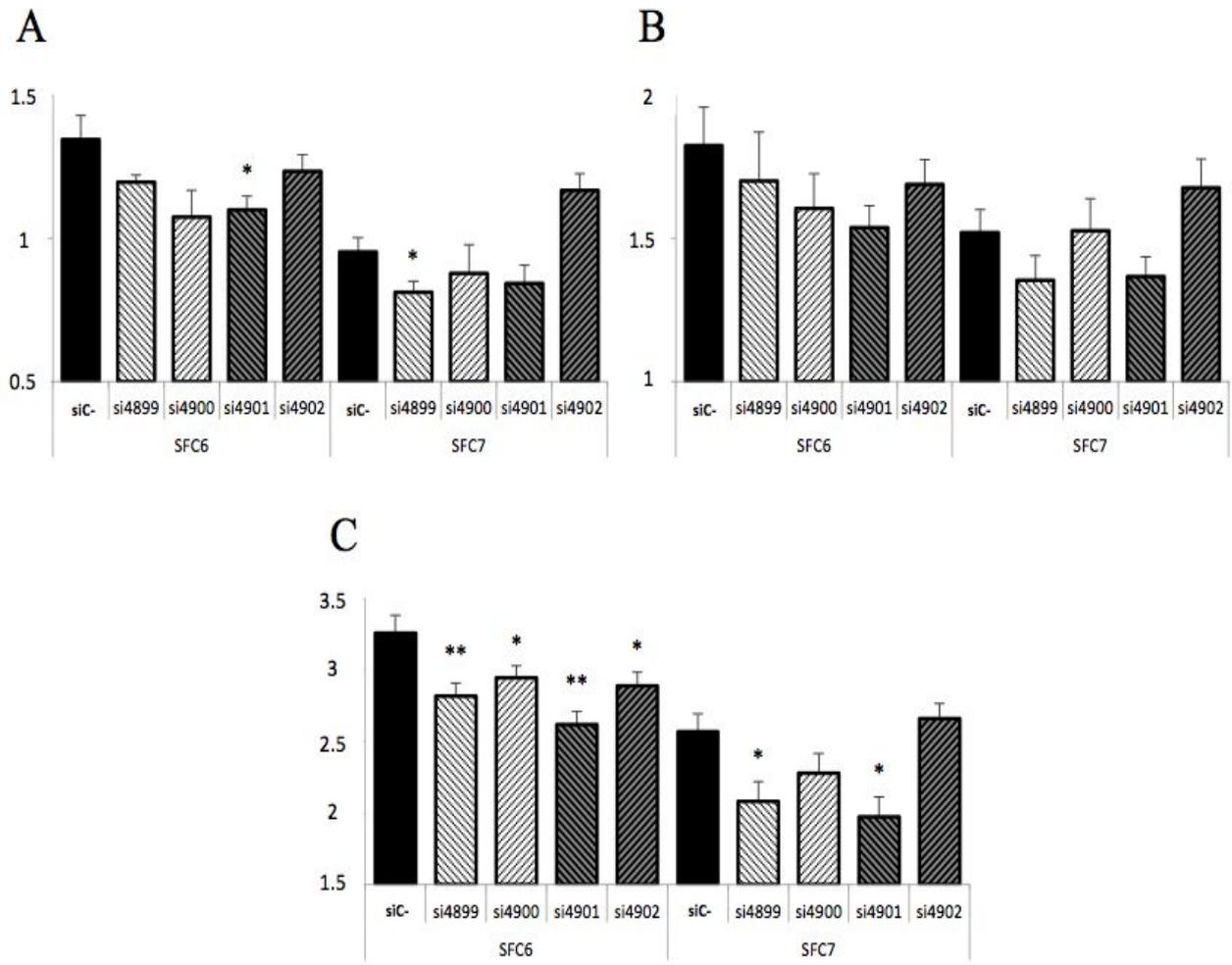


Figure 3





## **Article 4**



# High-order connectivity alterations in Sanfilippo C patient-specific neural networks.

Isaac Canals, Jordi Soriano, Javier G. Orlandi, Roger Torrent, Yvonne Richaud-Patin, Senda Jiménez-Delgado, Simone Merlin, Antonia Follenzi, Antonella Consiglio, Lluïsa Vilageliu, Daniel Grinberg, Ángel Raya.

## Summary

Induced pluripotent stem cell- (iPSC-) based technology has been successfully used to recapitulate phenotypic traits of dozens of human diseases in vitro. Patient-specific iPSC-based disease models are also expected to reveal early functional phenotypes that predate the cellular alterations observed in patients, although this remains to be formally proven. Here, we generated iPSC lines from two patients of Sanfilippo C, a lysosomal storage disorder caused by mutations in the *HGSNAT* gene that causes an inherited progressive neurodegenerative disease. Neurons differentiated from patient-specific iPSC lines recapitulated the lack of acetyl-CoA $\alpha$ -glucosaminide N-acetyltransferase activity, accumulation of glycosaminoglycans, and lysosomal dysfunction characteristic of the disease, which were not present in genetically corrected patient-specific iPSC-derived neurons. Moreover, neural networks organized in vitro from patient-derived cells showed early defects in connectivity as determined through generalized transfer entropy methods, and network-broad degradation over time when compared with networks of gene-corrected or control iPSC-derived neurons. Our findings establish the usefulness of iPSC-based technology to identify early functional phenotypes that can shed light onto the mechanisms that lead to brain dysfunction in these patients, as well as providing valuable readouts for screening of compounds that prevent, rather than reverting, the onset of neurodegeneration.



# High-order connectivity alterations in Sanfilippo C patient-specific neural networks.

Isaac Canals<sup>1,2,3</sup>, Jordi Soriano<sup>4</sup>, Javier G. Orlandi<sup>4</sup>, Roger Torrent<sup>3</sup>, Yvonne Richaud-Patin<sup>5,6</sup>, Senda Jiménez-Delgado<sup>5,6</sup>, Simone Merlin<sup>7</sup>, Antonia Follenzi<sup>7</sup>, Antonella Consiglio<sup>3,8</sup>, Lluïsa Vilageliu<sup>1,2,3</sup>, Daniel Grinberg<sup>1,2,3</sup>, Ángel Raya<sup>5,6,9</sup>.

<sup>1</sup> Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

<sup>2</sup> Centro de Investigación Biomédica En Red de Enfermedades Raras (CIBERER), Spain

<sup>3</sup> Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain

<sup>4</sup> Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona, Barcelona, Spain

<sup>5</sup> Control of Stem Cell Potency Group, Institut de Bioenginyeria de Catalunya (IBEC), Barcelona, Spain

<sup>6</sup> Centro de Investigación Biomédica En Red en Bioingeniería, Biomaterials y Nanomedicina (CIBER-BBN), Spain

<sup>7</sup> Health Sciences Department, Università' del Piemonte Orientale, Novara, Italy

<sup>8</sup> Department of Biomedical Science and Biotechnology, University of Brescia, Brescia, Italy

<sup>9</sup> Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

## Summary

Induced pluripotent stem cell- (iPSC-) based technology has been successfully used to recapitulate phenotypic traits of dozens of human diseases in vitro. Patient-specific iPSC-based disease models are also expected to reveal early functional phenotypes that predate the cellular alterations observed in patients, although this remains to be experimentally proven. Here, we generated iPSC lines from two patients of Sanfilippo C, a lysosomal storage disorder caused by mutations in the *HGSNAT* gene that causes an inherited progressive neurodegenerative disease. Neurons differentiated from patient-specific iPSC lines recapitulated the lack of acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase activity, accumulation of glycosaminoglycans, and lysosomal dysfunction characteristic of the disease, which were not present in genetically corrected patient-specific iPSC-derived neurons. Moreover, neural networks organized in vitro from patient-derived cells showed early defects in connectivity as determined through generalized transfer entropy methods, and network-broad degradation over time when compared with networks of gene-corrected or control iPSC-derived neurons. Our findings establish

the usefulness of iPSC-based technology to identify early functional phenotypes that can shed light onto the mechanisms that lead to brain dysfunction in these patients, as well as providing valuable readouts for screening of compounds that prevent, rather than reverting, the onset of neurodegeneration.

## Introduction

Sanfilippo syndrome, also known as muchopolysaccharidosis type III (MPS III), is a lysosomal storage disorder (LSD) with an autosomal recessive inheritance pattern. Four different subtypes have been described (A: OMIM 252900, B: OMIM 252920, C: OMIM 252930, D: OMIM 252940), which share clinical presentation, including early-onset severe central nervous system degeneration that typically results in death within the second and third decade of life (Valstar et al., 2008). Each subtype is caused by mutations in a different gene encoding enzymes involved in the degradation pathway of the glycosaminoglycan (GAG) heparan sulfate (Neufeld and Muenzer, 2001). The lack of activity of any of these enzymes leads to the accumulation of partially degraded heparan sulfate chains into lysosomes, which increase in their number and size. This accumulation causes an alteration in the lysosomal environment since heparan sulfate can bind to various hydrolases, reducing their activity and causing the secondary accumulation of gangliosides that may contribute to the central nervous system pathology (Lamanna et al., 2011; Walkley, 2004).

Sanfilippo syndrome type C (MPS IIIC) is caused by mutations in the *HGSNAT* gene, encoding acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase (EC 2. 3.1.78), a lysosomal membrane protein that catalyzes the acetylation of the terminal glucosamine residues of heparan sulfate prior to its hydrolysis by  $\alpha$ -N-acetyl glucosaminidase (Klein et al., 1978). The *HGSNAT* gene, located at chromosome 8p11.1 and containing 18 exons, was identified by two independent groups in 2006 (Fan et al., 2006; Hřebíček et al., 2006) and 64 different mutations have been identified since then (HGMD® Professional 2014.1). Until today, no animal or cellular model has been obtained for Sanfilippo type C.

The technology to reprogram somatic cells to pluripotency (Takahashi et al., 2006; Takahashi et al., 2007) has created new opportunities for generating *in vitro* models of disease based on disease-relevant cells differentiated from patient-specific induced pluripotent stem cell (iPSC) lines (recently reviewed by Trounson et al., 2012; Cherry and Daley, 2013; Inoue et al., 2014). This approach has been shown to be particularly useful in the case of congenital or early-onset monogenic diseases. In particular, iPSC-based models of various LSD have been recently established, including Gaucher disease (Mazzulli et al., 2011; Panicker et al., 2012; Tiscornia et al., 2013), Hurler syndrome (Tolar et al., 2011), Pompe disease (Higuchi et al., 2014; Huang et al., 2011), Sanfilippo B syndrome (Lemonnier et al., 2011) and Niemann-Pick type C1 (Maetzel et al., 2014; Trilck et al., 2013). In all these cases, disease-relevant cell types derived from patient-specific iPSC displayed morphologic, biochemical, and/or functional hallmarks of the disease, thus validating their use as *in vitro* models to investigate pathogenic mechanisms underlying disease progression and to screen for drugs capable of reverting LSD-related phenotypes.

In this study, we set out to test whether patient-specific iPSC-derived cells could be used to investigate the existence of early functional alterations prior to the appearance of disease-related phenotypes identified in patients. For this purpose, we generated iPSC from fibroblasts of Sanfilippo C syndrome patients (SFC-iPSC) and differentiated them into neurons, which recapitulate the pathological phenotypes observed *in vivo*, such as low acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase activity, accumulation of GAGs, and increase in lysosome size and number. Moreover, we found that neural networks organized *in vitro* from control iPSC-derived neurons grew in complexity over time, and this could be quantified in terms of connectivity using generalized transfer entropy methods. In contrast, networks of SFC-iPSC-derived neurons showed early defects in connectivity and network-broad degradation over time. The identification of early functional phenotypes in SFC-iPSC-derived neurons attests to the validity of iPSC-based technology to model pre-symptomatic stages of human diseases, thus widening the spectrum of potential applications of somatic cell reprogramming for biomedical research.

## Results

### Generation and characterization of patient-specific iPSC

Fibroblasts from two unrelated Spanish patients of Sanfilippo C syndrome (SFC6 and SFC7) and one healthy individual (WT) were used. Patient SFC6 was a compound heterozygote carrying a splicing mutation (c.633+1G>A), and a missense mutation (c.1334T>C; p.L445P). Patient SFC7 was homozygote for the most prevalent mutation in Spanish patients, affecting another splicing site (c.372-2A>G). The effect of these mutations on the splicing process and the transferase protein were described in a previous work (Canals et al., 2011).

Fibroblasts were reprogrammed at early passages (5-7) with retroviral delivery of *SOX2*, *KLF4* and *OCT4* (3F), or *SOX2*, *KLF4*, *OCT4* and *c-MYC* (4F) to generate 5-6 independent iPSC lines for each individual and for each factor combination. We selected clones displaying embryonic stem cell-like morphology and positive alkaline phosphatase staining. One clone representing each individual and reprogramming protocol (3F and 4F) was chosen to be thoroughly characterized and shown to be fully reprogrammed, as judged by integration of the reprogramming factors in the genome of the cells, silencing of the reprogramming transgenes and activation of endogenous pluripotency-associated factors, expression of pluripotency-associated transcription factors and surface markers, demethylation of *OCT4* and *NANOG* promoters, pluripotent differentiation ability *in vitro* and *in vivo*, and karyotype stability after more than 15 passages (Fig. 1A-G, Fig. S1A-G and Fig. S2).

Mutation analysis confirmed that SFC-iPSC bore the mutations present in the patients' fibroblasts (Fig. S2), with the same splicing effect (Fig. S3, see also Canals et al., 2011). SFC6-iPSC and SFC7-iPSC showed no detectable acetyl-CoA  $\alpha$ -glucosaminide N-acetyltransferase activity, consistent with the low enzyme activity levels found in patients' fibroblasts (1.78% and 3.02% of that of WT fibroblasts for SFC6 and SFC7, respectively, Fig. 2A-B). We also found that WT-iPSC showed a marked decrease in enzyme activity when compared to WT fibroblasts, consistent with the low lysosomal charge of iPSC compared with that of their parental fibroblasts (Fig. S4).

## Generation of patient-specific neurons

In order to reduce the variability associated to neural differentiation of iPSC (Falk et al., 2012), we derived neural progenitor cells from SFC-iPSC as spherical neural masses (SNM) that can be expanded and subsequently differentiated to neurons and glia (Cho et al., 2008). SNM derived from WT- and SFC-iPSC lines homogeneously expressed neural progenitor markers such as PAX6, NESTIN and SOX2, and proliferation markers such as Ki67 (Fig. 3A). Furthermore, when cultured in neuronal induction media supplemented with N2 and B27, SNMs generated Tuj1 positive cells, confirming their neurogenic capacity after 3, 6 and 9 weeks (Fig. 3B). Consistent with previous observations (Tiscornia et al., 2013), human SNM differentiated to heterogeneous cultures containing not only neurons positive for TUJ1 but also astrocytes positive for GFAP.

SNM derived from WT-iPSC showed enzyme activity of 1.211 nmol/h·mg of protein, while values below 0.01 nmol/h·mg of protein were obtained for patient-specific SNM (Fig. 2C), representing ~0.7% of the WT activity. Similarly, WT neurons showed an activity of 1.782 nmol/h·mg of protein, with no differences between 3, 6 and 9 weeks of differentiation, whereas values below 0.02 nmol/h·mg of protein were obtained for patient-specific neurons at 3, 6 and 9 weeks (Fig. 2D), which represents ~1% of the WT activity.

Total amount of GAGs accumulated over time in patients' neurons when compared to WT, reaching statistically significant differences after 9 weeks of differentiation (Fig. 2F). However, these differences were evident in patient's fibroblasts since the establishment of the culture, when they already doubled the total GAGs levels of WT fibroblasts (Fig. 2E). The splicing pattern of the *HGSNAT* gene in patient-specific neurons recapitulated the alterations seen in fibroblasts and iPSC cells (Fig. S3, see also Canals et al., 2011).

## Genetic complementation of patient-specific SNM

For this purpose, we generated a lentivirus (LV) carrying the correct cDNA for the *HGSNAT* gene (Fig. 4) under a CMV promoter. The vector also expressed GFP downstream of an IRES element. We subjected SNMs transduced with LV.CMV.HGSNAT.ires.GFP to a differentiation protocol for 3, 6 and 9 weeks. A LV carrying only the *GFP* gene was used as a control. SNMs showed GFP expression for both LVs after one week of transduction (Fig. 4). After 9 weeks neurons showed GFP expression (Fig. 4C) and a high activity of the enzyme, between 50- and 150-fold increase compared to the values of the WT transduced with the LV.CMV.GFP control vector (Fig 4D-E) indicating a long-term sustained activity.

## Storage alterations in lysosomes of iPSC-derived neurons detected by transmission electron microscopy (TEM)

Lysosome analysis by TEM revealed statistically significant differences in lysosomes size among control and patients, both in fibroblasts and in neurons at 3, 6 and 9 weeks. Images showed typical lysosomes for all control cultures while patients' samples showed larger and different vesicles, with an empty-like appearance (Fig. 5A). The vacuoles were analyzed by immunogold staining and they were LAMP1 positive (Fig. 5B), which led us to conclude that these vesicles were derived from lysosomes.

Regarding the size of lysosomes, an increase of around 30% for patient SFC6 and 40% for patient SFC7 was observed in fibroblasts. In neurons at 3 and 6 weeks, the increase of both patients' lysosomes was of around 20% compared to WT lysosomes. Nevertheless, at 9 weeks, the differences reached 80% for both patients (Fig. 5D).

When 9 weeks-neurons were analyzed after transduction with the *HGSNAT* cDNA, some of them presented lysosomes similar to those of WT (Fig. 5C). The proportion of corrected cells was similar to that of transduced cells, as shown by GFP fluorescence (data not shown).

## Neuronal activity in iPSC-derived neurons

A Calcium Fluorescence Imaging Assay was carried out to shed light on the differences in neuronal activity between healthy and affected cultures. We analyzed network activity of neurons derived from SFC-iPSC compared to WT-iPSC cultured for 3, 6 and 9 weeks.

Examples of typical calcium fluorescence traces at week 9 are provided in Fig. 6A. Traces are characterized by a background fluorescence combined with sharp peaks that correspond to depolarization events. The top traces in each condition correspond to GFP transduced cultures, and the bottom ones to both GFP and HGSNAT transduced cultures. WT-iPSC derived neurons showed repeating firing episodes of large amplitude that were not present in SFC6-iPSC neurons. SFC7-iPSC neurons showed a slightly richer activity than SFC6-iPSC, but the amplitude of neuronal firings was still low compared to WT-iPSC neurons. Interestingly, the inclusion of the HGSNAT gene significantly changed the activity of the mutant neurons, and both SCF6- and SCF7-iPSC neurons exhibited activity patterns that were similar to the WT ones.

The degradation in network activity due to the disorder, as well as its recovery after *HGSNAT* transduction, was quantified through the ‘network activity’ (see Experimental Procedures). We first considered the scenario of cultures that were not transduced to test the reliability of our analysis. As shown in Fig. 6B, network activity for WT healthy cultures was close to 1, meaning that on average most of the neurons were active. Activity was maintained within experimental error along the 6<sup>th</sup> and 9<sup>th</sup> weeks. This stability of WT measures allowed us to associate possible changes in activity solely to the disorder. Indeed, neuronal activity in the patients’ neurons showed a gradual decrease, although the loss of activity was sharper in SFC6-iPSC networks (Fig. 6B).

We next considered the cultures that were transduced by LV.CMV.GFP and LV.CMV.HGSNANT.iRes.GFP, and carried out identical measurements. The results are shown in Fig. 6C. For clarity, we compared the relative change in activity of the diseased cultures respect to their controls. LV.CMV.GFP transduced neurons showed the same trend as the untransduced ones within statistical error. Specifically, SFC7 and SFC6 neuronal

networks at week 9 decayed in activity by about 45% and 70%, respectively, compared to the control. On the other hand, *HGSNAT* transduced cultures showed a significant recovery already at week 3, with only a fall in activity by 20% in both SFC6 and SFC7. At week 6, activity was either maintained or started to improve, and by week 9 both culture types reached activity levels similar to the control case within statistical error.

To further characterize the differences between the transduced and the untransduced cultures, we also compared the fraction of active neurons present in the network. As shown in Fig. 6D, SFC6 and SFC7 exhibited at week 9 about 70% and 50% less active neurons than the control. However, after transduction with *HGSNAT*, both SFC6 and SFC7 maintained a fraction of active neurons comparable to the control.

We also provide in Fig. 6E the values for network activity at week 9. *LV.CMV.GFP* transduced SFC6 and SFC7 cultures show a significant loss of network activity, while the corresponding *LV.CMV.HGSNAT.ires.GFP* transduced counterparts reach activity levels indistinguishable from the normal condition.

### **Neuronal connectivity in iPS-derived neurons cultures**

Network connectivity was determined by identifying causal influences between neurons through Generalized Transfer Entropy (Stetter et al., 2012; Orlandi et al., 2014) (see Methods), an information-theoretic measure that allows us to draw the functional map of neuronal interactions in the network. Results are summarized in Fig. 7A. At week 3, the average degree, i.e. the average number of connections that a neuron establishes with others in the network, is lower in the affected cultures than in the WT, and pinpoints at a network-broad degradation due to the disorder already at early stages. Connectivity increases from week 3 to week 9 in WT cultures. The SFC6 cultures, however, exhibit a much lower connectivity at week 9, while the SFC7 cultures display connectivity in between WT and SFC6.

The action of *HGSNAT* transduction is well reflected in the connectivity, with both transduced SFC6 and SFC7 cultures showing a trend towards recovered connectivity

profiles, with an average connectivity similar to the one observed in the WT cultures. Indeed, transduced SFC6 cultures recover the average connectivity of WT cultures within statistical error, while SFC7 even surpasses it. We note that no statistical significance testing was performed in these analyses due to the low sample size (N=3 for each case) although the results were consistent within networks of the same preparation.

To illustrate the potential of the connectivity analysis to extract the connectivity features in healthy and affected cultures, we provide in Fig. 7B the connectivity maps of three representative neuronal cultures at weeks 3 and 9, and comparing the WT and SFC7 case before and after recovery. At week 3, WT and SFC7 cultures show similar connectivity with a slight increase in the number of connections for the WT culture. Neurons in the SFC7-*HGSNAT*-transduced culture are more connected than the other cultures. At week 9, the WT and SFC7-*HGSNAT*-transduced cultures display high connectivity and comparable network structures, with most of the neurons establishing a similar number of connections with other neurons. On the contrary, neurons in the standard SFC7 cultures are mostly disconnected. The few connected neurons show an inhomogeneous distribution of links, with a small subset of neurons retaining most of the connections.

### **Loss of active neurons could be explained by cell death via apoptosis and impairment in autophagy**

In order to explain the loss in the global number and in the percentage of active neurons in the SFC-iPSC-derived neural cell cultures, we analyzed the caspase 3 pattern in more than 100 cells for each line. WT-iPS derived neurons showed 2.1% of apoptotic cells, while SFC6-iPSC showed 13.6% and SFC7-iPSC 7.7% (Figure S5A-B).

LC3 distribution showed a clear increase in the signal in neurons derived from SFC-iPSC when compared to that of WT at any time point (Figure S5C).

## Discussion

This is, to our knowledge, the first report on the generation of a neuronal model of Sanfilippo C by reprogramming fibroblasts from two patients using the iPS cells technology. The generation of a neural model is relevant since the main features of the disease cannot be studied in fibroblasts and no animal model exist so far. The fact that we used samples from two patients validates the results and allows the detection of slight inter-individual differences.

The iPSC technology has been largely used to model different types of diseases, including those affecting the central nervous system (CNS) (Durnaoglu et al., 2011; Okano and Yamanaka, 2014). Some other LSD have been modeled using iPS cells (Huang et al., 2012; Panicker et al., 2012; Tiscornia et al., 2013; Trilck et al., 2013) and then used to differentiate to the cellular type of interest for each case. It is crucial in diseases affecting specific cell types in which patients' fibroblasts hardly can be considered a cellular model of the disease because of the significant differences in the cellular functions. An important routine test in iPSC technology consists in the introduction of the WT cDNA of the gene mutated in patients as an internal control to confirm that phenotypic features observed in the model are really due to the protein dysfunction.

We generated iPSC from two different Sanfilippo C patients. Enzyme activity is dispensable for reprogramming and iPSC maintenance, contrasting with other cases such as Fanconi, Sanfilippo B or Pompe. These iPSC showed a low lysosomal charge as proved by immunostaining with LAMP2, and also by TEM for any iPS cells. Moreover, for the control individual, a very low enzymatic activity was detected. Other lysosomal enzymatic activities were assayed and they also show lower levels when compared to fibroblasts (data not shown). A clear decrease in enzyme activity of WT-iPSC was previously described (Tiscornia et al., 2013). Taken together all these data suggest that the lysosomal machinery is working at a low level in the pluripotent cells or that the lysosomes are rapidly exocytosed in this cell type.

In this study, we have generated neurons, which also recapitulate the principal features of the disease. As expected, they showed lack of enzyme activity, same as

fibroblasts and iPSC. Neurons also displayed accumulation of GAGs and alteration of lysosomes. Except for the lack of enzyme activity, which occurs since the beginning, other alterations occur progressively: accumulation of GAGs, which with our technology, does not reach statistical significance vs control until 9 weeks (around 50% respect to control), and lysosomal alterations due to the HS storage evidenced by TEM / LAMP1 start to be noticed at 3-6 weeks, but are most evident at 9 weeks. Moreover, the appearance of the patients' lysosomes was clearly different from that of the control ones, as described for Sanfilippo B lysosomes in the mouse brain (Vitry et al., 2010).

This timeline of appearance of alterations closely mirrors disease progression seen in patients, and highlights current difficulties at predicting the extent of neurological decline: lack of enzyme activity is not predictive at all, accumulation of GAGs and alterations in lysosome size/number require invasive techniques that are only typically possible postmortem. The similarity of timelines in disease progression in patients and appearance of alterations in our model of SFC neuron cultures prompted us to investigate whether we could detect early functional alterations in our model predating the known pathological signs in patients.

For this purpose, we used calcium imaging to analyze neuronal function in patients' cells. An important decrease in spontaneous activity of SFC neurons compared to control ones was detected for patient SFC6 at 3 weeks of differentiation, and at 6 weeks for patient SFC7. Moreover, our results suggest that the fall in network activity is essentially caused by the gradual degeneration of neurons with time, a major feature of Sanfilippo's disease. The distinct behavior of SFC6 and SFC7 could be due to the particular features of their mutations. The mechanisms that lead to a halt in activity of these neurons remain unclear.

In order to shed some light into this issue, caspase 3 pattern was analyzed in the neurons and astrocytes cultures to clarify whether apoptosis is playing a role. Our results suggest a slight increase in the number of apoptotic cells for both patients, particularly SFC6, which presented a more pronounced decrease in neuronal activity. Furthermore, the increase of LC3 detection in the patients' neurons at all time points indicates an impairment in the autophagic pathway, which has been reported to contribute to neurodegeneration in LSD (Ballabio and Gieselmann, 2009). This block in autophagy can lead neurons to cell death (Settembre et al., 2008), correlating with the apoptotic

hypothesis. Further investigations should be carried out in order to tackle the cellular processes that induce this loss of active neurons. Our results suggest that calcium imaging is much more efficient and sensitive to detect alterations in the patient's neural cells, than other techniques that detect disease features only after 9 weeks of differentiation. Moreover, specific neuron activity can be measured.

It is well known that small differences in the network elements can have great impact in the network functionality and behavior. In our case, it could be small differences in the performance of all network elements (only measurable by TEM/immunofluorescence), or selective loss of a small number of neurons (caspase), or a combination of both. In any case, non-invasive analysis of the overall network connectivity appears as a sensitive, high-order surrogate

In addition to quantifying the changes in spontaneous activity, we carried out a complementary analysis in which we investigated the loss of network connectivity during the progression of the disorder. Recent advances in computational neuroscience show that major traits of the network's structural connectivity can be unveiled from the analysis of calcium imaging data (Stetter et al., 2012). Our results showed an increase in the WT-cultures connectivity, possibly signaling the gradual maturation of the network. SFC-derived cultures showed lower connectivity when compared to WT, with an inhomogeneous distribution and a small subset of neurons retaining most of the connections. This feature suggests that the disorder not only breaks down connectivity, but also disrupts the topology of the network, which could significantly affect its operation and functional traits. In initial steps of neuronal development (3 weeks), differences are not evident. However, SFC cultures showed a slight decrease in the connectivity when compared to WT cultures, especially SFC6. The relatively high connectivity of the SFC7 cultures compared to SFC6 may explain their concurrent high activity (Figure 6B), and highlights the lesser aggressiveness of the SFC7 mutation compared to SFC6 one. Hence, our calcium imaging data could be used in the future to investigate how the loss of neurons in Sanfilippo alters the network structural properties and risks its functionality.

Accordingly to these data, we conclude that the neural cell model hereby generated reproduces the major features of the Sanfilippo C syndrome, specially specific neuronal traits, and after the treatment with the WT cDNA as a proof of concept, it could be used as

a tool to test different possible therapeutic strategies. This is particularly relevant since neither animal nor cellular models exist for Sanfilippo C.

Our findings establish the usefulness of iPSC-based technology to identify early functional phenotypes that can shed light onto the mechanisms that lead to brain dysfunction in these patients, as well as providing valuable readouts for screening of compounds that prevent, rather than reverting, the onset of neurodegeneration. Moreover, the neuronal activity and connectivity analyses could be applicable to other neurodegenerative diseases in which iPSC-based models are available such as Parkinson's, Alzheimer, autism, and others. Further studies are needed to establish whether this technique would be able to detect differences in the network connectivity pattern before the disease onset and could lead to the development of *in vivo* analyses that could be used for the premature diagnosis of neurological-diseases affected patients and their follow-up after any treatment.

## **Experimental procedures**

### **Patients**

Studies were approved by the authors' Institutional Review Board and conducted under the Declaration of Helsinki. Patients were encoded to protect their confidentiality, and written informed consent obtained. The generation of human iPSCs was done following a protocol approved by the Spanish competent authorities (Commission on Guarantees concerning the Donation and Use of Human Tissues and Cells of the Carlos III Health Institute). The two patients, SFC6 and SFC7, have been previously described (Canals et al., 2011).

