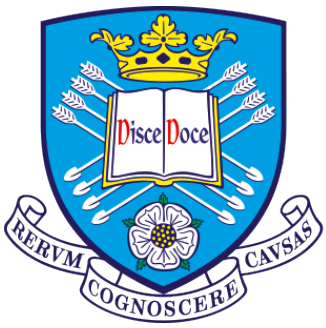


The University Of Sheffield.

Investigating the Role of the RNA Exporter Protein Gle1 in Spinal Muscular Atrophy

Emily Graves

Thesis Submitted for the degree of Doctor of Philosophy (PhD)

Sheffield Institute for Translational Neuroscience

January 2022

UKRI Medical Research Council

SITraN Sheffield Institute for Translational Neuroscience

Abstract

Background and rationale: The RNA exporter protein Gle1 is essential to cell survival. In humans it is present in two isoforms; hGle1a and hGle1b. The predominant isoform, hGle1b, is implicated in the nuclear export of mRNA molecules to the cytoplasm. hGle1a has been implicated in stress granule formation. Mutations within *GLE1* have been identified in cases of the foetal motor neuron disease lethal congenital contracture syndrome type 1 (LCCS1) (Nousiainen et al., 2008) as well as amyotrophic lateral sclerosis (ALS) (Kaneb et al., 2015). Spinal muscular atrophy (SMA) is a childhood form of motor neuron disease caused by a lack of survival motor neuron (SMN) protein. Clinical overlap between the phenotypes of LCCS1 and SMA has led us to investigate whether there are any common pathological mechanisms. Preliminary data from our lab demonstrates that Gle1 has potential to rescue the axonal growth defect in SMN-deficient motor neurons *in vitro.*

Aims: The aims of this project were therefore to: **i)** investigate whether Gle1 and SMN interact on a protein-protein level and where this interaction may occur in the cell. **ii)** evaluate the levels of Gle1 within SMA cell models. **iii)** assess whether overexpression of Gle1b can ameliorate disease phenotypes seen in SMA cell models.

Findings: Protein-protein interaction of Gle1 and SMN proteins has been shown by co-immunoprecipitations and proximity ligation assay. We report a reduction in the protein levels of Gle1 within SMA patient cell models. Finally, we looked at whether overexpression of Gle1b protein could ameliorate some known SMN disease phenotypes. Overexpression did not reduce DNA damage in SMA cells. Interestingly, Golgi morphological defects seen in SMA patient cells, was rescued by overexpression of Gle1b.

Conclusions: The clinical overlap seen between diseases caused by mutations in *GLE1* and *SMN1* combined with depleted levels of Gle1 in SMA patients suggests that there is a common pathway which is disrupted in these diseases. The interaction reported here between the two proteins could support this hypothesis. More research is needed to fully identify the pathway in which these two proteins interact and how disruption of this may lead to neurodegeneration.

Acknowledgements

I would first like to give enormous thanks to my supervisor, Professor Mimoun Azzouz, for giving me this opportunity and supporting me for the past four years. Your enthusiasm, patience and encouragement has guided me all through my PhD. I would also like to thank my secondary supervisor Dr Guillaume Hautbergue for his invaluable technical support, knowledge of RNA biology and experimental guidance from his team. Additional thanks go to Dr Laura Ferriauolo, for her valuable advice and help with the neuronal side of this project.

My PhD experience has been greatly enhanced by being part of the MRC DiMeN DTP, I am forever grateful for the opportunities, memories made, and the chance to meet and be part of a group of fantastic people.

I am indebted to all the members of SITraN, past and present, who have each given up their own time to help me during my PhD. Particular thanks go to Dr Eva Karyka, Dr Lara Marrone, Dr Christopher Webster, Miss Louise Whiteley, Mr Paolo Marchi, Dr Lydia Castelli, Dr Erica Lin and Dr Cleide dos Santos Souza for all of their technical support in the lab! Another huge thank you to the remainder of the Azzouz lab especially Dr Joe Scarrott, Dr Michela Pulix, Dr João Cruzerio, and Miss Hannah Thomas for always being there to brighten the long days in the lab.

I would not have got through this PhD without the support of my friends. Their constant words of encouragement, mutual despair and many, many laughs have always been just what was needed in that moment. The biggest thank you goes to them; Charlotte, Ale, Matt, Toby, Anushka, Mirinda, Lily, and Emily. I have found friends for life in all of you.

I would also like to thank all my friends outside of SITraN, for reminding me that there is a world outside of science and to enjoy it! Special thanks to Jade and Rosie for being some of the best housemates and always lending a sympathetic ear about failed experiments.

Finally, my family have always been my biggest supporters and I would not be where I am today without their love and support. Thank you for always being there and trying to understand I do on a daily basis. This PhD would never have been possible without you.

Table of Contents

List of Figures

List of Tables

Abbreviations

Axonal SMN	**a-SMN**
Adeno-associated virus	**AAV**
Adenosine Diphosphate	**ADP**
Amyotrophic lateral sclerosis	**ALS**
Antisense oligonucleotide	**ASO**
Adenosine Triphosphate	**ATP**
Ataxia-telangiectasia-mutated-and-Rad3-related kinase	**ATR**
Blood brain barrier	**BBB**
Bicinchronic acid	**BCA**
Bicaudal D2	**BICD2**
Bovine serum albumin	**BSA**
Chromosome 9 open reading frame 72	**C9orf72**
Calcium	**Ca^{2+}**
Checkpoint Kinase 1	**Chk1**
Central nervous system	**CNS**
Co-Immunoprecipitation	**co-IP**
Coat Protein I	**COPI**
Campthothecin	**CPT**
Cleavage stimulation factor (64kda subunit)	**CstF-64**
Cell Total Corrected Fluorescence	**CTCF**
DEAD box protein	**Dbp/DDX**
Diethyl pyrocarbonate	**DEPC**
Dulbecco's Minimum Essential media	**DMEM**
Dioxyribonucleic Acid	**DNA**
Days post fertilisation	**dpf**
Double Stranded DNA	**dsDNA**
Ethylenediaminetetraacetic acid	**EDTA**
Eurkaryote Initiation Factor 3	**eIF3**
Endoplasmic reticulum	**ER**
Erb-b2 receptor tyrosine kinase 3	**ERBB3**

Eurkaryote Release Factor 1	**eRF1**
Exocyst Complex component 6B	**EXOC6B**
Fluorescence Activated cell sorting	**FACS**
Foetal bovine serum	**FBS**
Phenylalanine-glycine	**FG**
Fluorescent in situ hybridisation	**FISH**
Fragile X mental retardation protein	**FMRP**
Fast skeletal muscle troponin activators	**FSTA**
Fused in sarcoma	**FUS**
Rad-GTPase-activating protein (SH3-domain) binding protein	**G3BP**
Glyceraldehyde 3-phosphate dehydrogenase	**GAPDH**
H/ACA ribonucleoprotein complex subunit 1	**GAR1**
Green fluorescent protein	**GFP**
Gle1 shuttling domain	**Gle1-SD**
Glycogen synthase kinase 3	**Gsk3**
Hank's Buffered Saline Solution	**HBSS**
Human embryonic kidney	**HEK293T**
Heterogenous nuclear ribonucleoprotein R	**hnRNP R**
Horseradish peroxidase	**HRP**
Immunocytochemistry	**ICC**
IGF-II mRNA binding protein	**IMP1**
Inositol polyphosphates	**InsPs**
Inositol hexakisphosphate	**IP_6**
Induced pluripotent stem cells	**iPSC**
Inverted terminal repeats	**ITRs**
Lethal arthrogryposis with anterior horn cell disease	**LAAHD**
Lethal congenital contracture syndrome	**LCCS1**
Long intergenic non-coding RNA	**lincRNA**
Lentivirus	**LV**
Mitogen activated protein kinase	**MAPK**
Makorin-2	**MKRN2**

Multiplicity of infection	**MOI**
Messenger RNA	**mRNA**
Messenger ribonucleoprotein	**mRNP**
Neurocalcin delta	**NCALD**
Non homologous end joining	**NHEJ**
Neural Progenitor Cell	**NP Cell**
Nuclear pore complex	**NPC**
Non template control	**NTC**
Nuclear pore proteins	**Nucleoporins**
Penicillin streptomycin	**P/S**
Poly(A) RNA granule	**PARG**
Phosphate Buffered Saline Solution	**PBS**
Polymerase Chain Reaction	**PCR**
Parkinson's Disease	**PD**
Polyethylenimine	**PEI**
Protein Inhibitor Complex	**PIC**
Proximity ligation assay	**PLA**
Plastin3	**Pls3**
Polyadenylated	**Poly(A)+**
Protein arginine methyltransferase 5	**PRMT5**
Proline rich	**Pro**
Phosphatidylinositol-4-phosphate	**PtdIns(4)P**
Polyvinylidene difluoride	**PVDF**
Qualitative PCR	**qPCR**
Recombinant adeno-associated virus	**rAAV**
RNA binding proteins	**RBP**
Rolling-circle amplification	**RCA**
RNA helicase A	**RHA/DHX9**
Radioimmunoprecipitation	**RIPA**
Ribonucleic Acids	**RNA**
RNA polymerase I	**RNA Pol I**

RNA polymerase II	**RNA Pol II**
Ribonucleoproteins	**RNP**
Ribosomal RNA	**rRNA**
Reverse Transcriptase PCR	**RT-PCR**
Saccharomyces cerevisiae	***S.cerevisiea***
Self-complimentary AAV9	**scAAV9**
Sodium dodecyl sulphate–polyacrylamide gel electrophoresis	**SDS-PAGE**
Senataxin	**SETX**
Smith-antigen	**sm**
Spinal muscular atrophy	**SMA**
Survival motor neuron 1	**SMN1**
SMN delta 7	**SMNΔ7**
small nucleolar RNPs	**snoRNP**
Small nuclear RNA	**snRNA**
Spliceosomal small nuclear riboproteins	**snRNP**
Superoxide dismutase (SOD1)	**SOD1**
Single stranded DNA	**ssDNA**
Tris/Acetate/EDTA	**TAE**
Tris Buffered Saline solution + Tween	**TBST**
TAR DNA-binding protein 43 (TDP-43)	**TDP-43**
Cytotoxic granule-associated RNA binding proteins	**TIA1/TIAR**
Ubiquitin-like modifier-activating enzyme 1	**UBA1**
Untransduced/Untreated	**UT**
Woodchuck Hepatitis Virus Posttranscriptional Response Element	**WPRE**
WD40-encoding RNA antisense to p53	**WRAP53β**
Wild type	**WT**
5'-3' Exoribonuclease 2	**XRN2**
Tyrosine/glycine rich domain	**YG box**
Zinc finger protein	**ZPR1**

1. Introduction

1.1 Spinal Muscular Atrophy

Spinal muscular atrophy (SMA) is a form of motor neuron disease which occurs predominantly in early childhood. Occurring in an autosomal, recessive manner, it is typically characterised by a loss of lower alpha motor neurons within the anterior horn of the spinal cord. This degeneration gradually progresses causing muscle wasting and progressive paralysis. In the most severe of cases this can cause respiratory failure and death before the age of two. A spectrum of SMA phenotypes has been identified, by which cases can be classified by severity (Zerres and Davies, 1999). The estimated prevalence of SMA in the general population is 1 in 6,000 to 1 in 10,000 live births, with the carrier frequency being between 1 in 40 to 1 in 60 (D'Amico et al., 2011). This makes SMA the most common genetic form of infant mortality.

1.1.1 Genetics of *SMN1*-related SMA

Genetic linkage studies narrowed down the region responsible for SMA to the long arm of chromosome 5, location 5q11.2 -13.3 (Brzustowicz et al., 1990; Melki et al., 1990). Research conducted in 1995 demonstrated that 95% of SMA cases studied were caused by a homologous deletion in the Survival Motor Neuron 1 (*SMN1)* gene, with the remaining 5% of cases being caused by point mutations within the same gene (Lefebvre et al., 1995). The developmental importance of this gene was confirmed through generation of a *SMN1* knockout mouse model in which the embryos did not survive until birth (Schrank et al., 1997). It is now known that 95% of SMA cases are caused through mutation or homozygous deletions of the *SMN1* gene, with the remaining cases being caused by mutations within other genes.

Uniquely, *SMN* is present twice in the human genome: a telomeric *SMN1* gene and a centrosomal *SMN2* gene, with *SMN1* occurring as the dominantly transcribed full length protein (Lefebvre et al., 1995; Rochette et al., 2001) (**Fig 1.1**). These genes are composed of nine exons which differ only by 5 nucleotides – *SMN2* has a C to T mutation in exon 7 affecting splicing of the protein. This results in exon 7 being skipped in the majority of transcripts, therefore producing truncated SMN protein (SMNΔ7) which is rapidly degraded (Lorson et al., 1999; Monani et al., 1999). Mutations within the *SMN1* gene result in a reduction in the levels

of full-length SMN protein, which cannot be fully compensated for by the small amounts of full length SMN produced by the *SMN2* gene (Feldkötter et al., 2002). SMN protein levels therefore correlate with disease severity, as demonstrated by a wide range of studies carried out *in vitro*, *in vivo* and within patient postmortem tissues (Crawford et al., 2012; Lefebvre et al., 1997; Simard et al., 2007; Sumner et al., 2006; Tiziano et al., 2009).

Chromosome 5 is one of the largest chromosomes within the human genome, duplication events occur predominantly within the 5q13.3 region where the *SMN* locus can be found (Melki et al., 1994; Schmutz et al., 2004). Due to this instability, humans can possess a variable copy number of the *SMN2* gene. This is a key determinant of disease severity, as a small amount of the mRNA transcribed from the *SMN2* gene (~10%) does retain exon 7 and can therefore produce full length SMN protein. SMNΔ7 protein has also been shown to be beneficial to the survival of a severe mouse model and so could still retain partial beneficial function (Le et al., 2005). Thus the higher the *SMN2* copy number, the less severe the form of SMA (Kolb and Kissel, 2015; Mailman et al., 2002).

However, there has been evidence to suggest that *SMN2* copy number is not the only factor at play in determining disease severity. Evidence suggests that SMN protein levels do not necessarily correlate with *SMN2* copy number (Crawford et al., 2012) and that there is a large overlap in the number of *SMN2* copies and disease severity (Wadman et al., 2020, 2018, 2017). This indicates the presence of other disease modifying genes, which will be discussed later in this chapter.

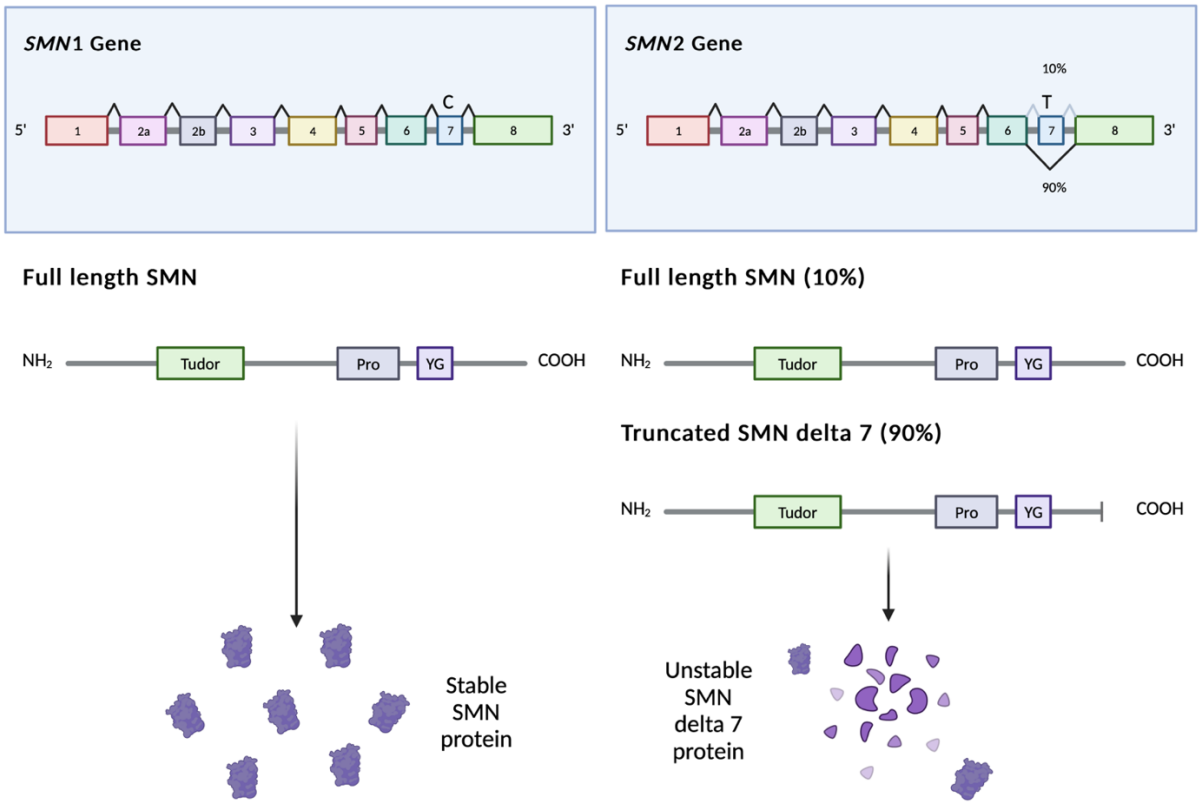


Figure 1.1: Genetic and protein structure of SMN1 and SMN2 genes. *SMN1 is a gene composed of 9 exons. A centrosomal duplicate gene, SMN2, differs only by a silent C to T nucleotide substitution in exon 7. This causes a splice variation removing exon 7, forming unstable SMN Δ7 protein. A small percentage of SMN2 is transcribed as full length SMN. The protein structure contains a Tudor domain, responsible for many of the proteins interactions, a Proline-Rich domain (Pro) and a tyrosine-glycine rich (YG) box. Figure created with biorender.com.*

1.1.2 Clinical features of SMA

SMA cases resulting from mutations within *SMN1* can be clinically classified as four distinct sub-groups of disease correlating inversely with *SMN2* gene copy number.

SMA type 0 is the most severe form of the disease, with symptoms being evident before birth. Neonates commonly present with severe weakness, hypotonia and a significant decrease in foetal movements. This reduced movement can result in infants being born with joint deformities and extremely weak muscle tone. Respiratory failure due to severe muscle weakness is the main cause of early death for these infants, with an average life expectancy of around 6 months old (Dubowitz, 1999; Macleod et al., 1999).

The most common form of SMA is SMA type I (frequently referred to as Werdnig-Hoffman disease). Infants with this condition show hypotonia prior to six months of age and are unable to sit unassisted or control their head movements. Slow growth often arises due to swallowing and feeding difficulties. Weakness occurring within intercostal muscles results in a 'bell-shaped' chest causing breathing pattern abnormalities, thus resulting in respiratory failure before the infant reaches two years of age (Finkel et al., 2014).

Development of muscle weakness between six to twelve months is the predominant characteristic of SMA Type II (also called Dubowitz disease). Infants presenting with this condition can sit without support during their development, yet never acquire the ability to stand or walk unaided. Muscle weakness, mainly within the legs, is progressive. Scolioisis (curvature of the spine), joint contractures and progressive respiratory muscle weakness can cause restrictive lung disease later in life. However, patients with this form of SMA frequently survive into their twenties or thirties with normal cognition (Kolb and Kissel, 2015).

SMA type III (also called Kugelberg-Welander disease) typically occurs after childhood, therefore patients develop the ability to walk and stand unaided. Progressive proximal muscle atrophy in the legs over time reduces the patient's ability to walk, leaving patients wheelchair-bound later in life. Patients with Type III SMA have a normal life expectancy and cognition (Zerres and Davies, 1999).

SMA type IV is the mildest presentation of SMA. Onset of disease occurs within adulthood presenting as mild to moderate muscle weakness, tremors, and mild respiratory issues. These patients have a normal life expectancy (Kolb and Kissel, 2015).

There are also cases of clinical SMA which are not linked to mutations within the *SMN1* gene, these cases make up less than 5% of infantile SMA and can result from mutations in a known number of genes including *Ubiquitin-like modifier-activating enzyme 1* (*UBA1)* and *Bicaudal D2* (*BICD2)*, among others. Symptoms include congenital hypotonia, arthrogryposis, progressive postnatal weakness and degeneration of the cells within the anterior horn of the spinal cord (Farrar and Kiernan, 2015).

1.1.3 Experimental Models of SMA

To evaluate the pathways involved in the pathogenesis of SMA, it is important to use a wide range of models which can act to recapitulate different aspects of disease. There are many advantages and disadvantages of each of these models and, as such, data generated from each should be assessed with this in mind.

In vitro cellular models

The main benefit of cellular models is that they are a human system, and so species-specific cellular functions do not have to be accounted for. However, there are disadvantages to looking at single populations of mono-layer cell growth as certain whole system aspects may be lost. Nonetheless, cellular *in vitro* models remain valuable in SMA research, especially with the advent of induced pluripotent stem cell (iPSC) technology allowing the growth of motor neuron specific cultures. These cultures can be seen to model the pathology which is characteristic of SMA; including motor neuron growth defects (Ebert and Svendsen, 2010; Zhang et al., 2017).

In addition to iPSC-derived cell studies, it is also of great benefit to study the patient fibroblasts from which these cells can be derived. Fibroblast cells can be easily obtained from patient biopsies and are easy to grow in culture; not requiring the expensive regents required for iPSC culture. However, as SMA is a disease which exclusively affects motor neurons, fibroblasts - usually derived from skin samples - can be a limited cell model. Despite this

caveat, fibroblasts are still highly useful for the study of multiple disease readouts in SMA (for example, DNA damage and golgi dysfunction (Custer et al., 2019; Karyka et al., 2022)).

Mouse models

Multiple mouse models of SMA have been developed over the years with a range of different disease severities. A well-established, and widely used model system for neurodegenerative diseases, mice possess closely related nervous systems to humans in an *in vivo* whole animal environment. A major caveat of these models to consider for SMA research is while humans possess two copies of the *SMN* gene, mice only have one. This *Smn* gene acts similarly to human *SMN1*, only producing full length transcripts.

Knockout of the mouse *Smn* is embryonic lethal (Schrank et al., 1997). Mice with a heterozygous deletion within the *Smn* gene display motor neuron loss from six months of age and are therefore more representative of SMN type III disease in humans (Jablonka et al., 2000). This correlates with a study which assesses the critical threshold of Smn protein expression needed to see SMA disease phenotypes – found to be around an 85% loss of protein levels (Bowerman et al., 2012). To recapitulate the severe, infantile phenotypes of SMA, more complex transgenic mice models have therefore been generated. Here I will discuss the main two models used in the field.

The addition of human *SMN2* gene into an *Smn* knockout background (*Smn*$^{-/-}$;*SMN2*) rescues the lethality caused by knockout of *Smn*. This transgenic cross produces mice which are comparable to wild type at birth (albeit smaller in size), but quickly lose motor neurons and can be distinguished from their WT littermates after 48 to 72 hours. Additionally, these mice show decreases in motor abilities and an average lifespan of 5 days (Hsieh-Li et al., 2000; Monani et al., 2000).

Additional expression of SMN Δ7 in this mouse model (*Smn*$^{-/-}$;*SMN2*;*SMN*$^{\Delta7/\Delta7}$) has become another widely used model of SMA. SMN Δ7 expression results in additional small amounts of full length SMN protein being produced and so reduces the severity of the phenotype. These mice still display gross motor defects and only survive for an average of 14 days (Le et al., 2005).

These mouse models can also be used to generate primary motor neuronal cultures *in vitro,* therefore allowing the study of these SMA phenotypes in cell culture rather than a whole model context. This allows quicker readouts and can indicate whether a therapy could have potential prior to conducting a full *in vivo* study.

Zebrafish models

Danio rerio (or zebrafish) are a commonly used model to study the developmental aspects of vertebrate biology. The use of zebrafish is popular among many researchers due to their quick life cycles, transparency of the embryos during development and the well characterised, simple neuromuscular circuits (Edens et al., 2015).

SMA mutant zebrafish models have been developed using antisense morpholino oligos which target endogenous zebrafish Smn protein. Similarly to mouse models, reduction of endogenous Smn results in motor neuron growth defects, and eventual death (McWhorter et al., 2003). In order to more accurately reflect human disease pathology, a transgenic line expressing the human *SMN2* gene within Smn knockout fish could also be seen to recapitulate human disease phenotypes. In this model *SMN2* produced mainly unstable SMN isoforms and small amounts of full length protein, minimally increasing survival from the Smn mutant lines (Hao et al., 2011).

1.1.4 Roles of Survival Motor Neuron Protein

SMN is a highly conserved protein made up of 294 amino acids, weighing approximately 38 kDa, which is ubiquitously expressed. The human full length SMN protein has multiple domains responsible for many of its functions. These include a N-terminal Gemin-2 and nucleic acid binding domain, a C-terminal tyrosine/glycine rich domain (YG) box as well as a proline rich (Pro) domain. A central Tudor domain is responsible for many of the protein-protein interactions mediated by SMN. Disease-causing mutations within *SMN1* have been found in all domains of the protein, indicating a universal importance in protein function (Singh et al., 2017) (**Fig 1.1**).

SMN can be found in both the nucleus and cytoplasm. Located in nuclear gems, speckles, and Cajal bodies, the nuclear role of SMN protein is involved in RNA processing, including

assembly of ribonucleoproteins (RNPs), splicing, and regulation of transcription. Cytosolic SMN has been implicated in the trafficking of mRNA (particularly in axons), stress granule formation and local translation. An axonal isoform (a-SMN), formed through the retention of intron 3, has been shown to be selectively expressed in the axons of developing motor neurons, promoting axonal growth (Setola et al., 2007).

The main roles of SMN can be seen to be summarised in **Figure 1.2**.

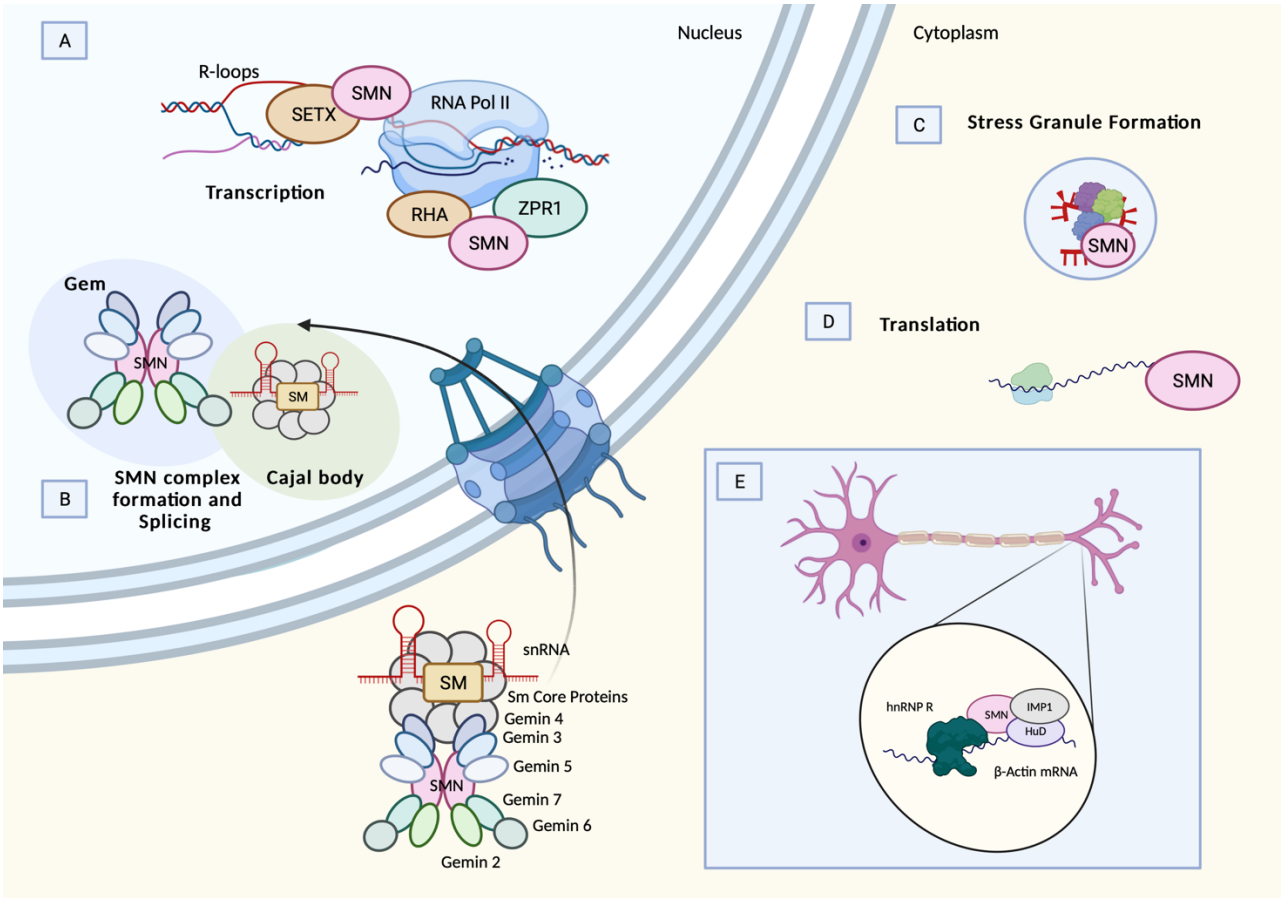


Figure 1.2: Roles of SMN protein. *(**A**) SMN interacts with RNA helicase A (RHA) and ZPR1 in the process of transcription. Transcription termination and R-loop resolution are mediated by SMN's interactions with RNA polymerase II (RNA pol II) and Senataxin (SETX). (**B**) SMN forms a complex with Gemins 2-8, presenting as nuclear gems. In the cytoplasm a complex of seven Smith-antigen (SM) proteins binds an snRNA, to form the snRNP. This process is modulated by the SMN complex, which then aids import into the nucleus. Maturation of the snRNP occurs within Cajal bodies, which then goes on to mediate splicing within the spliceosome. (**C**) SMN is required for stress granule formation (**D**) SMN-primed ribosomes are present in the*

*cytoplasm, responsible for the translation of specific transcripts related to SMA pathology. (**E**) Within neurons SMN is involved in the local translation of specific mRNAs, including B-actin, through interactions with IGF-II mRNA binding protein (IMP1), HuD and Heterogenous nuclear ribonucleoprotein R (hnRNP R). Figure created with biorender.com.*

Assembly of Ribonucleoproteins and pre-mRNA splicing

Full length SMN protein has been implicated in assembly of spliceosomal small nuclear riboproteins (snRNPs) (Liu and Dreyfuss, 1996) and so it has been suggested that *SMN* plays an important role in RNA metabolism which is then disrupted in SMA (Workman et al., 2012).

The SMN protein functions as part of a larger complex made up of Gemins 2-8 and Unrip. Uridine-rich snRNAs (U1, U2, U4, U5, U6, U11, U12, U4atac, U6atac) are transcribed in the nucleus, these are then exported into the cytoplasm where they form complexes with seven Smith-antigen (Sm) proteins. This forms an snRNP, with the Sm proteins in a ring-like structure around a central small nuclear RNA (snRNA) (Liu and Dreyfuss, 1996; Pellizzoni et al., 2002). The SMN-containing complex acts to modulate the production of snRNPs and their import into the nucleus (Li et al., 2014). This process allows maturation of the snRNP within Cajal bodies, a localisation which is mediated through WD40-encoding RNA antisense to p53 (WRAP53β) acting as a scaffold for SMN and coilin (a major component of Cajal bodies) to interact (Mahmoudi et al., 2010). Once mature snRNP's are then localised to nuclear speckles where they can carry out their role in the spliceosome through splicing of pre-mRNA (Lamond and Spector, 2003).

Depletion in the levels of SMN protein causes a reduction in snRNP biogenesis which directly correlates with the severity of SMA cases (Gabanella et al., 2007). This disruption of snRNPs causes substantial splicing defects across the genome, particularly in genes containing large numbers of introns (Zhang et al., 2008). However, these defects are often seen during a late stage of the disease, therefore indicating that this may be a downstream consequence of SMA disease pathology rather than a primary cause (Bäumer et al., 2009).

There is evidence to suggest that SMN also functions in the generation of small nucleolar RNPs (snoRNPs). Ribosomal RNA is transcribed within the nucleolus with post transcriptional processing mediated by snoRNPs (Pellizzoni et al., 2001a). SMN binds directly with fibrillarin and H/ACA ribonucleoprotein complex subunit 1 (GAR1), which are major components of snoRNP assembly within the nucleolus. Mutations in *SMN1* results in an accumulation of snoRNPs outside of the nucleolus, thus implicating SMN in snoRNP processing (Pellizzoni et al., 2001a). Loss of snoRNP chaperone protein Nopp140 within Cajal bodies is also seen in SMA cases (Renvoisé et al., 2009).

Transcriptional Regulation

Mediated by RNA Helicase A (RHA), SMN (alongside other components of the SMN complex) interacts with the C-terminal domain of RNA polymerase II (RNA pol II) (Pellizzoni et al., 2001b). There is also evidence to support interaction of SMN with known transcription factors, such as p53. In SMA patient fibroblasts, p53's localisation in Cajal bodies is disrupted due to the lack of full length SMN protein (Young et al., 2002). Transcriptional co-repressor SIN3A also binds directly to SMN protein (Zou et al., 2004).

RNA pol II-mediated transcription often results in formation of displaced single stranded DNA and DNA:RNA hybrid structures, known as R-loops. These prove to be a source of genomic instability and therefore must be resolved before they are otherwise processed to form DNA double strand breaks. During RNA pol II-mediated transcription, the C-terminal domain of RNA pol II is symmetrically dimethylated by protein arginine methyltransferase 5 (PRMT5). This region directly binds to the Tudor domain of SMN and therefore indirectly also recruits SMN's binding partner Senataxin (SETX) (Yanling Zhao et al., 2016), a helicase which acts to resolve R-loops (Skourti-Stathaki et al., 2011; Suraweera et al., 2009). This therefore places SMN upstream of Senataxin in the process of R-loop resolution and termination of transcription, through recruitment of the 5'-3' exoribonuclease 2 (XRN2) (Yanling Zhao et al., 2016).

Loss of either SMN or SETX results in increased R-loop formation, therefore causing increased levels of DNA damage (Jangi et al., 2017). In SMA cases, it is well charactered that DNA damage levels are elevated, without disruption to DNA repair pathways (Fayzullina and

Martin, 2016, 2014). Therefore, pointing towards genomic instability during transcription as a cause of DNA damage in SMA cases.

mRNA Transport

The localisation and translation of specific mRNAs in SMN-deficient primary neurons are dysregulated (Fallini et al., 2016, 2012; Rossoll et al., 2003). In a genome-wide analysis of SMN-linked mRNA complexes, around 30% were shown to localise to neuronal axons, thus indicating SMN's role in axonogenesis which may be causing neuronal-specific problems in SMA (Rage et al., 2013).

The first dysregulated transcript to be identified was a reduction in the levels of *β-Actin* in the axons and growth cones of motor neurons. Therefore, resulting in reduced growth cone size and axonal length. Heterogenous nuclear ribonucleoprotiein R (hnRNP R) association with *β-Actin* mRNA is dependent on SMN binding to hnRNP R, together these proteins localise *β-Actin* mRNA to motor neurons (Rossoll et al., 2003). Thus, emphasising the importance of SMN in axons of motor neurons.

HuD is another axonal specific RNA-binding protein which shows interaction with SMN (Akten et al., 2011). SMA patient mutations can reduce this interaction, resulting in lower levels of polyadenylated (poly(A)+) mRNA within axons of primary motor neurons, therefore indicating an important role of SMN in axonal trafficking of mRNA (Fallini et al., 2011). HuD is important in the process of axonal branching and dendrite formation, with HuD knockout zebrafish models displaying similar pathological phenotypes to those seen in SMN knockout models (Hao Le et al., 2017). Alongside HuD and SMN, IGF-II mRNA binding protein (IMP1) has also been implicated in RNA binding of certain transcripts important to motor neuronal function, including *β-Actin* (Kim et al., 2015), *Candidate plasticity-related gene 15* (*Cpg15*) (Akten et al., 2011) and *growth-associated protein 43* (*Gap 43*) (Fallini et al., 2016). These transcripts have been observed, in conjunction with RNA binding proteins hnRNP R and hnRNP Q, to be associated with microtubules travelling bidirectionally along axons (Fallini et al., 2014).

Axonal colocalization of SMN and the Golgi-associated Coat Protein I (COPI) complex has also been identified (Peter et al., 2011). SMN has been shown to bind the α-COP subunit, with

depletion of α-COP giving rise to mislocalisation of SMN protein from axonal growth cones to the trans-Golgi network (Ting et al., 2012). α-COP depletion has also been linked to axonal growth defects (Custer et al., 2013; Li et al., 2015). Complimentary with the role of the COPI complex in the trafficking of axonal mRNAs (Todd et al., 2013), together this supports the theory that SMN is involved in neuronal mRNA trafficking required for neuronal outgrowth and maturation.

More recently, SMA patient cells have been demonstrated to have an abnormal golgi morphology, with patients showing a more fragmented phenotype which is restored after SMN overexpression and dysregulated COPI-dependant trafficking. However, this did not result in increased endoplasmic reticulum (ER) stress (Custer et al., 2019).

Translational Regulation

SMN has been demonstrated to bind polyribosomes both *in vitro* and *in vivo* independently of RNA (Bernabò et al., 2017; Sanchez et al., 2013), with more recent research demonstrating that SMN is a ribosome-associated protein (Lauria et al., 2020). There is also a depletion of membrane associated ribosomes in SMA cases (Gabanella et al., 2016). By comparing the translatome of control and SMA mice, a substantial decrease in translation efficiency was observed within SMA cases. This corresponded to a decrease in the total levels of ribosomes (Bernabò et al., 2017).

Out of total cellular ribosomes, a small proportion these have been shown to be SMN-primed. These ribosomes are specifically associated with mRNAs which have been previously linked to SMA pathology (Lauria et al., 2020). This finding may help researchers to understand better why motor neurons are selectively sensitive in SMA disease progression.

Stress Granule Assembly

During the stress response the cell upregulates the transcription and translation of multiple stress related genes, while simultaneously silencing other unrelated genes. This results in a cytoplasmic pool of mRNA of which translation initiation is stalled, this dynamic structure is assembled, maintained and dissembled by a number of mRNA binding proteins, RNPs and protein complexes (Protter and Parker, 2016).

SMN has been shown to colocalise with known markers of stress granules: cytotoxic granule-associated RNA binding proteins (TIA1 and TIAR) and Rad-GTPase-activating protein (SH3-domain) binding protein (G3BP) (Hua and Zhou, 2004; Zou et al., 2011). Known SMN-binding partner, Fragile X mental retardation protein (FMRP) is also a component of stress granules (Linder et al., 2008). Reduction in SMN levels, as seen in SMA patients, can increase the cell's sensitivity to stress through the inhibition of stress granule formation (Zou et al., 2011). Coupled with the observation that SMN is present in stress granules prior to TIAR points towards SMN's function within stress granule assembly (Hua and Zhou, 2004).