### **Generation of iPSC**

Fibroblasts from one healthy two years old individual (Advancell) and fibroblasts from two patients were infected with retroviruses carrying human cDNA coding for KLF4, SOX2, and OCT4 or these three cDNAs and c-Myc as previously described (Raya et al.,

2009). Fibroblasts were maintained in DMEM (Sigma) supplemented with 10% FBS (Life Technologies) and 1% PenStrep (Life Technologies) before the infection. After the infection, fibroblasts were plated on irradiated human foreskin fibroblasts (ATCC) and maintained with hESC medium for 4-12 weeks until iPSc colonies appeared. One clone from each cell line was grown on iHFF until pass 15.

## **iPSC validation**

AP staining was performed using the Alkaline Phosphatase Blue Membrane Substrate Solution (Sigma). For immunocytochemistry cells were grown on HFF feeders layer for 6-10 days and then fixed in 4% PFA for 10 minutes. After EB formation, differentiation to the 3 germ layers was carried out. For endoderm, EBs were plated on 6-well plates previously treated with matrigel (BD Biosciences) for 1 hour at room temperature, and maintained for 28 days with EB medium. The same procedure was used for mesoderm, but using EB medium with 0.5 mM of Ascorbic Acid. For ectoderm differentiation, EBs were maintained in suspension for 10 days with N2B27 medium supplemented with FGF2, prepared as previously described (Sánchez-Danés et al., 2012). Then, EBs were plated on 6-well plates previously treated with matrigel for 1 hour at room temperature, and maintained for 21 days with N2B27 medium without FGF supplementation. Differentiated cells were fixed in 4% PFA for 10 minutes. Antibodies used are shown in Table S1. For nucleus staining DAPI (Invitrogen) at 0.5 µg/ml was used. The slides were mounted with PVA:DABCO mounting medium. Images were acquired with an SP2 confocal system (Leica) and analyzed with ImageJ software.

RT-qPCR analysis was done as previously described (Sánchez-Danés et al., 2012). All results were normalized to the average expression of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Transcript-specific primers used are shown in Table S2.

For karyotyping, iPSc cells were grown on matrigel and treated with colcemid (Life Technologies) at a final concentration of 20 ng/ml. Karyotyping analysis were carried out by Prenatal Genetics S.L. (Barcelona).

For promoter methylation, testing reprogramming genes integration and sequencing to prove that patients' iPSc were carrying mutations in HGSNAT gene, DNA was isolated using the QIAamp DNA Mini Kit (QIAGEN) following manufacturer's instructions. Bisulfite conversion of the promoters was carried out using the Methylamp DNA modification kit (Epigentek). Five clones of each promoter for each cell line were analyzed by sequencing. For testing genes integration, primers used are shown in Suppl. Table 2. For testing mutations, primers and sequencing conditions used were the same as previously described (Canals et al., 2011).

Severe combined immunodeficient (SCID) beige mice (Charles River Laboratories) were used to test the teratoma induction capacity of patient-specific iPS cells essentially as described (Raya et al., 2008). All animal experiments were conducted following experimental protocols previously approved by the Institutional Ethics Committee on Experimental Animals, in full compliance with Spanish and European laws and regulations.

### **HGSNAT activity and GAGs storage**

Fibroblasts, iPS cells, SNMs and neuronal from all the lines were harvested and enzyme activity was assayed as previously described (Canals et al., 2011). For GAGs quantification, cells were harvested using a cell scraper and extraction of GAGs content from cells was carried out as previously described (Tolar et al., 2011) GAGs quantification were performed using the Blyscan Assay Kit (Biocolor Ltd.) following the manufacturer's instructions.

### **HGSNAT splicing analysis**

RNA extraction from iPS and neural cells, RT-PCR, primers and cycloheximide conditions used were previously described (Canals et al., 2011).

## **iPSC differentiation to neural cells**

SNMs were obtained as previously described (Cho et al., 2008). SNMs were fixed in 4% PFA for 2 hours and characterized by immunostaining. Antibodies used are shown in Suppl. Table 1. For nucleus staining DAPI (Invitrogen) at 5 µg/ml was used. Mounting medium and imaging analysis were done as for in vitro differentiation test.

SNMs obtained from iPS-WT, iPS-SFC6 and iPS-SFC7, maintained in suspension, were then plated on slide-flasks, 6-well plates, 35mm plates or 10mm plates previously treated with matrigel for 1 hour at room temperature, and differentiated for 3, 6 or 9 weeks with N2B27 medium, without FGF supplementation, to obtain the neuronal cultures. The correct differentiation was assessed by immunostaining. Antibodies, DAPI and mounting medium were the same that for the ectoderm assay (see Supplemental Experimental Procedures).

## **Lentiviral production and transduction**

The complete WT cDNA of the HGSNAT gene into a vector carrying an IRES-GFP construction was obtained from Simone Merlin and Antonia Follenzi (Laboratory of Histology, Department of Medical Service, University of Piemonte Orientale “A. Avogadro”, Novara, Italy) and named pRRL-SIN-PPT-CMV-HGSNAT-ires-GFP-WPRE. High-titer VSV-pseudotyped LV stocks were produced in 293T cells by calcium phosphate-mediated transient transfection of the transfer vector pRRL-SIN-PPT-CMV-HGSNAT-ires-GFP-WPRE, the late generation packaging construct pMDL and the VSV envelope-expressing construct pMD2.G, and purified by ultra-centrifugation as previously described (Consiglio et al., 2004). Tritation of the virus was carried out using 293T cells and analyzed by FACS, obtaining  $2.9 \times 10^9$  TU/mL for the control virus and  $5.4 \times 10^8$  TU/mL for the virus carrying the correction. Transduction was carried out in the SNM step, using 1 µl of GFP virus and 3µl of HGSNAT-GFP virus.

## **Transmission electron microscopy**

Neuronal cultures were grown on 10mm plates as described above for 3, 6 and 9 weeks. Fibroblasts and iPSc were grown until subconfluence. Cells were fixed with 2.5% glutaraldehyd for 90 minutes and then collected and treated with 0.1M osmium for 2 hours. Dehydration was carried out with acetone and blocks were obtained with Epon. For cryoultramicrotomy and posterior immunogold staining, neurons, fibroblasts and iPSc were fixed with 4% PFA and 0.1% glutaraldehyd, washed with 0.15M glycine, treated with 12% gelatin, cryoprotected with 2.3M sacarose and cyrofixed with liquid nitrogen. Antibodies used are shown in Tabel S1. Images were acquired with a transmission microscopy JEOL1010 and analyzed with ImageJ software.

## **LC3, LAMP2 and Caspasa 3 immunostaining**

iPS cells cultures or neuronal cultures differentiated for 6 or 9 weeks were fixed in ice-cold 100% methanol or 4% PFA. Antibodies used are shown in Suppl. Table 1. DAPI, mounting medium and image analyses were performed as described in Supplemental Experimental Procedures.

## **Calcium Fluorescence Imaging**

We used calcium imaging (Orlandi et al., 2013; Takahashi et al., 2007; Takahashi et al., 2010; Tibau et al., 2013) to evaluate the differences in spontaneous activity between healthy and Sanfilippo's affected neuronal cultures. This technique uses a fluorescence probe to reveal the fast increase of calcium levels inside neurons upon firing. Several studies have shown that the emitted fluorescence captures well the action potentials elicited by the cells (Smetters et al., 1999; Sasaki et al., 2008; Chen et al., 2013). Although measurements of neuronal activity are more precise with patch clamp, electrodes or other techniques, calcium imaging allows the monitoring of a large population of neurons, simultaneously and non-invasively, which makes it particularly suitable for whole network analyses.

The studied cultures were prepared as described before. For each condition (WT, SFC6, and SFC7) we prepared 27 identical cultures. Nine of them were later transduced with HGSNAT-GFP, nine were transduced with GFP only and the rest remained untransduced. We then investigated the behavior of the cultures at 3, 6, and 9 weeks after plating. For each preset time, we measured the spontaneous activity in 3 HGSNAT-GFP-transfected, in 3 GFP-transfected and in 3 untransfected cultures, for 30 minutes.

Prior to imaging, the culture dish to study was first gently washed with 4ml PBS at room temperature to remove the original culture medium. Next, we incubated the cultures for 30 min in a solution that contained 1ml of recording medium (RM, consisting of 128 mM NaCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 45 mM sucrose, 10 mM glucose, and 0.01 M HEPES; treated to pH 7.4) and 4 µg/ml of the cell-permeant calcium sensitive dye Fluo-4-AM. At the end of incubation we washed the culture with 2 ml of fresh RM to remove residual free Fluo-4. This medium was discarded to place 4 ml of fresh RM, the final medium for actual recordings.

The culture dish was mounted on a Zeiss inverted microscope equipped with a CMOS camera (Hamamatsu Orca Flash 2.8) and an arc lamp for fluorescence. Grey-scale images of neuronal activity were acquired at intervals of 50 ms, with a size of 960x720 pixels and a spatial resolution of 2.90 µm/pixel. The latter settings provided a final field of view of 2.8x2.1 mm that contained between 150 and 300 neurons.

For each recorded culture we extracted the fluorescence amplitude of all neurons in the field of view as a function of time. Neuronal firing events were detected as a fast rise in the fluorescence signal, as illustrated in the traces of Figure 5A.

Fluorescence time series were finally analyzed to determine two network activity descriptors, namely the fraction of active neurons and the average activity of the network. Active cells were those that showed at least one firing event along the 30 min of recording. The fraction of active neurons was therefore the ratio between active cells and the total population monitored. The average activity of the network was determined as the total number of firing events observed along the recording divided by the total number of

monitored cells. The latter quantity is our main estimation of the health of the neuronal network.

Data was finally averaged among the 3 replicates of each time-point and transduction. Hence, our results are based on a statistics of at least 450 neurons per time-point and transduction.

## Network Reconstruction

Network reconstruction was carried out by identifying causal influences between neurons through Transfer Entropy (Schreiber et al., 2000; Vicente et al., 2011; Stetter et al., 2012). Transfer entropy (TE) is an information-theoretic measure that identifies the flow of information between two time traces. The measure is model free, i.e. does not require previous knowledge of the dynamics of the system, and is able to detect linear and non-linear interactions between any pair of traces. In a more formal description of TE, one considers two signals X and Y (any two neuronal fluorescence traces in our case), with the goal to assess the influence of X on Y (X->Y). TE measures the amount of uncertainty reduced in predicting the future of Y by taking into account both the pasts of Y and X, rather than the past of Y alone. Mathematically, this operation can be written as

$$TE_{X \rightarrow Y} = \sum P(Y_t, Y_t^k, X_t^k) \log \left( \frac{P(Y_t | Y_t^k, X_t^k)}{P(Y_t | Y_t^k)} \right)$$

where  $Y_t$  denotes the value of Y at time t and  $Y_t^k$  the past k-th values of Y.  $P(\cdot)$  is the probability of observing that particular sequence, whereas  $P(\cdot|\cdot)$  is the conditional probability. The sum is performed over all possible values of  $Y_t, Y_t^k, X_t^k$ .

There are different variations of the TE measure depending on the data under analysis, e.g. fMRI (Honey et al., 2007), MEG (Wibral, et al., 2011), spike trains (Ito et al., 2011), or fluorescence data as in our case. Here we use a version named ‘Generalized Transfer Entropy’ (GTE) (Stetter et al., 2012; Orlandi et al., 2014) that has been specifically developed for neuronal fluorescence signals (see Stetter et al., 2012 for details).

The analysis of the recorded spontaneous activity traces within the context of GTE was as follows. We initially computed the first derivative of the fluorescence trace, and the resulting values binned with  $n=3$  intervals for TE computation. TE was then applied to all pairs of neurons in the network. The assigns a score for every neuronal pair, however, only those pairs with a score above a given significance level were considered as a putative connection. Significance was established by comparing the raw Transfer Entropy scores and the bootstrapped versions (to account for bias) and those obtained by shuffling the data from only the presynaptic neuron. Bootstrapped versions were obtained by generating surrogates of the fluorescence data for every neuronal pair while preserving temporal correlations between the pairs. A paired Z-test was performed with the bootstrapped and shuffled scores, and only those values above the 97.5<sup>th</sup> ( $p < 0.002$ ) percentile were considered as putative connections.

Since the number of neurons monitored is different from culture to culture, the reconstructions were performed for a subset of  $N=85$  randomly chosen neurons in each network, and the resulting connectivity was averaged over 100 permutations. The final connectivity measure was then averaged over 2-3 different cultures.

## Statistical analysis

Differences between conditions respect to controls were evaluated using a U-Mann-Whitney non-parametric test, and statistical significance was set at  $p < 0.05$ , except for the network reconstruction experiment (see above).

## References

- Ballabio, A., and Gieselmann, V. (2009). Lysosomal disorders: from storage to cellular damage. *Biochim. Biophys. Acta* 1793, 684–96.
- Canals, I., Elalaoui, S.C., Pineda, M., Delgadillo, V., Szlago, M., Jaouad, I.C., Sefiani, A., Chabás, A., Coll, M.J., Grinberg, D., and Vilageliu, L. (2011). Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles. *Clin. Genet.* 80, 367–74.

- Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., and Kim, D.S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*. 499, 295-300.
- Cherry, A.B., and Daley, G.Q. (2013). Reprogrammed cells for disease modeling and regenerative medicine. *Annu. Rev. Med.* 64, 277-90.
- Cho, M.S., Hwang, D.Y., and Kim, D.W. (2008). Efficient derivation of functional dopaminergic neurons from human embryonic stem cells on a large scale. *Nat. Protoc.* 3, 1888-94.
- Consiglio, A., Gritti, A., Dolcetta, D., Follenzi, A., Bordignon, C., Gage, F.H., Vescovi, A.L., and Naldini, L. (2004). Robust in vivo gene transfer into adult mammalian neural stem cells by lentiviral vectors. *Proc. Natl. Acad. Sci. U. S. A.* 101, 14835-40.
- Durnaoglu, S., Genc, S., and Genc, K. (2011). Patient-specific pluripotent stem cells in neurological diseases. *Stem Cells Int.* 2011, 212487.
- Falk, A., Koch, P., Kesavan, J., Takashima, Y., Ladewig, J., Alexander, M., Wiskow, O., Taylor, J., Trotter, M., Pollard, S., Smith, A., and Brüstle, O. (2012) Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. *PLoS One* 7, e29597.
- Fan, X., Zhang, H., Zhang, S., Bagshaw, R.D., Tropak, M.B., Callahan, J.W., and Mahuran, D.J. (2006). Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). *Am. J. Hum. Genet.* 79, 738-44.
- Higuchi, T., Kawagoe, S., Otsu, M., Shimada, Y., Kobayashi, H., Hirayama, R., Eto, K., Ohashi, T., Nakauchi, H., and Eto, Y. (2014). The generation of induced pluripotent stem cells (iPSCs) from patients with infantile and late-onset types of Pompe disease and the effects of treatment with acid- $\alpha$ -glucosidase in Pompe's iPSCs. *Mol. Genet. Metab.* 112, 44-8.
- Honey, C.J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci.* 104, 10240-5.
- Hřebíček, M., Mrázová, L., Seyrantepe, V., Durand, S., Roslin, N., Nosková, L., Hartmannová, H., Ivánek, R., Cizková, A., Poupětová, H., Sikora, J., Urinovská, J., Stránecký, V., Zeman, J., Lepage, P., Roquis, D., Verner, A., Ausseil, J., Beesley, C., Maire, I., Poorthuis, B., van de Kamp, J., van Diggelen, O., Wevers, R., Hudson, T., Fujiwara, T., Majewski, J., Morgan, K., Knoch, S., and Pshezhetsky, A. (2006). Mutations in TMEM76 cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Am. J. Hum. Genet.* 79, 807-19.

- Huang, H.-P., Chen, P.-H., Hwu, W.-L., Chuang, C.-Y., Chien, Y.-H., Stone, L., Chien, C.-L., Li, L.-T., Chiang, S.-C., Chen, H.-F., Ho, H.-N., Chen, C.-H., and Kuo, H.-C. (2011). Human Pompe disease-induced pluripotent stem cells for pathogenesis modeling, drug testing and disease marker identification. *Hum. Mol. Genet.* 20, 4851–64.
- Huang, H.-P., Chuang, C.-Y., and Kuo, H.-C. (2012). Induced pluripotent stem cell technology for disease modeling and drug screening with emphasis on lysosomal storage diseases. *Stem Cell Res. Ther.* 3, 34.
- Inoue, H., Nagata, N., Kurokawa, H., and Yamanaka, S. (2014). iPS cells: a game changer for future medicine. *EMBO J.* 33, 409-17.
- Ito, S., Hansen, M.E., Heiland, R., Lumsdaine, A., Litke, A.M., and Beggs, J.M. (2011). Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS One* 6, e27431.
- Klein, U., Kresse, H., and von Figura, K. (1978). Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 75, 5185–9.
- Lamanna, W.C., Lawrence, R., Sarrazin, S., and Esko, J.D. (2011). Secondary storage of dermatan sulfate in Sanfilippo disease. *J. Biol. Chem.* 286, 6955–62.
- Lemonnier, T., Blanchard, S., Toli, D., Roy, E., Bigou, S., Froissart, R., Rouvet, I., Vitry, S., Heard, J.M., and Bohl, D. (2011). Modeling neuronal defects associated with a lysosomal disorder using patient-derived induced pluripotent stem cells. *Hum. Mol. Genet.* 20, 3653–66.
- Maetzel, D., Sarkar, S., Wang, H., Abi-Mosleh, L., Xu, P., Cheng, A.W., Gao, Q., Mitalipova, M., and Jaenisch, R. (2014). Genetic and chemical correction of cholesterol accumulation and impaired autophagy in hepatic and neural cells derived from Niemann-Pick type C patient-specific iPS cells. *Stem cell reports* 2, 866–80.
- Mazzulli, J.R., Xu, Y.-H., Sun, Y., Knight, A.L., McLean, P.J., Caldwell, G.A., Sidransky, E., Grabowski, G.A., and Krainc, D. (2011). Gaucher disease glucocerebrosidase and  $\alpha$ -synuclein form a bidirectional pathogenic loop in synucleinopathies. *Cell* 146, 37–52.
- Neufeld, E.F., and Muenzer, J. (2001). The Mucopolysaccharidoses.
- Okano, H., and Yamanaka, S. (2014). iPS cell technologies: significance and applications to CNS regeneration and disease. *Mol. Brain* 7, 22.
- Orlandi, J.G., Soriano, J., Alvarez-Lacalle, E., Teller, S., and Casademunt, J. (2013). Noise focusing and the emergence of coherent activity in neuronal cultures. *Nat. Phys.* 9, 582–90.
- Orlandi, J.G., Stetter, O., Soriano, J., Geisel, T., and Battaglia, D. (2014). Transfer entropy reconstruction and labeling of neuronal connections from simulated calcium imaging. *PLoS One* 9, e98842.

- Panicker, L., Miller, D., Park, T.S., Patel, B., Azevedo, J.L., Awad, O., Masood, M.A., Veenestra, T.D., Goldin, E., Stubblefield, B.K., Tayebi, N., Polumuri, S.K., Vogel, S.N., Sidransky, E., Zambidis, E.T., and Feldman, R.A. (2012). Induced pluripotent stem cell model recapitulates pathologic hallmarks of Gaucher disease. *Proc. Natl. Acad. Sci.* 109, 18054–9.
- Raya, A., Rodríguez-Pizà, I., Arán, B., Consiglio, A., Barri, P.N., Veiga, A., and Izpisúa Belmonte, J.C. (2008). Generation of cardiomyocytes from new human embryonic stem cell lines derived from poor-quality blastocysts. *Cold Spring Harb. Symp. on Quant. Biol.* 73, 127–35.
- Raya, A., Rodríguez-Pizà, I., Guenechea, G., Vassena, R., Navarro, S., Barrero, M.J., Consiglio, A., Castellà, M., Río, P., Sleep, E., González, F., Tiscornia, G., Garreta, E., Aasen, T., Veiga, A., Verma, I.M., Surrallés, J., Bueren, J., and Izpisúa Belmonte, J.C. (2009). Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* 460, 53–9.
- Sánchez-Danés, A., Consiglio, A., Richaud, Y., Rodríguez-Pizà, I., Dehay, B., Edel, M., Bové, J., Memo, M., Vila, M., Raya, A., and Izpisúa Belmonte, J.C. (2012). Efficient generation of A9 midbrain dopaminergic neurons by lentiviral delivery of LMX1A in human embryonic stem cells and induced pluripotent stem cells. *Hum. Gene Ther.* 23, 56–69.
- Sasaki, T., Takahashi, N., Matsuki, N., and Ikegaya, Y. (2008). Fast and accurate detection of action potentials from somatic calcium fluctuations. *J. Neurophysiol.* 100, 1668–76.
- Schreiber, T. (2000). Measuring information transfer. *Phys. Rev. Lett.* 85, 461–4
- Settembre, C., Fraldi, A., Jahreiss, L., Spampinato, C., Venturi, C., Medina, D., de Pablo, R., Tacchetti, C., Rubinsztein, D.C., and Ballabio, A. (2008). A block of autophagy in lysosomal storage disorders. *Hum. Mol. Genet.* 17, 119–29.
- Smetters, D., Majewska, A., and Yuste, R. (1999). Detecting action potentials in neuronal populations with calcium imaging. *Methods* 18, 215–21.
- Stetter, O., Battaglia, D., Soriano, J., and Geisel, T. (2012). Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput. Biol.* 8, e1002653.
- Takahashi, N., Sasaki, T., Usami, A., Matsuki, N., and Ikegaya, Y. (2007). Watching neuronal circuit dynamics through functional multineuron calcium imaging (fMCI). *Neurosci. Res.* 58, 219–25.
- Takahashi, N., Takahara, Y., Ishikawa, D., Matsuki, N., and Ikegaya, Y. (2010). Functional multineuron calcium imaging for systems pharmacology. *Anal. Bioanal. Chem.* 398, 211–8.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–72.

- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–76.
- Tibau, E., Valencia, M., and Soriano, J. (2013). Identification of neuronal network properties from the spectral analysis of calcium imaging signals in neuronal cultures. *Front. Neural Circuits* 7, 199.
- Tiscornia, G., Vivas, E.L., Matalonga, L., Berniakovich, I., Barragán Monasterio, M., Eguizábal, C., Gort, L., González, F., Ortiz Mellet, C., García Fernández, J.M., Ribes, A., Veiga, A., and Izpisua Belmonte, J.C. (2013). Neuronopathic Gaucher's disease: induced pluripotent stem cells for disease modelling and testing chaperone activity of small compounds. *Hum. Mol. Genet.* 22, 633–45.
- Tolar, J., Park, I., Xia, L., Lees, C.J., Peacock, B., Webber, B., McElmurry, R.T., Eide, C.R., Orchard, P.J., Kyba, M., Osborn, M.J., Lund, T.C., Wagner, J.E., Daley, G.Q., and Blazar, B.R. (2011). Hematopoietic differentiation of induced pluripotent stem cells from patients with mucopolysaccharidosis type I (Hurler syndrome). *Blood* 117, 839–47.
- Trilck, M., Hübner, R., Seibler, P., Klein, C., Rolfs, A., and Frech, M.J. (2013). Niemann-Pick type C1 patient-specific induced pluripotent stem cells display disease specific hallmarks. *Orph. J. Rare Dis.* 8, 144.
- Trounson, A., Shepard, K.A., and DeWitt, N.D. (2012). Human disease modeling with induced pluripotent stem cells. *Curr. Opin. Genet.* 22, 509-16.
- Valstar, M.J., Ruijter, G.J.G., van Diggelen, O.P., Poorthuis, B.J., and Wijburg, F.A. (2008). Sanfilippo syndrome: a mini-review. *J. Inherit. Metab. Dis.* 31, 240–52.
- Vicente, R., Wibral, M., Lindner, M., and Pipa, G. (2011). Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* 30, 45-67.
- Vitry, S., Bruyère, J., Hocquemiller, M., Bigou, S., Ausseil, J., Colle, M.-A., Prévost, M.-C., and Heard, J.M. (2010). Storage vesicles in neurons are related to Golgi complex alterations in mucopolysaccharidosis IIIB. *Am. J. Pathol.* 177, 2984–99.
- Walkley, S.U. (2004). Secondary accumulation of gangliosides in lysosomal storage disorders. *Semin. Cell Dev. Biol.* 15, 433–44.
- Wibral, M., Rahm, B., Rieder, M., Lindner, M., Vicente, R., and Kaiser, J. (2011). Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks. *Prog. Biophys. Mol. Biol.* 105, 80-97.

## Figure titles and legends

### Figure 1. iPS cells validation.

(A, left) Fibroblasts before being transduced with retroviruses carrying reprogramming factors. (Center) ESC-like colonies obtained after reprogramming. (Right) AP staining of the ESC-like colonies.

(B) Bisulphite genomic sequencing of the OCT4 and NANOG promoters showing demethylation in the 3 iPSC lines.

(C) RT-qPCR analyses of the expression levels of retroviral-derived reprogramming and endogenous genes.

(D) Karyotype of each line identical to that of parent fibroblasts [including the balanced Robertsonian translocation der(13;14)(q10;q10) of SFC6].

(E) Expression of pluripotency markers(OCT4, SOX2, NANOG –green-; SSEA3, SSEA4, TRA-1-81 –red-). Scale bar: 100  $\mu$ m.

(F) *In vitro* differentiation of the 3 iPS cell lines stained with different specific markers for each germ layer. Endoderm: a-fetoprotein (green), FOX2A (red); mesoderm: GATA4 (green), a-SMA (red); ectoderm: TUJ1 fetoprotein (green), GFAP (red). Scale bar: 100  $\mu$ m.

(G) Teratoma formation of the 3 iPS cell lines stained with different specific markers for each germ layer. Endoderm: a-fetoprotein (green), FOX2A (red); mesoderm: SOX9 (green), CS (red); ectoderm: TUJ1 fetoprotein (green), GFAP (red). Scale bar: 100  $\mu$ m.

### Figure 2. HGSNAT activity and GAGs storage.

(A-D) Enzyme activity expressed in nmol/h·mg protein for each differentiation stage. \*\* p<0.01 (WT vs. patients). \*\*\* p<0.001 (WT vs. patients).

(E-F) GAGs levels expressed in  $\mu$ g of GAGs/ $\mu$ g of DNA, in fibroblasts and neural cells (differentiated for 9 weeks).\*\* p<0.01 (WT vs. patients).

### Figure 3. Immunofluorescence analysis of SNM and differentiated neurons and astrocytes.

(A) SNM staining with specific markers for neural stem cells (Nestin and PAX6 –green-; Ki67 and SOX2 –red-). Scale bar: 100  $\mu$ m.

(B) Neurons and astrocytes stained with typical markers for these cell types (TUJ1 –green- for neurons and GFAP –red- for astrocytes). Scale bar: 100  $\mu$ m.

**Figure 4.** Transduction and enzyme activity recovery with a lentiviral vector bearing the WT *HGSNAT* cDNA.

(A) pRRL-SIN-PPT-CMV-HGSNAT-ires-GFP-WPRE vector.

(B) GFP fluorescence in SNMs, 3 days after transduction (upper image) and neural cells differentiated for 9 weeks (lower image).

(C) Enzyme activity expressed in log nmol/h·mg protein in neural cells after 9 weeks of differentiation. \*\*  $p < 0.01$  (transduced vs. corresponding culture transduced with empty vector).

(D) Enlarged graph of cultures transduced with empty vector. \*  $p < 0.05$  (WT vs. patients).

**Figure 5.** Electron microscopy analysis of lysosomes.

(A) Accumulation vesicles of neural cells after 9 weeks differentiation (black arrows). Scale bar: 2  $\mu\text{m}$ .

(B) Immunogold staining of accumulation vesicles with LAMP1 (black arrows indicate gold particles). Scale bar: 1  $\mu\text{m}$ .

(C) Lysosome-like vesicles in SFC6-iPS derived neural cells treated with the lentivirus carrying the *HGSNAT* cDNA (black arrows). Scale bar: 2  $\mu\text{m}$ .

(D) Comparative lysosome size in fibroblasts or in neural cells at 3, 6 or 9 weeks of differentiation. 200 lysosomes for each sample were analyzed. \*\*\*  $p < 0.001$  (patients vs. WT -arbitrarily set to 100%-).

**Figure 6.** Fluorescence calcium imaging recordings.

(A) Representative traces of neuronal spontaneous activity at week 9 of differentiation. Top traces correspond to GFP transduced cultures; bottom ones to both GFP and *HGSNAT* transduced cultures. Sharp increases in the fluorescence signal are neuronal firings. Arrow heads depict those firings events of low amplitude.

(B) Network activity (total number of firings per neuron monitored) for untransduced cultures, and for differentiation stages at 3, 6, and 9 weeks.

(C) Relative change in activity of both GFP and *HGSNAT* transduced SFC6 and SFC7 neuronal networks respect to their controls.

(D) Fraction of active neurons at week 9 of both GFP and *HGSNAT* transduced WT, SFC6 and SFC7 neuronal networks. Asterisks (\*) indicate statistical significance between patients and WT, with \* for  $p < 0.05$  and \*\* for  $p < 0.01$ . The hash symbols (#) indicate

statistical significance between GFP- and *HGSNAT-transduced cultures*, with # for  $p < 0.05$  and ## for  $p < 0.01$ .

(E) Corresponding network activity values at week 9 of both GFP and *HGSNAT* transduced WT, SFC6 and SFC7 neuronal networks. \*\*  $p < 0.01$  (patients vs. WT). ##  $p < 0.01$  (GFP- vs. *HGSNAT-transduced cultures*).

**Figure 7.** Connectivity analysis.

(A) Normalized neuronal connectivity for SFC6 (left) and SFC7 (right) patients, defined as the fraction of connections that a neuron, on average, establishes with the rest of the network. WT cultures at week 3 connect with about 8% of other neurons, a connectivity that increases to 18% at week 9. Standard SFC6 cultures exhibit a substantial decay in connectivity, while SFC7 cultures maintain their connectivity but below WT values. *HGSNAT* transduction in these cultures notably increases their connectivity. SFC6 treated cultures reach levels of connectivity at week 9 similar to WT, while SFC7 cultures even exceeds the WT values. For the analysis, each connectivity measure was averaged over 100 subsets of  $N=85$  neurons, randomly chosen within the same culture, and finally averaged over 2-3 different culture preparations. Error bars are not shown for clarity.