1.1.5 Disease Modifying Genes

A small number of individuals with a family history of SMA can appear fully asymptomatic despite having a homozygous mutation in the *SMN1* gene and the same copy number of *SMN2* as their affected siblings (Cobben et al., 1995; Hahnen et al., 1995; Prior et al., 2004). This therefore points to the hypothesis that other genes also have the capacity to act as disease modifiers. Several genetic modifiers of SMA disease severity have been identified within the literature; for example: *Plastin3* (*PLS3*), *neurocalcin delta* (*NCALD*) and Zinc Finger *ZPR1* (*ZPR1*) have all been shown to have protective properties in SMA patients (Ahmad et al., 2012; Hosseinibarkooie et al., 2016; Oprea et al., 2008; Riessland et al., 2017; Wadman et al., 2019).

Plastin 3 (PLS3)

A gene located on the X-chromosome, *PLS3*'s main function is involved in F-actin binding and bundling (Shinomiya, 2012), therefore involving the protein in all aspects of cellular function which rely on F-actin dynamics. This can include endocytosis, neurotransmission, axonal local translation and vesicle trafficking to name a few (Wolff et al., 2021). Plastin3 has also been shown to form actin-containing complexes with SMN (Oprea et al., 2008). This complex is also thought to contain hnRNP F and/or hnRNP H1/2 (Walsh et al., 2020).

An initial study in 2008 focussed on six families who together had eight female offspring that were fully asymptomatic despite inheriting the same *SMN1* and *SMN2* alleles as their male symptomatic siblings. This study established *PLS3* as a transcript which was consistently

upregulated in these asymptomatic individuals compared to their siblings, therefore making it a strong candidate as a SMA disease modifying gene (Oprea et al., 2008). Experimental data shows that Pls3 overexpression can rescue axonal growth defects caused by SMN reduction. It is thought that overexpression of Pls3 can ameliorate SMA pathogenesis through elevation of F-actin levels which can stabilize the axonal growth cone, thus aiding axonogenesis (Oprea et al., 2008). This has been confirmed in more recent studies which see dramatic increases in Pls3 protein in asymptomatic patient-derived iPSCs compared to symptomatic SMA patient cells, this protein can be seen in abundance in the growth cones of axons (Heesen et al., 2016).

Pls3 protein levels are dramatically reduced upon loss of SMN in zebrafish knockout models (Hao et al., 2012). This appears to be regulated at a translational level. Reduced protein levels of Pls3 have also been shown in a mouse model of SMA (Bowerman et al., 2009). Overexpression of Pls3 results in extended life spans and reduced disease severity within SMA mouse models (Ackermann et al., 2013; Alrafiah et al., 2018b; Kaifer et al., 2017)

These data together suggest that *PLS3* is a sex-specific, disease modifier of SMA functioning downstream of the SMN protein to stabilise axons and prevent axonal degeneration, which is disrupted in SMA patients.

Neurocalcin Delta (NCALD)

Neurocalcin Delta (NCALD) is a neuronal calcium (Ca^{2+}) sensor protein which is highly conserved in mammalian species. Present in high concentrations in cerebral neurons, spinal motor neurons and axonal growth cones (Iino et al., 1998; Wang et al., 2001), NCALD is highly implicated in endocytosis through interaction with actin and clathrin (Ivings et al., 2002). This therefore indicates NCALD's function is important in neuronal endocytosis-dependant mechanisms, including neurotransmitter release and axonal growth (Riessland et al., 2017). Upregulation of NCALD during development causes inhibition of neurite outgrowth (Yamatani et al., 2010).

Similarly to Pls3, transcriptome analysis of asymptomatic SMA patients established NCALD as a downregulated SMA disease modifier (Riessland et al., 2017). Low levels of SMN act to reduce voltage-dependant Ca^{2+} influx, NCALD acts to inhibit endocytosis at low Ca^{2+} levels.

Therefore, downregulation of NCALD can help to restore synaptic endocytosis and normal function. In support of this, knockdown of NCALD in SMA animal models can rescue motor neuron growth defects, neuromuscular junction dysfunction and improve motor function (Riessland et al., 2017).

Reductions in NCALD have been implicated in various neurodegenerative diseases, including Alzheimer's disease and schizophrenia (Upadhyay et al., 2019). Research demonstrates that full knockout of NCALD proves detrimental in mouse models, while reduction in the heterozygous mice appears to have no noticeable effect. Therefore supporting reduction of this gene as a possible therapeutic target for SMA (Upadhyay et al., 2019).

Zinc Finger Protein ZPR1 (ZPR1)

ZPR1 was first identified as a potential SMA disease modifier through genetic linkage analysis of nine SMA discordant families with unaffected siblings, with ZPR1 protein levels being higher in unaffected siblings than that of those displaying symptoms of SMA (Helmken et al., 2003). Reduced expression of ZPR1 in a mild SMA mouse model also displayed increased disease severity (Ahmad et al., 2012).

Although its function is yet to be fully understood, ZPR1 protein interacts with SMN and is necessary for its localisation to Cajal bodies and gems. This interaction is lost in SMA patient cells (Gangwani et al., 2001). *ZPR1* knockout mouse models proved to be embryonic lethal, with embryos demonstrating a loss of SMN localisation to these sub nuclear structures, alongside reduction in snRNP's (of which SMN is responsible for the production). ZPR1 knockdown in motor neurons also displayed SMA-like pathological phenotypes, such as growth cone defects, axonal growth reduction and increased apoptosis (Gangwani et al., 2005).

As the homozygous *ZPR1*$^{-/-}$ mice proved to be embryonic lethal, the heterozygous mice were studied in greater detail. These mice displayed neurogenerative phenotypes which led researchers to hypothesise that the reduction of ZPR1 seen in SMA patients is exacerbating the SMA neurodegenerative phenotypes and therefore could be a potential therapeutic target for SMN-independent treatment of patients (Doran et al., 2006).

Overexpression of ZPR1 in both *in vitro* and *in vivo* models of SMA displayed certain positive effects on SMA disease phenotype's; there was rescue axonal growth deficits of motor neurons, prevention of R-loop accumulation, and an increase in SMN protein levels (Ahmad et al., 2012; Kannan et al., 2020). ZPR1 has been shown to bind RNA pol II. In addition, *in vivo* association of ZPR1 and the genomic *SMN* locus have also been demonstrated, suggesting that ZPR1 is a component of the core transcription complex (Kannan et al., 2020). This therefore causes upregulation of *SMN* transcription, ameliorating disease phenotypes in an SMN-dependant manner. These findings indicate the importance of *ZPR1* as a protective modifier of disease and a new therapeutic target for the treatment of SMA.

1.1.6 Treatment of SMA

Developments within SMA therapy are increasingly exciting for the whole field of motor neuron diseases. As a disease with a clear monogenic cause which is now relatively well characterised, SMN-dependant therapies have been at the forefront of SMA research. In recent years, multiple ground-breaking disease modifying therapies have been approved for SMA patients. These treatments will not only be of great benefit to the SMA community but also set the precedent for similar therapies in other neurodegenerative diseases.

In addition to these therapies which focus on SMN-dependant mechanisms of disease, research is ongoing into SMN-independent therapeutic strategies as well as treatments which target pathways which are dysfunctional in disease.

SMN-dependant Therapies

The area of SMN dependant therapies is one which has been very successful in recent years. Nusinersen®, the first clinically approved treatment for SMA, is an antisense oligonucleotide (ASO) which binds to an intronic splice-silencing-site within the seventh intron of *SMN2* to prevent the exclusion of exon 7 from the *SMN2* mRNA transcript. This therefore gives rise to an increase in production of full length SMN and amelioration of SMA symptoms (Rigo et al., 2012). After promising results in clinical trial phases I and II (Chiriboga et al., 2016; Finkel et al., 2016), multiple phase III clinical trials were initiated, each assessing a different sub group of SMA patients: ENDEAR assessed infants under 7 months of age with Type I SMA (Finkel et

al., 2017), CHERISH focussed on older children (an average age of 4 years) with SMA type II (Mercuri et al., 2018), and finally NURTURE studied presymptomatic infants under 6 weeks of age with only 2 or 3 copies of *SMN2* (De Vivo et al., 2019). After significant improvement in all cases, the treatment was approved for clinical use. However, as an ASO this treatment is unable to cross the blood brain barrier (BBB) and so must be administered through lumbar puncture, an invasive and expensive procedure. It is also short-lived and therefore requires repeat doses.

Alternative therapies revolve around small molecules (e.g. Risdiplam®) which can act in a as a splice modifier of *SMN2,* similarly to Nusinsersen®. However these drugs are capable of crossing the BBB and therefore can be taken orally daily by the patient to increase full length SMN, providing a safer and more user friendly treatment option (Poirier et al., 2018; Singh et al., 2020). Risdiplam® received approval in 2020 to be administered to SMA patients of all ages (Dhillon, 2020).

Gene therapy for SMA has shown great promise in pre-clinical studies for several years (Benkhelifa-Ziyyat et al., 2013; Foust et al., 2010; Meyer et al., 2015; Valori et al., 2010). Still in its infancy, adeno-associated virus (AAV)-mediated gene therapy has previously not been approved for clinical use. Zolgensma® is a self-complimentary AAV9 (scAAV9) vector which contains the full length SMN gene. AAV9 has a strong tropism to the central nervous system (CNS), therefore meaning that the treatment can be injected intravenously and it will cross the BBB. Initial clinical trials with SMA type I patients demonstrated an unprecedented improvement in motor function and survival (Al-Zaidy et al., 2019), and so became approved for clinical use in patients under 2 years of age in 2019 (Hoy, 2019). Due to AAV's tropism for the liver, elevated levels of liver enzymes have been observed in patients administered this treatment, and so corticosteroid use is advised alongside treatment to prevent long term damage (Hoy, 2019; Mendell et al., 2017).

The extent to which these SMN-dependant therapies have a positive effect on SMA patients depends on the stage at which they are administered. For many patients, they offer a means to prevent further deterioration in motor function and a maintenance of quality of life rather than a restorative therapy. However, if treated early enough in the disease course, patients can retain much of their motor function (De Vivo et al., 2019; Finkel et al., 2016). This

highlights the importance of routine new-born screening for genetic disorders, as pre-symptomatic diagnosis of SMA gives individuals the best outcomes for treatment (Hensel et al., 2020).

SMN-independent therapies

Due to the multi-systemic nature of SMA, non-SMN related therapies may provide a therapeutic benefit to SMA patients.

Olesoxime is a small molecule which is known to act as a neuroprotective agent and was a potential therapy for type II and III SMA. Phase II clinical trials demonstrated that there was no adverse effects from Olesoxime treatment and motor function was preserved in patients (Bertini et al., 2017). However, long term follow up of this treatment did not lead to a significant benefit for these patients and so this was dropped as a potential SMA treatment (Muntoni et al., 2020).

Another avenue of therapy is through stimulation of muscle growth and function. Myostatin inhibitors are a group of molecules which prevent myostatin from functioning in its normal role to inhibit muscle over-growth within skeletal muscle. SRK-015 has demonstrated increased muscle function and mass in SMA mouse models (Long et al., 2019), a clinical variant is currently in development with Phase II clinical trials showing it is safe to use (Barrett et al., 2021).

Fast Skeletal Muscle Troponin Activators (FSTA)'s are another group of therapy's which target muscle weakness and neuromuscular dysfunction. CK-2127107 has been shown to be a safe therapy in its initial clinical trial data (Andrews et al., 2018), data from clinical trials involving SMA patients is still pending. Preliminary data from interim analysis indicate positive effects of this drug (Schorling et al., 2020).

1.2 *GLE1*: Potential new disease modifying gene

GLE1 is an essential gene which is highly conserved across species. In humans the *GLE1* gene is comprised of 16 exons which gives rise to two splice isoforms, named hGle1A and hGle1B. The full coding region of the gene, consisting of 698 amino acids, is present in the hGle1B while the hGle1A isoform is lacking a 43-amino acid region at the C-terminal domain. In addition hGle1A ends in a unique four amino acid sequence (Kendirgi et al., 2003) (**Fig 1.3**).

Initial studies in HeLa cells (Kendirgi et al., 2003), which have since been confirmed (Rayala et al., 2004), demonstrate that the hGle1B isoform is around 1000 times more abundant than hGle1A, making this the predominant isoform. Visualisation of GFP tagged isoforms has shown that these two isoforms differ in their localisation: comparison with wildtype hGle1 demonstrated that GFP-hGle1B localised at the Nuclear Pore Complex (NPC) similarly to the endogenous protein, with GFP-hGle1A being more diffuse within the cytoplasm (Kendirgi et al., 2003; Rayala et al., 2004).

The differing localisations of Gle1 isoforms therefore point towards a range of roles for human Gle1. hGle1b localises to the NPC through interaction with Nup42 (also known as hCG1) (Lin et al., 2018) and Nup155 (Rayala et al., 2004), these interactions can occur on both the nucleoplasmic and cytoplasmic faces of the NPC due to the abundance of these nucleoporins and Gle1's ability to shuttle between the nucleus and the cytoplasm (Kendirgi et al., 2005; Rayala et al., 2004). hGle1b has been shown to be required for nuclear export of poly(A)+ mRNA and coordination of transcriptional termination (Kendirgi et al., 2005; Sharma and Wente, 2020). Cytoplasmic hGle1A has been shown to be required for stress granule formation and initiation of translation (Aditi et al., 2019, 2015).

Studies into Gle1 have placed it as a modulator of Dbp/DDX (or DEAD-box) proteins. DEAD-box proteins are a family of RNA helicases which are regulated through ATP binding and hydrolysis thus changing their affinity for binding single and double stranded RNA (Jarmoskaite and Russell, 2011). N- and C-terminal extensions confer specificity of each DEAD-box protein to different functional protein or RNA targets, with humans having 37 different family members and 26 in *S.cerevisiea* (Gilman et al., 2017). Therefore giving each DEAD-box protein different functions in remodelling mRNP complexes, rearrangement of RNA or

forming/separating single or double stranded RNA (Jarmoskaite and Russell, 2011). Gle1 appears in many processes to bind to the DEAD-box protein to modulate its ATPase activity, with the C-terminal region of hGle1 responsible for DEAD-box protein modulation (Folkmann et al., 2013) (**Fig 1.3**).

Mutations within the *GLE1* gene have been shown to be responsible for severe embryonic motor neuron diseases (Lethal Congenital Contracture Syndrome (LCCS1) and Lethal arthrogryposis with anterior horn cell disease (LAAHD)), as well as Amyotrophic Lateral Sclerosis (ALS) (Kaneb et al., 2015; Nousiainen et al., 2008). The most severe of these mutations occurring within the coiled-coil domain of hGle1, this region is responsible for self-association alongside an aggregate prone N-terminal region. Disruptions to this process result in a mislocalisation of Gle1 protein and deficient Gle1 mediated processes (including mRNA export and stress granule dynamics) (Mason and Wente, 2020).

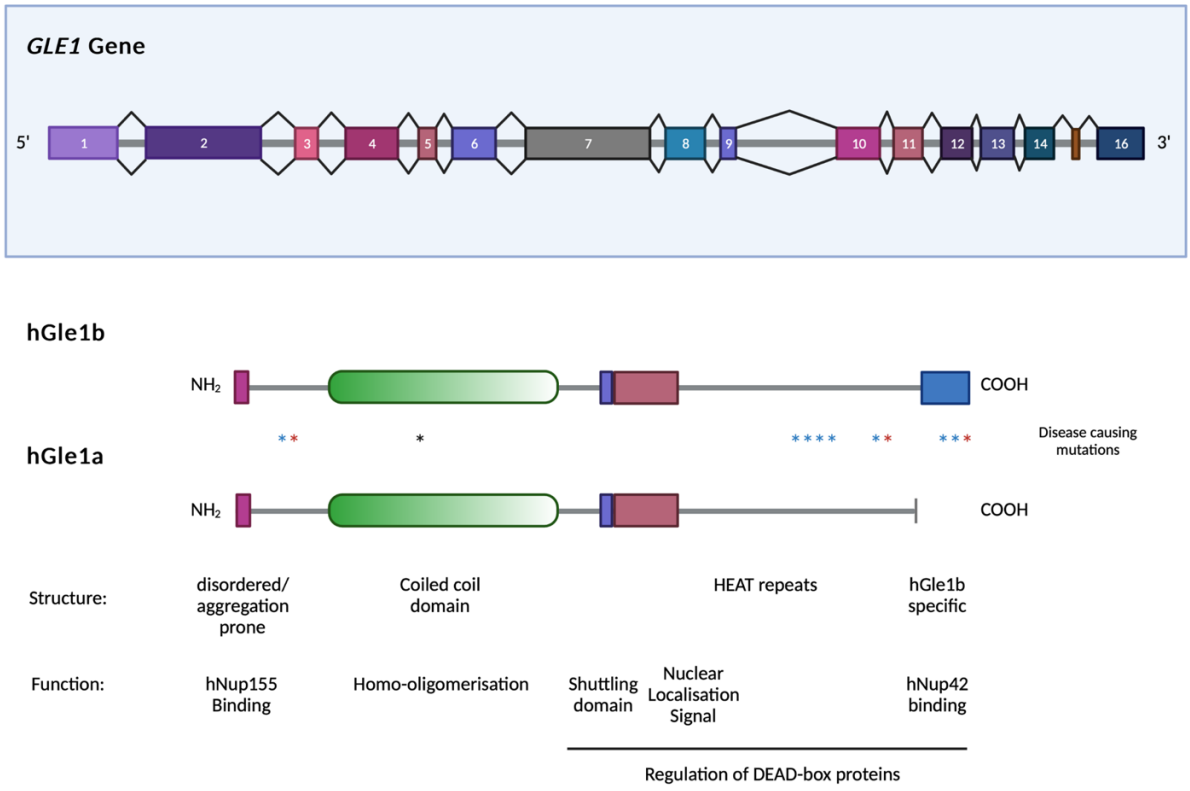


Figure 1.3: Genetic and protein structure of GLE1*. Genetic structure of the human GLE1 gene displaying its 16 exons. Human isoforms of the protein showing the protein domains: The N-terminal hNup155 binding domain which targets the protein to the NPC; The coiled coil domain which is responsible for protein-protein interactions; The shuttling domain which has a role in nucleocytoplasmic transport; The Nuclear localisation domain and the C-terminal hCG1 domain which is only present in the hGle1B isoform and is also involved in targeting the protein to the NPC. Stars correspond to locations of mutations seen in LCCS1 (black), LAAHD (blue) and ALS (red) patients. Figure is adapted from (Nousiainen et al., 2008), (Kaneb et al., 2015) and (Mason and Wente, 2020). Figure created with biorender.com.*

1.2.1 Nuclear Pore Complex

The NPC is one of the main routes of RNA nuclear export, this structure is remarkably conserved across all species of eukaryote both in structure and the proteins of which it is made, therefore demonstrating the evolutionary importance of this structure.

Each NPC is made up of thirty different nuclear pore proteins (nucleoporins) (Cronshaw et al., 2002). This group of proteins can be subdivided into three categories due to their chemical structures and roles within the NPC. The first of these being phenylalanine-glycine (FG) nucleoporins, these are rich in Phe-Gly repeat residues and act within the channel to facilitate active transport via direct interactions with soluble transport receptors. Nucleoporins that are lacking these FG repeats are important structural components of the NPC. Finally, the third variety are membrane proteins known as Nups which act to anchor the NPC within the nuclear membrane (Köhler and Hurt, 2007).

Scanning electron microscopy has helped to elucidate the structure of the NPC, again demonstrating the conservation across multiple eukaryotic species (Akey and Radermacher, 1993; Beck et al., 2004; Bui et al., 2013; Hinshaw et al., 1992; Kim et al., 2018). The core structure of the NPC is an eightfold symmetrical 'spoke complex' which spans the nuclear membrane. The interior of this structure forms the central transporter component of the NPC where proteins and complexes can be actively transported through specific interactions with the FG-repeat motifs (Kim et al., 2018; Stanley et al., 2017). This inner ring appears to be the region of the NPC which is most conserved between species (Obado et al., 2016). Two outer rings fuse together the inner and outer nuclear membranes and together with the inner ring form the central domain of the NPC (Alber et al., 2007). Attached to this core domain is a 'basket' like structure formed from eight rod shaped filaments which extend out into the nucleoplasm, eight cytoplasmic filaments reach into the cytoplasm of the cell. The majority of nucleoporins can be seen across the inner and outer structures symmetrically however there are certain proteins specific to each domain; these nucleoporins appear to have roles in the directional transport of molecules through the NPC (Hautbergue, 2017; Köhler and Hurt, 2007).

There are a number of important differences between the NPCs across eukaryotic species: the yeast NPC is much smaller than that of the vertebrate (50 mDa compared to 125 mDa) (Kim et al., 2018; Reichelt et al., 1990). Duplications appear in more evolutionarily advanced species, with both the inner and outer rings being expanded in vertebrates as compared to yeast, therefore explaining the difference in molecular weight. Human NPCs also have additional copies of certain Nups connecting the rings together (Nup157 or Nup170 is demonstrated to connect the outer and inner rings in humans) (Kim et al., 2018).

1.2.2 Roles of Gle1

Much of the preliminary research into the role of Gle1 was undertaken utilising the yeast *Saccharomyces cerevisiae*. Research into sequence homology between the human and yeast proteins demonstrated that the human orthologue does not contain the same leucine-rich nuclear export sequence seen in the yeast protein. Combined with the presence of two isoforms of human Gle1 compared to the one seen in yeast, indicates that the mechanism by which these proteins function may have evolved as systems became more complex. Regardless of this, the structural and functional similarities between much of the remaining regions of the gene across species still warrants the use of yeast as a tool in elucidating the role of Gle1 in RNA processing (Watkins et al., 1998). In this section, I will review the pathways as they were studied in yeast (*S.cerevisiea)*, before comparing how they differ in humans. A list of the names of the genes and protein isoforms in yeast and humans can be seen in **Table 1.1**. A summary of the roles which hGle1 has been established in can be seen in **Figure 1.4**.

Table 1.1: Genes which are studied in both yeast and human functions of Gle1

S.cerevisiea	**Humans**
Gle1	hGle1
Dbp5	DDX19
Nup42	NUP42 (hCG1)
Nup159	NUP214 (CAN)

Nup116	NUP98
Nup170/Nup157	NUP155
Ded1	DDX3

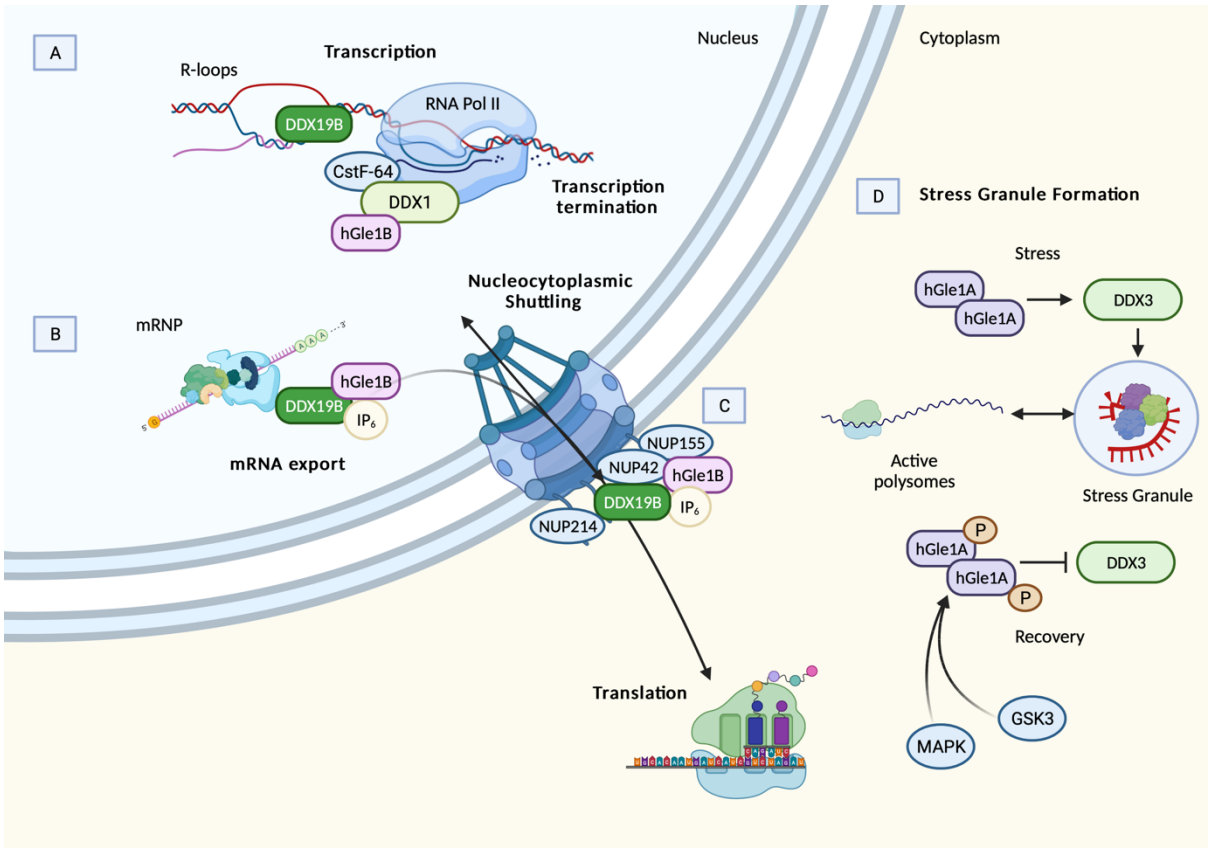


Figure 1.4: Current knowledge of Gle1's roles in humans. *(**A**) hGle1b modulates DDX1 in the cleavage of pre-mRNA by CstF-64, resulting in termination of transcription. DDX19B is implicated in R-loop resolution independently of hGle1b. (**B**) hGle1b activates DDX19B to bind RNA and export it through the nuclear pore complex (NPC). Inositol hexakisphophate (IP_6) binds Gle1 to activate the ATPase activity of DDX19B. (**C**) hGle1b is anchored to the cytoplasmic fibrils of the NPC by NUP155 and NUP42. DDX19B is anchored to the cytoplasmic fibrils by NUP214 and NUP42. Both hGle1b and DDX19B shuttle between the nucleus and the cytoplasm. (**D**) hGle1A is involved in stress granule formation in response to stress. Upon stress response, hGle1A activates DDX3 activity to silence translation and form stress granules. Phosphorylation of hGle1A by the Mitogen Activated Protein Kinase (MAPK)*

and glycogen synthase kinase 3 (GSK3) pathways result in no activation of DDX3 and so stress granules disassemble, causing translation to resume. Figure created with biorender.com.

mRNA Export

Early studies demonstrated that Gle1 is necessary for the export of poly(A)+ RNA in cell culture (Murphy and Wente, 1996; Strahm et al., 1999; Watkins et al., 1998). It has since been established that hGle1's nucleocytoplasmic shuttling activity is conferred by a 39-amino acid region (Kendirgi et al., 2003). This shuttling activity also means that hGle1 is a transient component of the NPC rather than being part of its stable conformation as previously thought (Kendirgi et al., 2003). Proteomic analysis supporting this alternative theory shows that hGle1 is not identified in an evaluation of NPC components despite showing association with all other known nucleoporins (Cronshaw et al., 2002).

Dbp5 (DDX19B in humans) is an ATP-dependant DEAD-box protein which is required for poly(A)+ RNA export (Tseng et al., 1998). Initial yeast studies demonstrated that Dbp5 continuously shuttles between the nucleus and cytoplasm, displaying interactions with Gle1 (Hodge et al., 1999). In addition, further study demonstrated that Dbp5 interacts with mRNA during the process of transcription as part of the transcriptional complex (Estruch and Cole, 2003; Zhao et al., 2002).

Inositol polyphosphates (InsPs) are a group of small, soluble molecules which undergo multiple rounds of hydrolysis and phosphorylation. They have been heavily implicated in the export of mRNA from the nucleus, with cytoplasmic inositol hexakisphosphate (IP_6) being required for Gle1-mediated mRNA export in *S.cerevisiea* (Miller et al., 2004; York et al., 1999).

IP_6-bound Gle1 acts to stimulate Dbp5's ATPase activity. Dbp5 alone has weak enzymatic ability and so is unable to bind mRNA effectively. Presence of Gle1 causes a significant increase in Dbp5 activity, suggesting that activation of Dbp5 at the NPC by Gle1 is required to export mRNA from the nucleus (Alcázar-Román et al., 2006; Weirich et al., 2006). Dbp5 is recruited to the NPC by Nup159 binding (Schmitt et al., 1999). This interaction is necessary to

control Dbp5's activation by IP_6-bound Gle1. Nup159 binding causes release of Dbp5-bound ADP, the ATPase activity of Dbp5 is then stimulated by Gle1-IP_6 and so activates Dbp5 for RNA loading (Noble et al., 2011). This conformational shift allows multiple messenger ribonucleoprotein (mRNP) molecules to be remodelled and exported successfully by a single Dbp5 enzyme at the NPC (Folkmann et al., 2011; Noble et al., 2011).

In humans, spatial regulation of this process is controlled through a variety of interactions: A sequence of 29-amino acids at the N-terminal of hGle1 recognises Nup155 and aids localisation of Gle1 to the NPC. Cells which are lacking hGle1B display mRNA export defects (Aditi et al., 2015), indicating that this interaction alone is not sufficient to recruit both isoform hGle1A and hGle1B to the NPC (Rayala et al., 2004). Therefore, this spatial regulation is reliant upon the C-terminal domain specific to hGle1b as a second binding region. This 43 amino acid sequence binds to Nup42 (also known as hCG1 in humans) and localises hGle1B to the NPC (Kendirgi et al., 2005). In yeast, Nup42 and Nup159 not only bind to Gle1 and Dbp5 but through phenylalanine-glycine (FG) repeat domains they act to position the mRNP in preparation for the remodelling activity of Dbp5 (Adams et al., 2014), however this function is yet to be confirmed in humans.

Evolutionary conservation within this pathway in yeast and humans has been shown though research showing that in the presence of IP_6, a Nup42-Gle1-Dbp5 trimeric complex is formed. This acts to stimulate the activating effect of Gle1 on Dbp5, rather than just being for localisation to the NPC (Adams et al., 2017).

In contrast, more recently it has been shown that in humans Gle1 stimulation of DDX19 does not require IP_6. This indicates structural changes between yeast Dbp5 and human DDX19 in which the IP_6 coordinating residues have been lost. This results in a lack of IP_6 seen in crystal structures of Gle1-Nup42-DDX19 (Lin et al., 2018). This study also implicates Nup42 as an important modifier of Gle1 thermostabilty. *GLE1* mutations seen in human disease demonstrated instability at 37 ^{O}C, where the severity of the disease correlated with the level of thermoinstabilty observed (Lin et al., 2018). They therefore attribute the changes in activation levels seen in the previous paper to changes in Gle1 thermostability rather than direct effects of IP_6 binding. The 2017 paper by Adams and colleagues also focusses much more on the *S.cerevisiae* model rather than human cells, demonstrating where the pathways

may have diverged during evolution. More research is required to fully understand whether IP_6 is still involved in this pathway in humans.

Makorin-2 (MKRN2) is a RNA-binding E3 ubiquitin ligase which has been found to be a negative regulator of Gle1-mediated RNA export, with evidence that MRKN2 binds RNA, Gle1 and Nup42 (Wolf et al., 2020).

Translation Regulation

Multiple studies have identified additional roles for Dbp5, Gle1 and IP_6 which are unrelated to their function in mRNA export (Bolger et al., 2008; Gross et al., 2007). In studies using yeast, it has been evidenced that Gle1 is involved in both the initiation and termination of translation. During the termination of translation Gle1 has been shown to interact with the termination factor Sup45/eukaryote release factor 1 (eRF1). Gle1 mutants also show defects in both of these processes confirming that this protein is necessary for these actions (Bolger et al., 2008). Dbp5 has also been implicated, alongside Gle1, in recruitment of eRF1 through mRNP remodelling in order to allow correct binding of the proteins and thus an efficient termination (Gross et al., 2007). Gle1 is the main target of IP_6 and this association is seen to be necessary to regulate termination of translation as well as mRNA export, thus also making IP_6 necessary in the activation of Dbp5 activity during both of these processes (Alcazar-Roman et al., 2010).

Initiation is carried out through associations between Gle1 and subunits of eukaryotic initiation factor 3 (eIF3). Unlike termination, translation initiation is regulated solely by Gle1, with mutations in IP_6 and Dbp5 having no impact on this process. Therefore demonstrating that Gle1 has additional solo functions (Bolger et al., 2008). It was later found that this is due to the ability of Gle1 to modulate another DEAD-box protein: Ded1 (DDX3 in humans). In contrast to its effect on Dbp5, Gle1 acts to directly inhibit Ded1 ATPase activity independent of IP_6 (Bolger and Wente, 2011). Depending on the condition, Ded1 can act to either promote translation when active or when inactive it can repress this pathway. Excess levels of Ded1 can also repress translation. Gle1 therefore acts as a modulator of Ded1 which can fine tune the rate of translation depending on the situation (Aryanpur et al., 2017).

In humans, polysome profiling of cells treated with siRNA against hGle1 demonstrated an increased monosome (80S) peak alongside reduced polysomes, as is also observed in yeast Gle1 depleted cells, therefore indicating that Gle1's role in translation initiation is conversed in humans (Aditi et al., 2015). Interaction of hGle1 with DDX3 was also confirmed to be conserved, with translation defects seen in Gle1 depleted cells being rescued by DDX3 overexpression. Thus, implicating hGle1 in DDX3 modulation of translation (Aditi et al., 2015).

Stress Granule Formation

The human Gle1a isoform of the gene is not necessary for mRNA transport and therefore has distinct cellular roles to hGle1b (Aditi et al., 2015). As explained previously, Yeast Gle1 modulates the expression of DEAD-box protein Ded1 and interacts with translation initiation factor eIF3 and also functions in translation termination (Bolger et al., 2008; Bolger and Wente, 2011). Due to the dynamic and closely linked nature of translation regulation and stress granules, it was investigated whether hGle1a's role was related to these pathways (Aditi et al., 2015). Upon cellular stress it could be seen that hGle1a was recruited to stress granules. Depletion of hGle1 lead to impaired formation of stress granules and increased translation during a stress response, thus implicating Gle1 in the assembly of stress granules (Aditi et al., 2015). This process is thought to be as a result of Gle1a's ability to modulate activity of DDX3 to distribute mRNPs between translation and stress granule formation (Aditi et al., 2015).

More recently, it has been demonstrated that hGle1a's phosphorylation state can influence stress granule formation. Gle1a is hyperphosphorylated in response to cellular stress, and this acts to modulate stress granule dynamics. Basally phosphorylated Gle1a self-associates and promotes stress granule assembly through DDX3 activation. Mitogen activated protein kinase (MAPK) and Glycogen synthase kinase 3 (Gsk3) activity, triggered by cellular stress, causes phosphorylation of Gle1a. This phosphorylation state inhibits DDX3 ATPase activity and therefore causes stress granule disassembly (Aditi et al., 2019).

hGle1 oligomerization, through the coiled coil and N-terminal aggregate prone region, has been shown to be required for regulation of stress-related translation and stress granule formation (Mason and Wente, 2020).

The survival of cells once exposed to toxic stressors, such as sodium arsenite, has also been shown to be dependent on hGle1a. Depletion of Gle1a resulted in reduced survival while overexpression showed increased survival rates. This therefore supports hGle1a's role in stress granule dynamics (Glass and Wente, 2019).

Nuclear Roles of hGle1b

At the time this PhD was started, very few nuclear roles for Gle1 had been investigated. During this PhD project new roles for hGle1 have been described, such as hGle1b's role in the termination of transcription.

DDX19B was recently implicated in the resolution of DNA damage associated R-loops formed during replication and transcription of RNA, therefore maintaining genomic integrity. Depletion of DDX19 in mammalian cells correlated with an increase in R-loops, this increase was shown to be significantly higher than that of depleted Gle1 (Hodroj et al., 2017a). Upon DNA damage DDX19 re-localises to the nucleus from the NPC and resolves R-loops, independent of other known R-loop helicases (e.g. SETX). This study implicates DDX19 in the resolution of R-loops through Ataxia-telangiectasia-mutated-and-Rad3-related kinase and its downstream effector Checkpoint Kinase 1 (ATR-Chk1 pathway), therefore linking the processes of mRNA nuclear export, transcription and replication stress (Hodroj et al., 2017a, 2017b).

Additional research into depletion of Gle1 and DDX19 demonstrated that downregulation of these proteins resulted in increased levels of DNA damage marker γH2AX, with Gle1 showing the most significant increases. Evaluation of cytoplasmic mRNA following siRNA knockdown identified specific subsets of mRNA showing changes, among others there was a decrease in expression of DNA repair factors, again with Gle1 showing the most prominent effects (Okamura et al., 2018). There were also mitotic progression defects seen in cells depleted of Gle1, but not DDX19, with a specific deficit at the G2/M progression checkpoint related with DNA damage. In addition there was evidence of scattered chromosomes, indicating that spindle and centrosome formation was impaired (Okamura et al., 2018). Gle1 has been shown to be localised to the centrosome and basal bodies, localising at the centriole around G2/M phase of the cell cycle, with depletion of Gle1 resulting in microtubule dysfunction (Jao et al.,

2017). Therefore, indicating an additional nuclear role specific to hGle1b which is unrelated to its role as an activator of mRNA nuclear export.

The theory of Gle1 being involved in DNA damage pathways was evaluated by Sharma and colleagues. Upon disruption of Gle1 shuttling between the nucleus and cytoplasm, they observe an increase in R-loop staining within the nucleus. However, they discount the role of Gle1 within the DNA damage response due to the lack of γH2AX staining seen when Gle1 shuttling is blocked. This therefore suggests that the increase in R-loops observed is linked to the process of transcriptional termination where R-loops can form due to the disruption to pre-mRNA cleavage processes (Sharma and Wente, 2020).

Multiple DEAD-box proteins are implicated in the process of transcriptional termination; DDX1 coordinates pre-mRNA cleavage alongside Cleavage stimulation factor (64kDa subunit) (CstF-64) (Bléoo et al., 2001). DHX9 (also known as RNA Helicase A) and SETX are responsible for the termination of transcription through removing R-loops (Chakraborty et al., 2018; Skourti-Stathaki et al., 2011). The process by which these enzymes coordinate was still unanswered, combined with an undetermined nuclear role for hGle1 researchers decided to investigate whether Gle1 could be involved in the modulation of these factors (Sharma and Wente, 2020).