(B) Structure of representative SFC7 networks at weeks 3 (top) and 9 (bottom) as reconstructed through Transfer Entropy. Both WT (left) and *HGSNAT-transduced cultures* (right) show a uniform increase in connectivity, i.e. most of the neurons share a similar number of connections. The diseased networks (center) display a different behavior in which a small number of neurons retain most of the connections, particularly at week 9. This 'topological shift' highlights severe functional changes that may compromise the operability of the network. In all the depicted networks, circles show the actual position of the neurons in the culture, and color-coded according to their relative connectivity. For clarity, the number of neurons in each network is limited to 150, which are randomly chosen from the original set, and only those connections with  $p < 0.002$  are represented.

## Supplementary Tables and Supplementary Figures Titles and Legends

**Table S1.** Primary and secondary antibodies used.

Mouse IgG anti-OCT3/4	Santa Cruz Biotechnology	1:60
Rabbit IgG anti-SOX2	Fisher Scientific	1:100
Goat IgG anti-NANOG	R&D Systems	1:25
Rat IgM anti-SSEA3	Hybridoma Bank	1:3
Mouse IgG anti-SSEA4	Hybridoma Bank	1:3
Mouse IgM anti-TRA-1-81	Millipore	1:200
Rabbit IgG anti- $\alpha$ -1-fetoprotein	Dako	1:400
Goat IgG anti-FOXA2	R&D Systems	1:50
Rabbit IgG anti-GATA4	Santa Cruz Biotechnology	1:50
Mouse IgG anti- $\alpha$ -SMA	Sigma	1:400
Mouse IgG anti-CS	Sigma	1:400
Goat IgG anti-SOX9	R&D Systems	1:20
Mouse IgG anti-TUJ1	Covance	1:500
Rabbit IgG anti-GFAP	Dako	1:500
Mouse IgG anti-Ki67	Dako	1:100
Rabbit IgG anti-Nestin	Chemicon	1:250
Rabbit IgG anti-PAX6	Covance	1:100
Mouse IgG anti-SOX2	R&D Systems	1:50
Mouse IgG anti-LAMP1	BD Biosciences	1:10
Mouse IgG anti-LAMP2	Hybridoma Bank	1:100
Rabbit IgG anti-LC3	Cell Signaling	1:100
Rabbit IgG anti-Caspasa 3	Cell Signaling	1:400
Donkey anti-rabbit IgG Cy2	Jackson ImmunoResearch	1:200
Donkey anti-goat IgG Cy3	Jackson ImmunoResearch	1:200
Donkey anti-mouse IgG Cy2	Jackson ImmunoResearch	1:200
Donkey anti-rabbit IgG Cy3	Jackson ImmunoResearch	1:200
Donkey anti-mouse IgG Cy3	Jackson ImmunoResearch	1:200
Donkey anti-rat IgM Cy3	Jackson ImmunoResearch	1:200
Donkey anti-goat IgG Cy2	Jackson ImmunoResearch	1:200
Donkey anti-mouse IgM Cy3	Jackson ImmunoResearch	1:200
Donkey anti-mouse IgG A488	Jackson ImmunoResearch	1:200
Donkey anti-rabbit IgG A488	Jackson ImmunoResearch	1:200
Donkey anti-mouse IgM Cy2	Jackson ImmunoResearch	1:200
Goat anti-mouse IgG 12 $\eta$ m	Jackson ImmunoResearch	1:30

**Table S2.** Primers used.

qPCR Total OCT4 Forward	5'-GGAGGAAGCTGACAACAATGAAA-3'
qPCR Total OCT4 Reverse	5'-GGCCTGCACGAGGGTTT-3'
qPCR Total SOX2 Forward	5'-TGCGAGCGCTGCACAT-3'
qPCR Total SOX2 Reverse	5'-TCATGAGCGTCTTGGTTTTCC-3'
qPCR Total KLF4 Forward	5'-CGAACCCACACAGGTGAGAA-3'
qPCR Total KLF4 Reverse	5'-GAGCGGGCGAATTTCCAT-3'
qPCR Total c-MYC Forward	5'-AGGGTCAAGTTGGACAGTGTCA-3'
qPCR Total c-MYC Reverse	5'-TGGTGCATTTTTCGGTTGTTG-3'
qPCR Trans OCT4 Forward	5'-TGGACTACAAGGACGACGATGA-3'
qPCR Trans OCT4 Reverse	5'-CAGGTGTCCC GCCATGA-3'
qPCR Trans SOX2 Forward	5'-GCTCGAGGTTAACGAATTCATGT-3'
qPCR Trans SOX2 Reverse	5'-GCCCCGGCGGCTTCA-3'
qPCR Trans KLF4 Forward	5'-TGGACTACAAGGACGACGATGA-3'
qPCR Trans KLF4 Reverse	5'-CGTCGCTGACAGCCATGA-3'
qPCR Trans c-MYC Forward	5'-TGGACTACAAGGACGACGATGA-3'
qPCR Trans c-MYC Reverse	5'-GTTCTGTGTTGGTGAAGCTAACGT-3'
qPCR NANOG Forward	5'-ACAACCTGGCCGAAGAATAGCA-3'
qPCR NANOG Reverse	5'-GGTCCCAGTCGGGTTTCCAC-3'
qPCR CRIPTO Forward	5'-CGGAACTGTGAGCACGATGT-3'
qPCR CRIPTO Reverse	5'-GGGCAGCCAGGTGTCATG-3'
qPCR REX1 Forward	5'-CCTGCAGGCGGAAATAGAAC-3'
qPCR REX1 Reverse	5'-GCACACATAGCCATCACATAAGG-3'
qPCR GAPDH Forward	5'-GCACCGTCAAGGCTGAGAAC-3'
qPCR GAPDH Reverse	5'-AGGGATCTCGCTCCTGGAA-3'
Bisulfite seq OCT4 Forward	5'-GGATGTTATTAAGATGAAGATAGTTGG-3'
Bisulfite seq OCT4 Reverse	5'-CCTAAACTCCCCTTCAAATCTATT-3'
Bisulfite seq NANOG Forward	5'-AGAGATAGGAGGGTAAGTTTTTTTT-3'
Bisulfite seq NANOG Reverse	5'-ACTCCCACACAACTAACTTTTATTC-3'
Integration KLF4 Forward	5'-AATTACCCATCCTTCCTGCC-3'
Integration KLF4 Reverse	5'-TTAAAAATGCCTCTTCATGTGTA-3'
Integration OCT4 Forward	5'-TAAGCTTCCAAGGCCCTCC-3'
Integration OCT4 Reverse	5'-CTCCTCCGGTTTTTGCTCC-3'
Integration SOX2 Forward	5'-AGTACAACCTCCATGACCAGC-3'
Integration SOX2 Reverse	5'-TCACATGTGTGAGAGGGGC-3'
Integration c-MYC Forward	5'-TCCACTCGGAAGGACTATCC-3'
Integration c-MYC Reverse	5'-TTACGCACAAGAGTTCCGTAG-3'

**Figure S1.** 4 factors iPS cells validation.

(A) ESC-like colonies of WT, SFC6 and SFC7 lines and AP staining.

(B) RT-qPCR analyses of the expression levels of retroviral-derived reprogramming and endogenous genes.

(C) Karyotype of each line identical to that of parent fibroblasts [including the balanced Robertsonian translocation der(13;14)(q10;q10) of SFC6].

(D) Expression of pluripotency markers (OCT4, SOX2, NANOG –green–; SSEA3, SSEA4, TRA-1-81 –red–). Scale bar: 100  $\mu$ m.

(E) *In vitro* differentiation of the 3 iPS cell lines stained with different specific markers for each germ layer. Endoderm:  $\alpha$ -fetoprotein (green), FOX2A (red); mesoderm: GATA4 (green),  $\alpha$ -SMA (red); ectoderm: TUJ1 (green), GFAP (red). Scale bar: 100  $\mu$ m.

(F) Teratoma formation of the 3 iPS cell lines stained with different specific markers for each germ layer. Endoderm:  $\alpha$ -fetoprotein (green), FOX2A (red); mesoderm: SOX9 (green), CS (red); ectoderm: TUJ1 (green), GFAP (red). Mouse injected for the 4 factors WT line died before the development of teratomas. Scale bar: 100  $\mu$ m.

**Figure S2.** Integration of the reprogramming factors and testing patients' lines for the mutations.

(A) PCR analysis of the presence of the reprogramming genes in the DNA of the iPS cells for the 3 factors and the 4 factors lines.

(B) Sequence of the 3 factors and 4 factors lines to test the presence of the mutations in the patients' cells.

**Figure S3.** Splicing pattern in neurons and iPS cells.

(A) Splicing pattern for c.633+1G>A mutation in neurons (WT and SFC6 lines) with and without cycloheximide (CHX) and iPS cells (WT, SFC6 and SFC7).

(B) Splicing pattern for c.372-2G>A mutation in neurons (WT and SFC7 lines) with and without cycloheximide (CHX) and iPS cells (WT, SFC6 and SFC7).

**Figure S4.** LAMP2 staining.

(A) Immunofluorescence with LAMP2 antibody in fibroblasts. Scale bar: 100  $\mu$ m.

(B) Immunofluorescence with LAMP2 antibody in iPS cells. Scale bar: 100  $\mu$ m.

**Figure S5.** Apoptosis and autophagy analyses.

(A) Example of a positive cell for caspase 3 staining. Scale bar: 25  $\mu\text{m}$ .

(B) Differences in the percentage of apoptotic cells in the WT- and SFC-iPS derived neural cells for 6 weeks.

(C) Autophagy analysis by LC3 immunofluorescence in TUJ1-positive neurons. Scale bar: 100  $\mu\text{m}$ .

Figure 1

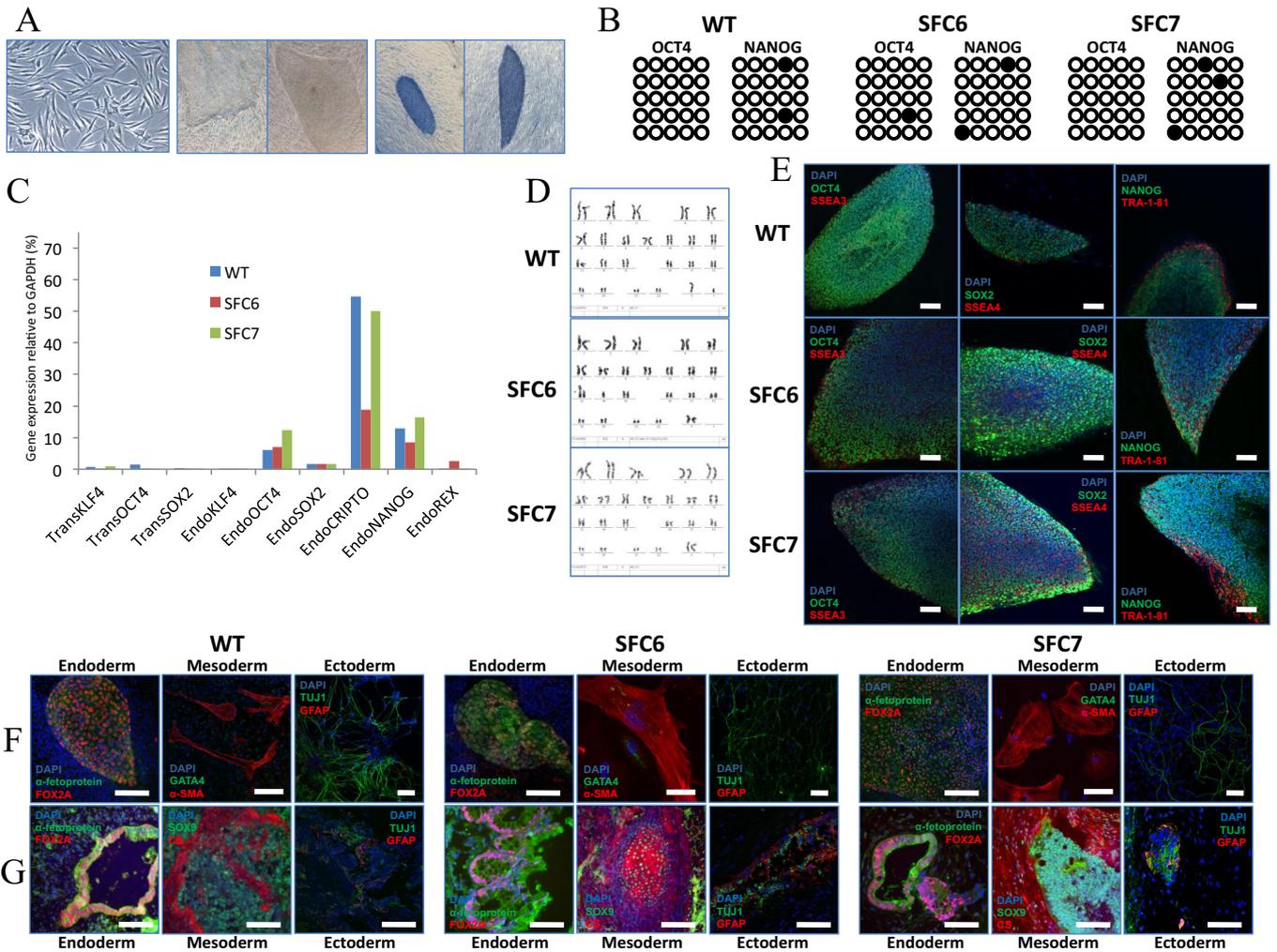


Figure 2

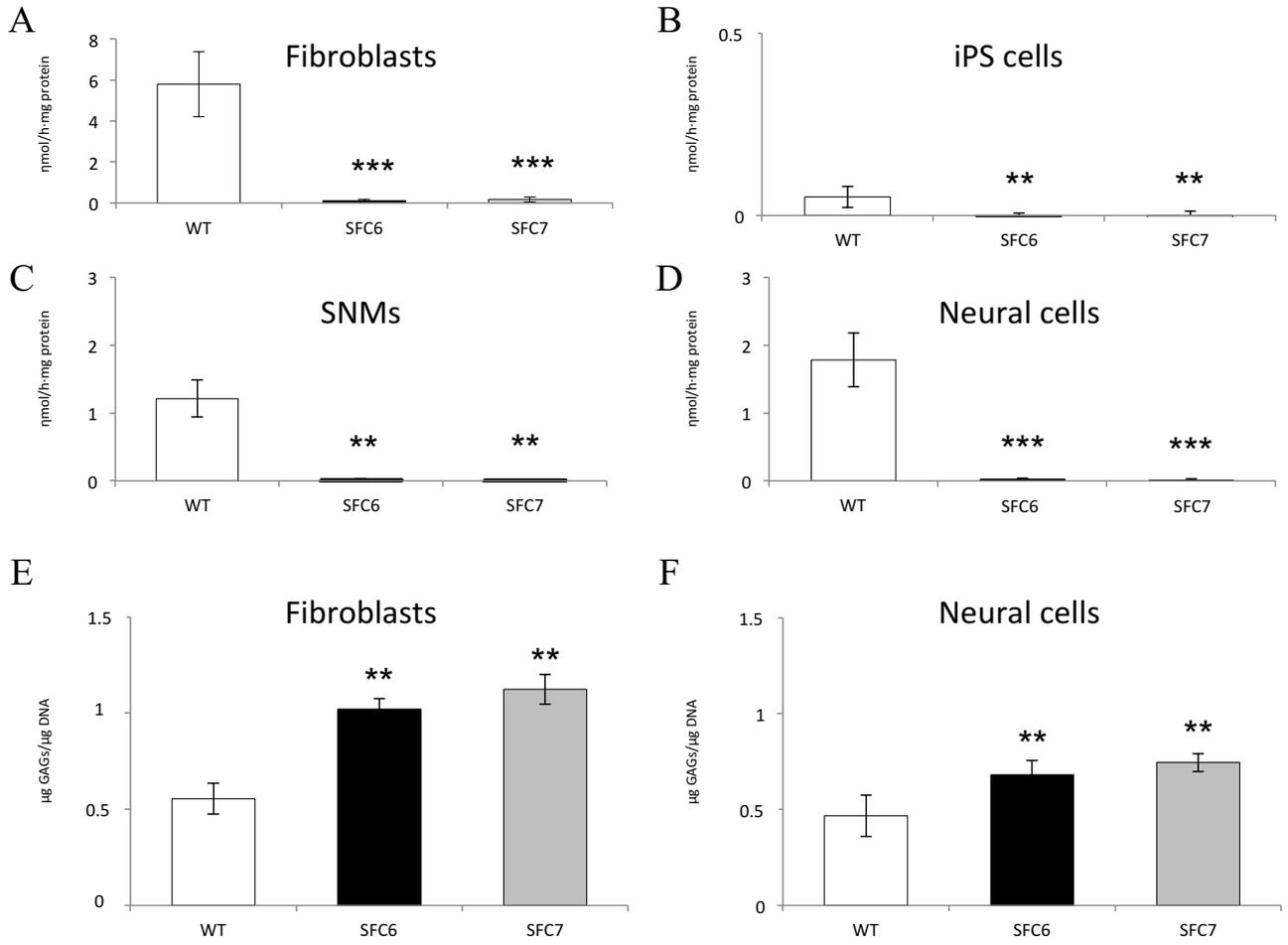


Figure 3

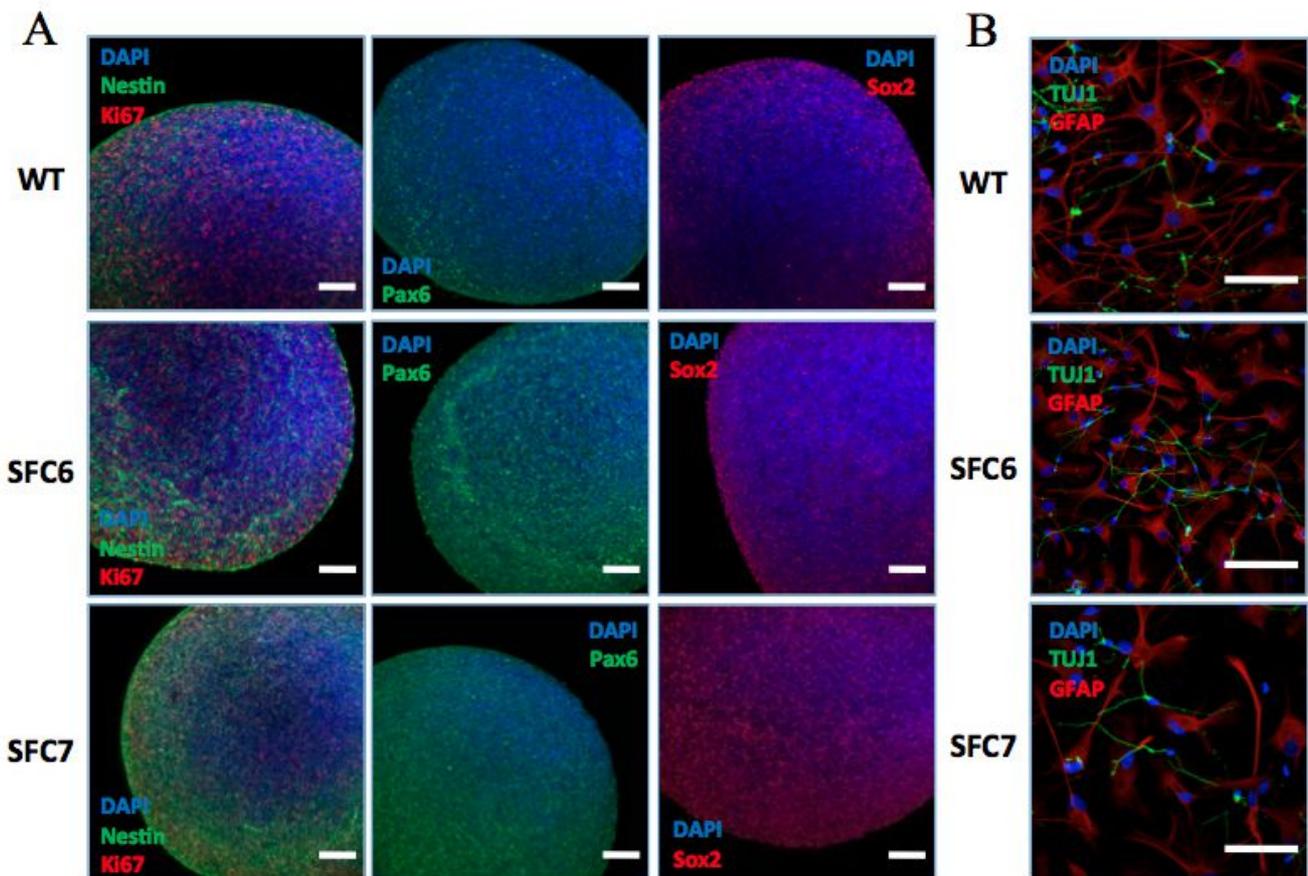


Figure 4

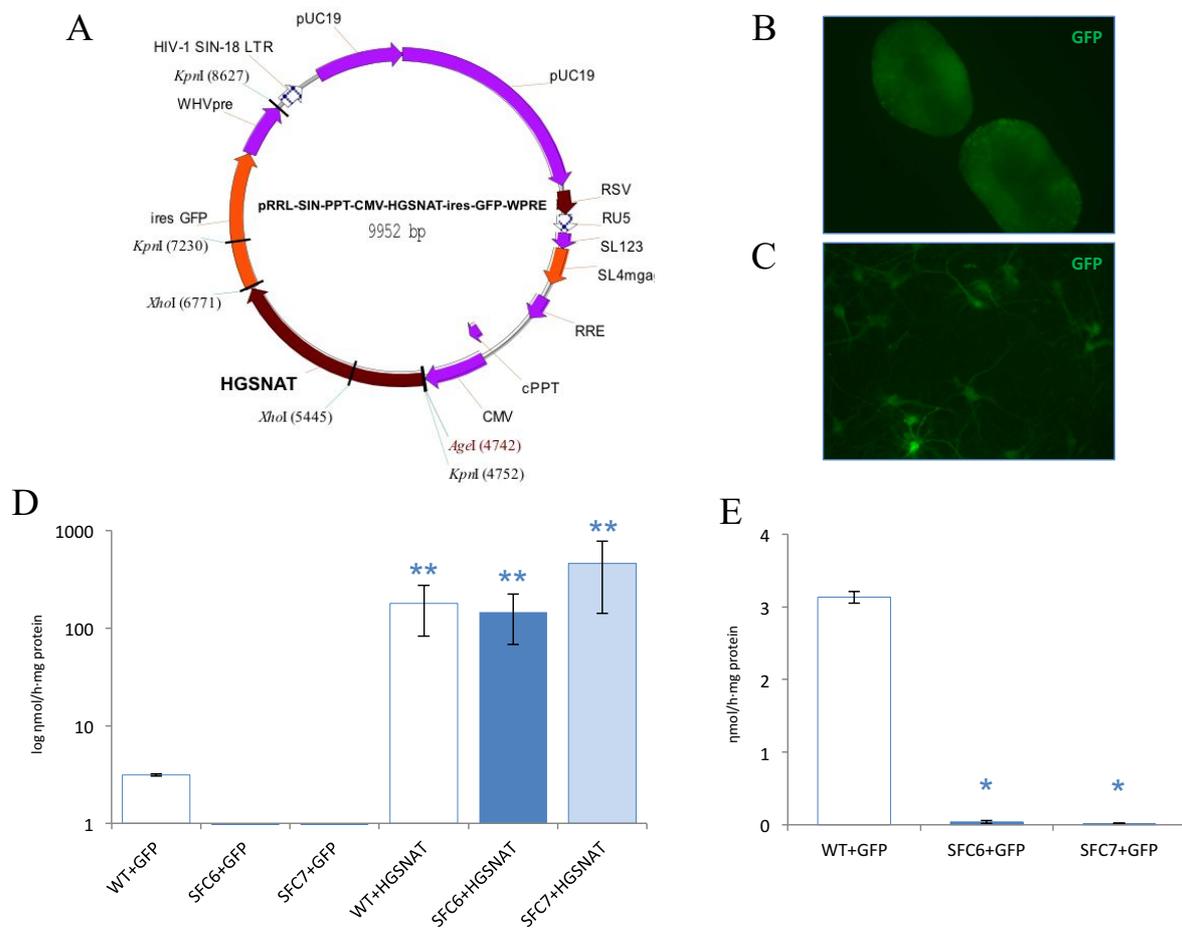


Figure 5

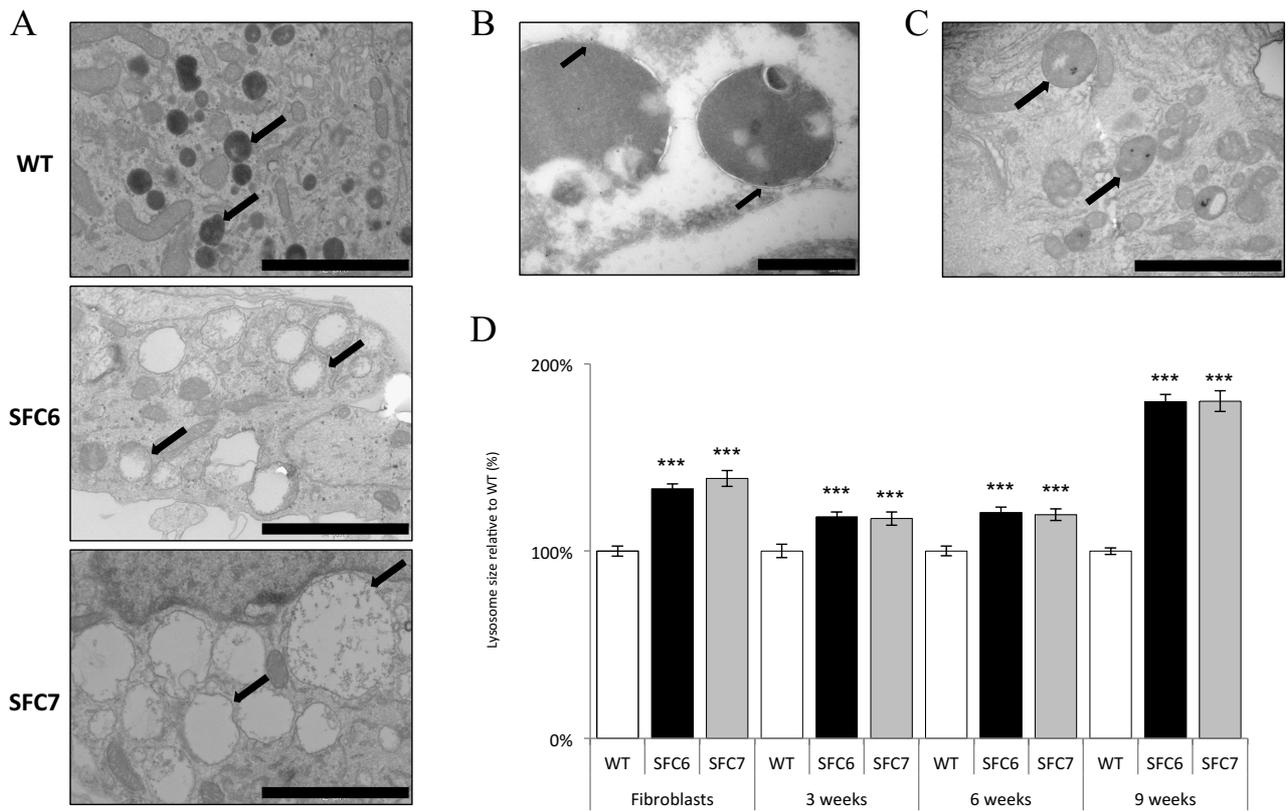


Figure 6

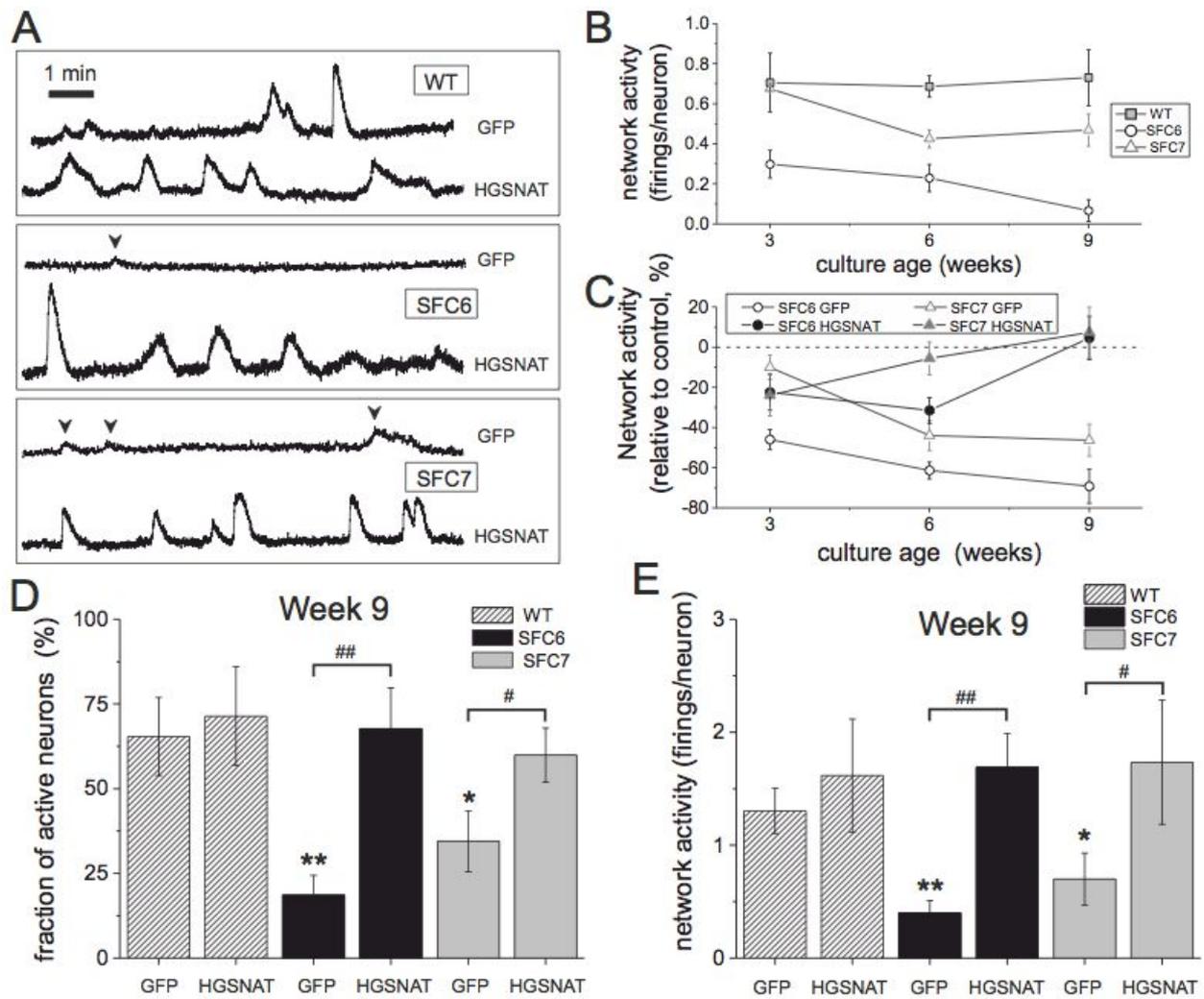


Figure 7

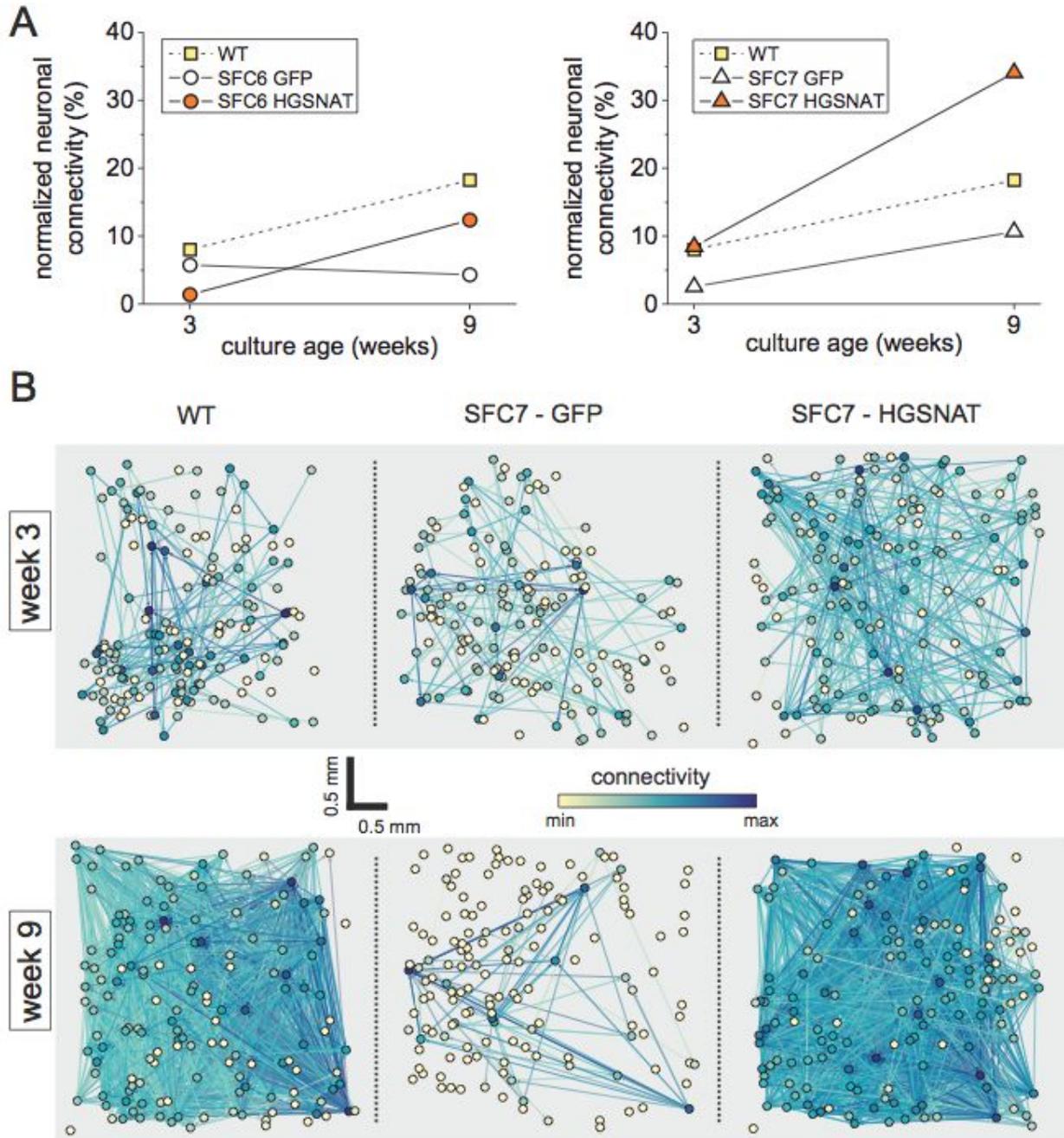


Figure S1

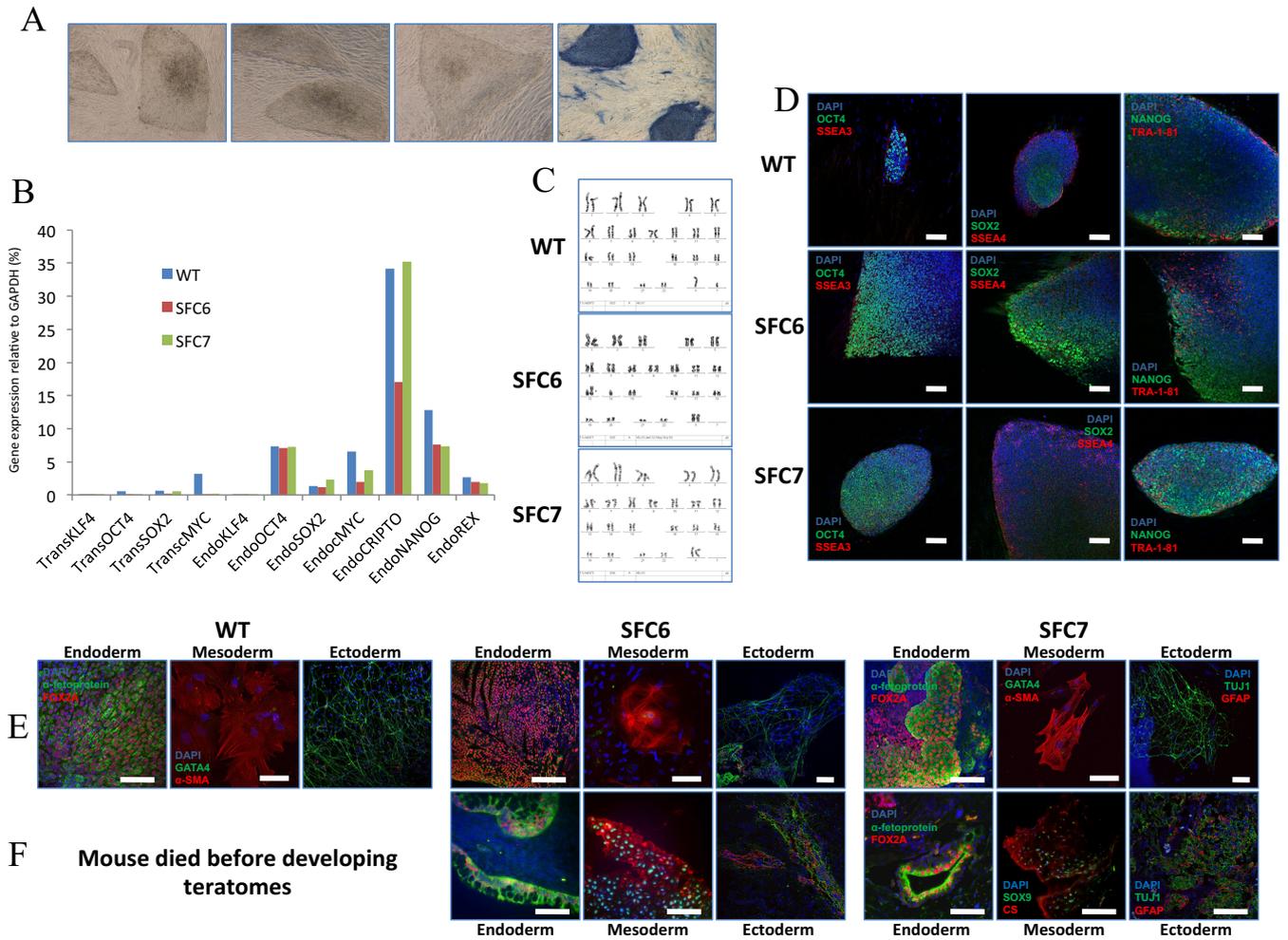


Figure S2

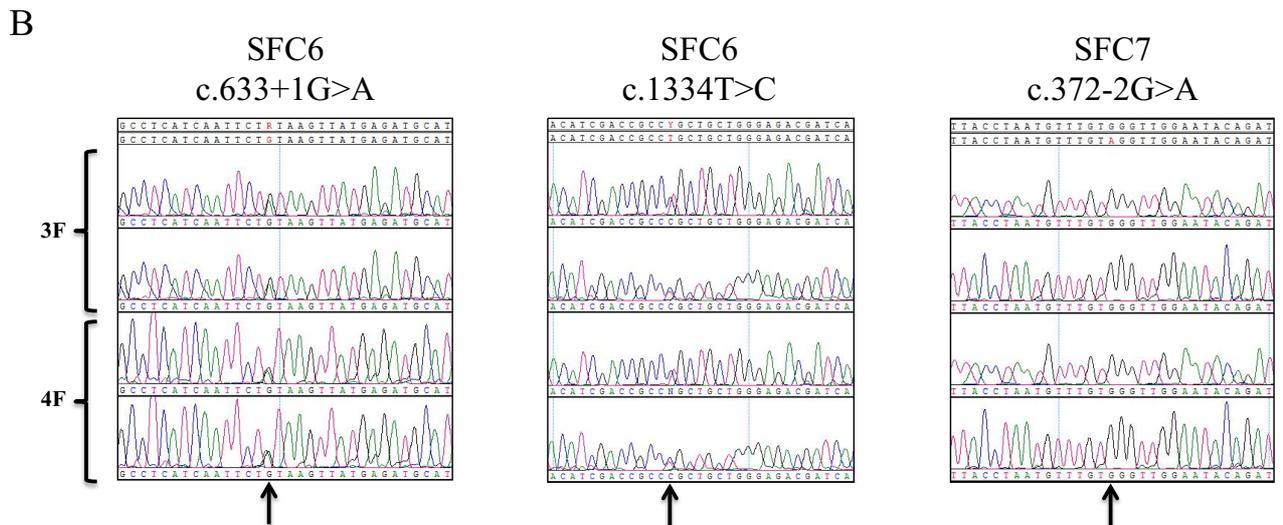
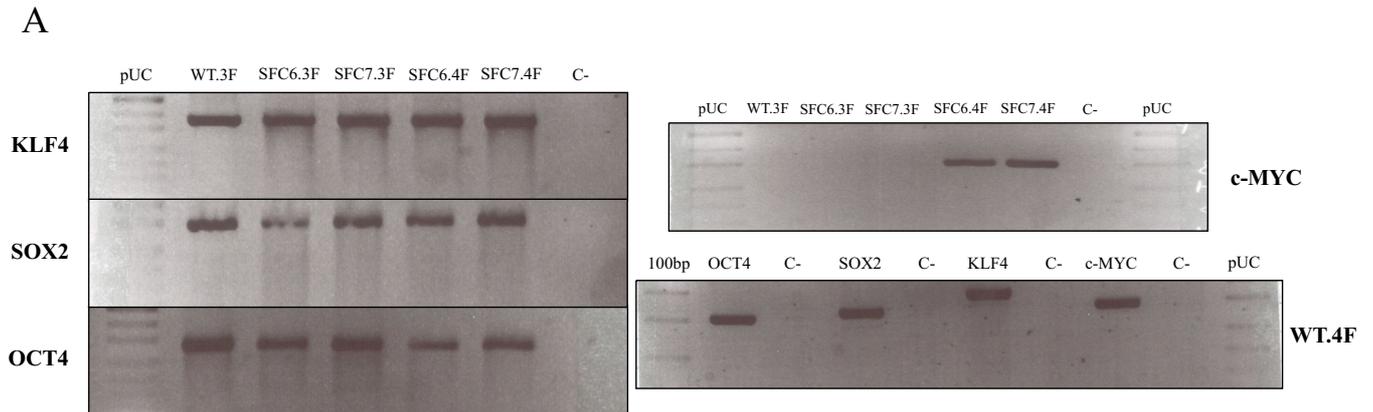


Figure S3

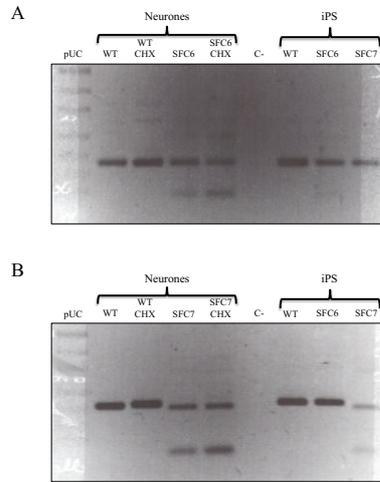


Figure S4

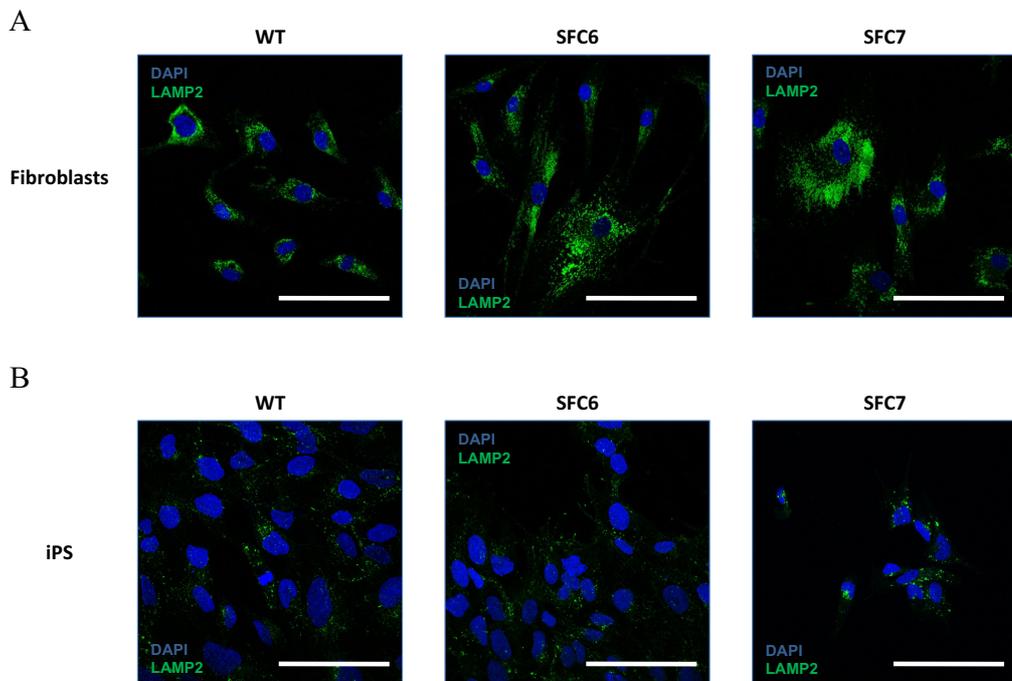
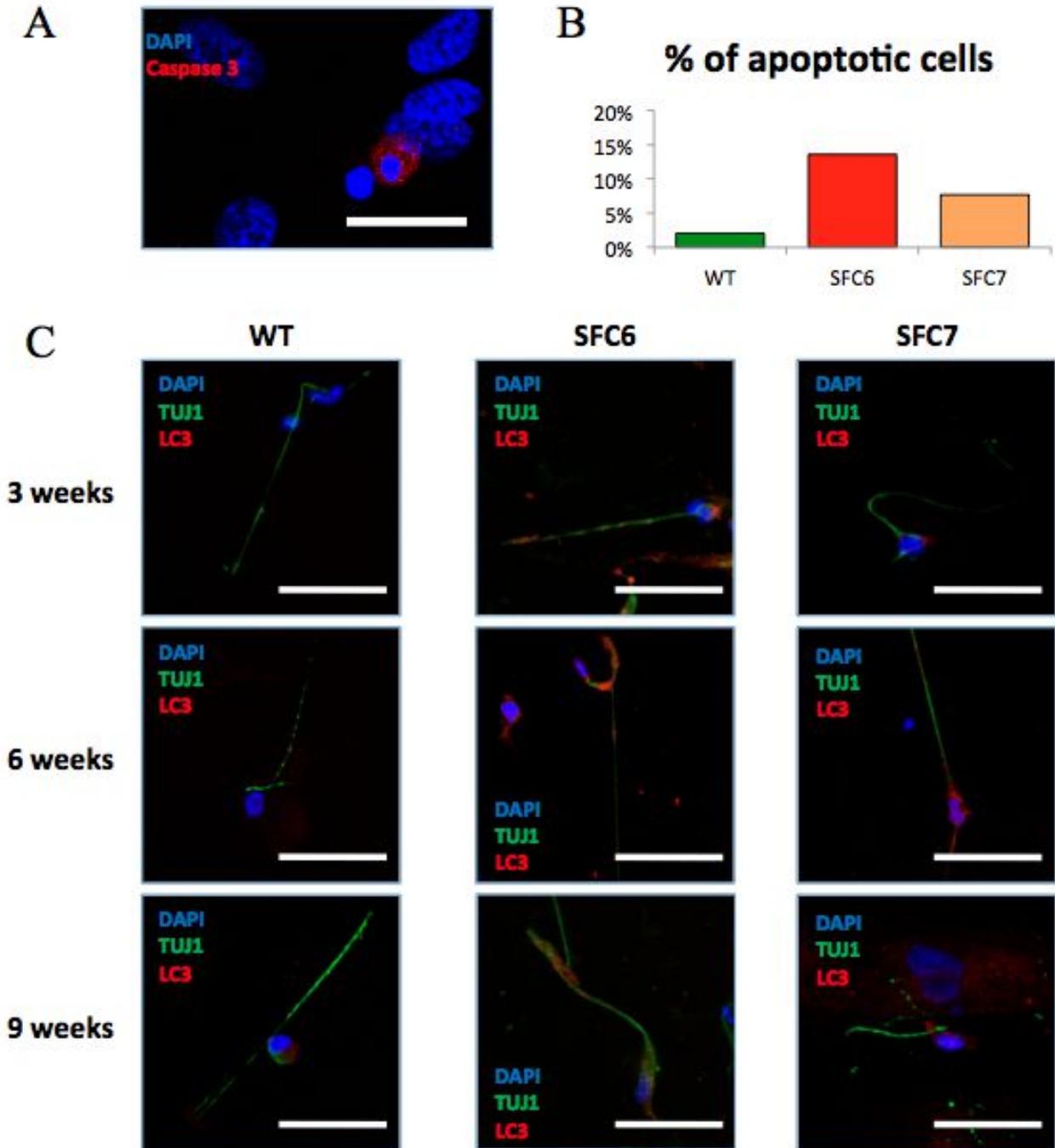


Figure S5



**Annex**



## Mutation analysis in patients and relatives not included in article 1.

Some Sanfilippo patients were analyzed after the publication of article 1. These patients are presented in this annex. Materials and methods used were the same as described in article 1. Patients' origin and mutations found are described in table 1.

**Table 1. Mutational analysis.** Results of the mutational analysis in patients of Sanfilippo C not published in article 1.

Patient	Origin	Allele 1	Allele 2
<b>SFC 11</b>	Sweden	c.1622C>T (p.S541L)	c.1622C>T (p.S541L)
<b>SFC13</b>	France	c.1150C>T (p.R384*)	c.1542+4dupA
<b>SFC14</b>	Poland	c.234+1G>A	c.1150C>T (p.R384*)
<b>SFC15</b>	Poland	c.641delG	c.1150C>T (p.R384*)

Furthermore, samples from relatives of some of the patients described in article 1 were sequenced to assess the carrier status. The relationship to the patients, origins and results are shown in table 2.

**Table 2. Patients' relatives analysis.** Carrier analysis of several relatives of Sanfilippo C patients performed during this thesis.

Familiar	Patient	Origin	Result
Aunt	SFC1	Spain	Non-carrier
Brother	SFC4	Spain	Non-carrier
Sister	SFC6	Spain	Non-carrier
Sister	SFC10	Argentina	c.161C>T carrier
Brother	SFC10	Argentina	c.821-31_821-13del carrier
Sister	SFC10	Argentina	c.161C>T carrier
Sister	SFC10	Argentina	Non-carrier
Aunt	SFC10	Argentina	Non-carrier
Uncle	SFC10	Argentina	c.161C>T carrier
Sister	SFC12	Spain	c.372-2G>A carrier
Aunt	SFC12	Spain	Non-carrier
Uncle	SFC12	Spain	Non-carrier



# **DISCUSSION**



Since 2006, when the gene responsible for Sanfilippo C, *HGSNAT*, was identified by two independent groups (Fan et al., 2006; Hrebicek et al., 2006), several mutational analyses have been performed in different populations (Coutinho et al., 2008; Fan et al., 2006; Fedele et al., 2007; Feldhammer et al., 2009a; Hřebíček et al., 2006; Huh et al., 2013; Ouesleti et al., 2011; Ruijter et al., 2008).

Mutational analyses are of great importance for different reasons. First of all, they allow us to know the molecular change that causes the disease in patients. Initially, an analysis of 11 patients from Spain (seven patients), Argentina (one) and Morocco (three) was performed (article 1). Additionally, four other affected children have been analyzed (annex). As a whole, 15 patients of different origins were studied.

A total of 13 mutations have been detected during this thesis, seven of which were not previously described. These mutations included five missense mutations (p.A54V, p.L113P, p.G424V, p.L445P and p.S541L), four splicing mutations (c.234+1G>A, c.372-2A>G, c.633+1G>A and c.1378-1G>A), one nonsense mutation (p.R384\*), two small deletions (c.641delG and c.821-31 821-13del) and one 1-bp duplication (c.1542+4dupA).

One important aspect of mutational analysis is that it makes possible to perform carrier analysis and prenatal diagnosis, which are more reliable than other tests based on enzyme assays. This is of great importance in terms of genetic counselling. During this thesis, some carrier analyses have been done for different patients' relatives (annex). In all cases we were able to detect the carrier status of the relative. Parents of some of the patients were also analyzed and were shown to be carriers for one of the mutated alleles, as expected (patients SFC1, SFC4, SFC6, SFC8, SFC10, SFC12 and SFC14).

For some diseases, genotype-phenotype correlations can be established. They allow to predict the prognosis of the disease and could be also important for the choice of current or future available therapeutic options. In the case of Sanfilippo C syndrome, only two mutations are suggested to be associated with an attenuated phenotype, p.G262R and p.S539C (Ruijter et al., 2008), but considering that these two mutations were only found in compound heterozygosity in two sibs, this association should be taken with caution. All patients analyzed in this thesis presented a typical severe phenotype with an early onset age, so clear genotype-phenotype correlations could not be established.

Even so, some patients carrying mutation c.372-2G>A showed a slightly slower progression of the disease. This mutation gives rise to a protein with the loss of 4 amino acids (due to the use of a cryptic acceptor ss), which retains some enzyme activity. This mutated protein was analyzed (article 2) and a significant but very low activity was detected. However, the trafficking of the enzyme to the lysosome was altered, as it was reported for other missense mutations (Feldhammer et al., 2009b). This low residual enzyme activity may be the basis of the slower progression of the disease found in patients carrying this mutation, although the genetic background could also play a role.

Mutational analysis can also provide useful information about allelic variability in different populations. For some disorders and populations, prevalent mutations have been described, accounting for a large number of alleles. In these cases, a direct test of prevalent mutations can save time and money in a routine diagnostic protocol. Regarding Sanfilippo C syndrome, some prevalent mutations were described in different populations such as c.852-1G>A (30% of Italian alleles), c.525dupT (62.5% of Portuguese alleles), p.R344C and p.S518F (26.5% and 35.3% of Dutch alleles, respectively) (Coutinho et al., 2008; Fedele et al., 2007; Ruijter et al., 2008).

In this work, one mutation was found to be the most prevalent in Spanish patients, c-372-2G>A, previously described in a Portuguese patient in heterozygosis (Coutinho et al., 2008). In this work, this mutation was identified in four Spanish patients, one heterozygous and three homozygous, representing the 50% (7/14) of all Spanish alleles of our series. If one additional Spanish patient, not carrier of this mutation, analyzed in a previous work (Hrebicek et al., 2006) is included, this mutation would account for 43.8% of all described Spanish mutant alleles. Mutation c.234+1G>A was found in homozygosis in one Spanish patient of our series and it was also present in homozygosis in another Spanish patient previously reported (Hrebicek et al., 2006). Taking all the data together, it represents the 25% of the Spanish mutant alleles. Overall, these two mutations account for 68.8% of total Spanish alleles, meaning that probably two out of three mutations in Spanish patients could be easily identified analyzing only these two prevalent mutant alleles. Moreover, in Moroccan patients, mutation c.234+1G>A represents four out of six alleles of our series and 12 out of 14 of the alleles described in other reports (Hrebicek et al., 2006; Ruijter et al., 2008), representing the 85.7% of the mutant alleles in Morocco and meaning that probably 4 out of 5 alleles can be easily diagnosed testing Moroccan patients

for this prevalent mutation. It is worth noting that taking in consideration all the studies published, splicing mutations account for around 20% of all MPS IIIC mutations, while in Spanish patients they represent more than 85%. This is due to the fact that the two more prevalent Spanish mutations are splicing mutations accounting for almost 70% of Spanish mutated alleles (figure 32). However it should be note that the number of patients is low.

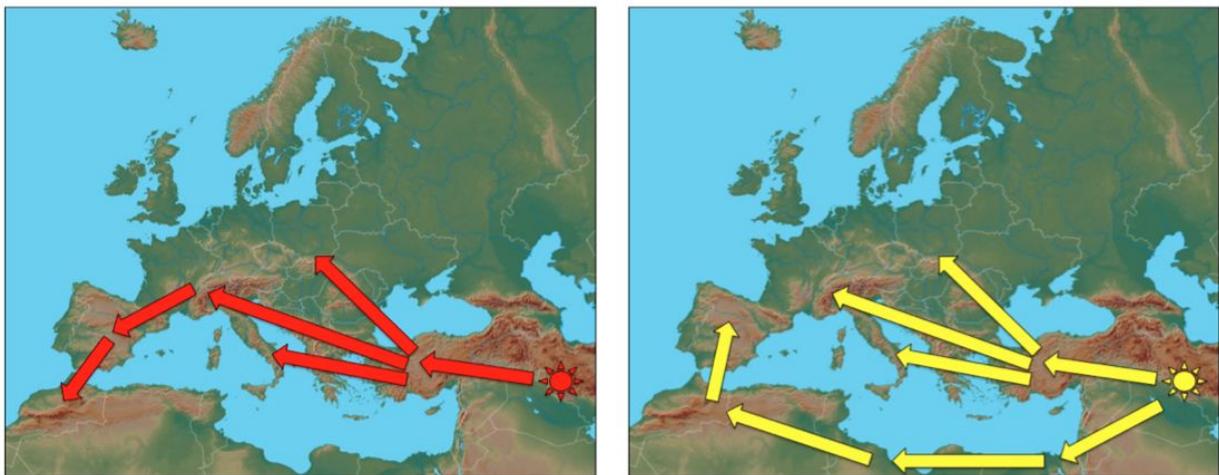


**Figure 32. Percentage of alleles for each mutation in Spanish patients.** Splicing mutations are indicated in red shades while blue shades correspond to missense mutations. Splicing mutations account for more than 85% of the total.

Furthermore, a novel single nucleotide polymorphism (SNP) was identified in intron 5 (c.564-98T>C) in all patients bearing the c.234+1G>A from Spanish and Moroccan origin. This change was not found in either 50 Spanish healthy individuals nor in 50 Moroccan, suggesting that this change could be a rare variant associated with that mutation. The possible effect of this change in the splicing process was analyzed and no alteration in the pattern was found.

An haplotype analysis including 14 SNPs in the *HGSNAT* gene allowed us to assess the possible single origin for each recurrent mutation. In the case of the c.234+1G>A mutation, it was first described to be associated with the missense change p.P237Q. This change was initially considered as a pathogenic mutation (Hrebicek et al., 2006), but later works showed high enzyme activity levels for the protein carrying this variant (Feldhammer et al., 2009b). At present, the c.234+1G>A mutation has been shown to be associated to that polymorphism in all Moroccan and Spanish alleles. However, in patients

from other origin such as Turkish, French, Italian and Polish, this association has not been found. This fact, together with the results of the haplotype analysis of our series of patients, strongly suggested a single origin of this mutation for Spanish and Moroccan patients. Moreover, the presence of the rare new SNP in intron 5 in our series of patients carrying this mutation, confirms this hypothesis. It is a reasonable idea since the two countries are closed and recent studies have proved the genetic exchange among North African populations and South European populations in the last centuries (Botigué et al., 2013). It is probable that the mutation originated in a Middle East ancestor and introduced into the South Europe, arriving to Spain, where it was associated with the p.P237Q and then introduced to North Africa. The fact that the p.P237Q variant was not found in more than 100 Spanish and 100 Moroccan healthy individuals may support the hypothesis that the polymorphism appeared in an allele carrying the mutation (figure 33, left image). The same explanation could be accepted for the rare SNP in the intron 5. Alternatively it can be hypothesized that the mutation appeared in the Middle East and expanded both in Europe and North Africa arriving to Morocco. Then the p.P237Q appeared in the mutated allele and was transmitted to Spain (figure 33, right image).