Co-localisation between Gle1 and DDX1 and R-loops was observed through proximity ligation assays (PLA). Upon blockage of Gle1 shuttling between the nucleus and cytoplasm, the colocalization of Gle1 and DDX1, as well as interaction of DDX1 and CstF-64, were reduced. Therefore, indicating that Gle1 is required for DDX1 association with CstF-64. When Gle1 is prevented from shuttling mRNA cleavage is therefore disrupted, causing elongated 3'-UTRs of mRNAs by prolonged contact with RNA pol II. Transcripts which are modulated by Gle1 are specifically prone to R-loop formation during transcription, explaining the increase in R-loop staining when Gle1 shuttling is blocked. This role of Gle1 in transcription termination is independent of DDX19B (Sharma and Wente, 2020).

1.2.3 Mutations within *GLE1*

Table 1.2: Mutations within the GLE1 gene that have been implicated in disease. *The paper in which each mutation was first characterised have been referenced. In some cases, a predicted effect of the mutation was not suggested.*

Mutation	Disease	Position and nucleotide change	Type of mutation and region effected	Predicted effect	Reference
Fin_{Major} T144_E145insPFQ	LCCS1	c.432-10A>G	3 new amino acids in the coiled coil domain (intron 3)	Disrupts the coiled coil domain	(Nousiainen et al., 2008)
p.R569H	LAAHD (heterozygous with Fin_{major})	c.1706G>A	Point mutation in exon 12. R to H substitution	Likely to damage the C-terminal domain structure	(Nousiainen et al., 2008)
p.R569H	LAAHD (homozygous)	c.1706G>A	Point mutation in exon 12. R to H substitution	Likely to damage the C-terminal domain structure	(Ellard et al., 2015)
p. V617M	LAAHD (heterozygous with Fin_{major})	c.1849G>A	Missense point mutation in exon 13		(Nousiainen et al., 2008)
p. V617M	LAAHD (homozygous)	c.1849G>A	Missense point mutation in exon 13		(Ellard et al., 2015)

p.I684T	LAAHD (heterozygous with Fin_{major})	c. 2051T>C	Missense mutation in exon 16	Likely to damage the C-terminal domain structure	(Nousiainen et al., 2008)
p.I684T	LAAHD (homozygous)	c. 2051T>C	Missense mutation in exon 16	Likely to damage the C-terminal domain structure	(Paakkola et al., 2017)
p.[Ser693Phe]	LAAHD (homozygous)	c.2078c>t	Exon 16		(Said et al., 2017)
p.[Arg569His]	LAAHD (heterozygous with p.[Arg584Trp])	c.1706G>A	Exon 12		(Said et al., 2017)
p.[Arg584Trp]	LAAHD (heterozygous with p.[Arg569Trp])	c.1750C>T	Exon 12		(Said et al., 2017)
p.[D34_K107del]	LAAHD (heterozygous with p.[V238_Nfs*2])	c.100-7_100-3delTCTCT	In frame deletion of 74 amino acids in exon 2	Decrease the strength of splice acceptor site in exon 2/increase change of exon skipping	(Smith et al., 2017)

p.[V238_Nfs*2]	LAAHD (heterozygous with p.[D34_K107del])	c.1882-2A>G	Exclusion of exon 14	Abolish exon 14 splice acceptor site	(Smith et al., 2017)
p.[Arg603Leu]	LAAHD (heterozygous with p.[Gly666Val])	c.1808G>T	Nonconservative amino acid Substitution in exon 12/13	Disrupts IP6 binding	(Tan et al., 2017)
p.[Arg603Leu]	LAAHD (homozygous)	c.1808G>T	Nonconservative amino acid Substitution in exon 12/13	Disrupts IP_6 binding	(Cerino et al., 2020)
p.[Gly666Val]	LAAHD (heterozygous with p.[Arg603Leu])	c.1997G>T	Conservative amino acid substitution in exon 16	Disrupts C-terminal domain	(Tan et al., 2017)
p.S70X	ALS	c.209C>A	Nonsense mutation in exon 2		(Kaneb et al., 2015)
hGle1-IVS14-2A>C	ALS	c.1965-2A>C	Splice site mutation in intron 14	Novel 88 amino acid domain replaces hCG1 binding domain	(Kaneb et al., 2015)
p.R697C	ALS	c.2089C>T	Missense mutation in exon 16		(Kaneb et al., 2015)

LCCS1/LAAHD

Lethal congenital contracture syndrome 1 (LCCS1) is an autosomal recessive condition which occurs during gestation and leads to prenatal death. Paralysis of the foetus arises from a lack of skeletal muscle, anterior horn neurons and severe atrophy of the ventral spinal cord (Herva et al., 1985). Lethal arthrogryposis with anterior horn cell disease (LAAHD) is a milder, clinically similar phenotype with overlapping neuropathological and fetal akinesia deformation sequence phenotypes with LCCS1 thus indicating a common origin between these two disorders (Nousiainen et al., 2008; Vuopala et al., 1995). This has since established to be a result of mutations within the *GLE1* gene (Mäkelä-Bengs et al., 1998; Nousiainen et al., 2008) (**Table 1.2, Fig 1.3**).

Mutations linked to LCCS1

Sequence analysis of Finnish patient genomes allowed the specific mutation to be identified: a homozygous A to G substitution (c.432-10A>G) within the third intron of *GLE1,* referred to as Fin_{Major}. The resulting effect of this substitution is an illegitimate splice acceptor site, giving rise to an additional three amino acids (proline, phenylalanine and glutamine; T144_E145insPFQ) being inserted within the coiled coil domain of the gene (Nousiainen et al., 2008). This region of the protein is important as this allows self-association, giving rise to multimers and disk-like structures *in vitro*. Therefore, when this region is mutated, these structures cannot form correctly. *In vitro* (HeLa cells) the mutated Fin_{Major} protein disrupts mRNA export and displays less efficient nucleocytoplasmic shuttling than WT Gle1 (Folkmann et al., 2013), thought to be a result of this misfolding.

Fluorescently tagged GFP-Gle1 proteins have also demonstrated a significant loss of NPC localisation in the presence of LCCS1 and LAAHD causing mutations, thus indicating that defects during the mRNP remodelling process can have disease-causing effects (Folkmann et al., 2014).

Functional characterisation through animal models is an area of *GLE1* research which hasn't been very well explored. This is in part due to Gle1's essential role in RNA nuclear export,

therefore knock out models result in a phenotype causing prenatal death. Thus, meaning that behavioural and post-natal functional consequences of losing this gene cannot be studied.

An example of a *GLE1* knockdown model was shown using zebrafish containing both a *GLE1* insertional mutant and ASO knockdown. This model successfully demonstrated that disruption of *GLE1* function resulted in a comparable phenotype to that seen in LCCS1 patient foetuses. These embryos could be identified from their wildtype siblings from a phenotype which could be seen from two days post fertilisation (dpf). However, as in human LCCS1 patients, the maximum survival was 9 dpf (Jao et al., 2012; Seytanoglu et al., 2016). In contrast to the initial hypothesis proposed in earlier papers (Nousiainen et al., 2008), suggesting that the phenotype seen in LCCS1 is due to problems during motor neuron development from dysfunctional mRNA processing. This model demonstrates that a deficit of Gle1 impacts the rapidly dividing neuronal precursors rather than post mitotic motor neurons. It is proposed that this could instead be due to the required high rate of protein synthesis not being met because of insufficient Gle1 activity (Jao et al., 2012). This zebrafish model also displayed a reduced number of spinal cord motor neurons and abnormal arborisation of motor axons (Jao et al., 2012).

Utilisation of the zebrafish disease model has also allowed researchers to evaluate the health of other cell populations besides motor neurons. In other forms of LCCS, causative genes known to be implicated in the development and function of Schwann cells (Seytanoglu et al., 2016). For example; Erb-B2 Receptor Tyrosine Kinase 3 (ERBB3) – the gene known to cause LCCS2 (Narkis et al., 2007) - has been shown to be important in multiple animal models for the survival, migration and proliferation of Schwann cells (Lyons et al., 2005; Riethmacher et al., 1997). A gene expression study, using LCCS foetus spinal cord, demonstrated that there was oligodendrocyte dysfunction present in patient populations (Pakkasjärvi et al., 2006). Studies using zebrafish *gle1-/-* mutant embryos showed that defective Schwann cell development was observed despite normal expression of myelin basic protein, therefore indicating that Gle1 is required for the formation of myelinated Schwann Cells (Seytanoglu et al., 2016). They hypothesise that this could be due to Gle1's role in the export and/or transport of myelin basic protein mRNA within differentiating Schwann cells (Seytanoglu et al., 2016).

Mutations linked to LAAHD

Lethal arthrogryposis with anterior horn cell disease (LAAHD) is a disease which possesses a clinical phenotype similar, yet milder, to that of LCCS1 resulting in perinatal death (Vuopala et al., 1995).

To investigate whether these disorders have a common origin a mutation screen was carried out: of the 12 patients evaluated, all were heterozygous for the Fin_{Major} mutation responsible for LCCS1. Therefore, demonstrating that there was a clear genetic overlap between the two diseases with the greater severity of LCCS1 arising from homozygosity of the Fin_{Major} mutation affecting the coiled coil domain, while heterozygotes still possess one functional isoform of the gene which can act to compensate.

Alongside this Fin_{Major} mutation, half the patients possessed a missense point mutation within exon 13 (c.1849G>A, p.V617M) and the other a missense mutation within exon 16 (c.2051T>C, p.I684T) (Nousiainen et al., 2008). These mutations are hypothesised to have an effect on the C-terminal domain of Gle1 and therefore affects localisation to the NPC, especially in the case of the p.I684T mutation as this disrupts the Nup42 binding domain of the protein (Kendirgi et al., 2003; Nousiainen et al., 2008).

More recently, cases of individuals who are homozygous for these mutations have been identified. Two Finnish siblings who were clinically diagnosed with LAAHD after birth were found to have a homozygous p.I684T mutation in *GLE1,* postmortem studies carried out using patient fibroblasts showed significant mislocalisation of Gle1 protein with decreased amounts in the nucleus as compared to control fibroblasts. Therefore supporting the previous hypothesis that this localisation is important to correct function of Gle1 (Paakkola et al., 2017).

To date, genetic analysis of many LAAHD patients have been carried out in order to identify multiple other mutations occurring within the *GLE1* gene (**Table 1.2, Fig 1.3**) (Cerino et al., 2020; Said et al., 2017; Smith et al., 2017; Tan et al., 2017; Yates et al., 2020). In contrast to previous studies, these mutations are found to occur in a heterozygous manner with each other, but not however with Fin_{Major}. The conservation of the coiled coil domain of Gle1 in

both isoforms is thought to result in less severe symptoms and prolonged survival (Said et al., 2017). It has been suggested that this extended survival could be attributed to early medical care, including tracheostomy and ventilation (Tan et al., 2017), however, it raises the question whether these mutations should be grouped separately as a *GLE1*-associated disease (Said et al., 2017; Smith et al., 2017). Although this is disputed as so few patients have demonstrated this prolonged survival (one in each case) and so to categorise them separately may make clinical diagnoses more complicated (Paakkola et al., 2017). Four more patients displaying increased survival past 6 months have also recently been diagnosed (Yates et al., 2020).

In recent studies, two siblings have been observed with a homozygous c.1808G>T [p.(Arg603Leu)] mutation (Cerino et al., 2020), which had only previously been observed alongside the c.1997G>T [p.(Gly666Val)] mutation (Tan et al., 2017). This homozygous mutation gave rise to a clinically milder disease than seen previously and was predicted to be within the domain of the protein which is responsible for IP_6 binding. Therefore, perhaps giving some indication into which protein domains are more essential than others (Cerino et al., 2020).

Amyotrophic Lateral Sclerosis

The most prevalent form of MND is amyotrophic lateral sclerosis (ALS) which is characterised by a relentless neurodegeneration of upper and lower motor neurons, thus resulting in muscle weakness and atrophy causing progressive paralysis, eventually leading to death. The cause of ALS is still largely unknown with 90% of cases occurring on a sporadic basis. There is, however, some genetic overlap with between both sporadic and familial forms with mutations demonstrated to be present in multiple genes including: *superoxide dismutase (SOD1), TAR DNA-binding protein 43 (TDP-43), fused in sarcoma (FUS)* and *chromosome 9 open reading frame 72 (C9ORF72)* (Renton et al., 2013). A hallmark of these genetic mutations is the aggregates that they form (Ling and Song, 2010; Neumann et al., 2006). These proteins are all DNA/RNA-binding proteins implicated in numerous stages of RNA metabolism, indicating that dysfunction in the processing of RNA could be causative in the degeneration of motor neurons (Ling and Song, 2010). Mutations within the *C9orf72* gene have also been hypothesised to result in ALS due to RNA-mediated toxicity or haploinsufficiency (DeJesus-Hernandez et al., 2011).

Mutations linked to ALS

As Gle1 plays a crucial role in the export of RNA from the nucleus, genomic DNA from 933 ALS cases (both sporadic (760) and familial (173)) was evaluated for mutations within *GLE1* compared to 190 controls. Within this study there were three mutant variants in the ALS patients not seen in the control group: a nonsense mutation (c.209C>A, p.S70X) in a sporadic ALS patient, a splice site mutation (c.1965-2A>C, hGle1-IVS14-2A>C) in a familial ALS patient and a missense mutation (c.2089C>T, p.R697C) (Kaneb et al., 2015). Evaluation of these mutants indicates that the deleterious mutants cause haploinsufficiency through nonsense mediated decay of the mutated isoforms.

The missense mutant (p.R697C) results in a loss of the C-terminal Nup42 binding domain, this protein isoform therefore cannot localise to the NPC (Kaneb et al., 2015) which as described previously is known to give rise to detrimental effects.

Further study shows that the deletional mutant hGle1-IVS14-2A>C can, in fact, regulate stress granule formation and mRNA export and so mimics the functions of both hGle1A and hGle1B. This data suggests that this mutation in *GLE1* results in disease pathology as it disrupts hGle1 homeostasis and dynamics of mRNA export rather than through haploinsuffiency as previously hypothesised, demonstrating the importance of the Nup42 binding domain for regulation of hGle1 (Aditi et al., 2016). *GLE1*'s role in stress granule formation is also important here as it has been observed that hGle1-IVS14-2A>C forms cytoplasmic aggregates similar to those seen in neuronal inclusions of other ALS patients therefore potentially linking inclusion formation to translation misregulation (Aditi et al., 2016). Continued research in this area could prove vital to understanding the detrimental effects of these inclusions in multiple neurodegenerative diseases.

A more recent Chinese study has also identified seven *Gle1* loss of function mutations within a population of 628 ALS patients (mostly sporadic). These are mutations which either affect the coiled coil domain, disrupt the hCG binding domain resulting in mislocalisation or cause a total loss of the protein (Li et al., 2021).

1.3 Aims of PhD

There is a high degree of similarity between the pathological features seen in LCCS1 and SMA, with RNA processing being implicated as a potential disrupted pathway due to the roles Gle1 and SMN respectively play. With common roles in multiple aspects in this pathway, including transcriptional regulation, stress granule formation and transport of mRNA there is potential that Gle1 and SMN could be functioning together in a shared pathway. Uncovering more about the roles of these proteins and how they function would increase our understanding of these diseases, and therefore may help in the development of novel disease modifying treatments as seen with *Plastin3* and *ZPR1*.

Previous studies in the Azzouz laboratory have examined whether Gle1 can act as a neuroprotective factor in SMA cell models. Preliminary data suggests that lentiviral mediated overexpression of both Gle1a and Gle1b was sufficient to restore axonal growth defects present in SMN-deficient primary motor neurons *in vitro* (**Figure 1.5, Appendix 1, unpublished data**).

Advances published during this PhD demonstrated that knockdown of Gle1 *in vitro* can result in a reduction in *SMN1* mRNA within the cytoplasm (Okamura et al., 2018) and that there is evidence of a reduction in *GLE1* mRNA in SMA cases (Alrafiah et al., 2018a). However, there is limited information about what the relationship between Gle1 and SMN protein may be.

Our preliminary data combined with this research underlies the research aims that this thesis will set out to answer:

1. Do SMN and Gle1 interact together?
2. Is Gle1 expression changed in SMA cell models?
3. Can Gle1b act as a neuroprotective protein to rescue SMA disease phenotypes?

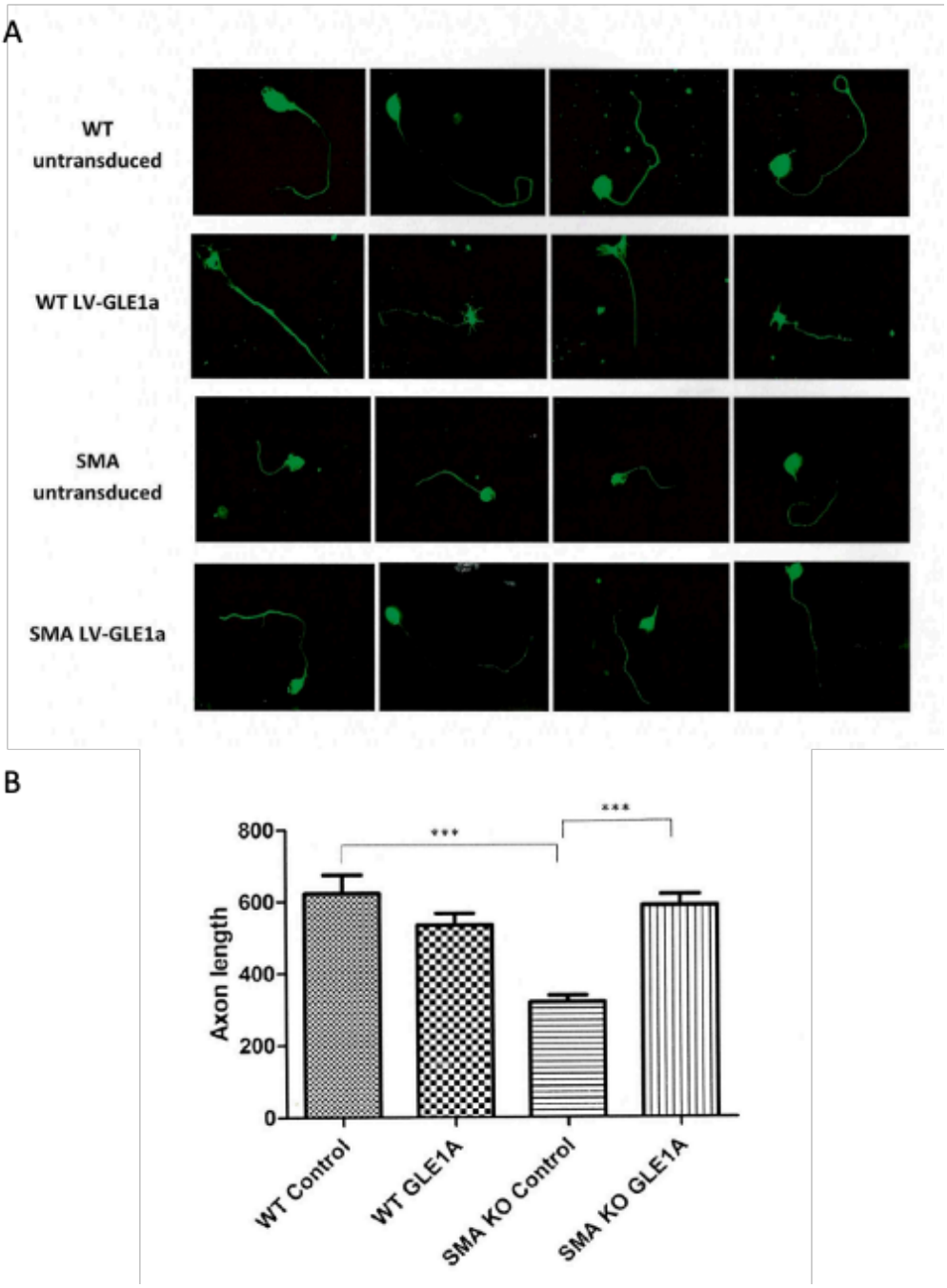


Figure 1.5: Overexpression of Gle1a in primary cultured motor neurons from SMN delta 7 mice show rescue of axonal growth defects. *(**A**) Representative images alpha-tubulin staining in primary motor neuronal cultures. Primary cultures taken from WT or SMN Δ7 mice 6 days post treatment with LV-Gle1a (MOI 50). (**B**) Axonal length (µm) of motor neurons. Data displayed as mean +/- SEM. 40 cells per repeat (One way ANOVA and Bonferroni test, ***, p<0.001, n=4)* ***Data taken from MSc thesis by Sameehan Uday Mahjani, 2010***