**Figure 33. Origin and expansion of the c.234+1G>A allele.** Two most probable hypothesis.

Regarding the c.372-2G>A mutation, all homozygous patients carried the same haplotype. In the heterozygous patient the genotype of the only heterozygous SNP was

consistent with this common haplotype but, since parents' DNA was not available, phases could not be established. So, the results suggested a single origin for this mutation in the Spanish patients. This mutation was found in heterozygosis in one Portuguese patient (Coutinho et al., 2008). It could be speculated that this mutation had a single origin in the Iberian population but it was not possible to confirm this hypothesis since DNA of this Portuguese patient was not available.

In mutation analysis, once a change has been identified and in genetic diseases that the protein is known, functional analysis can be performed, providing useful information about the pathogenicity of the mutation. The expression *in vitro* of different mutated alleles has been carried out for several Sanfilippo C mutations (Durand et al., 2010; Fedele and Hopwood, 2010; Feldhammer et al., 2009b), showing that most of the mutated proteins had negligible activity levels. One exception is mutation p.W403C that has been described to retain around 25% of WT protein activity level. This mutation has only been found as a complex allele with the polymorphism p.A615T, which has been reported to decrease enzyme activity to 50% of the WT enzyme. The association of both changes in the same allele results in a low enzyme activity level that promotes the development of the disease (Fedele and Hopwood, 2010). It seems that each change alone is not pathogenic, but their combination in one allele causes a higher decrease of the enzyme activity and the consequent pathology.

An *in vitro* expression study in COS-7 cells was performed for all the missense mutations found in our first series of patients (p.A54V, p.L113P, p.G424V and p.L445P). In the study, two previously described prevalent mutations in Dutch patients (p.R344C and p.S518F) and the p.P237Q variant were also included. Both Dutch alleles were used as negative controls since they were reported to have negligible enzyme activity (Fedele and Hopwood, 2010; Feldhammer et al., 2009b). In contrast, the p.P237Q vector was used as a positive control due to its previously reported high activity level (Fedele and Hopwood, 2010; Feldhammer et al., 2009b).

The results showed an activity of  $95.5 \pm 25.7$  nmol/h/mg for the WT protein, whereas for the endogenous enzyme was  $8.6 \pm 2.2$  nmol/h/mg. These results validated the experimental model since the activity of the transfected human WT enzyme is much higher than the endogenous activity in COS-7 cells. All mutant proteins showed an enzymatic

activity between 0 and 1.2% of the WT protein, confirming their pathogenic effect. For the p.P237Q variant, an enzyme activity of around 92% was found, demonstrating that this change is a polymorphism, as suggested in previous works (Fedele and Hopwood, 2010; Feldhammer et al., 2009b). It has been proposed that enzyme activities below 10-20% are associated with the development of the LSDs (Lachmann, 2010) and these results are in agreement with this hypothesis.

Unfortunately, it was not possible to analyze the trafficking of the missense mutations since there was no appropriate antibody to perform the experiments. It would have been interesting to confirm the suggested theory that proteins bearing missense mutations responsible for Sanfilippo C syndrome are retained into the ER due to an impaired trafficking (Durand et al., 2010; Feldhammer et al., 2009b). In this regard, glucosamine has been assayed as a chaperone to improve the correct folding and trafficking of the mutated proteins leading to a partial restore of the enzymatic activity (Feldhammer et al., 2009b). Since the recovery, despite significant, does not reach necessary threshold of about 10% of WT activity (Sawkar et al., 2002; Suzuki et al., 2009), it might be appropriate to evaluate the effect of the use of glucosamine as a chaperone in combination with another therapy such as SRT, which would decrease the HS synthesis.

The use of siRNAs as a SRT, knocking down the HS synthetic pathway may represent a good tool for this purpose. There are different enzymes that participate in this pathway, mainly products of the *EXT* gene family. Since the *EXT1* and *EXT2* genes, responsible for the HS chain elongation, are not good targets because when mutated, they cause exostoses, we decided to inhibit the *EXTL2* and *EXTL3* genes. They code for enzymes that start the elongation process. The goal was to target this step in order to decrease the HS chain formation. Similar strategies had been tested before, one using siRNAs to inhibit genes coding enzymes involved in the linkage region formation (Dziedzic et al., 2010), and another one using shRNAs to inhibit the *EXTL2* and *EXTL3* genes (Kaidonis et al., 2010).

Two siRNAs to inhibit *EXTL2* and two siRNAs to inhibit *EXTL3* were tested in fibroblasts from two different patients (article 3). High inhibition levels of the mRNAs, of around 90%, were detected from three to 14 days. These inhibition levels were higher than those obtained in previous works and were maintained for a longer time, indicating that these siRNAs were more efficient.

After three days of siRNA treatment, GAG synthesis was inhibited for both patients between 30 and 60% depending on the siRNA used. Two facts are important to be remarked: first, siRNAs targeting *EXTL2* showed better results (50 to 60%) than those targeting *EXTL3* (30 to 50%) suggesting that *EXTL2* could be a better therapeutic target; and second, real HS inhibition level could be higher taking into account that all GAGs are being measured and not only HS. After seven and 14 days, GAG synthesis was decreased even in untreated cultures. It may be due to the confluent status of the cultures because of the high growing rate of these cells. Probably, siRNAs could function better in non-growing cells such as neurons, the affected cell type in Sanfilippo syndrome.

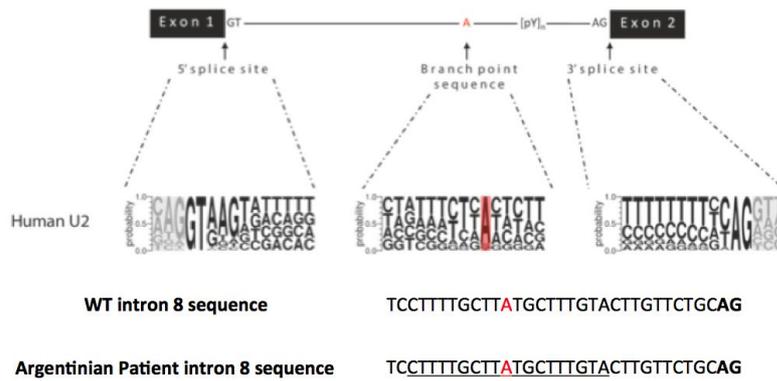
An increase of the GAG amounts with time was detected, which is consistent with the progressive storage nature of the disease due to inability of patients to degrade HS. After treatment, slight decreases between 10 to 20% were obtained for most of the siRNAs. Decreases were more significant at 14 days, where GAG amounts were higher and, thus, reduction was also higher. Again, it should be noted that total GAG amounts was quantified, not only HS, suggesting that HS reduction could be underestimated.

The results confirmed the potential of siRNAs targeting the *EXTL* genes to be used as a SRT approach. Fibroblasts of two different patients were used and results were similar in both cases, suggesting that siRNA therapy could be useful for any Sanfilippo patient regardless of which mutations he/she is carrying. Moreover, the potential of these siRNAs to maintain *EXTL* genes inhibition for at least 14 days was demonstrated, improving the results of previous works. This study opens the possibility to assay long-term therapies using siRNAs targeting *EXTL* genes (especially *EXTL2*) and also encourages the assessment of other RNAi therapies such as shRNAs, based in these siRNAs, as a SRT for Sanfilippo syndrome of any subtype, since all of them accumulate HS independently of the mutated gene. A recent work suggests that *EXTL2* is responsible for ending the elongation of HS chains (Nadanaka et al., 2013). In that study, knockout mice for *Extl2* showed higher production of GAGs. This work and previous results, in which *EXTL2* inhibition resulted in the decrease of GAG synthesis and storage (article 3; Dziejczak et al., 2010; Kaidonis et al., 2010), are in disagreement with this suggestion. Further studies should be done in order to elucidate the exact role of *EXTL2* in the HS biosynthesis.

Considering that the results of the use of siRNAs for SRT show only a partial reduction but not a complete elimination of the storage, this therapy should be combined with another therapy to improve the excretion (for instance, the overexpression of TFEB) or the degradation of HS inside the lysosomes (for example, using glucosamine as a chaperone).

Going back to the characterization of mutations identified in this work, the effect of the mutations affecting the ss on the splicing process was studied. Moreover, the goal was to assess whether the intronic deletion found in heterozygosis in the Argentinian patient (c.821-31\_821-13del) included the branch point. Since fibroblasts were not available for this patient, a minigene analysis was performed, an approach widely used in these situations. Minigene assays present some advantages and some disadvantages. On the one hand, it does not correspond to the natural context for the splicing process since the gene is not complete and thus, some sequences that could be important for splicing regulation may be lacking. On the other hand, it allows us to work in the absence of a second allele, making the results clearer.

Exon 9 of the *HGSNAT* gene and the normal and mutated intronic flanking regions (including the c.821-31\_821-13del in the patient case) were cloned in a vector (pGLB1) previously developed in the group containing exon 7, intron 7, exon 8, intron 8 and exon 9 of the *GLB1* gene, more concretely inside intron 7. The results suggested that this minigene represented a good model to study the splicing process and that the mutation promoted exon 9 skipping. WT minigene showed the inclusion of *HGSNAT* exon 9 in the resulting transcript, validating this experiment. In contrast, the mutated minigene showed the exclusion of this exon, suggesting that this intronic deletion promotes the exon 9 skipping and demonstrating its pathogenicity. This result confirms the hypothesis that the deletion contains the branch point as well as some nucleotides of the polypyrimidine tract (figure 34), essential sequences for the splicing process.



**Figure 34. Intronic deletion.**

Consensus splicing sequences and deletion identified in the patient from Argentina (underlined), affecting the branch point (the invariable A nucleotide indicated in red), near the 3' ss (in bold).

For the other splicing mutations affecting the conserved ss (c.234+1G>A, c.372-2A>G, c.633+1G>A and c.1378-1G>A) fibroblasts were available. Patients' cells were grown either in the presence or absence of cycloheximide (CHX), an inhibitor of the NMD process, since PTC could be generated as a consequence of the aberrant splicing caused by the mutations.

In the case of mutations affecting the 3' ss (c.372-2A>G and c.1378-1G>A), results showed alterations in the splicing process. For the c.1378-1G>A mutation, two transcripts were detected. The first transcript corresponded to exon 14 skipping, as expected. This exon contains 87 nucleotides, which means that the reading frame is conserved and so this mRNA would not be a substrate for NMD. This transcript would give rise to a protein lacking 29 amino acids. Due to this big loss, it is highly probable that it would not retain any enzyme activity. Furthermore, this protein could present an incorrect folding process affecting its trafficking to the lysosome. The other transcript corresponded to the use of an alternative acceptor site generated by the mutation. Since the AG is changed to AA, this new A, together with the adjacent first exonic G (frequent in most of the exons) of exon 14, are used as an alternative AG ss (article 1). Consequently, the G is now a part of the intron, and the mRNA lacks one nucleotide, changing the frame and generating a PTC. As expected, this transcript presented NMD, since its levels increased upon CHX treatment.

Regarding the c.372-2A>G mutation, two transcripts were detected. The first transcript was produced by the exon 4 skipping, which generates a frame-shift and a PTC and was shown to be subjected to NMD. In the second transcript, the use of an alternative ss, located 12 nucleotides downstream of the normal acceptor site, gives rise to a protein

lacking four amino acids of the first lysosomal loop. This alternative ss is not used in the absence of the mutation. As mentioned above, the enzyme activity of this protein lacking four amino acids was very low (article 2), suggesting that a therapy enhancing this splicing would not be enough to obtain an improvement in patients symptomatology. As in the case of missense mutations, this mutation could be a good target for a chaperone treatment using glucosamine. To check this hypothesis, the potential of this therapy was tested in fibroblasts of one patient carrying this mutation in homozygosis and a slight increase in the enzyme activity when compared to untreated cells was obtained (article 2), but still far from the WT levels. As discussed before, for many LSD an enzyme activity about the 10% of the WT activity is thought to be enough to restore the normal phenotype but we did not reach this percentage. These results suggest that molecules with a better chaperone effect should be obtained, and/or the chaperone treatment should be applied in combination with another approach, such as SRT, as pointed out above, to get a synergic effect.

The two cases affecting 5' ss (c.234+1G>A and c.633+1G>A) promote exon skipping (exon 2 and exon 6, respectively), as expected. Both skipping processes originate transcripts with a PTC, but only in the second case the NMD was detected. Since the PTC is generated in the exon 3 (out of 18 exons), NMD should have taken place, although is not the first time that a transcript with a PTC is not degraded by the NMD mechanism (reviewed in Holbrook et al., 2004).

The feasibility of overexpressing modified U1 snRNAs to correct defects in the splicing process due to mutations in different 5' ss of the *HGSNAT* gene was analyzed (article 2). This approach is based on the modification of the U1 snRNA sequence to improve the complementarity to the mutated 5' ss. To our knowledge, this was the first attempt to develop this type of treatment in cells of Sanfilippo C patients, though it was tested for many other disorders (see introduction).

A total of nine different modified U1 snRNAs were designed, as well as the U1-WT, four for patients SFCP and SFC3 (homozygous for the c.234+1G>A mutation), two for patient SFC6 (heterozygous for the c.633+1G>A mutation and a missense mutation) and three for patient SFC13 (heterozygous for c.1542+4dupA mutation and a nonsense mutation). In each set, modified U1 snRNAs presented different degrees of complementarity to the mutated sequence (figure 35).

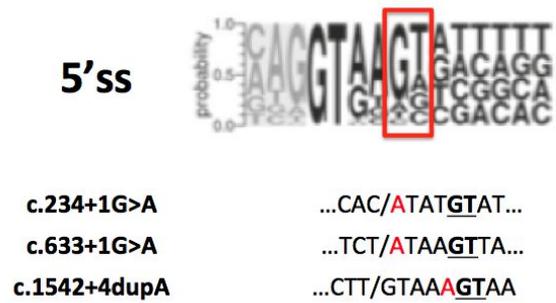
### Modified U1 snRNA complementarity to the mutated 5'ss

c.234+1G>A	U1-WT < U1-sup1 < U1-sup2 < U1-sup3 < U1-sup4
c.633+1G>A	U1-WT < U1-sup5 < U1-sup6
c.1542+4dupA	U1-WT < U1-sup7 < U1-sup8 < U1-sup9

**Figure 35. Modified U1 snRNAs used.** Different modified U1 snRNAs for each mutation presented different degrees of complementarity to the mutated 5' ss

First, modified U1 snRNA overexpression was tested on minigenes bearing the WT and mutant ss, in order to evaluate their effect on the splicing process. Minigene assays are very useful as a first approximation in splicing studies as discussed above. Six different minigenes were constructed, two for each mutation containing the WT or the mutated sequence, respectively. Overexpression of all modified U1 snRNAs and WT U1 snRNA showed not to interfere with the normal splicing pattern in all WT minigenes. Modified U1 snRNAs were overexpressed together with the specific mutated minigenes. For all three mutations, WT splicing pattern was not recovered, but some modified U1 snRNAs enhanced, partially or completely, an alternative splicing process using the cryptic "GT" donor site situated at intronic positions +5 and +6 (or +6 and +7 in the case of c.1542+4dupA mutation) (figure 36). This enhancement was observed particularly for modified U1 snRNAs that presented higher complementary to the mutated site. The reason why these modified U1 snRNAs designed to better recognise the mutated sequence were enhancing the use of a cryptic ss remains unclear. It has been suggested that the U6 snRNA also plays an important role in the recognition of 5' ss (Brackenridge et al., 2003; Hwang and Cohen, 1996). The fact that in this study U6 was not modified could explain the lack of success of this strategy.

**Figure 36. Alternative ss near the natural one.** Mutations in 5' ss (in red) promotes the use of alternative sites (underlined and in bold and boxed in red in the consensus sequence) in positions +5 and +6 (c.234+1G>A and c.633+1G>A) or +6 and +7 (c.1542+4dupA).



Despite the negative results, afterwards we decided to try all modified U1 snRNAs in fibroblasts, due to the fact that minigenes do not reproduce exactly human cellular conditions. First of all, the lack of alteration of the normal splicing process in WT fibroblasts after U1 snRNAs overexpression was confirmed. Later, modified U1 snRNAs were overexpressed in patients' cell lines. In patients SFC6 and SFC13 cells, neither the corrected nor the alternative transcripts were detected after overexpression of all modified U1 snRNAs. It is not clear why the alternative ss are being used in minigene studies but not in fibroblasts.

In patients SFCP and SFC3, both carrying the c.234+1G>A mutation, when U1-sup4 was overexpressed, a partial recovery of the normal splicing process was obtained, together with the alternative splicing using the cryptic site at intronic positions +5 and +6. The percentage of correctly spliced mRNA was around 50%. It was a surprisingly positive result since +1 and +2 mutations, affecting the high conserved nucleotides in the 5' ss have been reported to be extremely difficult to be corrected. Until now, only one previous work has showed a partial correction of a +1 mutation (Hartmann et al., 2010). However, in this case, the mutated allele already produced some degree of the normal spliced transcript before treatment and the alternative "GT" at positions +5 and +6 presented a low score accordingly to different splicing predictors. Therefore, this study is the first case where a correction was obtained in a +1 mutation that did not produce any of the WT spliced mRNA when untreated. In order to test the effect of this partial recovery of correctly spliced mRNA in the protein function, enzyme activity was measured in patients' cells after U1-sup4 treatment, but no increase in the activity level was detected. It is possible that a higher recovery of the splicing process is needed to promote an increase in the enzyme activity. Taking into account that the alternative splicing is also enhanced, it would be

interesting to check whether the combination of modified U6 snRNAs and modified U1 snRNAs could improve specifically the use of the mutated site and not the alternative one.

Further studies are needed to improve the efficacy and to test the toxicity and side effects of U1 overexpression in order to confirm the feasibility of the use of modified U1 snRNAs for therapy. A recent work where modified U1 snRNAs introduced with AAV were overexpressed in mice liver showed that, at low viral charges, toxicity was not detected in treated mice (Balestra et al., 2014). This was the first *in vivo* example, in which modified U1 snRNAs could represent a possible therapeutic approach.

An interesting conclusion that this work points out is the importance of the presence of an additional "GT" dinucleotide, usually at positions +5 and +6, which depending on the context, may be used as an alternative ss. In previous reports where mutations were partially or totally corrected, this alternative site presented a low score in accordance with different predictors, in contrast to the cases described in this study.

Until now, no effective therapy for Sanfilippo C patients has been developed. Different approaches have been tested in the last years, mainly for subtypes A and B, such as ERT, SRT or gene therapy (widely exposed in the introduction). A therapy for Sanfilippo C syndrome should allow the complete degradation of HS intermediates or their elimination from the lysosome. A crucial aspect when developing a new therapy is the availability of cellular and animal models of the disease to assess its safety and efficacy. Considering that Sanfilippo C is a neurodegenerative disorder, the most appropriate cellular model would be a neuronal model, and any animal model developed should present neuronal affectation. As mentioned in the introduction, currently neither a cellular nor an animal model of Sanfilippo C exists, so the development of a neuronal model would represent a significant accomplishment in the field.

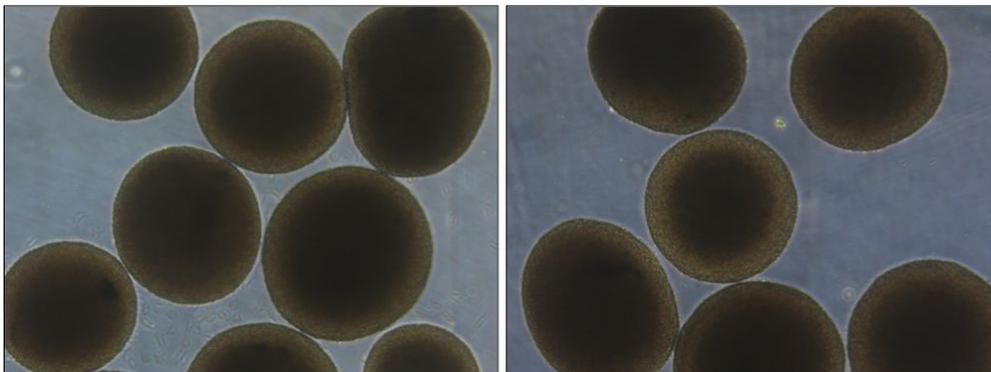
The last part of this thesis consisted in the development of a neuronal model for Sanfilippo C syndrome using the iPSC technology (article 4). Patients' fibroblasts were reprogrammed to obtain induced pluripotent stem cells that were later differentiated to neurons. A neuronal model was previously generated for subtype B (Lemonnier et al., 2011) but this is the first neuronal model for Sanfilippo C syndrome. Fibroblasts from a healthy individual as a control, and from two different patients, SFC6 (heterozygous for

c.633+1G>A and p.L445P) and SFC7 (homozygous for c.372-2G>A) were used. The choice of the patients was based on the availability, on the frequency and type of mutation. At this moment, after the results of the splicing correction, the use of any patient carrying the c.234+1G>A instead of the c.633+1G>A mutation would have been a better option.

Retroviruses were used to introduce KLF4, OCT4 and SOX2 in the patients' fibroblasts. It has to be noted that contrary to what was described for Sanfilippo B (Lemonnier et al., 2011) and Pompe disease (Huang et al., 2011), to supply patients' fibroblasts with the defective enzyme was not needed to achieve the reprogramming. Despite this, differences in the efficiency and in the time of reprogramming were observed among control and patients. Typical colonies for all the lines were obtained and the iPSC validation for one clone of each line was performed. Reprogrammed lines using four factors (using also c-MYC) were also carried out, but we decided to work with the three-factor lines to develop our model to avoid the use of this last reprogramming factor, an oncogen. Even more, as the reprogramming factors are integrated in the genome, if fewer are used, fewer problems derived from integration processes will take place. However, a recent report indicates that four-factor lines are better reprogrammed and would be more appropriate to model diseases in comparison with classical three-factor lines (Habib et al., 2013).

Patients' iPSC did not show any disease feature except for the lack of enzyme activity. The enzyme activity for the control cells was also low, although it was significantly higher than that of the patients, a feature that has been previously reported for other LSD-derived iPSC (Tiscornia et al., 2013). Other lysosomal enzyme activities such as  $\beta$ -hexosaminidase were tested in iPSC, showing also lower levels when compared to other cell types. Stem cells are known to contain a small cytoplasmatic fraction. In this work, a low lysosomal charge, as observed with confocal and transmission electron microscopy, was detected, which could explain this low enzyme activity in the control line. These results contrasted with those obtained for the Sanfilippo B model, where iPSC showed phenotypic traits of the disease, such as higher lysosomal charge in cells derived from patients when compared to cells derived from healthy controls, HS accumulation and GA disorganization (Lemonnier et al., 2011). The later development of the disease and the slower progression in Sanfilippo C subtype, compared to subtype B, may explain these differences in the iPSC lines.

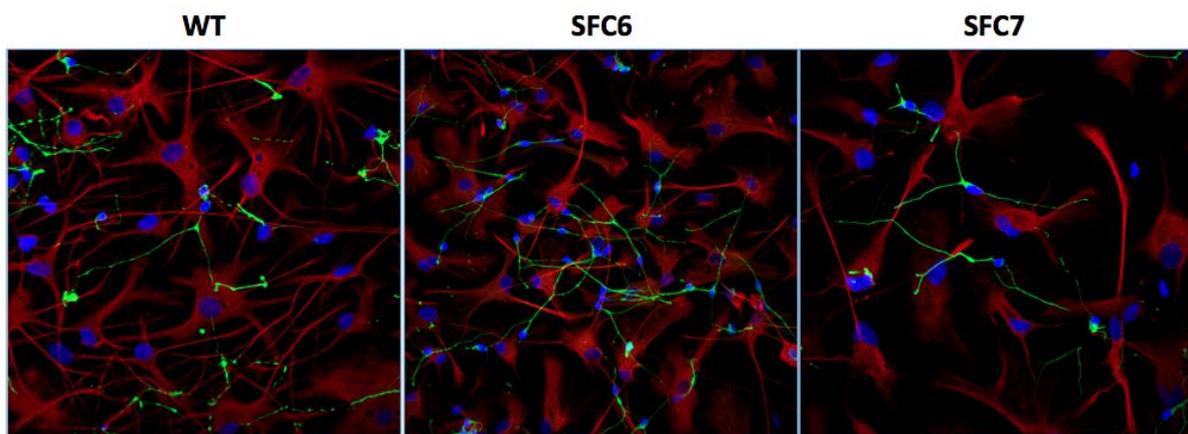
To obtain patient-derived neurons, iPSC were differentiated to NSC, which were maintained as spherical neural masses (SNMs) in the adequate medium (figure 37). They represented a constant source of neurons that allowed us to perform different studies. Control-derived SNMs showed a higher enzyme activity level when compared to iPSC but still low compared to that of differentiated cells (fibroblasts and iPSC-derived neurons). As expected, patient-derived SNMs did not show any activity. These NSC did not present an increase in the GAG content, contrasting with the results obtained for Sanfilippo B model, where NSC accumulated GAGs. Again, these differences could explain the more severe phenotype in Sanfilippo B patients when compared to Sanfilippo C. In a recent work, accumulation of undigested disaccharides resulted in an alteration in the cell polarisation and migration in the NSC derived from Sanfilippo B-iPSC (Bruyère et al., 2014), suggesting that Sanfilippo syndrome could present an alteration in the neurodevelopment of affected children. However, the results could only be corroborated in the mouse model of MPS IIIB for adult neurogenesis and migration, when HS is already highly accumulated, but not in newborn mice.



**Figure 37. SNMs in suspension.** No morphological differences were observed between SNMs derived from WT-iPSC (left image) and SFC-iPSC (right image).

As Sanfilippo syndrome is a progressive storage disease and the goal was to assure that this model recapitulated the main features of the disorder, we decided to differentiate at 3, 6 and 9 weeks. These time points were decided after collaborators modelling Parkinson disease (Sánchez-Danés et al., 2012), an adulthood neurological disorder, found clear disease features at 75 days (almost 11 weeks). Taking into account that Sanfilippo is

a childhood disorder, we thought that 9 weeks could be enough to see the phenotype in these neurons. The other time points were used to study the disease progression. The protocol to differentiate iPSC to neurons was not specific for any type of neuron, since it is assumed that in Sanfilippo C patients all types of neurons are affected. These cultures also present astrocytes after NSC differentiation. Astrocytes have been reported to contribute to the disease development in LSDs (Di Malta et al., 2012), so its presence would not alter or mask the phenotypic features in the neurons. The differentiation process did not show any clear difference between control and patients, although patients' neurons seemed to be slightly larger and less numerous, especially for SFC6 (figure 38). In the Sanfilippo B model, the differentiation rate was similar for all lines (Lemonnier et al., 2011). Thus, it seems that in Sanfilippo syndrome the development of the neurons is not altered in the earlier stages.



**Figure 38. Neuronal cultures.** Neurons (in green) from WT (left image), SFC6 (middle image) and SFC7 (right image) were obtained after SNM-differentiation together with astrocytes (in red).

Patient-derived neurons presented lack of enzyme activity at all time points, a significant increase of around 50% in the GAG storage after 9 weeks of differentiation and an increase of around 80% in the lysosomes size after 9 weeks when compared to WT neurons. These are the main features of Sanfilippo syndrome. There are two important facts that must be remarked. Firstly, since all GAGs were being quantified and patients accumulated mainly HS, the increase of that one would be higher. Secondly, at 3 and 6 weeks of differentiation, lysosomes showed an increase of around 20% in their size, and a

trend of increase in GAG storage (without reaching significance), supporting the idea that the neuronal model mimics the progressive storage nature of the disease. Importantly, it was demonstrated that these phenotypes were due to the lack of enzyme activity, since the introduction of the WT *HGSNAT* cDNA with a lentivirus previous to the differentiation, prevented the development of any disease feature. Taking all the data together, we could conclude that this represents a good neuronal model of Sanfilippo C syndrome. In the Sanfilippo B model, together with the alterations and storage in the lysosomes, they detected a disorganization and severe affectation of the GA (Lemonnier et al., 2011). Even so, it seems that the protein trafficking was correct. That brings up the question whether these GA alterations are the cause (as they suggest) or the consequence of the lysosome affectation and the succeeding impaired functionality of many cellular organelles.

Considering that the major symptomatic feature of the disease is the severe neurodegeneration, one of the major goals of this work was to study the neuronal function and the organization of the neural network of this model, which was performed using a calcium imaging experiment. This technique provides simultaneous information about hundreds of neurons that could be analyzed at a network scale, while other techniques can only be applied in a small number of neurons. Previous studies have pointed out that results about cell action potentials obtained with these different techniques are similar, suggesting that either calcium imaging or patch-clamp could be used (Chen et al., 2013; Sasaki et al., 2008; Smetters et al., 1999).

Results showed an important and progressive decrease in the neuronal activity and in the number of active neurons over the time for both patients, especially for SFC6. An increase of the apoptotic cells was also detected for both patients, suggesting that the decrease in the number of active neurons could be due to a neuronal death. Again, SFC6 showed a higher number of apoptotic cells, supporting the idea of a most severe phenotype for this patient. Autophagy pathway showed to be impaired in both patients at all time-points. It has been reported that autophagy contributes to the neurodegeneration via cell death in LSD (Ballabio and Gieselmann, 2009; Settembre et al., 2008). All these data suggest the idea that neurodegeneration in Sanfilippo C patients is due to an impaired autophagy pathway that activates cell death via apoptosis. The death of neurons promotes the loss of neuronal activity in the neural network of patients' cells, and could explain the severe neurological phenotype in Sanfilippo C patients. Cultures transfected with the WT

*HGSNAT* cDNA showed a progressive recovery in neuronal activity, demonstrating that the phenotype observed was due to the enzyme dysfunction. This test was the most accurate and sensitive to detect alterations in this cellular model, since differences can be detected already at 3 weeks for at least one of the patients.

Additionally, the progressive loss of network connectivity due to neurons alterations was explored. The results pointed out a gradual increase of connectivity in WT networks, probably due to its maturation. On the contrary, SFC-derived cultures showed lower connectivity, especially after 9 weeks of differentiation, and a loss of the network homogeneity with a small number of neurons retaining most of the connections. These results suggested that Sanfilippo disease promotes a fall in the connectivity and a disruption in the neuronal network structure, which could alter functionality. Once again, SFC6 showed a heavier phenotype compared to SFC7. Cultures treated with the WT *HGSNAT* cDNA showed a recovery trend, demonstrating one more time that the alterations found were due to the enzymatic dysfunction.

In summary, a neuronal model for Sanfilippo C syndrome that recapitulates the major features of the disease was established, highlighting the specific neuron phenotypes of this neurodegenerative disorder such as impaired neural activity and loss of network connectivity and structure. This is the first model to study the molecular mechanisms that lead to disease in the affected cell type and the first platform to assay possible therapeutic approaches. This is relevant since fibroblasts, used for routine diagnosis purposes, are not suitable to study neurological features. It should be noted that cells from two different patients were used and results were similar although some inter-individual differences were detected, which could be due to the nature of mutations or to the genetic background. Moreover, the fact that only one clone for each line was used, does not allow to assure that these divergences were not caused by differences in the reprogramming process. It has also been shown that the introduction of the WT *HGSNAT* cDNA into the affected neurons could be a good therapeutic tool for Sanfilippo C syndrome. Finally, neuronal activity and connectivity analyses allowed us to detect early phenotypic trends that could be further investigated in order to be used in the screening of drugs and treatments that could prevent the onset of neurodegeneration. Moreover, improved neuronal activity analyses *in vitro* could help the development of future *in vivo* analyses for the premature diagnosis of patients and their follow-up after any treatment.

The use of iPSC is still nowadays the preferable option to develop a neuronal model of diseases. Despite other strategies have been developed in the recent years such as direct conversion from fibroblasts to neurons (Vierbuchen et al., 2010) or the direct reprogramming to iNSC (Ring et al., 2012), the use of iPSC present advantages over them. Direct conversion does not pursue to obtain a permanent cell source, which means that conversion should be done constantly. For obvious reasons, this is not the best option to model a disease, but it must be considered an interesting tool for future cell replacement approaches and regenerative medicine. On the other hand, reprogramming to iNSC avoids the step of potential tumorigenic cells and, thus, it is also an interesting approach for regenerative medicine. In this case, a constant cell source is obtained, but restricted to neural lineage, which means that any phenotype affecting other cell lineages could not be studied. Taking into account all these considerations, iPSC still represent the best option to obtain human cellular models for several disorders, which could be used as a tool to study the molecular mechanisms causing diseases and as a platform in drug screening assays.



# **CONCLUSIONS**



- \* A mutational analysis in 15 Sanfilippo C syndrome patients from Spain, Morocco, France, Argentina, Sweden and Poland has been performed, identifying a total of 13 mutations, seven of which were novel. These mutations were: five missense mutations (p.A54V, p.L113P, p.G424V, p.L445P and p.S541L), four splicing mutations (c.234+1G>A, c.372-2A>G, c.633+1G>A and c.1378-1G>A), one nonsense mutation (p.R384\*), 2 small deletions (c.641delG and c.821-31 821-13del) and one duplication (c.1542+4dupA).
- \* Missense mutations (p.A54V, p.L113P, p.G424V, p.L445P and p.S541L) have been characterized using an *in vitro* expression assay in COS-7 cells, revealing negligible enzyme activities, and thus, demonstrating their pathogenicity.
- \* Mutations affecting the splicing process have been characterized at the mRNA level showing the splicing defects promoted by each mutation either in fibroblasts (c.234+1G>A, c.372-2A>G, c.633+1G>A and c.1378-1G>A) or using a minigene construct (c.821-31 821-13del).
- \* A single common origin has been suggested after haplotype studies for the c.234+1G>A mutation in Spanish and Moroccan patients and for the c.372-2G>A mutation in Spanish patients.
- \* Modified U1 snRNA sup4, totally complementary to the mutated 5' splice site, has been shown to partially restore the normal splicing process in patients' fibroblasts carrying the c.234+1G>A mutation in homozygosis. However, no improvement in the enzymatic activity was detected after treatment.
- \* Glucosamine has been shown to partially correct the protein misfolding of the c.372-2G>A mutated protein, lacking four amino acids, and to slightly increase the enzymatic activity.

- \* Small interfering RNAs targeting the *EXTL* genes, especially those targeting *EXTL2*, has been shown to delay the substrate accumulation in Sanfilippo C patients' fibroblasts suggesting their possible application for a substrate reduction therapy.
- \* Induced pluripotent stem cells from two different Sanfilippo C patients and one healthy control have been obtained and validated.
- \* These induced pluripotent stem cells have been differentiated to neurons that recapitulate the main features of the disease such as lack of enzyme activity, accumulation of glycosaminoglycans inside the lysosomes, increased size of these organelles, impaired autophagy and increased number of apoptotic cells.
- \* Specific analysis of the neuronal cultures revealed alterations in the neuronal activity patterns and affection in the neuronal network development and maintenance, demonstrating that the model mimics the patients' brain dysfunction.

# REFERENCES



**A**

- Aagaard, L., Rossi, J.J., **2007**. RNAi therapeutics: principles, prospects and challenges. *Adv. Drug Deliv. Rev.* 59, 75–86.
- Aartsma-Rus, A., van Ommen, G.J., **2007**. Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA* 13, 1609–24
- Aldenhoven, M., Boelens, J.J., de Koning, T.J., **2008**. The clinical outcome of Hurler syndrome after stem cell transplantation. *Biol. Blood Marrow Transplant.* 14, 485–98.
- Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., Jacobson, A., **2004**. A faux 30-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* 432, 112–8.
- Amrani, N., Sachs, M.S., Jacobson, A., **2006**. Early nonsense: mRNA decay solves a translational problem. *Nat. Rev. Mol. Cell Biol.* 7, 415–25.
- Applegarth, D., Toone, J., Lowry, R., **2000**. Incidence of inborn errors of metabolism in British Columbia, 1969–1996. *Pediatrics* 105, e10.
- Aronovich, E.L., Johnston, J.M., Wang, P., Giger, U., Whitley, C.B., **2001**. Molecular basis of mucopolysaccharidosis type IIIB in emu (*Dromaius novaehollandiae*): an avian model of Sanfilippo syndrome type B. *Genomics* 74, 299–305.

**B**

- Baehner, F., Schmiedeskamp, C., Krummenauer, F., Miebach, E., Bajbouj, M., Whybra, C., Kohlschütter, A., Kampmann, C., Beck, M., **2005**. Cumulative incidence rates of the mucopolysaccharidoses in Germany. *J. Inherit. Metab. Dis.* 28, 1011–7.
- Balestra, D., Faella, A., Margaritis, P., Cavallari, N., Pagani, F., Bernardi, F., Arruda, V.R., Pinotti, M., **2014**. An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. *J. Thromb. Haemost.* 12, 177–85.
- Ballabio, A., Gieselmann, V., **2009**. Lysosomal disorders: from storage to cellular damage. *Biochim. Biophys. Acta* 1793, 684–96.
- Banecka-Majkutewicz, Z., Jakóbkiewicz-Banecka, J., Gabig-Cimińska, M., Węgrzyn, A., Węgrzyn, G., **2012**. Putative biological mechanisms of efficiency of substrate reduction therapies for mucopolysaccharidoses. *Arch. Immunol. Ther. Exp.* 60, 461–8.
- Baralle, M., Baralle, D., Conti, L. De, Mattocks, C., Whittaker, J., Knezevich, A., Ffrench-Constant, C., Baralle, F.E., **2003**. Identification of a mutation that perturbs NF1 gene splicing using genomic DNA samples and a minigene assay. *J. Med. Genet.* 40, 220–2.

- Barbosa, I., Garcia, S., Barbier-Chassefière, V., Caruelle, J.-P., Martelly, I., Papy-García, D., **2003**. Improved and simple micro assay for sulfated glycosaminoglycans quantification in biological extracts and its use in skin and muscle tissue studies. *Glycobiology* 13, 647–53.
- Bartel, D.P., **2004**. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–97.
- Bartolomeo, R., Polishchuk, E. V, Volpi, N., Polishchuk, R.S., Auricchio, A., **2013**. Pharmacological read-through of nonsense ARSB mutations as a potential therapeutic approach for mucopolysaccharidosis VI. *J. Inherit. Metab. Dis.* 36, 363–71.
- Berger-Plantinga, E.G., Vanneste, J. a L., Groener, J.E.M., van Schooneveld, M.J., **2004**. Adult-onset dementia and retinitis pigmentosa due to mucopolysaccharidosis III-C in two sisters. *J. Neurol.* 251, 479–81.
- Bernards, R., **2006**. Exploring the uses of RNAi-Gene knockdown and the Nobel Prize. *N. Engl. J. Med.* 355, 2391–3.
- Bernier, V., Lagacé, M., Bichet, D.G., Bouvier, M., **2004**. Pharmacological chaperones: potential treatment for conformational diseases. *Trends Endocrinol. Metab.* 15, 222–8.
- Bhattacharyya, R., Gliddon, B., Beccari, T., Hopwood, J.J., Stanley, P., **2001**. A novel missense mutation in lysosomal sulfamidase is the basis of MPS III A in a spontaneous mouse mutant. *Glycobiology* 11, 99–103.
- Bhaumik, M., Muller, V.J., Rozaklis, T., Johnson, L., Dobrenis, K., Bhattacharyya, R., Wurzelmann, S., Finamore, P., Hopwood, J.J., Walkley, S.U., Stanley, P., **1999**. A mouse model for mucopolysaccharidosis type III A (Sanfilippo syndrome). *Glycobiology* 9, 1389–96.
- Bielicki, J., Hopwood, J., Anson, D., **1996**. Correction of Sanfilippo A skin fibroblasts by retroviral vector-mediated gene transfer. *Hum. Gene Ther.* 1970, 1965–70.
- Boelens, J.J., Prasad, V.K., Tolar, J., Wynn, R.F., Peters, C., **2010**. Current international perspectives on hematopoietic stem cell transplantation for inherited metabolic disorders. *Pediatr. Clin. North Am.* 57, 123–45.
- Botigué, L.R., Henn, B.M., Gravel, S., Maples, B.K., Gignoux, C.R., Corona, E., Atzmon, G., Burns, E., Ostrer, H., Flores, C., Bertranpetit, J., Comas, D., Bustamante, C.D., **2013**. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11791–6.
- Boustany, R.-M.N., **2013**. Lysosomal storage diseases - the horizon expands. *Nat. Rev. Neurol.* 9, 583–98.
- Brackenridge, S., Wilkie, A.O.M., Sreaton, G.R., **2003**. Efficient use of a “dead-end” GA 5’ splice site in the human fibroblast growth factor receptor genes. *EMBO J.* 22, 1620–31.

- Brooks, D. A., Muller, V.J., Hopwood, J.J., **2006**. Stop-codon read-through for patients affected by a lysosomal storage disorder. *Trends Mol. Med.* 12, 367–73.
- Bruyère, J., Roy, E., Ausseil, J., Lemonnier, T., Teyre, G., Bohl, D., Etienne-Manneville, S., Lortat-Jacob, H., Heard, J.M., Vitry, S., **2014**. Heparan sulphate saccharides modify focal adhesions: implication in mucopolysaccharidosis neuropathophysiology. *J. Mol. Biol.*
- Bunge, S., Ince, H., Steglich, C., Kleijer, W.J., Beck, M., Zaremba, J., van Diggelen, O.P., Weber, B., Hopwood, J.J., Gal, A., **1997**. Identification of 16 sulfamidase gene mutations including the common R74C in patients with mucopolysaccharidosis type IIIA (Sanfilippo A). *Hum. Mutat.* 10, 479–85.
- Busse, M., Feta, A., Presto, J., Wilén, M., Grønning, M., Kjellén, L., Kusche-Gullberg, M., **2007**. Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation. *J. Biol. Chem.* 282, 32802–10.
- Busse-Wicher, M., Wicher, K.B., Kusche-Gullberg, M., **2014**. The extostosin family: proteins with many functions. *Matrix Biol.* 35, 25–33.

## C

- Cartegni, L., Chew, S.L., Krainer, A.R., **2002**. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.* 3, 285–98.
- Chanda, S., Ang, C.E., Davila, J., Pak, C., Mall, M., Lee, Q.Y., Ahlenius, H., Jung, S.W., Südhof, T.C., Wernig, M., **2014**. Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Reports* 3, 282–96.
- Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R. A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., **2013**. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300.
- Ciechanover, A., **2005**. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat. Rev. Mol. Cell Biol.* 6, 79–86.
- Coppa, G., Galeotti, F., Zampini, L., Galeazzi, T., Padella, L., Santoro, L., Maccari, F., Gabrielli, O., Volpi, N., **2013**. Mild mental retardation and low levels of urinary heparan sulfate in a patient with the attenuated phenotype of mucopolysaccharidosis type IIIA. *Clin. Biochem.* 46, 688–90.
- Coutinho, M.F., Lacerda, L., Prata, M.J., Ribeiro, H., Lopes, L., Ferreira, C., Alves, S., **2008**. Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene. *Clin. Genet.* 74, 194–5.

Cox, T., Lachmann, R., Hollak, C., Aerts, J., van Weely, S., Hrebicek, M., Platt, F., Butters, T., Dwek, R., Moyses, C., Gow, I., Elstein, D., Zimran, A., **2000**. Novel oral treatment of Gaucher's disease with N-butyldoxynojirimycin (OGT 918) to decrease substrate biosynthesis. *Lancet* 355, 1481–5

Crawley, A.C., Marshall, N., Beard, H., Hassiotis, S., Walsh, V., King, B., Hucker, N., Fuller, M., Jolly, R.D., Hopwood, J.J., Hemsley, K.M., **2011**. Enzyme replacement reduces neuropathology in MPS IIIA dogs. *Neurobiol. Dis.* 43, 422–34.

## D

Davidson, B.L., McCray, P.B., **2011**. Current prospects for RNA interference-based therapies. *Nat. Rev. Genet.* 12, 329–40.

De Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., Appelmans, F., **1955**. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochem. J.* 60, 604–17.

De Duve, C., Wattiaux, R., **1966**. Functions of lysosomes. *Annu. Rev. Physiol.* 28, 435–92.

Dehay, B., Bové, J., Rodríguez-Muela, N., Perier, C., Recasens, A., Boya, P., Vila, M., **2010**. Pathogenic lysosomal depletion in Parkinson's disease. *J. Neurosci.* 30, 12535–44.

Delgado, V., O'Callaghan, M.D.M., Artuch, R., Montero, R., Pineda, M., **2011**. Genistein supplementation in patients affected by Sanfilippo disease. *J. Inherit. Metab. Dis.* 34, 1039–44.

Delgado, V., O'Callaghan, M.D.M., Gort, L., Coll, M.J., Pineda, M., **2013**. Natural history of Sanfilippo syndrome in Spain. *Orphanet J. Rare Dis.* 8, 189.

Deng, Y., Wang, C.C., Choy, K.W., Du, Q., Chen, J., Wang, Q., Li, L., Chung, T.K.H., Tang, T., **2014**. Therapeutic potentials of gene silencing by RNA interference: principles, challenges, and new strategies. *Gene* 538, 217–27.

Desnick, R.J., Schuchman, E.H., **2012**. Enzyme replacement therapy for lysosomal diseases: lessons from 20 years of experience and remaining challenges. *Annu. Rev. Genomics Hum. Genet.* 13, 307–35.

DeVincenzo, J.P., **2012**. The promise, pitfalls and progress of RNA-interference-based antiviral therapy for respiratory viruses. *Antivir. Ther.* 17, 213–25.

Di Malta, C., Fryer, J.D., Settembre, C., Ballabio, A., **2012**. Astrocyte dysfunction triggers neurodegeneration in a lysosomal storage disorder. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2334–42.

Di Natale, P., Balzano, N., Esposito, S., Villani, G.R., **1998**. Identification of molecular defects in Italian Sanfilippo A patients including 13 novel mutations. *Hum. Mutat.* 11, 313–20.

- Downs-Kelly, E., Jones, M.Z., Alroy, J., Cavanagh, K.T., King, B., Lucas, R.E., Baker, J.C., Kraemer, S.A., Hopwood, J.J., **2000**. Caprine mucopolysaccharidosis IIID: a preliminary trial of enzyme replacement therapy. *J. Mol. Neurosci.* 15, 251–62.
- Durand, S., Feldhammer, M., Bonneil, E., Thibault, P., Pshezhetsky, A. V, **2010**. Analysis of the biogenesis of heparan sulfate acetyl-CoA:alpha-glucosaminide N-acetyltransferase provides insights into the mechanism underlying its complete deficiency in mucopolysaccharidosis IIIC. *J. Biol. Chem.* 285, 31233–42.
- Dziedzic, D., Wegrzyn, G., Jakóbkiewicz-Banecka, J., **2010**. Impairment of glycosaminoglycan synthesis in mucopolysaccharidosis type IIIA cells by using siRNA: a potential therapeutic approach for Sanfilippo disease. *Eur. J. Hum. Genet.* 18, 200–5.

## E

- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T., **2001**. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494–8.
- Ellinwood, N.M., Ausseil, J., Desmaris, N., Bigou, S., Liu, S., Jens, J.K., Snella, E.M., Mohammed, E.E.A., Thomson, C.B., Raoul, S., Joussemet, B., Roux, F., Chérel, Y., Lajat, Y., Piraud, M., Benchaouir, R., Hermening, S., Petry, H., Froissart, R., Tardieu, M., Ciron, C., Moullier, P., Parkes, J., Kline, K.L., Maire, I., Vanier, M.-T., Heard, J.-M., Colle, M.-A., **2011**. Safe, efficient, and reproducible gene therapy of the brain in the dog models of Sanfilippo and Hurler syndromes. *Mol. Ther.* 19, 251–9.
- Ellinwood, N.M., Vite, C.H., Haskins, M.E., **2004**. Gene therapy for lysosomal storage diseases: the lessons and promise of animal models. *J. Gene Med.* 6, 481–506.
- Ellinwood, N.M., Wang, P., Skeen, T., Sharp, N.J.H., Cesta, M., Decker, S., Edwards, N.J., Bublot, I., Thompson, J.N., Bush, W., Hardam, E., Haskins, M.E., Giger, U., **2003**. A model of mucopolysaccharidosis IIIB (Sanfilippo syndrome type IIIB): N-acetyl-alpha-D-glucosaminidase deficiency in Schipperke dogs. *J. Inherit. Metab. Dis.* 26, 489–504.
- Emre, S., Terzioglu, M., Tokatli, A., Coskun, T., Ozalp, I., Weber, B., Hopwood, J.J., **2002**. Sanfilippo syndrome in Turkey: identification of novel mutations in subtypes A and B. *Hum. Mutat.* 19, 184–5.
- Esko, J., Zhang, L., **1996**. Influence of core protein sequence on glycosaminoglycan assembly. *Curr. Opin. Struct. Biol.* 6, 663–70.
- Esko, J.D., Selleck, S.B., **2002**. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* 71, 435–71.

## F

- Fan, X., Tkachyova, I., Sinha, A., Rigat, B., Mahuran, D., **2011**. Characterization of the biosynthesis, processing and kinetic mechanism of action of the enzyme deficient in mucopolysaccharidosis IIIC. *PLoS One* 6, e24951.
- Fan, X., Zhang, H., Zhang, S., Bagshaw, R.D., Tropak, M.B., Callahan, J.W., Mahuran, D.J., **2006**. Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). *Am. J. Hum. Genet.* 79, 738–44.
- Fedele, A.O., Filocamo, M., Di Rocco, M., Sersale, G., Lübke, T., Di Natale, P., Cosma, M.P., Ballabio, A., **2007**. Mutational analysis of the HGSNAT gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Hum. Mutat.* 28, 523.
- Fedele, A.O., Hopwood, J.J., **2010**. Functional analysis of the HGSNAT gene in patients with mucopolysaccharidosis IIIC (Sanfilippo C Syndrome). *Hum. Mutat.* 31, E1574–86.
- Feldhammer, M., Durand, S., Mrázová, L., Boucher, R.-M., Laframboise, R., Steinfeld, R., Wraith, J.E., Michelakakis, H., van Diggelen, O.P., Hřebíček, M., Kmoch, S., Pshezhetsky, A. V, **2009a**. Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: alpha-glucosaminide N-acetyltransferase (HGSNAT) gene. *Hum. Mutat.* 30, 918–25.
- Feldhammer, M., Durand, S., Pshezhetsky, A. V, **2009b**. Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C. *PLoS One* 4, e7434.
- Fernandez Alanis, E., Pinotti, M., Dal Mas, A., Balestra, D., Cavallari, N., Rogalska, M.E., Bernardi, F., Pagani, F., **2012**. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. *Hum. Mol. Genet.* 21, 2389–98.
- Fischer, A., Carmichael, K.P., Munnell, J.F., Jhabvala, P., Thompson, J.N., Matalon, R., Jezyk, P.F., Wang, P., Giger, U., **1998**. Sulfamidase deficiency in a family of dachshunds: a canine model of mucopolysaccharidosis IIIA (Sanfilippo A). *Pediatr. Res.* 44, 74–82.
- Fire, A., Xu, S., Montgomery, M., Kostas, S., Driver, S., Mello, C., **1998**. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–11.
- Fraldi, A., Hemsley, K., Crawley, A., Lombardi, A., Lau, A., Sutherland, L., Auricchio, A., Ballabio, A., Hopwood, J.J., **2007**. Functional correction of CNS lesions in an MPS-IIIA mouse model by intracerebral AAV-mediated delivery of sulfamidase and SUMF1 genes. *Hum. Mol. Genet.* 16, 2693–702.
- Fransson, L., Belting, M., Cheng, F., Jönsson, M., Mani, K., Sandgren, S., **2004**. Novel aspects of glypican glycobiology. *Cell. Mol. Life Sci.* 61, 1016–24.

- Freeman, C., Hopwood, J., **1992**. Human alpha-L-iduronidase: catalytic properties and an integrated role in the lysosomal degradation of heparan sulphate. *Biochem. J.* 282, 899–908.
- Fu, H., Bartz, J.D., Stephens, R.L., McCarty, D.M., **2012**. Peripheral nervous system neuropathology and progressive sensory impairments in a mouse model of Mucopolysaccharidosis IIIB. *PLoS One* 7, e45992.
- Fu, H., Kang, L., Jennings, J.S., Moy, S.S., Perez, A., Dirosario, J., McCarty, D.M., Muenzer, J., **2007**. Significantly increased lifespan and improved behavioral performances by rAAV gene delivery in adult mucopolysaccharidosis IIIB mice. *Gene Ther.* 14, 1065–77.
- Futerman, A.H., van Meer, G., **2004**. The cell biology of lysosomal storage disorders. *Nat. Rev. Mol. Cell Biol.* 5, 554–65.
- ## G
- Gabrielli, O., Coppa, G.V., Bruni, S., Villani, G.R., Pontarelli, G., Di Natale, N.P., **2005**. An adult Sanfilippo type A patient with homozygous mutation R206P in the sulfamidase gene. *Am. J. Med. Genet.* 133, 85–9.
- Garbuzova-Davis, S., Klasko, S.K., Sanberg, P.R., **2009**. Intravenous administration of human umbilical cord blood cells in an animal model of MPS III B. *J. Comp. Neurol.* 515, 93–101.
- Garbuzova-Davis, S., Willing, A., Desjarlais, T., Sanberg, C., Sanberg, P., **2005**. Transplantation of human umbilical cord blood cells benefits an animal model of Sanfilippo syndrome type B. *Stem Cells Dev.* 394, 384–94.
- Gille, L., Nohl, H., **2000**. The existence of a lysosomal redox chain and the role of ubiquinone. *Arch. Biochem. Biophys.* 375, 347–54.
- Giugliani, R., Federhen, A., Silva, A.A., Bittar, C.M., Moura de Souza, C.F., Brinckmann, C., Netto, O., Mayer, F.Q., Baldo, G., Matte, U., **2012**. Emerging treatment options for the mucopolysaccharidoses. *Res. Reports Endocr. Disord.* 2, 53–64.
- Glaus, E., Schmid, F., Da Costa, R., Berger, W., Neidhardt, J., **2011**. Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells. *Mol. Ther.* 19, 936–41.
- Gliddon, B., Hopwood, J., **2004**. Enzyme-replacement therapy from birth delays the development of behavior and learning problems in mucopolysaccharidosis type IIIA mice. *Pediatr. Res.* 56, 65–72.
- Gong, F., Jemth, P., Escobar Galvis, M.L., Vlodaysky, I., Horner, A., Lindahl, U., Li, J., **2003**. Processing of macromolecular heparin by heparanase. *J. Biol. Chem.* 278, 35152–8.

González, F., Boué, S., Izpisua Belmonte, J.C., **2011**. Methods for making induced pluripotent stem cells: reprogramming à la carte. *Nat. Rev. Genet.* 12, 231–42.

## H

Habib, O., Habib, G., Choi, H.W., Hong, K.S., Do, J.T., Moon, S.H., Chung, H.M., **2013**. An improved method for the derivation of high quality iPSCs in the absence of c-Myc. *Exp. Cell Res.* 319, 190–200.

Hamilton, A.J., Baulcombe, D.C., **1999**. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–2.

Hartmann, L., Neveling, K., Borkens, S., Schneider, H., Freund, M., Grassman, E., Theiss, S., Wawer, A., Burdach, S., Auerbach, A.D., Schindler, D., Hanenberg, H., Schaal, H., **2010**. Correct mRNA processing at a mutant TT splice donor in FANCC ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs. *Am. J. Hum. Genet.* 87, 480–93.

Hassiotis, S., Beard, H., Luck, A., Trim, P.J., King, B., Snel, M.F., Hopwood, J.J., Hemsley, K.M., **2014**. Disease stage determines the efficacy of treatment of a paediatric neurodegenerative disease. *Eur. J. Neurosci.* 39, 2139–50.

Haurigot, V., Marcó, S., Ribera, A., Garcia, M., Ruzo, A., Villacampa, P., Ayuso, E., Añor, S., Andaluz, A., Pineda, M., García-Fructuoso, G., Molas, M., Maggioni, L., Muñoz, S., Motas, S., Ruberte, J., Mingozzi, F., Pumarola, M., Bosch, F., **2013**. Whole body correction of mucopolysaccharidosis IIIA by intracerebrospinal fluid gene therapy. *J. Clin. Invest.*

Heldermon, C.D., Qin, E.Y., Ohlemiller, K.K., Herzog, E.D., Brown, J.R., Vogler, C., Hou, W., Orrock, J.L., Crawford, B.E., Sands, M.S., **2013**. Disease correction by combined neonatal intracranial AAV and systemic lentiviral gene therapy in Sanfilippo Syndrome type B mice. *Gene Ther.* 20, 913–21.

Hemsley, K.M., Beard, H., King, B.M., Hopwood, J.J., **2008**. Effect of high dose, repeated intra-CSF injection of sulphamidase on neuropathology in MPS IIIA mice. *Genes Brain Behav.* 7, 740–53.

Hemsley, K.M., Hopwood, J.J., **2005**. Development of motor deficits in a murine model of mucopolysaccharidosis type IIIA (MPS-IIIa). *Behav. Brain Res.* 158, 191–9.

Hemsley, K.M., Luck, A.J., Crawley, A.C., Hassiotis, S., Beard, H., King, B., Rozek, T., Rozaklis, T., Fuller, M., Hopwood, J.J., **2009a**. Examination of intravenous and intra-CSF protein delivery for treatment of neurological disease. *Eur. J. Neurosci.* 29, 1197–214.

Hemsley, K.M., Norman, E.J., Crawley, A.C., Auclair, D., King, B., Fuller, M., Lang, D.L., Dean, C.J., Jolly, R.D., Hopwood, J.J., **2009b**. Effect of cisternal sulfamidase delivery in MPS IIIA Huntaway dogs - a proof of principle study. *Mol. Genet. Metab.* 98, 383–92.

- Héron, B., Mikaeloff, Y., Froissart, R., Caridade, G., Maire, I., Caillaud, C., Levade, T., Chabrol, B., Feillet, F., Ogier, H., Valayannopoulos, V., Michelakakis, H., Zafeiriou, D., Lavery, L., Wraith, E., Danos, O., Heard, J.-M., Tardieu, M., **2011**. Incidence and natural history of mucopolysaccharidosis type III in France and comparison with United Kingdom and Greece. *Am. J. Med. Genet.* 155A, 58–68.
- Hers, H., **1963**. alpha-Glucosidase deficiency in generalized glycogen-storage disease (Pompe's Disease). *Biochem. J.* 86, 11–6.
- Hertel, K., **2014**. Spliceosomal pre-mRNA splicing, *Methods in Molecular Biology*.
- Higuchi, T., Kawagoe, S., Otsu, M., Shimada, Y., Kobayashi, H., Hirayama, R., Eto, K., Ida, H., Ohashi, T., Nakauchi, H., Eto, Y., **2014**. The generation of induced pluripotent stem cells (iPSCs) from patients with infantile and late-onset types of Pompe disease and the effects of treatment with acid- $\alpha$ -glucosidase in Pompe's iPSCs. *Mol. Genet. Metab.* 112, 44–8.
- Hilleren, P., Parker, R., **1999**. mRNA surveillance in eukaryotes: kinetic proofreading of proper translation termination as assessed by mRNP domain organization? *RNA* 5, 711–9
- Holbrook, J. A., Neu-Yilik, G., Hentze, M.W., Kulozik, A.E., **2004**. Nonsense-mediated decay approaches the clinic. *Nat. Genet.* 36, 801–8.
- Hollak, C.E.M., Wijburg, F.A., **2014**. Treatment of lysosomal storage disorders: successes and challenges. *J. Inherit. Metab. Dis.* 37, 587–98.
- Hoogerbrugge, P.M., Brouwer, O.F., Bordigoni, P., Ringden, O., Kapaun, P., Ortega, J.J., O'Meara, A., Cornu, G., Souillet, G., Frappaz, D., **1995**. Allogeneic bone marrow transplantation for lysosomal storage diseases. *Lancet* 345, 1398–402.
- Hopwood, J.J., **2005**. Prenatal diagnosis of Sanfilippo syndrome. *Prenat. Diagn.* 25, 148–50.
- Hřebíček, M., Mrázová, L., Seyrantepe, V., Durand, S., Roslin, N., Nosková, L., Hartmannová, H., Ivánek, R., Cizková, A., Poupetová, H., Sikora, J., Urinovská, J., Stránecký, V., Zeman, J., Lepage, P., Roquis, D., Verner, A., Ausseil, J., Beesley, C., Maire, I., Poorthuis, B., van de Kamp, J., van Diggelen, O., Wevers, R., Hudson, T., Fujiwara, T., Majewski, J., Morgan, K., Knoch, S., Pshezhetsky, A., **2006**. Mutations in TMEM76 cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Am. J. Hum. Genet.* 79, 807–19.
- Huang, H.-P., Chen, P.-H., Hwu, W.-L., Chuang, C.-Y., Chien, Y.-H., Stone, L., Chien, C.-L., Li, L.-T., Chiang, S.-C., Chen, H.-F., Ho, H.-N., Chen, C.-H., Kuo, H.-C., **2011**. Human Pompe disease-induced pluripotent stem cells for pathogenesis modeling, drug testing and disease marker identification. *Hum. Mol. Genet.* 20, 4851–64.
- Huang, H.-P., Chuang, C.-Y., Kuo, H.-C., **2012**. Induced pluripotent stem cell technology for disease modeling and drug screening with emphasis on lysosomal storage diseases. *Stem Cell Res. Ther.* 3, 34.