2. Materials and Methods

2.1 Materials

2.1.1 Expression Constructs

Table 2.1: Plasmid DNA. *A list of the plasmids which were generated and used for this project*

Plasmid	Obtained from
AAV Gle1b	In house (Cloned by Matthew Wyles)
AAV Gle1a	In house (Cloned by Matthew Wyles)
LV Gle1b	In house (Cloned by Matthew Wyles)
LV VOS	Modified by Dr Kurt de Vos at SITraN
LV FLAG Gle1b	In house (Chapter 3)
pcFLAG Gle1b	In house (Cloned by Dr Guillaume Hautbergue)
LV GFP	In house
pEGFP	Dr Eva Karyka
pEGFP-SMN	Dr Eva Karyka
pEGFP-Gle1b	In house (Chapter 3)
pCMVdelta8.2	Dr Nicole Deglon
pDM2.G	Dr Nicole Deglon
pRSV-Rev	Dr Nicole Deglon
pHelper	plamidfactory.com; PF1346-120619
pAAV 2-9 packaging plasmid	University of Pennsylvania; PF1347-120217

2.1.2 Viral Vectors

Table 2.2: Viral vectors. *A list of the viral vectors which were generated and used for this project*

Virus	Produced
LV GFP	In house (Chapter 3)

LV FLAG Gle1b	In house (Chapter 3)
AAV Gle1b	In house (Chapter 3)
LV Gle1b	In house
AAV mCherry	Vector builder
LV SMN (FL)	In house (Made by Dr Eva Karyka)

2.1.3 Antibodies

Table 2.3: Antibodies. *A list of the antibodies which were used during this project*

	Target Protein	Host Species	Concentration WB	Concentration ICC	Source
Primary Antibody	GLE1	Rabbit	-	1:1000	Abcam Ab96007
	GLE1	Rabbit	1:1000	-	ProteinTech 26466-1-AP
	FLAG M2	Mouse	1:1000	1:1000	Merck F1804
	GAPDH	Mouse	1:1000	-	Millipore CB1001-500
	Tubulin	Mouse	1:5000	-	Sigma T9026
	SSRP1	Mouse	1:1000	-	Abcam ab26212
	TUJ1	Chicken	1:1000	-	RayBiotech 119-15313
	AAV VP1, 2, 3	Rabbit	1:1000	-	ARP 03-61084
	GFP	Mouse	1:1000	-	Abcam ab1218
	SMN	Mouse	1:1000	1:1000	BD 610646

	γH2AX	Rabbit	-	1:1000	Cell Signalling 2577
	γH2AX	Mouse	-	1:1000	Merk Millipore 05-636
	RNA/DNA hybrids (S9.6)	Mouse	-	1:1000 **Methanol/ Acetone fixation	Kerafast ENH001
	Nucleolin	Mouse	-	1:1000	Abcam ab13541
	RNA Helicase A	Rabbit	-	1:1000	Abcam ab26271
	DNA-Pk	Mouse	1:200	-	ThermoFisher MA5-13238
	GOLPH3	Rabbit	1:1000	-	ProteinTech 19112-1-AP
	TGN46	Sheep	-	1:1000	BioRad AHP500GT
	GM130	Mouse	-	1:500	BD Biosciences 610823
	MAP2	Guinea Pig	-	1:1000	Synaptic Systems 188004
	Coilin	Rabbit	-	1:1000	ProteinTech 10976-1-AP
Secondary antibody	Anti-rabbit horse rabbit peroxidase (HRP)	Goat	1:3000	-	Thermo 32460

	Anti-mouse HRP	Goat	1:3000	-	Biorad 1706516
	Anti-chicken HRP	Goat	1:3000	-	Abcam ab97135
	Anti-mouse Alexa 488	Goat	-	1:1000	Invitrogen
	Anti-rabbit Alexa 488	Goat	-	1:1000	Invitrogen
	Anti-mouse Alexa 568	Goat	-	1:1000	Invitrogen
	Anti-rabbit Alexa 568	Goat	-	1:1000	Invitrogen
	Anti-Guinea Pig Alexa 647	Goat	-	1:1000	Invitrogen
	Hoeshct 33342	-	-	1:2000	Thermofischer

2.1.4 Primer Sequences

Table 2.4 Primers. *A list of the primers which were used during this project*

Primer Name	**Purpose**	**Sequence (5' to 3')**
LV FLAG-Gle1 F	Cloning	GCATCGGCTAGCGCCACCATGGACTACAAAGAC
LV FLAG-Gle1 R	Cloning	GATTCGACGCGTTCAGGAGCGCCAGAAGGAGGA
pEGFP-Gle1 insert F	Cloning	CTGATCATAATCAGCCATACCAC
pEGFP-Gle1 insert R	Cloning	GGGAATTCGAAGCTTGAGCTC

pEGFP vector F	Cloning	agctcaagcttcgaattcccATGCCGTCTGAGGGTCGC
pEGFP vector R	Cloning	gtatggctgattatgatcagTCAGGAGCGCCAGAAGGAG
Gle1 Seq 1	Sequencing	CTACGCAGTTCCGACAAAGG
Gle1 Seq 3	Sequencing	AGCAGAGAGTCAAGCTGAGG
Gle1 Seq 5	Sequencing	GGAGGGAATGGCTTTGGAAGAC
pCMV	Sequencing	ATAGCGGTTTGACTCACGGG
PGK	Sequencing	TCGCACACATTCCACATCCACC
WPRE F	Titration of LV	CCCGTACGGCTTTCGTTTTC
WPRE R	Titration of LV	CAAACACAGAGCACACCACG
18S F	Titration of LV	ATGGCCGTTCTTAGTTGGTG
18S R	Titration of LV	CGCTGAGCCAGTCAGTGTAG
HGH Poly A F	Titration of AAV	CTCCTGGCCCTGGAAGTTG
HGH Poly A F	Titration of AAV	ACTTGCCCCTTGCTCCATAC
Gle1 F	qPCR	CCCTACGCAGTTCCGACAAA
Gle1 R	qPCR	TTCCACTTTGATCCCCCGTG
GAPDH F	qPCR	CAACTTTGGTATCGTGGAAGGAC
GAPDH R	qPCR	ACAGTCTTCTGGGTGGCAGTG
Fos F	qPCR	GCCTCTCTTACTACCACTCACC
Fos R	qPCR	AGATGGCAGTGACCGTGGGAAT
ZNPF91-CNTF F	qPCR	GCCATCGTGATACAGAGAACACC
ZNPF91-CNTF R	qPCR	CTAATGCCACCTGGAGACTGATG
RPL13a7 F	qPCR	CCTCAAGGTGTTTGACCGCATC
RPL13a7 R	qPCR	TACTTCCAGCCAACCTCGTGAG

2.1.5 Cells used in project

Table 2.5: Immortalised Cell Lines. *A list of the immortalised cell lines which were used during this project*

Cell	Obtained from
HEK293T	ATCC
Hela	ATCC

Table 2.6: Fibroblast Cells. *A list of the fibroblast cells which were used during this project*

Cell Reference	Age	Sex	Genotype	Obtained from
GM08680	5 months	Male	Apparently Healthy	Coriell Institute
GM00498	3 Years	Male	Apparently Healthy	Coriell Institute
GM09677	2 Year	Male	SMA Type I	Coriell Institute
GM00232	7 months	Male	SMA Type I	Coriell Institute

Table 2.7: iPSC/Neural Progenitor Cells/Motor Neurons. *A list of the iPSC/Neural Progenitor Cells/Motor Neurons which were used during this project*

Cell Reference	**Origin**	Age	Sex	Genotype	Obtained from
Miff1	Foreskin fibroblast	Foetal	M	Apparently Healthy	University of Sheffield
Cs14	Fibroblast	30-35	F	Apparently Healthy	Ceders-Senai
GM	Fibroblast	55	M	Apparently Healthy	Coriell
SMA 77	Fibroblast	2 Years	M	SMA Type I	Coriell
SMA 84	Fibroblast	6 months	F	SMA Type I	Ceders-Senai

2.1.6 Equipment List

***Table 2.8: Equipment* List.** *A list of the Equipment which was used during this project*

Equipment	Manufacturer
NanoDrop 1000	LabTech
G;Box gel imaging system	SynGene
Pherastar Plate reader	BMG LabTech
PCR machine	G-Storm
CFX96 RealTime System C1000 Touch Thermal Cycler	BIO-RAD
Confocal Microscope	Leica SP5 microscope system
OPERA PHENiX High Throughput imaging system	Perkin-Elmer
Optima L-100K Ultracentrifuge	Beckmann Coulter

2.2 Cell Culture

All cells were grown in an incubator at 37 ^{O}C with 95 % humidified air atmosphere and 5 % CO_2. Experiments were plated by cell counting using a haemocytometer and Trypan Blue.

2.2.1 Immortalised Cell Culture and Maintenance

Human embryonic kidney (HEK293T) cells and HeLa cells were cultured in Dulbecco's Minimum Essential media (DMEM) (Lonza BE12-741F) containing 10 % heat inactivated foetal bovine serum (FBS) (Life Science Production 5-001A-H1-BR) and 1 % Penicillin Streptomycin (P/S) (Lonza DE17-603E). To maintain cell growth, cells were split twice weekly in a 1:10 dilution using trypsin (Lonza BE02-007E).

2.2.2 Fibroblast Cell Culture and Maintenance

Human fibroblast cell lines from Coriell were cultured in DMEM containing 10 % heat inactivated FBS, 1 % P/S, and 50 ug/ml uridine. To maintain cell growth, cells were split twice weekly in a 1:3 dilution using trypsin. A wash in 0.5M Ethylenediaminetetraacetic acid (EDTA)

in Hanks buffered saline solution (HBSS) was performed pre-trypsinisation to loosen tight junctions and allow easier passaging of the cells to minimise stress.

2.2.3 Cortical Neuron Preparation and Culture

Primary cortical neuron cultures were carried out by Dr Christopher Webster.

For primary cortical neuron culture, dissection of E16 wild type (WT) embryos from pregnant females was carried out. Tissue was removed from the cortex and digested in 0.25% trypsin in HBSS without calcium or magnesium (GIBCO) at 37^{O}C for 15 minutes. Triturating medium was added to the suspension and cells were then manually dissociated, using three fire-burnt Pasteur pipettes which has progressively smaller openings. Dissociated cortical neurons were then plated at a density of 4,500,000 cells per plate with a poly-D-lysine (Sigma) coating. Neurobasal medium (Life Technologies) supplemented with 2% B27 (Life Technologies), 0.5 mM GlutaMax (Life Technologies) and 100 U/ml penicillin with 100 µg/ml streptomycin (Lonza) was used to maintain cells.

2.2.4 iPSC Cell Culture

iPSC Cell Culture was carried out by Dr Cleide dos Santos Souza

iPSC cells derived from patient and control fibroblasts were converted prior to this project starting. For the purposes of this project these cells were differentiated into neural progenitor cells following the protocol as published in (Du et al., 2015). Briefly, iPSCs were cultured onto 6 well plates coated in matrigel, in a basal media composed of 48% knockout DMEM/F12 (Gibco), 48% Neurobasal media (Gibco), 0.5% B-27 (Gibco), 1% GlutaMax (Gibco), 1% Penicillin/Streptomycin (Lonza), 0.5% N-2 supplement (Gibco). From days one to six, 2 µM DMH-1 (Tocris) and 2 µM SB 431542 (Tocris) were also added to the culture medium, this was refreshed by 50% daily. From days seven to twelve, 1 µM CHIR 99021 (Tocris), 2 µM DMH-1, 2µM SB 431542, 0.1 µM Retanoic acid (STEMCELL Technologies) and 0.5 µM PUR (Tocris) were added to the basal media and refreshed 50% daily.

2.2.5 Neural Progenitor Cell Culture and Differentiation into Motor Neurons

Neural Progenitor cells (NP Cell) were fully formed at day 12 of iPSC conversion. At this point, cells could be maintained at NP cell stage and expanded in basal media containing 3 µM CHIR 99021, 2 µM DMH-1, 2 µM SB 431542, 0.1 µM Retinoic Acid, 0.5 µM PUR and 0.5 µM VPA (Merck). Cells were split when they reached 80% confluency through dissociation with Accutase (Thermo) for 8 minutes at 37 ^{O}C. Cells were spun at 400xg for 8 minutes and resuspended in NPC expansion media with 10 µM Rock Inhibitor (1:1000) (Tocris). These were then split at a ratio of 1:4 and plated onto 6 well plates pre-coated in 10 mg/µl polyornithine (1:1000) (Sigma) (Overnight at room temperature) and matrigel (one hour at room temperature). 24 hours after plating cells were media changed into NP cell expansion media without Rock inhibitor. Cells were then media changed every second day and split when needed.

For differentiation into motor neurons, cells at a density of around 70% - 80% confluency were fully media changed into day 13 to 18 media: basal media containing 0.5 µM retinoic acid and 0.1 µM PUR. This was changed 50% daily. On day 19, cells were changed into day 19-28 media: basal media containing 0.5 µM Retinoic acid, 0.1 µM PUR, 0.1 µM compound E (Tocris), 10 ng/ml BDNF (Thermo), 10 ng/ml CNTF (Thermo), and 10 ng/ml IGF-1 (Thermo). Cells were dissociated as previously on day 21 and plated onto pre-coated plates as needed for experiments at a density of 10,000 per well (96 well plate), 250,000 per well (24 well plate), 2,000,000 per well (6 well plate). These cells were maintained in day 19 to 28 media until processing for experiments at day 33 with a 50% media change every second day.

Cells transduced with virus were treated on day 28 in minimal media volume. Media was topped up six hours post transduction and changed 50% every two days.

Generation of iPSC-derived motor neurons was performed using established protocols which have been validated to produce a highly enriched (>90%) population of mature motor neurons (Du et al., 2015), with staining for neuronal markers MAP2 and CHAT. This has been extensively validated in house by Dr Cleide dos Santos Souza, showing motor neuron populations of ~90% positive for MAP2, Tuj1, CHAT and NeuN.

2.2.6 Transfection

Transfections of DNA into the cells were carried out following the ratio of 1 DNA to 3 PEI (**Table 2.9**). For each transfection condition the corresponding amount of DNA and Polyethylenimine (PEI) was added to separate eppendorfs containing equal volumes of serum free media (DMEM without FBS added). These were left to rest for 5 minutes before combining the contents of the tubes, after immediately vortexing for 10 seconds the solution was then left for 15 minutes before adding dropwise to each well. Cells were left for 48 hours prior to processing.

Table 2.9: Different ratios of PEI to DNA used in cell culture plates

Plate	DNA (µg)	PEI (µg)	Serum free media (SFM) (µl)
6 well	2	6	600
12 well	1	3	300
24 well	0.5	1.5	150

2.3 Biochemical Methods

2.3.1 Whole cell lysis

For protein extraction, cells were plated in six well plates. Cells were washed once in ice cold Phosphate Buffered Saline (PBS) before being scraped in 500 µl PBS. Cells were pelleted at 400xg for 5 mins at 4 ^{O}C. 200 µl radioimmunoprecipitation (RIPA) lysis buffer (Tris-HCl 50mM, NP-40 1%, Na-deoxycholate 0.5%, EDTA 2mM, NaCl 150mM, SDS 1%) was used to extract proteins with 1X protease inhibitor cocktail (PIC) (Sigma). The lysate was left on ice for 30 minutes before being spun at 13,000*xg* for 10 minutes at 4 $^{\circ}$C. The supernatant was collected and stored at -20 ^{O}C prior to protein quantification.

2.3.2 Bicinchronic acid (BCA) Assay for protein quantification

Protein concentration was measured using the bicinchoninic acid assay (BCA assay). This technique involved setting up a range of bovine serum albumin (BSA) (Pierce Thermoscientific) standard concentrations (**Table 2.10**) to create a linear regression fit in which the unknown protein concentrations can be assessed.

Table 2.10: Volumes of BSA and protein lysis buffer needed to create BSA standards

BSA (mg/ml)	Volume BSA (µl)	Volume RIPA buffer (µl)
2.0	50	0
1.5	37.5	12.5
1.0	25	25
0.75	18.8	31.3
0.5	12.5	37.5
0.25	6.3	43.5
0.125	3.1	46.9
0	0	50

Two 20 µl replicates of each of these standards were pipetted into a 96 well plate. A suitable dilution of each protein sample extracted previously was also pipetted in duplicate into the same 96 well plate. The Pierce BCA reagents were combined in a volume sufficient to add 200 µl of solution to each well of both standards and samples; this was done in a 50:1 ratio of Reagent A to Reagent B. The plate was then covered to protect the pigments from light and placed at 37 ^{O}C for 30 minutes. The PheraStar plate reader was then used to read the absorbance levels of the plate at 560 nm.

A range of protein concentrations ranging from 10 µg to 40 µg were used to load western blots depending on the experiment.

Samples were diluted using milliQ H_2O and 4x lamelli buffer (10% Glycerol, 60 mM Tris/HCl (pH 6.8), 2% SDS, 0.01% bromophenol blue, 1.2% beta-mercaptoethanol). Samples should then be denatured by heating to 95 ^{O}C for 5 minutes.

2.3.3 Nuclear/cytoplasmic fractionation

One 6 well plate was required per condition. On the day of harvest, each well was washed with 1x PBS. For the total fractions one well was scraped into 250 µl 1x Reporter Lysis Buffer (Promega) and left on ice to lyse for 10 minutes. This was then spun at 13,000*xg*/4 ^{O}C/5 minutes. 20 µl was retained for western blot validation of the fractionation and 750 µl of Trizol (Ambion) was added to the remainder for RNA extraction.

Protease and Ribosafe RNase inhibitors (Bioline) were added to both the hypotonic (10mM HEPES (pH7.9), 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT) and reporter lysis buffers in the ratios of 1:50 and 1:1000 respectively.

Cytoplasmic fractionation:

1 ml 1x PBS was added to each remaining well and a large cut P1000 tip was used to knock off cells gently. This was then spun at 400*xg* for 4 minutes to pellet the cells. The pellet was then quickly washed in 500 µl 1x Hypotonic lysis buffer. Without adding any physical force, two sizes of a cut P1000 tip were used to resuspend the cell pellet in 400 µl hypotonic lysis buffer. The cells were left to lyse on ice for 10 minutes after which they were spun at 1500*xg*/3

minutes/4^{O}C. The supernatant was then transferred to a fresh tube and the pellet retained for the nuclear fractions. The supernatant was spun again at 3500*xg*/8 minutes/4^{O}C, transferred to a fresh tube and spun once more at 17,000*xg*/1 minute/4^{O}C. This final supernatant was the cytoplasmic fraction and 20 µl was kept aside to validate fractionation by western blot, while 750 µl of trizol was added to the remainder (250 µl). This was then stored on ice with the total fractions.