- Huh, H.J., Seo, J.Y., Cho, S.Y., Ki, C.-S., Lee, S.-Y., Kim, J.-W., Park, H.-D., Jin, D.-K., **2013**. The first Korean case of mucopolysaccharidosis IIIC (Sanfilippo syndrome type C) confirmed by biochemical and molecular investigation. *Ann. Lab. Med.* 33, 75–9.
- Humbel, R., Chamoles, N.A., **1972**. Sequential thin layer chromatography of urinary acidic glycosaminoglycans. *Clin. Chim. Acta.* 40, 290-3.
- Hwang, D.Y., Cohen, J.B., **1996**. U1 snRNA promotes the selection of nearby 5' splice sites by U6 snRNA in mammalian cells. *Genes Dev.* 10, 338–50.

## J

- Jakóbkiewicz-Banecka, J., Piotrowska, E., Narajczyk, M., Barańska, S., Wegrzyn, G., **2009**. Genistein-mediated inhibition of glycosaminoglycan synthesis, which corrects storage in cells of patients suffering from mucopolysaccharidoses, acts by influencing an epidermal growth factor-dependent pathway. *J. Biomed. Sci.* 16, 26.
- John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., Marks, D.S., **2004**. Human microRNA targets. *PLoS Biol.* 2, e363.
- Jolly, R.D., Marshall, N.R., Perrott, M.R., Dittmer, K.E., Hemsley, K.M., Beard, H., **2011**. Intracisternal enzyme replacement therapy in lysosomal storage diseases: routes of absorption into brain. *Neuropathol. Appl. Neurobiol.* 37, 414–22.
- Joo, M.K., Yhee, J.Y., Kim, S.H., Kim, K., **2014**. The potential and advances in RNAi therapy: chemical and structural modifications of siRNA molecules and use of biocompatible nanocarriers. *J. Control. Release.*

## K

- Kaidonis, X., Liaw, W.C., Roberts, A.D., Ly, M., Anson, D., Byers, S., **2010**. Gene silencing of EXTL2 and EXTL3 as a substrate deprivation therapy for heparan sulphate storing mucopolysaccharidoses. *Eur. J. Hum. Genet.* 18, 194–9.
- Kaji, T., Kawashima, T., Sakamoto, M., **1991**. Rhodamine B inhibition of glycosaminoglycan production by cultured human lip fibroblasts. *Toxicol. Appl. Pharmacol.* 111, 82–9
- Kan, S., Troitskaya, L. A., Sinow, C.S., Haitz, K., Todd, A.K., Di Stefano, A., Le, S.Q., Dickson, P.I., Tippin, B.L., **2014**. Insulin-like growth factor II peptide fusion enables uptake and lysosomal delivery of  $\alpha$ -N-acetylglucosaminidase to mucopolysaccharidosis type IIIB fibroblasts. *Biochem. J.* 458, 281–9.
- Keeling, K.M., Wang, D., Dai, Y., Murugesan, S., Chenna, B., Clark, J., Belakhov, V., Kandasamy, J., Velu, S.E., Baasov, T., Bedwell, D.M., **2013**. Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression. *PLoS One* 8, e60478.

- Keeling, K.M., Xue, X., Gunn, G., Bedwell, D.M., **2014**. Therapeutics based on stop codon readthrough. *Annu. Rev. Genomics Hum. Genet.* 15, 1–24.
- Kim, B.T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U., Sugahara, K., **2001**. Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4-N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 98, 7176–81.
- Kim, K.H., Dodsworth, C., Paras, A., Burton, B.K., **2013**. High dose genistein aglycone therapy is safe in patients with mucopolysaccharidoses involving the central nervous system. *Mol. Genet. Metab.* 109, 382–5.
- Kim, S.U., **2014**. Lysosomal storage diseases: Stem cell-based cell- and gene-therapy. *Cell Transplant.*
- Kitagawa, H., Shimakawa, H., Sugahara, K., **1999**. The tumor suppressor EXT-like gene EXTL2 encodes an 1,4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region: the key enzyme for the chain initiation of heparan sulfate. *J. Biol. Chem.* 274, 13933–7.
- Kleijer, W.J., Karpova, E.A., Geilen, G.C., Keulemans, J.L., Huijmans, J.G., Tsvetkova, I.V., Voznyi, Y.V., van Diggelen, O.P., **1996**. Prenatal diagnosis of Sanfilippo A syndrome: experience in 35 pregnancies at risk and the use of a new fluorogenic substrate for the heparin sulphamidase assay. *Prenat. Diagn.* 16, 829-35.
- Klein, U., Kresse, H., von Figura, K., **1978**. Sanfilippo syndrome type C: deficiency of acetyl-CoA:alpha-glucosaminide N-acetyltransferase in skin fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 75, 5185–9.
- Kloska, A., Jakóbkiewicz-Banecka, J., Narajczyk, M., Banecka-Majkutewicz, Z., Węgrzyn, G., **2011**. Effects of flavonoids on glycosaminoglycan synthesis: implications for substrate reduction therapy in Sanfilippo disease and other mucopolysaccharidoses. *Metab. Brain Dis.* 26, 1–8.
- Kohan, R., Cismondi, I., Oller-Ramirez, A., Guelbert, N., Anzolini, V., Alonso, G., Mole, S., Dodelson de Kremer, R., Noher de Halac, I., **2011**. Therapeutic approaches to the challenge of neuronal ceroid lipofuscinoses. *Curr. Pharmacol. Biotechnol.* 12, 867–83.
- Kreuger, J., Kjellén, L., **2012**. Heparan sulfate biosynthesis: regulation and variability. *J. Histochem. Cytochem.* 60, 898–907.
- Krivit, W., Sung, J.H., Shapiro, E.G., Lockman, L.A., **1995**. Microglia: the effector cell for reconstitution of the central nervous system following bone marrow transplantation for lysosomal and peroxisomal storage diseases. *Cell Transplant.* 4, 385–92.
- Kurihara, M., Kumagai, K., Yagishita, S., **1996**. Sanfilippo syndrome type C: a clinicopathological autopsy study of a long-term survivor. *Pediatr. Neurol.* 14, 317–21.

## L

- Lachmann, R., **2010**. Treatments for lysosomal storage disorders. *Biochem. Soc. Trans.* 38, 1465–8.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., **2001**. Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–8.
- Lamanna, W.C., Lawrence, R., Sarrazin, S., Esko, J.D., **2011**. Secondary storage of dermatan sulfate in Sanfilippo disease. *J. Biol. Chem.* 286, 6955–62.
- Langford-Smith, A., Wilkinson, F.L., Langford-Smith, K.J., Holley, R.J., Sergijenko, A., Howe, S.J., Bennett, W.R., Jones, S. A., Wraith, J., Merry, C.L., Wynn, R.F., Bigger, B.W., **2012**. Hematopoietic stem cell and gene therapy corrects primary neuropathology and behavior in mucopolysaccharidosis IIIA mice. *Mol. Ther.* 20, 1610–21.
- Lau, A., Shamsani, N., Winner, L., **2013**. Neonatal bone marrow transplantation in MPS IIIA mice. *JIMD Rep.* 8, 121–32.
- Lau, A.A., Hannouche, H., Rozaklis, T., Hassiotis, S., Hopwood, J.J., Hemsley, K.M., **2010a**. Allogeneic stem cell transplantation does not improve neurological deficits in mucopolysaccharidosis type IIIA mice. *Exp. Neurol.* 225, 445–54.
- Lau, A.A., Hopwood, J.J., Kremer, E.J., Hemsley, K.M., **2010b**. SGSH gene transfer in mucopolysaccharidosis type IIIA mice using canine adenovirus vectors. *Mol. Genet. Metab.* 100, 168–75.
- Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., **2001**. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–62.
- Le Hir, H., Moore, M.J., Maquat, L.E., **2000**. Pre-mRNA splicing alters mRNP composition: Evidence for stable association of proteins at exon–exon junctions. *Genes. Dev.* 14, 1098–108
- Lee, R.C., Ambros, V., **2001**. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–4.
- Lefrancois, S., Zeng, J., Hassan, A., Canuel, M., Morales, C.R., **2003**. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. *EMBO J.* 22, 6430–7.
- Lemonnier, T., Blanchard, S., Toli, D., Roy, E., Bigou, S., Froissart, R., Rouvet, I., Vitry, S., Heard, J.M., Bohl, D., **2011**. Modeling neuronal defects associated with a lysosomal disorder using patient-derived induced pluripotent stem cells. *Hum. Mol. Genet.* 20, 3653–66.
- Li, H.H., Yu, W.H., Rozengurt, N., Zhao, H.Z., Lyons, K.M., Anagnostaras, S., Fanselow, M.S., Suzuki, K., Vanier, M.T., Neufeld, E.F., **1999**. Mouse model of Sanfilippo syndrome type B produced by targeted disruption of the gene encoding alpha-N-acetylglucosaminidase. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14505–10.

Lin, H.-Y., Lin, S.-P., Chuang, C.-K., Niu, D.-M., Chen, M.-R., Tsai, F.-J., Chao, M.-C., Chiu, P.-C., Lin, S.-J., Tsai, L.-P., Hwu, W.-L., Lin, J.-L., **2009**. Incidence of the mucopolysaccharidoses in Taiwan, 1984-2004. *Am. J. Med. Genet.* 149A, 960–4.

## M

Maetzel, D., Sarkar, S., Wang, H., Abi-Mosleh, L., Xu, P., Cheng, A.W., Gao, Q., Mitalipova, M., Jaenisch, R., **2014**. Genetic and chemical correction of cholesterol accumulation and impaired autophagy in hepatic and neural cells derived from Niemann-Pick type C patient-specific iPSCs. *Stem cell reports* 2, 866–80.

Malinová, V., Węgrzyn, G., Narajczyk, M., **2012**. The use of elevated doses of genistein-rich soy extract in the gene expression-targeted isoflavone therapy for Sanfilippo disease patients. *JIMD Rep.* 5, 21–5.

Malinowska, M., Wilkinson, F.L., Bennett, W., Langford-Smith, K.J., O’Leary, H.A., Jakobkiewicz-Banecka, J., Wynn, R., Wraith, J.E., Węgrzyn, G., Bigger, B.W., **2009**. Genistein reduces lysosomal storage in peripheral tissues of mucopolysaccharide IIIB mice. *Mol. Genet. Metab.* 98, 235–42.

Malinowska, M., Wilkinson, F.L., Langford-Smith, K.J., Langford-Smith, A., Brown, J.R., Crawford, B.E., Vanier, M.T., Gryniewicz, G., Wynn, R.F., Wraith, J.E., Węgrzyn, G., Bigger, B.W., **2010**. Genistein improves neuropathology and corrects behaviour in a mouse model of neurodegenerative metabolic disease. *PLoS One* 5, e14192.

Malm, G., Månsson, J.-E., **2010**. Mucopolysaccharidosis type III (Sanfilippo disease) in Sweden: clinical presentation of 22 children diagnosed during a 30-year period. *Acta Paediatr.* 99, 1253–7.

Mangas, M., Nogueira, C., Prata, M.J., Lacerda, L., Coll, M.J., Soares, G., Ribeiro, G., Amaral, O., Ferreira, C., Alves, C., Coutinho, M.F., Alves, S., **2008**. Molecular analysis of mucopolysaccharidosis type IIIB in Portugal: evidence of a single origin for a common mutation (R234C) in the Iberian Peninsula. *Clin. Genet.* 73, 251–6.

Marchetto, M.C., Brennand, K.J., Boyer, L.F., Gage, F.H., **2011**. Induced pluripotent stem cells (iPSCs) and neurological disease modeling: progress and promises. *Hum. Mol. Genet.* 20, R109–15.

Marshall, N.R., Hassiotis, S., King, B., Rozaklis, T., Trim, P.J., Duplock, S.K., Winner, L.K., Beard, H., Snel, M.F., Jolly, R.D., Hopwood, J.J., Hemsley, K.M., **2014**. Delivery of therapeutic protein for prevention of neurodegenerative changes: comparison of different CSF-delivery methods. *Exp. Neurol.*

Matalonga, L., Arias, A., Coll, M.J., Garcia-Villoria, J., Gort, L., Ribes, A., **2014**. Treatment effect of coenzyme Q10 and an antioxidant cocktail in fibroblasts of patients with Sanfilippo disease. *J. Inherit. Metab. Dis.* 37, 439–46.

- Mattioli, C., Pianigiani, G., De Rocco, D., Bianco, A.M.R., Cappelli, E., Savoia, A., Pagani, F., **2014**. Unusual splice site mutations disrupt FANCA exon 8 definition. *Biochim. Biophys. Acta* 1842, 1052–8.
- Mazzulli, J.R., Xu, Y.-H., Sun, Y., Knight, A.L., McLean, P.J., Caldwell, G.A., Sidransky, E., Grabowski, G.A., Krainc, D., **2011**. Gaucher disease glucocerebrosidase and  $\alpha$ -synuclein form a bidirectional pathogenic loop in synucleinopathies. *Cell* 146, 37–52.
- McIntyre, C., Byers, S., Anson, D., **2010**. Correction of mucopolysaccharidosis type IIIA somatic and central nervous system pathology by lentiviral-mediated gene transfer. *J. Gene Med.* 12, 717–28.
- McIntyre, C., Derrick Roberts, A.L., Ranieri, E., Clements, P.R., Byers, S., Anson, D.S., **2008**. Lentiviral-mediated gene therapy for murine mucopolysaccharidosis type IIIA. *Mol. Genet. Metab.* 93, 411–8.
- Medina, D.L., Fraldi, A., Bouche, V., Annunziata, F., Mansueto, G., Spampanato, C., Puri, C., Pignata, A., Martina, J.A., Sardiello, M., Palmieri, M., Polishchuk, R., Puertollano, R., Ballabio, A., **2011**. Transcriptional activation of lysosomal exocytosis promotes cellular clearance. *Dev. Cell.* 21, 421–30.
- Meikle, P.J., Hopwood, J.J., Clague, A.E., Carey, W.F., **1999**. Prevalence of lysosomal storage disorders. *JAMA* 281, 249–54.
- Meng, X.-L., Shen, J.-S., Kawagoe, S., Ohashi, T., Brady, R.O., Eto, Y., **2010**. Induced pluripotent stem cells derived from mouse models of lysosomal storage disorders. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7886–91.
- Meyer, A., Kossow, K., Gal, A., Steglich, C., Mühlhausen, C., Ullrich, K., Braulke, T., Muschol, N., **2008**. The Mutation p.Ser298Pro in the sulphamidase gene (SGSH) is associated with a slowly progressive clinical phenotype in mucopolysaccharidosis type IIIA (Sanfilippo A syndrome). *Hum. Mutat.* 29, 770.
- Michelakakis, H., Dimitriou, E., Tsagaraki, S., Giouroukos, S., Schulpis, K., Bartsocas, C.S., **1995**. Lysosomal storage diseases in Greece. *Genet. Couns.* 6, 43–7.
- Montfort, M., Vilageliu, L., Garcia-Giralt, N., Guidi, S., Coll, M.J., Chabas, A., Grinberg, D., **1998**. Mutation 1091delC is highly prevalent in Spanish Sanfilippo syndrome type A patients. *Hum. Mutat.* 12, 274–9.
- Moog, U., van Mierlo, I., van Schrojenstein Lantman-de Valk, H., Spaapen, L., Maaskant, M., Curfs, L., **2007**. Is Sanfilippo type B in your mind when you see adults with mental retardation and behavioral problems? *Am. J. Med. Genet.* 145C, 293–301.
- Moskot, M., Montefusco, S., Jakóbkiewicz-Banecka, J., Mozolewski, P., Węgrzyn, A., Di Bernardo, D., Węgrzyn, G., Medina, D.L., Ballabio, A., Gabig-Cimińska, M., **2014**. The phytoestrogen genistein modulates lysosomal metabolism and transcription factor EB (TFEB) activation. *J. Biol. Chem.* 289, 17054–69.

Muchowski, P.J., Wacker, J.L., **2005**. Modulation of neurodegeneration by molecular chaperones. *Nat. Rev. Neurosci.* 6, 11–22.

Muenzer, J., **1986**. Mucopolysaccharidoses. *Adv. Pediatr.* 33, 269–302.

Murrey, D.A., Naughton, B.J., Duncan, F.J., Meadows, A.S., Ware, T.A., Campbell, K.J., Bremer, W.G., Walker, C.M., Goodchild, L., Bolon, B., La Perle, K., Flanigan, K.M., McBride, K.L., McCarty, D.M., Fu, H., **2014**. Feasibility and safety of systemic rAAV9-hNAGLU delivery for treating mucopolysaccharidosis IIIB: toxicology, biodistribution, and immunological assessments in primates. *Hum. Gene Ther. Clin. Dev.* 25, 72–84.

## N

Nadanaka, S., Zhou, S., Kagiya, S., Shoji, N., Sugahara, K., Sugihara, K., Asano, M., Kitagawa, H., **2013**. EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a xylose kinase-dependent manner. *J. Biol. Chem.* 288, 9321–33.

Nelson, J., Crowhurst, J., Carey, B., Greed, L., **2003**. Incidence of the mucopolysaccharidoses in Western Australia. *Am. J. Med. Genet.* 123A, 310–3.

Neufeld, E.F., Muenzer, J., **2001**. The Mucopolysaccharidoses.

Nissim-Rafinia, M., Kerem, B., **2005**. The splicing machinery is a genetic modifier of disease severity. *Trends Genet.* 21, 480–3.

Nudelman, I., Rebibo-Sabbah, A., Cherniavsky, M., Belakhov, V., Hainrichson, M., Chen, F., Schacht, J., Pilch, D.S., Ben-Yosef, T., Baasov, T., **2009**. Development of novel aminoglycoside (NB54) with reduced toxicity and enhanced suppression of disease-causing premature stop mutations. *J. Med. Chem.* 52, 2836–45.

## O

Ohkubo, A., Kondo, Y., Suzuki, M., Kobayashi, H., Kanamori, T., Masaki, Y., Seio, K., Nagai, K., Sekine, M., **2013**. Chemical synthesis of U1 snRNA derivatives. *Org. Lett.* 15, 4386–9.

Ohmi, K., Greenberg, D.S., Rajavel, K.S., Ryazantsev, S., Li, H.H., Neufeld, E.F., **2003**. Activated microglia in cortex of mouse models of mucopolysaccharidoses I and IIIB. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1902–7.

Ortolano, S., Viéitez, I., Navarro, C., Spuch, C., **2014**. Treatment of lysosomal storage diseases: recent patents and future strategies. *Recent Pat. Endocr. Metab. Immune Drug Discov.* 8, 9–25.

Ouesleti, S., Brunel, V., Ben Turkia, H., Dranguet, H., Miled, A., Miladi, N., Ben Dridi, M.F., Lavoine, A., Saugier-veber, P., Bekri, S., **2011**. Molecular characterization of MPS IIIA, MPS IIIB and MPS IIIC in Tunisian patients. *Clin. Chim. Acta* 412, 2326–31.

## P

Palmucci, S., Attinà, G., Lanza, M.L., Belfiore, G., Cappello, G., Foti, P.V., Milone, P., Di Bella, D., Barone, R., Fiumara, A., Sorge, G., Ettorre, G.C., **2013**. Imaging findings of mucopolysaccharidoses: a pictorial review. *Insights imaging* 4, 443–59.

Panicker, L., Miller, D., Park, T.S., Patel, B., Azevedo, J.L., Awad, O., Masood, M.A., Veenestra, T.D., Goldin, E., Stubblefield, B.K., Tayebi, N., Polumuri, S.K., Vogel, S.N., Sidransky, E., Zambidis, E.T., Feldman, R.A., **2012**. Induced pluripotent stem cell model recapitulates pathologic hallmarks of Gaucher disease. *Proc. Natl. Acad. Sci.* 109, 18054–9.

Park, I., Arora, N., Huo, H., Maherali, N., Shimamura, A., Lensch, M.W., Cowan, C., Daley, G.Q., **2008**. Disease-specific induced pluripotent stem (iPS) cells. *Cell* 134, 877–86.

Pastore, N., Blomenkamp, K., Annunziata, F., Piccolo, P., Mithbaokar, P., Maria Sepe, R., Vetrini, F., Palmer, D., Ng, P., Polishchuk, E., Iacobacci, S., Polishchuk, R., Teckman, J., Ballabio, A., Brunetti-Pierri, N., **2013**. Gene transfer of master autophagy regulator TFEB results in clearance of toxic protein and correction of hepatic disease in alpha-1-anti-trypsin deficiency. *EMBO Mol. Med.* 5, 397-412.

Patterson, M.C., Vecchio, D., Prady, H., Abel, L., Wraith, J.E., **2007**. Miglustat for treatment of Niemann-Pick C disease: a randomised controlled study. *Lancet Neurol.* 6, 765–72.

Peltz, S.W., Morsy, M., Welch, E.M., Jacobson, A., **2013**. Ataluren as an agent for therapeutic nonsense suppression. *Annu. Rev. Med.* 64, 407–25.

Peters, C., Steward, C.G., **2003**. Hematopoietic cell transplantation for inherited metabolic diseases: an overview of outcomes and practice guidelines. *Bone Marrow Transplant.* 31, 229–39.

Picanço-Castro, V., Moreira, L.F., Kashima, S., Covas, D.T., **2014**. Can pluripotent stem cells be used in cell-based therapy? *Cell. Reprogram.* 16, 98–107.

Pinotti, M., Rizzotto, L., Balestra, D., Lewandowska, M.A., Cavallari, N., Marchetti, G., Bernardi, F., Pagani, F., **2008**. U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. *Blood* 111, 2681–4.

Pinto, R., Caseiro, C., Lemos, M., Lopes, L., Fontes, A., Ribeiro, H., Pinto, E., Silva, E., Rocha, S., Marcão, A., Ribeiro, I., Lacerda, L., Ribeiro, G., Amaral, O., Sá Miranda, M.C., **2004**. Prevalence of lysosomal storage diseases in Portugal. *Eur. J. Hum. Genet.* 12, 87–92.

Piotrowska, E., Jakóbkiewicz-Banecka, J., Barańska, S., Tylki-Szymańska, A., Czartoryska, B., Wegrzyn, A., Wegrzyn, G., **2006**. Genistein-mediated inhibition of glycosaminoglycan synthesis as a basis for gene expression-targeted isoflavone therapy for mucopolysaccharidoses. *Eur. J. Hum. Genet.* 14, 846–52.

- Platt, F.M., Boland, B., van der Spoel, A.C., **2012**. The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. *J. Cell Biol.* 199, 723–34.
- Poorthuis, B.J., Wevers, R.A., Kleijer, W.J., Groener, J.E., de Jong, J.G., van Weely, S., Niezen-Koning, K.E., van Diggelen, O.P., **1999**. The frequency of lysosomal storage diseases in the Netherlands. *Hum. Genet.* 105, 151–6.
- Poupetová, H., Ledvinová, J., Berná, L., Dvoráková, L., Kozich, V., Elleder, M., **2010**. The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. *J. Inherit. Metab. Dis.* 33, 387–96.
- Prayle, A., Watson, A., Fortnum, H., Smyth, A., **2010**. Side effects of aminoglycosides on the kidney, ear and balance in cystic fibrosis. *Thorax* 65, 654–8.

## Q

- Quiviger, M., Arfi, A., Mansard, D., Delacotte, L., Pastor, M., Scherman, D., Marie, C., **2014**. High and prolonged sulfamidase secretion by the liver of MPS-IIIa mice following hydrodynamic tail vein delivery of antibiotic-free pFAR4 plasmid vector. *Gene Ther.*

## R

- Rana, T.M., **2007**. Illuminating the silence: understanding the structure and function of small RNAs. *Nat. Rev. Mol. Cell Biol.* 8, 23–36.
- Reczek, D., Schwake, M., Schröder, J., Hughes, H., Blanz, J., Jin, X., Brondyk, W., Van Patten, S., Edmunds, T., Saftig, P., **2007**. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell* 131, 770–83.
- Ring, K.L., Tong, L.M., Balestra, M.E., Javier, R., Andrews-Zwilling, Y., Li, G., Walker, D., Zhang, W.R., Kreitzer, A.C., Huang, Y., **2012**. Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. *Cell Stem Cell* 11, 100–9.
- Roberts, A.L.K., Fletcher, J.M., Moore, L., Byers, S., **2010**. Trans-generational exposure to low levels of rhodamine B does not adversely affect litter size or liver function in murine mucopolysaccharidosis type IIIa. *Mol. Genet. Metab.* 101, 208–13.
- Roberts, A.L.K., Rees, M.H., Klebe, S., Fletcher, J.M., Byers, S., **2007**. Improvement in behaviour after substrate deprivation therapy with rhodamine B in a mouse model of MPS IIIa. *Mol. Genet. Metab.* 92, 115–21.

- Roberts, A.L.K., Thomas, B.J., Wilkinson, A.S., Fletcher, J.M., Byers, S., **2006**. Inhibition of glycosaminoglycan synthesis using rhodamine B in a mouse model of mucopolysaccharidosis type IIIA. *Pediatr. Res.* 60, 309–14.
- Robertson, D.A., Callen, D.F., Baker, E.G., Morris, C.P., Hopwood, J.J., **1988**. Chromosomal localization of the gene for human glucosamine-6-sulphatase to 12q14. *Hum. Genet.* 79, 175–8.
- Robinson, A.J., Zhao, G., Rathjen, J., Rathjen, P.D., Hutchinson, R.G., Eyre, H.J., Hemsley, K.M., Hopwood, J.J., **2010**. Embryonic stem cell-derived glial precursors as a vehicle for sulfamidase production in the MPS-III A mouse brain. *Cell Transplant.* 19, 985–98.
- Robinton, D.A., Daley, G.Q., **2012**. The promise of induced pluripotent stem cells in research and therapy. *Nature* 481, 295–305.
- Roca, X., Akerman, M., Gaus, H., Berdeja, A., Bennett, C.F., Krainer, A.R., **2012**. Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes Dev.* 26, 1098–109.
- Roca, X., Krainer, A., **2009**. Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. *Nat. Struct. Mol. Biol.* 16, 176–82.
- Roca, X., Krainer, A.R., Eperon, I.C., **2013**. Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes Dev.* 27, 129–44.
- Roy, E., Bruyère, J., Flamant, P., Bigou, S., Ausseil, J., Vitry, S., Heard, J.M., **2012**. GM130 gain-of-function induces cell pathology in a model of lysosomal storage disease. *Hum. Mol. Genet.* 21, 1481–95.
- Rozaklis, T., Beard, H., Hassiotis, S., Garcia, A.R., Tonini, M., Luck, A., Pan, J., Lamsa, J.C., Hopwood, J.J., Hemsley, K.M., **2011**. Impact of high-dose, chemically modified sulfamidase on pathology in a murine model of MPS IIIA. *Exp. Neurol.* 230, 123–30.
- Ru, M.H. de, Boelens, J.J., Das, A.M., Jones, S.A., van der Lee, J.H., Mahlaoui, N., Mengel, E., Offringa, M., O'Meara, A., Parini, R., Rovelli, A., Sykora, K.-W., Valayannopoulos, V., Vellodi, A., Wynn, R.F., Wijburg, F.A., **2011**. Enzyme replacement therapy and/or hematopoietic stem cell transplantation at diagnosis in patients with mucopolysaccharidosis type I: results of a European consensus procedure. *Orphanet J. Rare Dis.* 6, 55.
- Ruijter, G.J.G., Valstar, M.J., van de Kamp, J.M., van der Helm, R.M., Durand, S., van Diggelen, O.P., Wevers, R.A., Poorthuis, B.J., Pshezhetsky, A. V, Wijburg, F.A., **2008**. Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in the Netherlands. *Mol. Genet. Metab.* 93, 104–11.
- Ruijter, J. de, Ru, M.H. de, Wagemans, T., Ijlst, L., Lund, A.M., Orchard, P.J., Schaefer, G.B., Wijburg, F.A., Vlies, N. van, **2012a**. Heparan sulfate and dermatan sulfate derived disaccharides are sensitive markers for newborn screening for mucopolysaccharidoses types I, II and III. *Mol. Genet. Metab.* 107, 705–10.

Ruijter, J. de, Valstar, M.J., Narajczyk, M., Wegrzyn, G., Kulik, W., Ijlst, L., Wagemans, T., van der Wal, W.M., Wijburg, F.A., **2012b**. Genistein in Sanfilippo disease: a randomized controlled crossover trial. *Ann. Neurol.* 71, 110–20.

Ruzo, A., Garcia, M., Ribera, A., Villacampa, P., Haurigot, V., Marcó, S., Ayuso, E., Anguela, X.M., Roca, C., Agudo, J., Ramos, D., Ruberte, J., Bosch, F., **2012a**. Liver production of sulfamidase reverses peripheral and ameliorates CNS pathology in mucopolysaccharidosis IIIA mice. *Mol. Ther.* 20, 254–66.

Ruzo, A., Marcó, S., García, M., Villacampa, P., Ribera, A., Ayuso, E., Maggioni, L., Mingozi, F., Haurigot, V., Bosch, F., **2012b**. Correction of pathological accumulation of glycosaminoglycans in central nervous system and peripheral tissues of MPSIIIA mice through systemic AAV9 gene transfer. *Hum. Gene Ther.* 23, 1237–46.