Nuclear fractionation:

The pellet formed from the first cytoplasmic spin was washed four times with hypotonic lysis buffer, resuspended each time with a cut P1000 tip and spun at 1500*xg*/3 minutes/4^{O}C. After the final wash, 400 µl of 1x reporter lysis buffer was added and the pellet fully resuspended by drawing the cells through a 25G needle ten times. This was then lysed on ice for 10 minutes before being spun at 13,000*xg*/5 minutes/4^{O}C. As with the cytoplasmic and total fractions 25 µl of supernatant was taken for fractionation validation by western blot and 750 µl Trizol was added to the remaining supernatant (250 µl).

RNA extraction was carried out as stated later in the methods chapter (**Section 2.3.6**).

For western blot analysis 15 µl of each sample was loaded with 5.8 µl 4x Laemelli buffer and 3 µl of reporter buffer (1x reporter for total and nuclear, 5x reporter for cytoplasmic).

Samples were boiled for 5 minutes at 95 ^{O}C before loading to a polyacrylamide gel for western blotting.

2.3.4 Western blotting

Precast gradient gels (4%-20%) (Biorad) were soaked in 1x running buffer for 10 minutes prior to gel loading. A BioRad western blotting kit was used to run the sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).

2.5 µl of SDS PAGE – prism ultra-protein ladder (ab116028) was used where necessary, appropriate volumes of sample loaded in laemelli buffer was added to each well. Empty wells

were filled with 1X laemelli buffer to ensure even running of the gel. Gels were run at 150V for 90 minutes in 1x running buffer (25 mM Tris, 3.5 mM SDS, 20 mM glycine).

Transfer of proteins from the gel to a polyvinylidene difluoride (PVDF) membrane, (preactivated in 100% methanol) was done in a transfer buffer (47.9 mM Tris, 38.6 mM Glycine, 20% Methanol) run at 250 mA for 90 minutes on ice.

The quality of the transfer and protein loading was assessed through staining with Ponceau dye (0.1% Ponceau S, 5% acetic acid). Following this, the membrane was washed with ddH_2O several times before being placed into 5% milk (Marvel)/Tris Buffered Saline with Tween 20 (TBST) blocking solution (20 mM Tris, 137 mM NaCl, 0.2% Tween 20, pH 7.6) on a shaker for one hour at room temperature.

Incubation of membranes with primary antibody occurred overnight at 4 ^{O}C on a haematology mixer. Antibodies were diluted to the correct dilution (**Table 2.3**) in 5% milk in TBST. Membranes were washed three times in TBST for 10 minutes each at room temperature on a haematology mixer. The membranes were then incubated with secondary antibodies conjugated to horseradish peroxidase (HRP) (**Table 2.3**) for an hour at room temperature again on a haematology mixer. Following this the membranes were then washed again in TBST three times for 10 minutes.

The membranes were developed using the ECL chemiluminescence detection kit (Pierce ECL, ThermoFisher Scientific). Equal parts of Reagents A and B were mixed, and membranes were incubated for 1 minute prior to imaging. Imaging was carried out using a G box and Gene Snap program (Syngene).

Blots were quantified using GeneTools (Syngene). Data was normalised by loading control and plotted relative to the average of the control in most cases.

2.3.5 Immunoprecipitation

GFP-fusion proteins were generated for GFP-tagged immunoprecipitation with the Chromotek GFP-Trap Magnetic Agarose Kit. Protocol carried out as described by manufacturer.

For each condition one 10cm dish was used. Cells were scraped in 1x PBS and pelleted at 400*xg*/4 minutes/4^{O}C. The pellet was resuspended in 200 µl of lysis buffer (Chromotek) (+PIC) and left for 30 minutes on ice, pipetting up and down every 10 minutes. Each condition was carried out both with and without RNase A (10 µM) (Sigma). The cell lysate was then centrifuged at 13,000*xg*/10 minutes/4^{O}C. The supernatant was transferred to a fresh tube and 300 µl of dilution buffer (Chromotek) (+PIC) was added. 50 µl of this diluted lysate was retained for western blot confirmation of the input fraction.

The magnetic agarose beads were equilibrated in dilution buffer prior to use. This was done by gentle resuspension of the beads using a cut P200 tip. 25 µl of bead slurry was then added to 500 µl dilution buffer. The beads were washed three times in 500 µl dilution buffer, each time separating the beads using a magnetic rack. The protein lysate was then added to the equilibrated beads and rotated end over end overnight at 4 ^{O}C.

Beads were removed with a magnet. Remaining supernatant was discarded and the beads were resuspended in 500 µl wash buffer. Beads were separated with a magnet at least three times, checking to see that the supernatant was clear. The beads were transferred to a fresh tube and supernatant was removed.

Immunocomplexes from the beads were then dissociated through resuspension in 100 µl 2x laemelli buffer and boiling at 95^{O}C for 5 minutes. The beads were separated using a magnet and the supernatant was analysed through SDS-page and immunoblotting to check for protein binding.

2.3.6 RNA extraction

RNA extraction for this project was carried out in two ways depending on the experiment.

Whole cell RNA extraction

For whole cell RNA extraction, the RNeasy kit from QIAGEN was used. This was carried out following the manufacturers protocol.

Cells were grown in a 6 well plate, up to 3 wells per condition were used. Cells were scraped and pelleted at 400*xg*/4 minutes/4^{O}C in ice cold 1x PBS. 350 µl lysis Buffer RLT was added to the pellet. An equal volume of 70% ethanol was then added to the lysate, mixed well by pipetting. This sample was then transferred to a RNeasy spin column housed in a 2 ml collection tube and spun at 8000*xg*/15 seconds. All flow-through was discarded,

The column was then washed with 700 µl Buffer RW1, again centrifuging at 8000xg/15 seconds and discarding the flow through. This step is repeated with 500 µl Buffer RPE and then again adding a further 500 µl Buffer RPE for 2 minutes at 8000*xg*. The column was then placed in a fresh 1.5 ml collection tube and 30-50 µl RNase Free water was added directly to the spin column membrane. The RNA was eluted by spinning at 8000*xg*/1 minute and the sample was nanodropped to measure RNA concentration and purity.

Nuclear/Cytoplasmic Fraction RNA extraction

The second method of RNA extraction follows the fractionation experiments. This method uses a Phenol-Chloroform gradient. After addition of 3 parts Trizol to 1 part lysate in the previous protocol, the sample was left at room temperature for 10 minutes. A volume of chloroform equal to one fifth of the total volume was added to each sample and the tubes shaken vigorously for 15 seconds. This was then left for a further 10 minutes at room temperature before being spun at 12,000*xg*/10 minutes/4^{O}C. This formed a gradient of the RNA, precipitate, and phenol waste, carefully the top layer containing the RNA was removed and added to a fresh tube. To this 1 µl of Glycogen, one tenth of the total volume 3M NaAc and equal volume isopropanol was added and the mixture left at -20^{O}C overnight.

The following day the samples were spun at 13,000*xg*/20 minutes/4^{O}C. The supernatant was discarded, and the pellet washed with 500 µl of 70% Ethanol (in DEPC water), after which the samples were spun for a further 10 minutes at 13,000*xg*/4^{O}C. Following removal of the ethanol supernatant, the pellets were left to airdry for 10 minutes and were then resuspended in 25 µl DEPC water. To prepare the samples for cDNA generation the RNA was treated with DNase through addition 2.5 µl 10X DNase buffer and 1 µl DNase. This reaction mix was then incubated at 37 ^{O}C for 20 minutes. The DNase was inactivated through addition

of 1 µl 25 mM EDTA and heating at 75 ^{O}C for 10 minutes. The RNA's quality and concentration were then assessed by nanodrop.

2.3.6 Reverse transcriptase PCR (RT-PCR) generation of cDNA

cDNA was generated from high quality RNA samples prior to analysis by qPCR. This was done with the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) following the manufacturers protocol: Depending on the experiment, a range of RNA concentrations from 500 ng to 2000 ng was converted into cDNA. Per reaction the following volumes of reagents were combined as in **Table 2.11.**

Table 2.11: Reaction mix for stage one of cDNA synthesis

Reagent	Volume (µl)
RNA	500 ng – 2000 ng
50 ng/ µl Random Hexamers	1
10 mM dNTP mix	1
DEPC-treated H_20	Up to 10

This reaction was then incubated at 65^{O}C for 5 minutes before leaving on ice for at least 1 minute.

A cDNA synthesis master mix (**Table 2.12**) was then made up to add to each RNA/primer sample. A control reaction for each sample was also made which contained no reverse transcriptase enzyme. 10 µl of this master mix was added to each sample and the reaction was run on a thermocycler according to the parameters in **Table 2.13**.

Table 2.12: Master mix for stage two of cDNA synthesis

Reagent	Volume per sample (µl)
10X RT buffer	2
25 mM $MgCl_2$	4
0.1 M DTT	2
RNase OUT (40U/ µl)	1
SuperScript III RT (200 U/ µl)	1

Table 2.13: Thermocycler conditions for cDNA synthesis

Temperature Step (^{O}C)	Time (minutes)
25	10
50	50
85	5

Following incubation, 1 µl of RNase H was added to each sample and further incubated for 20 minutes at 37 ^{O}C. cDNA was then stored at -20 ^{O}C for future use.

2.3.7 Quantitative PCR (qPCR)

qPCR was used to analyse the levels of certain mRNA transcripts present in the cDNA samples previously generated. Each sample was run in triplicate to ensure technical accuracy.

Primer concentrations were optimised through standard PCR reactions in the following reaction mix and thermocycler conditions as seen in **Table 2.14**.

Table 2.14: Reaction mix for Primer Optimisation PCR reaction

Reagent	Volume (µl)
5x FIREpol master mix (Solis Biodyne)	2
Template cDNA (2000 ng)	1
Forward primer	1
Reverse primer	1
DEPC H_20	5

Primer concentrations were diluted from a 100 µM stock to make a series of 1/10, 1/20, 1/40, 1/80 dilutions. Therefore giving 1 µM, 0.5 µM, 0.25 µM, and 0.125 µM in the final 10 µl reaction mix. These reactions were run in a thermocycler according to the parameters in **Table 2.15**.

Table 2.15: Thermocycler conditions for primer optimisation

Temperature	Time	Cycle
110	Heated lid	
95	5 mins	
95	30 seconds	Cycle x35
60	1 minute	
72	1 minute	
10	Store	

PCR products were then run on a 2.5% agarose gel as descripted later in this chapter to check for clear products of the correct size and no primer dimers.

Once the primers were checked and the appropriate dilution was selected for optimal amplification, the standard curve of each primer was checked to ensure that the amplification of the product is linear across a whole range of template concentrations. Thus, ensuring that even low levels of the product can be detected accurately. Template cDNA was serially diluted from a 2000 ng stock 5-fold- six times, therefore giving a range of dilutions from 2000 ng to 0.128 ng assuming 100% conversion during cDNA production. The qPCR reaction mix was prepared as outlined in **Table 2.16**. Each sample was run in duplicate and with a no template control (NTC) for each reaction. These reactions were run on the thermocycler program outlined in Table **2.17.**

Table 2.16: Reaction mix for cDNA Optimisation qPCR reaction

Reagent	Volume (µl)
2X Brilliant III Syber green (Aglient)	5
Forward Primer	1
Reverse Primer	1
DEPC H_20	2
cDNA/NTC	1, added per well

Table 2.17: Thermocycler conditions for cDNA Optimisation

Step	Temperature	Time	Cycles
PCR initial activation	95	10 mins	
Denaturation	95	30 sec	40 cycles
Annealing	60	30 sec	
Extension	72	1 min	
Melt curve analysis	95	1 min	
Melt curve analysis	60	5 sec	
Melt curve analysis	Ramp	(+0.5^{O}C/5s)	
Melt curve analysis	95	-	
Storage	10	Hold	

qPCR results were checked using Bio-Rad CFX Maestro program. Melt curves were checked for clean peaks, single products and the standard curve must show an R^2 of >0.9 with an efficiency between 90 and 110%. Template concentration chosen which gives a CT value between 20 and 30. This optimal concentration was then used to run the experimental qPCR where each sample was run in triplicate with appropriate non template controls and non-reverse transcriptase controls in place. Reaction mixtures and thermocycler steps remain identical to those used during optimisation.

Relative gene expression values were quantified using the ΔΔCt method (Livak and Schmittgen, 2001). GAPDH was used as a housekeeping gene to normalise cDNA concentration levels for each target gene. Levels of gene expression were then plotted relative to the average of the control.

2.3 Molecular Biology Techniques

2.3.1 Cloning of expression constructs: Gibson assembly cloning

Primers were designed complementary to the plasmid and fragment of interest with overlaps of ~30 base pairs on either end. These primers were then used to generate DNA segments through PCR using the following thermocycler conditions (**Table 2.18**).

Table 2.18: Thermocycler conditions for amplification

Temperature	Time	Cycle
110	Heated lid	
95	5 mins	
95	30 seconds	Cycle x35
60	1 minute	
72	1 minute	
10	Store	

2.3.3 Agarose gel electrophoresis

PCR products and DNA digestions were used for separation of linear DNA fragments through agarose gel electrophoresis. 1 % agarose gels were prepared by adding 1 g of agarose powder into 100 ml 1x Tris/Acetate/EDTA (TAE) buffer (40 mM Tris, 20 mM NaOAc, 1 mM EDTA (pH 8.0). This was dissolved through applying heat to the solution, upon generation of a clear solution the mixture was cooled, and 100 ng/ml ethidium bromide was added. This was then poured into a gel tray containing 20-well combs and left to set.

Upon loading, the gel was transferred into a tank of 1x TAE buffer and samples (containing 6x Loading buffer (NEB)) were added to each well. A ladder of the appropriate molecular weight was loaded alongside the DNA samples (Bioline: DNA hyperladders). Gels were left to run at 120 V for 30 minutes before imaging with a GENI UV light imaging system (Syngene).

2.3.4 DNA gel extraction and Purification

DNA bands of the correct size were visualised using a UV trans illuminator. These were carefully cut from the gel using a sharp blade and transferred into an Eppendorf tube. This DNA was then extracted using the Wizard® SV Gel and PCR Cleanup kit (Promega) following the manufacturer's protocol.

2.3.5 Gibson Assembly

Fragments were then assembled in a reaction mix, using the Gibson Assembly Master mix (2x) (NEB), in the quantities suggested by the manufacturer (NEB). Samples were incubated in a thermocycler at 50 ^{O}C for 15 minutes and then stored on ice or at -20 ^{O}C.

2.3.6 Transformation of NEB stable 3 competent cells

Working aseptically, 10 µl of competent cells (NEB) were left to thaw on ice for five minutes before carefully adding 1 µl of the plasmid of interest. This was then incubated on ice for 30 minutes before being placed for 30 seconds into a water bath heated to 37 ^{O}C to heat shock the cells. Cells were immediately placed back on ice for a further 5 minutes prior to 1 ml of outgrowth media (NEB) being added. This mixture was then left at 37 ^{O}C with agitation for 45 minutes, meanwhile a plate of the corresponding bacterial resistance (carbenicillin in this case) was also left at 37 ^{O}C to warm up. 200 µl of the grown bacteria is then added to the plate and spread using silica beads. This is then left at 37 ^{O}C overnight for the colonies to grow.

2.3.7 Plasmid growth

Following the transformation, individual colonies were picked using a p200 pipette tip and added to a tube containing an appropriate volume of LB broth (**Table 2.19**) with a 1:1000 dilution of carbenicillin or kanamycin (depending on the plasmids resistance gene). This was again left overnight at 37 ^{O}C to grow.

Table 2.19: Volumes of LB broth required per each QIAgen DNA kit

Amount of DNA required	DNA Prep	Volume of LB broth (ml)
<20 ug	MiniPrep	5
Up to 250 ug	MidiPrep	50
Up to 2.5 mg	MegaPrep	500

2.3.8 DNA preparation

MiniPrep/MidiPrep/MegaPrep

On the third day, the plasmid was purified using a QIAgen kit. This protocol was carried out using the instructions provided with this kit. For Midi and MegaPreps a vacuum was used instead of spin columns.

DNA concentration was quantified using a Nanodrop according to the manufacturer's protocol.

2.3.9 Restriction digest

Plasmids were digested to check for correct insertions. 1 ug of DNA was digested as standard using restriction enzymes determined by the plasmid's unique restriction sites. These were located using A Plasmid Editor software (ApE). Typically, a 10 µl digestion mixture was made up, using appropriate buffers for the enzymes selected.

2.3.10 DNA Sequencing

Sanger sequencing of DNA was performed by the DNA Sequencing service at the University of Sheffield Core Genomics Facility.

2.3.11 Genomic DNA extraction

Genomic DNA was extracted using the GenElute Mammalian Genomic DNA Miniprep kit (Sigma). The protocols for cultured cell preparation were followed as stated in the manufacturer's handbook.

2.4 Viral production

2.4.1 Lentiviral production

HEK293T cells were used for lentiviral production. Twenty 10cm dishes were seeded with $3x10^6$ cells per dish. The following day the cells were transfected with a calcium chloride transfection mix containing the four lentiviral component plasmids (**Table 2.20**), calcium chloride (0.5M $CaCl_2$) and 2x HBS.

Table 2.20: Plasmids required for LV transfection

Plasmid	Role	Concentration (µg)
pCMV.delta8.2	Packaging	260
pRSV-Rev	Packaging	60
pMD.G	Envelope	75
SIN-W-PGK	Transgene	260

5 ml of 0.5M $CaCl_2$ was added to 5ml of the DNA mixture dropwise. 10 ml of HBS (2X) was then gently aerated (using a 5ml strippette) while adding the DNA/$CaCl_2$ mixture dropwise with a P1000. The mixture was then left for 10 minutes to allow a fine white precipitate to form. 1 ml of this mixture was then added to each 10cm dish dropwise, gently swirling the plate to ensure the whole surface is covered. The plates were then incubated for 6 hours before a full media change was performed, and the plates incubated for a further 48 hours.

At this point the virus was harvested in a tissue culture hood. All the media was combined into a T175 tissue culture flask and filtered into a fresh T175 using a 0.45 µm filter and a 50 ml syringe. This media was then split equally across 6 Beckman tubes, using a balance to ensure equal loading of the paired tubes in the SW28 Beckman hanging rota. This was spun at 19,000*xrpm*/90 minutes/4 ^{O}C in a Beckman Ultracentrifuge.

Supernatant was discarded into virkon, removing as much of the media as possible without disturbing the viral pellet. Each pellet was resuspended with 280 µl of 1% BSA in PBS, left for 1 hour on ice before combining into one tube. This was thoroughly mixed to ensure that virus is evenly distributed and then aliquoted for long term storage at -80 ^{O}C.

2.4.2 Lentiviral Titration and Validation

HEK293T cells were transduced with set volumes of a control virus whose titre is known by Flourescence activated cell sorting (FACS) (LV-GFP produced by Nelly Berrueta Ramirez and Dr Eva Karyka; titre of 4.71 x10^8 TU/ml) and the newly made virus. Media was changed 6 hours after transduction. 72 hours post transduction, cells were harvested, and genomic DNA was extracted (as described in **Section 2.3.11**). Titre of this virus was analysed by qPCR using a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) primer and an 18S control primer for normalisation. The qPCR mixture was made as described in **Table 2.21** and run in a thermocycler according to the parameters in **table 2.22**.

Table 2.21: Reaction mix for LV titration by qPCR

Reagent	Volume (µl)
SYBER green master mix (QuantiFast)	5
Forward primer 5 µM	1
Reverse primer 5 µM	1
Nuclease free water	2
DNA (10 ug/µl)	1

Table 2.22: Thermocycler conditions for LV titration

Step	Temperature	Time	Cycles
PCR initial activation	95	5 mins	
Denaturation	95	10 sec	40 cycles
Annealing	60	30 secs	
Melt curve analysis	65	5 sec	
Melt curve analysis	Ramp	(+0.5^{O}C/5s)	
Melt curve analysis	95	-	
Storage	10	Hold	

Relative gene expression values were quantified using the ΔΔCt method (Livak & Schmittgen, 2001) and the titre worked out compared to that of the known virus (LV-GFP).