## S

Saftig, P., Klumperman, J., **2009**. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat. Rev. Mol. Cell Biol.* 10, 623–35.

Sánchez-Alcudia, R., Pérez, B., Pérez-Cerdá, C., Ugarte, M., Desviat, L.R., **2011**. Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia. *Mol. Genet. Metab.* 102, 134–8.

Sánchez-Danés, A., Richaud-Patin, Y., Carballo-Carbajal, I., Jiménez-Delgado, S., Caig, C., Mora, S., Di Guglielmo, C., Ezquerro, M., Patel, B., Giralt, A., Canals, J.M., Memo, M., Alberch, J., López-Barneo, J., Vila, M., Cuervo, A.M., Tolosa, E., Consiglio, A., Raya, A., **2012**. Disease-specific phenotypes in dopamine neurons from human iPS-based models of genetic and sporadic Parkinson's disease. *EMBO Mol. Med.* 4, 380–95.

Sanfilippo, S.J., Podosin, R., Langer, L.O.Jr, Good, R.A., **1963**. Mental retardation associated with acid mucopolysacchariduria (heparitin sulfate type). *J. Pediatr.* 63, 837–8.

Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., Banfi, S., Parenti, G., Cattaneo, E., Ballabio, A., **2009**. A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–7.

Sarrazin, S., Lamanna, W.C., Esko, J.D., **2011**. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* 3, a004952.

Sasaki, T., Takahashi, N., Matsuki, N., Ikegaya, Y., **2008**. Fast and accurate detection of action potentials from somatic calcium fluctuations. *J. Neurophysiol.* 100, 1668–76.

Savas, P.S., Hemsley, K.M., Hopwood, J.J., **2004**. Intracerebral injection of sulfamidase delays neuropathology in murine MPS-III A. *Mol. Genet. Metab.* 82, 273–85.

- Sawkar, A.R., Cheng, W.-C., Beutler, E., Wong, C.-H., Balch, W.E., Kelly, J.W., **2002**. Chemical chaperones increase the cellular activity of N370S beta-glucosidase: a therapeutic strategy for Gaucher disease. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15428–33.
- Schmid, F., Hiller, T., Korner, G., Glaus, E., Berger, W., Neidhardt, J., **2013**. A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs. *Hum. Gene Ther.* 24, 97–104.
- Schultz, M.M.L., Tecedor, L., Chang, M., Davidson, B.L.B., **2011**. Clarifying lysosomal storage diseases. *Trends Neurosci.* 34, 401–10.
- Scott, H., Blanch, L., Guo, X., Freeman, C., Orsborn, A., Baker, E., Sutherland, G., Morris, C., Hopwood, J., **1995**. Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. *Nat. Genet.* 11, 465–7.
- Scuderi, C., Stecca, C., Iacomino, A., Steardo, L., **2013**. Role of astrocytes in major neurological disorders: the evidence and implications. *IUBMB Life* 65, 957–61.
- Settembre, C., Fraldi, A., Jahreiss, L., Spampinato, C., Venturi, C., Medina, D., de Pablo, R., Tacchetti, C., Rubinsztein, D.C., Ballabio, A., **2008**. A block of autophagy in lysosomal storage disorders. *Hum. Mol. Genet.* 17, 119–29.
- Shayman, J., **2014**. Thematic review series: Recent advances in the treatment of lysosomal storage diseases. *J. Lipid Res.* 55, 993–4.
- Shyu, A.-B., Wilkinson, M.F., van Hoof, A., **2008**. Messenger RNA regulation: to translate or to degrade. *EMBO J.* 27, 471–81.
- Sivakumur, P., Wraith, J.E., **1999**. Bone marrow transplantation in mucopolysaccharidosis type IIIA: a comparison of an early treated patient with his untreated sibling. *J. Inher. Metab. Dis.* 22, 849–50.
- Smetters, D., Majewska, A., Yuste, R., **1999**. Detecting action potentials in neuronal populations with calcium imaging. *Methods* 18, 215–21.
- Sorrentino, N.C., D’Orsi, L., Sambri, I., Nusco, E., Monaco, C., Spampinato, C., Polishchuk, E., Saccone, P., De Leonibus, E., Ballabio, A., Fraldi, A., **2013**. A highly secreted sulphamidase engineered to cross the blood-brain barrier corrects brain lesions of mice with mucopolysaccharidosis type IIIA. *EMBO Mol. Med.* 5, 675–90.
- Spampinato, C., Feeney, E., Li, L., Cardone, M., Lim, J.-A., Annunziata, F., Zare, H., Polishchuk, R., Puertollano, R., Parenti, G., Ballabio, A., Raben, N., **2013**. Transcription factor EB (TFEB) is a new therapeutic target for Pompe disease. *EMBO Mol. Med.* 5, 691–706.
- Stone, J.E., **1998**. Urine analysis in the diagnosis of mucopolysaccharide disorders. *Ann. Clin. Biochem.* 35, 207–25.

- Susani, L., Pangrazio, A., Sobacchi, C., Taranta, A., Mortier, G., Savarirayan, R., Villa, A., Orchard, P., Vezzoni, P., Albertini, A., Frattini, A., Pagani, F., **2004**. TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA. *Hum. Mutat.* 24, 225–35.
- Sutherland, L., Hemsley, K., Hopwood, J., **2008**. Primary culture of neural cells isolated from the cerebellum of newborn and adult mucopolysaccharidosis type IIIA mice. *Cell. Mol. Neurobiol.* 28, 949–59.
- Suzuki, Y., **2014**. Emerging novel concept of chaperone therapies for protein misfolding diseases. *Proc. Japan Acad.* 90, 145–62.
- Suzuki, Y., Ogawa, S., Sakakibara, Y., **2009**. Chaperone therapy for neuronopathic lysosomal diseases: competitive inhibitors as chemical chaperones for enhancement of mutant enzyme activities. *Perspect. Medicin. Chem.* 3, 7–19.
- ## T
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., **2007**. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–72.
- Takahashi, K., Yamanaka, S., **2006**. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–76.
- Tanaka, A., Kimura, M., Lan, H., Takaura, N., Yamano, T., **2002**. Molecular analysis of the  $\alpha$ -N-acetylglucosaminidase gene in seven Japanese patients from six unrelated families with mucopolysaccharidosis IIIB (Sanfilippo type B), including two novel mutations. *J. Hum. Genet.* 47, 484–7.
- Tanner, G., Glaus, E., Barthelmes, D., Ader, M., Fleischhauer, J., Pagani, F., Berger, W., Neidhardt, J., **2009**. Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. *Hum. Mutat.* 30, 255–63.
- Tardieu, M., Zérah, M., Husson, B., de Bournonville, S., Deiva, K., Adamsbaum, C., Vincent, F., Hocquemiller, M., Broissand, C., Furlan, V., Ballabio, A., Fraldi, A., Crystal, R.G., Baugnon, T., Roujeau, T., Heard, J.-M., Danos, O., **2014**. Intracerebral administration of adeno-associated viral vector serotype rh.10 carrying human SGSH and SUMF1 cDNAs in children with mucopolysaccharidosis type IIIA disease: Results of a phase I/II trial. *Hum. Gene Ther.* 25, 506–16.
- Thompson, J.N., Jones, M.Z., Dawson, G., Huffman, P.S., **1992**. N-acetylglucosamine 6-sulphatase deficiency in a Nubian goat: a model of Sanfilippo syndrome type D (mucopolysaccharidosis IIID). *J. Inherit. Metab. Dis.* 15, 760–8.
- Tiscornia, G., Vivas, E.L., Izpisua Belmonte, J.C., **2011**. Diseases in a dish: modeling human genetic disorders using induced pluripotent cells. *Nat. Med.* 17, 1570–6.

- Tiscornia, G., Vivas, E.L., Matalonga, L., Berniakovich, I., Barragán Monasterio, M., Eguizábal, C., Gort, L., González, F., Ortiz Mellet, C., García Fernández, J.M., Ribes, A., Veiga, A., Izpisua Belmonte, J.C., **2013**. Neuronopathic Gaucher's disease: induced pluripotent stem cells for disease modelling and testing chaperone activity of small compounds. *Hum. Mol. Genet.* 22, 633–45.
- Tolar, J., Park, I., Xia, L., Lees, C.J., Peacock, B., Webber, B., McElmurry, R.T., Eide, C.R., Orchard, P.J., Kyba, M., Osborn, M.J., Lund, T.C., Wagner, J.E., Daley, G.Q., Blazar, B.R., **2011**. Hematopoietic differentiation of induced pluripotent stem cells from patients with mucopolysaccharidosis type I (Hurler syndrome). *Blood* 117, 839–47.
- Trilck, M., Hübner, R., Seibler, P., Klein, C., Rolfs, A., Frech, M.J., **2013**. Niemann-Pick type C1 patient-specific induced pluripotent stem cells display disease specific hallmarks. *Orphanet J. Rare Dis.* 8, 144.
- Tsunemi, T., Ashe, T.D., Morrison, B.E., Soriano, K.R., Au, J., Roque, R.A., Lazarowski, E.R., Damian, V.A., Masliah, E., La Spada, A.R., **2012**. PGC-1 $\alpha$  rescues Huntington's disease proteotoxicity by preventing oxidative stress and promoting TFEB function. *Sci. Transl. Med.* 4, 142ra97.
- Turunen, M., Olsson, J., Dallner, G., **2004**. Metabolism and function of coenzyme Q. *Biochim. Biophys. Acta* 1660, 171–99.

## U

- Ueno, M., Yamada, S., Zako, M., Bernfield, M., Sugahara, K., **2001**. Structural characterization of heparan sulfate and chondroitin sulfate of syndecan-1 purified from normal murine mammary gland epithelial cells. Common phosphorylation of xylose and differential sulfation of galactose in the protein linkage region tetrasaccharide sequence. *J. Biol. Chem.* 276, 29134–40.

## V

- Valstar, M.J., Marchal, J.P., Grootenhuis, M., Colland, V., Wijburg, F.A., **2011**. Cognitive development in patients with Mucopolysaccharidosis type III (Sanfilippo syndrome). *Orphanet J. Rare Dis.* 6, 43.
- Valstar, M.J., Neijs, S., Bruggenwirth, H.T., Olmer, R., Ruijter, G.J.G., Wevers, R.A., van Diggelen, O.P., Poorthuis, B.J., Halley, D.J., Wijburg, F.A., **2010**. Mucopolysaccharidosis type IIIA: clinical spectrum and genotype-phenotype correlations. *Ann. Neurol.* 68, 876–87.
- Valstar, M.J., Ruijter, G.J.G., van Diggelen, O.P., Poorthuis, B.J., Wijburg, F.A., **2008**. Sanfilippo syndrome: a mini-review. *J. Inherit. Metab. Dis.* 31, 240–52.

- van de Kamp, J.J., Niermeijer, M.F., von Figura, K., Giesberts, M.A., **1981**. Genetic heterogeneity and clinical variability in the Sanfilippo syndrome (types A, B, and C). *Clin. Genet.* 20, 152–60
- Velasco, I., Salazar, P., Giorgetti, A., Ramos-Mejía, V., Castaño, J., Romero-Moya, D., Menendez, P., **2014**. Generation of neurons from somatic cells of healthy individuals and neurological patients through induced pluripotency or direct conversion. *Stem Cells* 32, 2811–7.
- Vellodi, A., Young, E., New, M., Pot-Mees, C., Hugh-Jones, K., **1992**. Bone marrow transplantation for Sanfilippo disease type B. *J. Inherit. Metab. Dis.* 15, 911–8.
- Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., Wernig, M., **2010**. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463, 1035–41.
- Villani, G.R.D., Follenzi, A., Vanacore, B., Domenico, C.D.I., Naldini, L., Natale, P.D.I., **2002**. Correction of mucopolysaccharidosis type IIIb fibroblasts by lentiviral vector-mediated gene transfer. *Biochem. J.* 364, 747–53.
- Vitry, S., Bruyère, J., Hocquemiller, M., Bigou, S., Ausseil, J., Colle, M.-A., Prévost, M.-C., Heard, J.M., **2010**. Storage vesicles in neurons are related to Golgi complex alterations in mucopolysaccharidosis IIIB. *Am. J. Pathol.* 177, 2984–99.

## W

- Wahl, M.C., Will, C.L., Lührmann, R., **2009**. The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–18.
- Walkley, S.U., **2004**. Secondary accumulation of gangliosides in lysosomal storage disorders. *Semin. Cell Dev. Biol.* 15, 433–44.
- Wang, D., Belakhov, V., Kandasamy, J., Baasov, T., Li, S.-C., Li, Y.-T., Bedwell, D.M., Keeling, K.M., **2012**. The designer aminoglycoside NB84 significantly reduces glycosaminoglycan accumulation associated with MPS I-H in the Idua-W392X mouse. *Mol. Genet. Metab.* 105, 116–25.
- Wang, G.S., Cooper, T.A., **2007**. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* 8, 749–61.
- Weber, B., Kamp, J. Van De, Kleijer, W., Guo, X.-H., Blanch, L., Diggelen, O. Van, Wevers, R., Poorthuis, B., Hopwood, J., **1998**. Identification of a common mutation (R245H) in Sanfilippo A patients from the Netherlands. *J. Inherit. Metab. Dis.* 21, 416–22.
- Welling, L., Marchal, J.P., van Hasselt, P., van der Ploeg, A.T., Wijburg, F.A., Boelens, J.J., **2014**. Early umbilical cord blood-derived stem cell transplantation does not prevent neurological deterioration in mucopolysaccharidosis type III. *JIMD Rep.*

Wilkinson, F.L., Holley, R.J., Langford-Smith, K.J., Badrinath, S., Liao, A., Langford-Smith, A., Cooper, J.D., Jones, S.A., Wraith, J.E., Wynn, R.F., Merry, C.L.R., Bigger, B.W., **2012**. Neuropathology in mouse models of mucopolysaccharidosis type I, IIIA and IIIB. *PLoS One* 7, e35787.

Wolfe, B., Ghomashchi, F., Kim, T., Abam, C., Sadilek, M., Jack, R., Thompson, J., Scott, C., Gelb, M., Turecek, F., **2012**. New substrates and enzyme assays for the detection of mucopolysaccharidosis III (Sanfilippo Syndrome) types A, B, C, and D by tandem mass spectrometry. *Bioconjug. Chem.* 23, 557–64.

## Y

Yogalingam, G., Hopwood, J.J., **2001**. Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications. *Hum. Mutat.* 18, 264–81.

Yu, J.Y., DeRuiter, S.L., Turner, D.L., **2002**. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6047–52.

Yu, W.H., Zhao, K.W., Ryazantsev, S., Rozengurt, N., Neufeld, E.F., **2000**. Short-term enzyme replacement in the murine model of Sanfilippo syndrome type B. *Mol. Genet. Metab.* 71, 573–80.

## Z

Zak, B.M., Crawford, B.E., Esko, J.D., **2002**. Hereditary multiple exostoses and heparan sulfate polymerization. *Biochim. Biophys. Acta* 1573, 346–55.

Zhang, L., David, G., Esko, J.D., **1995**. Repetitive Ser-Gly sequences enhance heparan sulfate assembly in proteoglycans. *J. Biol. Chem.* 270, 27127–35.

Zhang, J., Sun, X., Qian, Y., LaDuca, J.P., Maquat, L.E., **1998**. At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. *Mol. Cell. Biol.* 18, 5272–83.

Zhao, H., Li, H., Bach, G., Schmidtchen, A., Neufeld, E., **1996**. The molecular basis of Sanfilippo syndrome type B. *Proc. Natl. Acad. Sci.* 93, 6101–5.

Zheng, Y., Ryazantsev, S., Ohmi, K., Zhao, H.Z., Rozengurt, N., Kohn, D.B., Neufeld, E.F., **2004**. Retrovirally transduced bone marrow has a therapeutic effect on brain in the mouse model of mucopolysaccharidosis IIIB. *Mol. Genet. Metab.* 82, 286–95.

Zhou, Y., Zhang, C., Liang, W., **2014**. Development of RNAi technology for targeted therapy - a track of siRNA based agents to RNAi therapeutics. *J. Control. Release* 193, 270–81.

## WEBSITES

<http://www.ensembl.org/index.html>

<http://www.glycoforum.gr.jp/science/word/proteoglycan/PGA11E.html>

<http://www.minoryx.com/see-tx/>

<http://watcut.uwaterloo.ca/webnotes/Pharmacology/rnaNovelTargets.html>



**RESUM**



La Síndrome de Sanfilippo C és una malaltia d'acumulació lisosòmica causada per mutacions en el gen *HGSNAT*, localitzat en el cromosoma 8, que presenta una herència autosòmica recessiva. Aquest gen codifica una proteïna de membrana del lisosoma, acetil-CoA  $\alpha$ -glucosaminida N-acetiltransferasa, que té la funció d'acetilar les glucosamines terminals de l'heparà sulfat durant la seva degradació com a pas previ indispensable a l'acció del següent enzim de la via. L'heparà sulfat és una molècula que es localitza en la matriu extracel·lular formant part dels proteoglicans i que participa en nombrosos i importants processos cel·lulars. La disfunció de la proteïna HGSNAT promou l'acumulació d'intermediaris de la degradació de l'heparà sulfat dins del lisosoma, causant una alteració de múltiples processos cel·lulars que afecta sobretot a les neurones i que provoca la progressiva i severa neurodegeneració que apareix com principal símptoma en els pacients.

Aquesta tesi representa un important treball per l'estudi de les bases moleculars de la Síndrome de Sanfilippo C. En primer lloc s'ha realitzat un anàlisi mutacional que ha permès identificar les mutacions causants de la malaltia en 15 pacients de diferents orígens. S'han trobat un total de 13 mutacions diferents, set de les quals no havien estat descrites prèviament. S'ha observat la patogenicitat de les mutacions de canvi d'aminoàcid identificades, comprovant que l'activitat enzimàtica s'ha reduït dràsticament. També s'ha analitzat com les mutacions que afecten a regions d'*splicing* conservades promouen defectes en aquest important procés regulador de l'expressió gènica. S'han identificat dues mutacions prevalents en la població espanyola que expliquen gairebé un 70% dels al·lels mutats i, mitjançant anàlisis d'haplotips, s'ha pogut suggerir un origen comú per cada una d'elles en la població estudiada.

Per altra banda, s'han assajat diverses estratègies terapèutiques com un primer pas en la cerca d'una teràpia efectiva, encara no existent per aquesta malaltia. S'ha vist que la utilització de U1 snRNAs modificats que presenten una major homologia pels llocs donadors d'*splicing* podria representar una bona opció terapèutica. També s'ha pogut veure que la glucosamina podria ser utilitzada com a xaperona per tal de promoure el correcte plegament d'aquelles proteïnes que presenten mutacions de canvi d'aminoàcid o la pèrdua d'alguns d'ells. Finalment s'ha vist que l'ús de siRNAs per inhibir la via de síntesi de l'heparà sulfat, concretament els gens de la família *EXTL*, en especial *EXTL2*, podria representar una bona opció com a teràpia de reducció de substrat.

Per últim, en el transcurs d'aquesta tesi s'ha obtingut un model neuronal de la malaltia, fet que representa un gran avenç en el seu estudi, ja que a fins ara no es disposa de cap altre model cel·lular o animal per la Síndrome de Sanfilippo C. Per assolir aquest objectiu, s'han obtingut cèl·lules pluripotents induïdes a partir de fibroblasts de dos pacients, les quals posteriorment s'han diferenciat a neurones. S'ha comprovat que aquestes neurones presenten els símptomes característics de la malaltia com són la falta d'activitat enzimàtica, una acumulació d'heparà sulfat, l'augment en el nombre i mida dels lisosomes, una alteració del procés d'autofàgia i un augment en el nombre de cèl·lules apoptòtiques. Mitjançant experiments específics que han permès estudiar l'activitat neuronal dels cultius obtinguts, s'ha vist un decaïment progressiu de l'activitat en les neurones dels pacients i problemes en el manteniment i desenvolupament de les xarxes neuronals. Aquest model permetrà estudiar més detalladament les bases moleculars, cel·lulars i cerebrals de la malaltia per tal de poder aprofundir en el coneixement i les possibilitats terapèutiques de la Síndrome de Sanfilippo C.

# **AGRAÏMENTS**



No voldria començar sense agrair a la Lluïsa i el Dani la oportunitat que em van donar en el seu moment, ja llunyà, quan vaig aparèixer pel departament amb la intenció d'estar-hi un curs com a alumne intern mentre acabava la carrera. M'han demostrat la seva paciència, que la bondat no està renyida amb l'exigència, que sempre s'ha d'esperar el millor i treballar per evitar el pitjor, i que els resultats sempre arriben quan un s'hi deixa la pell. Molts anys i vivències després aquest interminable viatge s'acosta al final i sempre guardaré magnífics records dels "jefes", tant en la part científica com en la part personal.

Evidentment, en aquest recorregut, molts han estat els companys d'aventures que d'una manera o altra, han fet que el "llarg patiment" de la tesi valgui la pena en molts sentits i als quals estaré eternament agraït. A la Laura, perquè mai hauria imaginat una germana de tesi més meravellosa, ha estat un privilegi compartir aquests anys amb tu, aprenent en el laboratori i en el treball "animístic". Ets una de les millors persones que conec i t'admiro. Eres una "admirada". A la Noe, perquè tot i que sempre la burxo buscant la confrontació, molts cops he sentit a dir que jo només els faig això a les persones que realment aprecio, així que espero que mai m'ho hagi tingut en compte. Sempre has estat amb mi quan ho he necessitat, i això és el que un espera d'una bona amista. A la Bàrbara, perquè és qui sempre sap com fer-me riure, tot i que moltes vegades ho aguanti per no deixar que s'ho cregui en excés. M'has donat l'alegria necessària en tot moment. A la Jenny, perquè veien-te nerviosa a tu, moltes cops m'he relaxat i he fet broma (tot i que d'altres m'he estressat encara més), no siguis tan patidora perquè tu vals molt dins i fóra del lab. Estic segur que tindràs molts èxits en la vida. A la Marta, perquè compensa amb la seva tranquil·litat, i permet que les coses flueixin més fàcilment, tot i que últimament m'ha semblat veure que potser ho portes per dins. Espero que tinguis molta sort en la part final de la tesi. A l'Anna i el Raül, per ser els millors primers mestres, sempre heu estat un referent per mi. Al Freddy, perquè ha estat el company perfecte durant aquests anys, en els bons i els mals moments, tots ells inoblidables. Sé que per molta distància que ens separi i molt temps que passem sense veure'ns, en ell tinc un amic per tota la vida. El teu suport en tot ha estat i serà fonamental. Muito obrigado.

Moltes gràcies a tots els altres que han format part d'aquests magnífics anys al departament, la Mònica per retallar-me el nom amb tanta gràcia com quan parla en anglès, el Ricart per picar-me constantment, la Neus perquè és magnífica i sempre està disposada a

ajudar, l'Edu perquè sempre tindrà un acudit dolent per explicar, l'Ona perquè és un encant i "hablará catalán", la Patri per la seva alegria, la Lídia per la seva gracia, el Bru per tots els anàlisis barcelonistes post-partit, la Susana per viure amb passió fins i tot una "immuno", les "mosques", la gent del lab gran de la primera planta i als de la segona planta per totes les festes i aventures viscudes, al Jon, l'Amanda, l'Esther, la Pilar, la Sheila i la Marina, perquè son gent estupenda i m'encanta veure'ls amb qualsevol excusa, i en general a totes aquelles persones que he conegut durant la meva estància al departament i m'han demostrat el seu valor. Si els nombrés a tots ompliríem una altra tesi.

Tampoc vull oblidar als meus companys i amics del "meu altre laboratori". Ha estat un plaer poder treballar i aprendre de tots vosaltres. Gràcies a l'Àngel i l'Antonella per la seva inestimable ajuda i els seus magnífics consells. Ha estat un plaer col·laborar i aprendre de vosaltres. A l'Yvonne per tota la seva ajuda, la seva paciència amb mi i les meves bromes, i sobretot per tindre la capacitat de saber treure un somriure de mi sota qualsevol circumstància. A la Senda per no tindre mai un no per resposta i estar sempre disposada a ajudar i a mostrar el seu preciós i encantador somriure. Al "Sergi", perquè moltes han estat les converses i els riures que hem compartit durant el seu procés de catalanització. A la Claudi i l'Isil pels seus constants ànims i alegria, la vida és més bonica amb gent com vosaltres al costat. Al Roger pels seu humor especial i acudits, són únics. A les Adrianes perquè no se si el fet de compartir nom les ha fet igual de pacients, rialleres i escandalitzades per qualsevol broma, però amb les dues ha estat un plaer coincidir. Al "Charlie" perquè és un nano collonut i sempre ens arrossega a conèixer llocs diferents en la nostra pròpia ciutat, però no per això menys interessants. A la Carla pel patiment compartit diferenciant a cultius neuronals amb esperança de que mostrin activitat. Al JuanLu per ensenyar-me a parlar mexicà amb tanta paciència. Al Juan perquè no hi ha ningú que amenitzi millor una tarda al sol amb la seva guitarra. A la Cristina per no enfadar-se mai amb mi i sempre respondre amb un somriure li digui la tonteria que li digui. Al Kike per ser tan divertit i tan bona persona. I a tots aquells i aquelles que han passat per aquell laboratori, ha estat genial compartir aquest temps amb vosaltres.

A la colla d'amics del poble, per estar sempre al costat. A l'Albert, perquè hem crescut junts i sempre ens hem tingut al costat, en els bons i els mals moments, des de que

érem uns marrecs que jugaven a bàsquet. No podré oblidar els càstigs escoltant la Xuxa al cotxe quan perdiem els partits fora de casa, ni els estius treballant per fàbriques o empreses de mala mort i anant a dormir escoltant els acudits del Eugenio, ni els dijous de festes universitàries, ni tots els viatges dels estius bojos, ni un incomptable nombre de vivències que han fet de nosaltres les persones que som avui. Al Marc, pels infinits moments compartits, fent el piti a la “curva”, gaudint del seu “refranero” popular, repetint fins al cansament els diàlegs de les nostres “pelis” preferides, decidint viatges en impulsos incontrolables, per les hores que hem passat jugant al Pro i guanyant incomptables títols, per tot i més, moltes gràcies. Al Ferran, tot i que desaparegui constantment, i només ens gratifiqui amb la seva presència amb comptagotes, sempre ha sabut fer que ens pixéssim de riure, ja sigui d’ell o amb ell, qui ens diria que seria el primer de nosaltres en assentar cap... A Giselle, perquè la hemos querido como a una más del grupo desde el primer día que llegó, una noche veraniega de hace ya más de 10 años. Sus simpáticos “punyales” encajaron bien rápido en el grupo y a pesar de la distancia que hay ahora, nuestro amor por ella no menguara jamás. A la Miriam, per ser una mica la mare del grup, la metgessa pot ser la més jove, però també la que sembla tenir més seny, i això sempre ens ha anat bé, sobretot als més esbojarrats. Ets un solet. A Núria, por ser “uno” más, eres de lo más basto que he conocido, pero eso te hace única y especial, la risa vamos.

Als amics de la carrera, perquè vam compartir els millors anys. La Belén, perquè ha estat com tenir una germana, tot i que a la gent li costi d’entendre. Sempre et portaré al curassó. A la Raquel i l’Anna, perquè tot i que ara ens veiem molt menys, segurament per culpa meva, sou estupendes. Al Pol, perquè és un crack amb gorra, l’Helena per haver despertat a temps de la seva “empanadeta”, al Llu perquè és genial, la Maria perquè sempre ens hem entès i ajudat, la Bàrbara perquè sap com animar-me en moments baixos, la Su perquè hem acabat compartint i patint tesi al departament, i a tota la resta, que amb els anys hem perdut contacte però mai els oblidaré.

A tots aquells que he conegut vivint a Barcelona. Ja fa més de 6 anys que vaig vindre cap aquí, i sempre he pensat que va ser una gran decisió, per l’experiència i sobretot per la gent coneguda durant aquest temps. A les meves primeres companyes de pis, per mimar-me com si encara estigués vivint amb la mare. A la Marta per sempre tindre temps per explicar-

nos les penes i les alegries, al Xavi pels “singstars” i les hores distrets en consoles i sopars, a l’Andreas, per totes les nits de festa. A Àlex, perquè jamás habría pensado que haría gran amistad con mi casero, pero a personas como él es imposible no cogerles un cariño especial. A l’Eric i al Dani, perquè sou genials, totes les nits de festa han estat un no parar de riure i passar-ho genial. A més sempre que ha fet falta posar-se una mica seriós, he sabut que podia comptar amb vosaltres. Sou uns “cracks”. I als meus últims companys de pis, Néstor, Manu, Pablo i Robert, perquè tot i que només han estat uns mesos, hem creat un ambient únic i ens ho hem passat genial.

També voldria donar les gràcies a l’associació MPS España i en especial al Jordi i la Mercè pel seu inestimable suport, la seva força i les seves enormes ganes d’ajudar i mirar sempre endavant. Han estat un gran exemple per mi.

A la família, catalana i gallega. Als meus tiets i tietes, cosins, cosines, padrina i tots, perquè del primer fins a l’últim, tots són genials, perquè m’heu cuidat sempre com al petit de la casa i això no té preu. Per uns cosins que han estat grans amics. Us estimo a tots. Als que ja no hi són, la iaia, que mai l’oblidaré, sabatilla en mà per fer que ens portéssim bé, i l’avi, per les seves idees fixes i passades, però que em va ensenyar valors i lliçons que sempre portaré amb mi. A toda la familia gallega, porque aunque haya llegado tarde, siempre me he sentido querido como uno más por todos. Pai, tías, tíos, primas, primos, tita y mi genial avoa, no os imagináis cuanta alegría ha significado el poder entrar en vuestras vidas. Nunca es tarde si la dicha es buena. Os quiero.

I sobretot a la meva mare, a la qual mai podré tornar tot el que ha fet i ha donat per mi, per tots els seus sacrificis, pel seu amor, per la seva comprensió, pels seus ànims. Perquè tot i que moltes vegades no ens entenem i acabem discutint per tonteries, passi el que passi, fins i tot quan creguis que no ho demostro, sempre t’estimaré.