Lentivirus expression and optimal multiplicity of infection (MOI) dosage was confirmed through western blot analysis and immunocytochemistry.

2.4.3 Small scale AAV production

A confluent T175 flask of HEK293T cells was split 1:3 into two fresh flasks, these were transfected 24 hours later. 3 ml of serum free media was put into two tubes labelled DNA and PEI respectively. 40 µg pHelper (plamidfactory.com; PF1346-120619), 20 µg 2-9 packaging plasmid (University of Pennsylvania; PF1347-120217) and 20 µg of the plasmid of interest was added to the DNA tube while 240 µl of PEI (1 mg/ml) was added to the PEI tube. After thorough mixing these were left to incubate for 5 minutes at room temperature, at which point the contents of the DNA tube were transferred to the PEI tube and immediately vortexed for 10 seconds. After a further five-minute incubation, 3 ml of the mixture was added to each T175 flask of HEK cells.

48 hours post transfection the virus was collected. The media was removed, and cells were washed with 20 ml of PBS per flask. Using a scraper, the cells were scraped into PBS and split evenly between four 15 ml falcon tubes. Here the cells were pelleted by spinning them down at 500 g for 5 minutes. The supernatant was discarded and the pellets were resuspended in 150 µl PBS. At this point, the contents of the four tubes were combined and transferred to a 1.5 ml Eppendorf tube. The contents were vortexted for 10 seconds and then snap-frozen in dry ice. The lysates were then thawed in the 37 ^{O}C water bath and vortexed for another 10 seconds. This process was repeated three times in total before being passed through a 19 gauge needle 10 times. Finally, the cell debris was pelleted through centrifugation at max speed at 4 ^{O}C for 15 minutes. After this, the supernatant (containing the virus) was passed into a new Eppendorf and stored at 4 ^{O}C.

To validate the viability of the small scale AAV production 10 µl of the viral prep was mixed with 10 µl 2x laemelli buffer and was ran on a 10% SDS PAGE gel. This blot was probed with a rabbit anti-AAV primary antibody to assess the integrity of the viral capsid.

A 12-well plate of HEK cells was also transduced with an increasing amount of the virus (e.g. 25 µl, 50 µl and 100 µl). Protein was extracted from these cells 48 hours post transduction as

previously and then ran on a 10% SDS-PAGE gel to assess whether the virus is both functional and the transgene is correctly expressed.

2.3.4 Large scale AAV production

Transfection of virus plasmids

HEK293T cells were used for large scale production of Adeno-Associated Virus. Thirty T175 flasks were seeded 16 to 20 hours before transfection. Two hours prior to transfection the medium was changed to serum free DMEM (supplemented with 0.5% Penicillin/Streptomycin).

DNA is prepared according to **Table 2.23**:

Table 2.23: Quantities of Plasmids required for AAV transfection

Plasmid	Ratio	Amount per plate (µg)	Amount per 30 plate transfection (µg)
Helper plasmid (pHelper)	2	26	780
Packaging plasmid (pAAV9)	1	13	390
AAV transgene	1	13	390

Six 50 ml falcon tubes were set up each containing 15.5 ml of serum free DMEM, three for PEI and three for DNA. To each of the PEI tubes, 1560 µl of PEI was added. To each of the DNA tubes one third of the DNA mixture was added. Each PEI tube was added to one of the DNA falcons and then immediately vortexed for 10 seconds. Transfection mixes were left at room temperature for 15 minutes.

3.3 ml of DNA complexes were added to each T175 flask of HEK293T cells and swirled to ensure full coverage of the monolayer. Transfected cells are returned to the incubator and left for 5 days.

Harvest and concentration

Following incubation, the supernatant was collected and placed in sterile 500 ml bottles. Benzonase (Sigma, 250 U/µl) is added to each bottle at the final concentration of 12.5 U/ml. The bottles were then incubated at 37 ^{O}C for 2 to 3 hours, mixing occasionally by inversion. The solution was split between eight 50 ml falcons and cell debris were collected by a short spin at 3850*xg*/10 minutes. The supernatant was then transferred to the top part of a filtering unit (ThermoScientific Nalgene Rapid flow filter unit) attached to a 500 ml duran bottle to which a vacuum is applied.

This cleared supernatant was then concentrated using Amicon Ultra-15 centrifugal 100k filters (Millipore). 15 ml of supernatant was added to each filter and centrifuged at 3800*xg*/4^{O}C for increasingly longer spins (starting at 10 mins, extending gradually to 1 hour) until the final volume of 27 to 28ml is achieved. This should be kept on ice.

AAV purification by iodixanol density gradient:

All iodixanol gradient solutions were prepared prior to starting the purification (**Table 2.24**).

Table 2.24: Iodixanol Gradient Solution Preparation

Percentage Iodixanol	**Iodixanol 60% stock**	**5M NaCl**	**5 x PBS-MK ($MgCl_2$ and KCl)**	**H_2O**	**Phenol Red**
15%	12.5ml	10ml	10ml	17.5ml	-
25%	20.8ml	-	10ml	19.2ml	100µl
40%	33.3ml	-	10ml	6.7ml	-
54%	45ml	-	-	5ml	100µl

Using two Quick-seal 39 ml tubes (Beckman Coulter #344326), the iodixanol gradients and concentrated virus solutions were added in the following order using disposable syringes and a 100 mm, 18G blunt-end needle (Hamilton #7750-09), ensuring all layers are kept separate:

1. Virus solution (14 ml)
2. 4 ml of 15% iodixanol
3. 9 ml of 25% iodixanol
4. 9 ml of 40% iodixanol

5. 5 ml 54% iodixanol
6. Any left-over space in the tube is filled to the rim of the neck with PBS (avoiding bubbles).

The tubes were then sealed using a heat-sealing device, ensuring they are fully sealed to avoid leakage. These tubes were placed in the type 70Ti rotor (Beckman Coulter) and spun at 69,000*xrpm*/1 hour 20 minutes/18^{O}C using the maximum acceleration but no brake.

After centrifugation, there was a clear fraction in the 40% layer which contained the virus. To isolate this fraction, the tube was clamped into a retort stand and a 19G needle used to puncture a hole in the top of the tube in order to introduce air into the tube. A second 19G needle is inserted approximately 1 cm from the bottom of the tube, towards the top of the 60% iodixanol layer. This needle should be inserted in such a way that it allows flow-through of the contents of the tube so that the viral fraction can be collected in a drop wise manner. This fraction is collected in 250-500 µl samples until the entire of the 40% layer and the first half of the 25% layer have been collected (approx. 45-50 eppendorfs). This is then repeated with the second tube.

Fraction analysis:

SYPRO ruby staining was used to analyse the capsid of the virus and show whether the virus has formed correctly and if the fractions were clean or dirty. 6 µl of each fraction was diluted with 6 µl water and 3 µl 4x Laemmli loading buffer, boiled for 5 minutes at 95 ^{O}C and then run on a 10% polyacrylamide gel by SDS-PAGE. Viral capsid proteins are visualised through SYPRO (Sigma) ruby staining of the gel following the manufacturer's instructions.

Fractions which appear with no high molecular weight background are classed as pure 'high quality' virus which can be used for *in vivo* work after being pooled together and mixed homogenously, while the 'Low-quality' virus can also be pooled for *in vitro* use.

Concentration and desalting of AAV preparations:

Following pooling of high and low quality AAV, the fractions were concentrated and desalted using a Amicon Ultra Centrifual filter device (Millipore UFC9100008). 15 ml of PBS containing 35 mM NaCl (sterile filtered) was added to each filter in a 50 ml falcon. This was spun at

3000*xg*/15 minutes. The pooled samples were then added to the filter, topped up to 15 ml with additional PBS+35 mM NaCl and spun for a further 15-20 minutes. The solution was then pipetted up and down five times on each side to ensure the virus was not stuck to the filter. A further 15 ml of PBS+35 mM NaCl was added to the column and spun again as previously. This process was repeated until the original fraction volume was exchanged 10 times at least. After thorough resuspension the viruses were aliquoted and stored at -80 ^{O}C long term, with one aliquot being kept at -20 ^{O}C for validation.

Evaluation of vector purity:

The formation of the viral capsid in the final purified, concentrated, and desalted virus is then validated through both western blotting and SYPRO ruby dye of the virus directly loaded into laemelli buffer and boiled for 5 minutes at 95^{O}C prior to loading in a 10% acrylamide gel. It is suggested that two volumes are run of each virus, so here we chose 1 µl and 5 µl.

2.3.5 AAV Titration and Validation

For working out the titre of each AAV virus, the virus samples are directly compared through qPCR to a linearised copy of the plasmid containing the transgene used in the transfection. The AAV Gle1b plasmid linearised by MluI was used in this case.

A series of serial dilutions (from 10^{-1} to 10^{-8}) of the linearised plasmid and virus are run by qPCR as previously for the LV titration.

2.4 Microscopy

2.4.1 Immunofluorescence

For immunostaining, 24-well plates were set up prior to the addition of cells. 13 mm coverslips were left in 70% IMS for 15 minutes before being rinsed in sterile water three times and placed one per well. These were then coated with 250 µl 0.5mg/ml Gelatin in sterile PBS (Sigma) and left at 37 ^{O}C for 1 hour. This was then removed, and wells were rinsed three times with sterile PBS prior to plating cells at a density of 45,000 cells per well for HEK cells, 25,000 cells per well for HeLa cells, and 20,000 cells per well for fibroblasts.

48 hours after transfection, media was removed, and each well was washed with PBS twice. Cells were fixed through the addition of 200 µl 4% PFA per well for 15 minutes at room temperature, after which they were washed with a further two PBS washes. If not staining immediately, cells were stored at 4 ^{O}C in PBS.

To begin the immunocytochemistry, cells were permeabilised by the addition of 200 µl 0.5 % triton X-100 in PBS for 10 minutes at room temperature. This was then removed and cells washed twice with PBS. 200 µl of 3 % BSA in PBS was then added per well to block cells for 30 minutes at room temperature. Following this, 50 µl of the primary antibody solution (**Table 2.3**) was pipetted onto parafilm in a humidity chamber and the coverslip was removed from the well and kept at 4 ^{O}C overnight.

The following day, the coverslips were replaced into their respective wells and washed three times with PBS. 200 µl of secondary antibody (**Table 2.3**) and Hoescht diluted in 3% BSA in PBS was added to each well and left at room temperature for 1 hour. Cells were then washed a further three times in PBS before being mounted in Fluromount (Sigma F4680) on glass slides. These were then left in a cool dark place to dry overnight before being stored at 4 ^{O}C.

2.4.2 Poly (A)+ RNA Fluorescent in Situ Hybridisation (FISH)

Cells were fixed and permeabilised as previously. Hybridisation mix was made up of 1 ml Hybridisation buffer (20% formamide, 2x SSC 10% dextran sulphate and 1% BSA), 50 µl ssDNA and 1 µl Cy3-Oligo(dT)18 at 1 µg/µl. This was added to the cells and incubated at 37 ^{O}C for 2

hours. Immunofluorescence for cytoplasmic markers was carried out following this stage as previously.

All solutions for this protocol were made with either ddH_20 or PBS treated with Diethyl pyrocarbonate (DEPC) (Biochmica, Panreac Applichem) to remove any contaminating RNase enzymes.

2.4.3 Confocal Imaging

Cells on coverslips were imaged using a Leica SP5 confocal microscope, using the 63X oil objective.

2.4.3 Quantification of immunofluorescence

Quantification of staining intensity was worked out through the Cell Total Corrected Fluorescence (CTCF) method on ImageJ as shown in (Walker et al., 2017). This method is explained in **Figure 2.1**. This was done blinded with the help of Paolo Marchi and Louise Whiteley. Here, the intensity of the nuclear staining would be worked out through the following formula: **Corrected Total Cell Fluorescence (CTCF) = Integrated density – (Nuclear Area x Mean fluorescence of the background).** This formula uses the integrated density of the region of interest (ROI) which is the sum intensity of the pixels in the area of interest. The mean fluorescence of the background was worked out by taking the average integrated densities of ten samples from the background of the image, this value was then multiplied by the area of the ROI (nucleus in this case) and the total subtracted from the integrated density value. From this value, the average intensity of the stain can be worked out accounting for any background intensity across the area of the ROI (the nucleus in this case).

For FISH cytoplasmic quantification a sample from a perinuclear region was taken and worked out in the same way as previously.

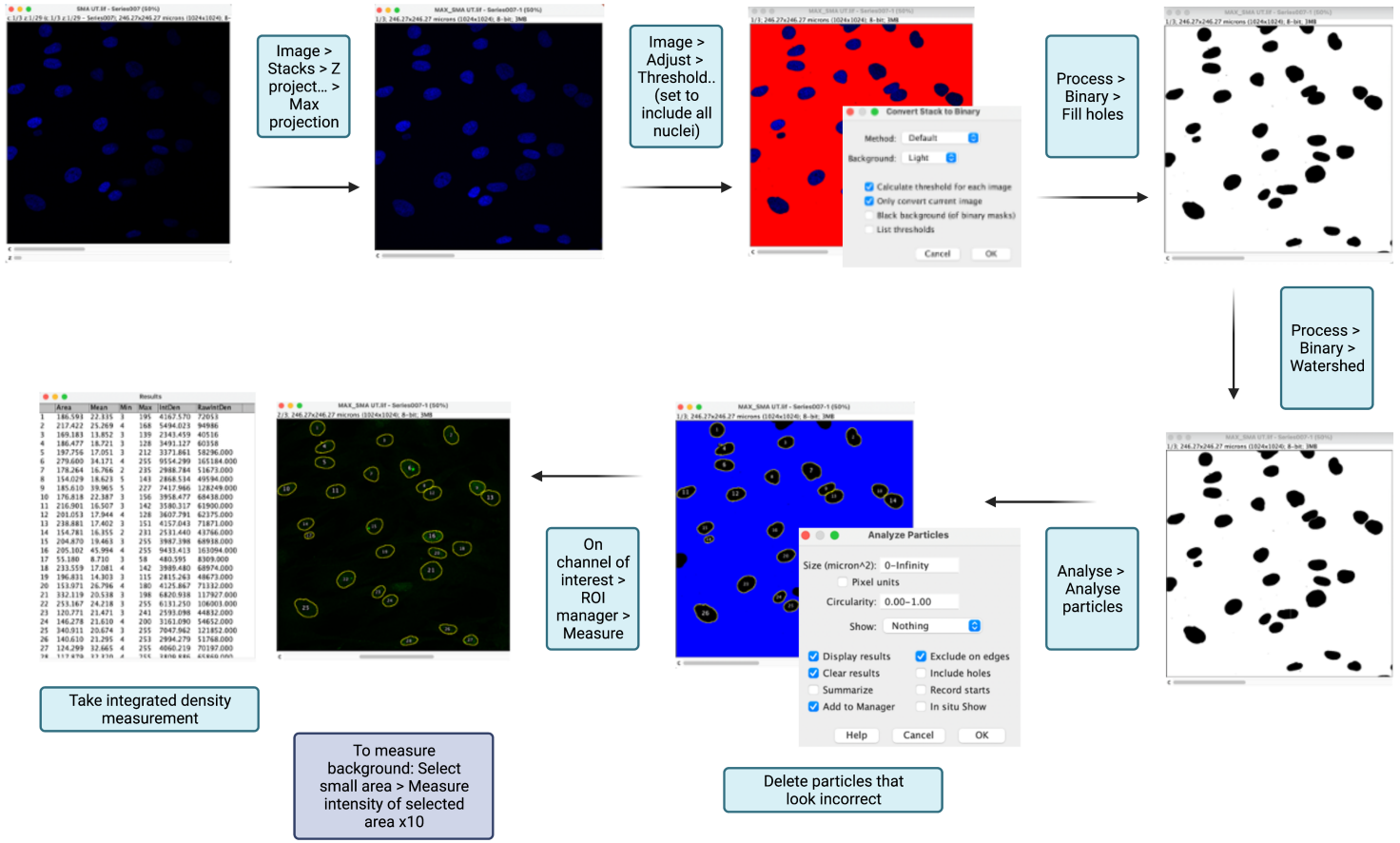


Figure 2.1: Method of analysis for immunofluorescence imaging using ImageJ Fiji. *(1) The image stack is compressed to its maximum projection. (2) A threshold is applied to the DAPI channel to select the nuclei as a region of interest (ROI). (3) Any holes in the nuclei are filled to select the whole area (4) A watershed is applied to separate any adjoining cells into two regions of interest (5) These ROI's are turned into numbered areas which can then be measured individually (6) The integrated density of the ROI's is measured on the channel of interest (in this case, the 488 channel). (7) The background integrated density is established by measuring an average of 10 small ROIs determined by drawing a small area outside of the original ROI.*

2.4.4 Categorisation of Golgi stain

Golgi staining was categorised depending on its morphology as shown in (Custer et al., 2019). This was done in a blinded manner with help from Paolo Marchi and Matthew Roach. Cells were presented alongside a reference of what each category represented.

2.4.5 Proximity ligation assay (PLA)

For the proximity ligation assay (PLA) (Sigma) a Duolink® In Situ Orange Starter kit Mouse/Rabbit was used. Protocol was adapted from manufacturer's protocol as follows. Cells were fixed and permeabilised as previously described in **section 2.4.1**. Blocking was carried out through incubation with PLA blocking buffer (from kit) on parafilm in a humidity chamber at 37 ^{O}C for 30 minutes. Primary antibody incubation was carried out overnight at 4 ^{O}C in PLA antibody dilutant (from kit). Following antibody incubation, cells were washed three times with Wash Buffer A (from kit). PLA secondary probes (anti-mouse and anti-rabbit) were prepared in antibody dilutant, and cells were incubated in a humidity chamber for 1 hour at 37 ^{O}C. Following an additional three washes in wash buffer A, cells were then incubated in a ligation mixture containing ligase and ligase buffer (from kit) for 30 minutes at 37 ^{O}C. Again, cells were washed three times using wash buffer A. Incubation with the amplification mix (Amplification enzyme and buffer (from kit)) was then incubated for 100 minutes at 37 ^{O}C. Cells were then washed in Wash Buffer B three times, before incubation with Hoescht (1:5,000) for 10 minutes at room temperature with wash buffer B (0.01X). Finally, cells were washed twice with 0.01X wash buffer B and mounted on glass slides as previously. Antibodies for immunofluorescence cytoplasmic staining were added at the same time as primary antibody incubation, secondaries were done alongside PLA secondary probes.

2.5.6 Quantification of PLA

Cells were imaged in Z-stacks which allowed full view of the cell's nucleus to exclude any foci which occurred outside of the cell. Foci were counted using the ImageJ cell counter plug in developed by Dr Kurt de Vos at SITraN. Foci number was divided by cell count to work out foci per nucleus/cytoplasm.

2.5.7 Imaging on the Opera Phenix High Content Imaging System

Immunofluorescence in motor neurons was done using 96 well optical plates (Gibco). These were imaged on the High Throughput Opera Phenix imaging system (Perkin Elmer) using a 40x water objective. For each condition at least 2 replicate wells were used per condition, with at least 20 fields of view per well. Cell count per condition was >200 cells.

2.5.8 Columbus Analysis

Columbus Software (Perkin Elmer) was used in the image analysis setting to analyse data obtained from the Opera Phenix imaging system. Image analysis pipelines (**Figure 2.2**) were generated to obtain the mean nuclear fluorescence levels from the 568 and 488 channels respectively. Data was taken as an average from at least 2 technical replicates in each repeat.

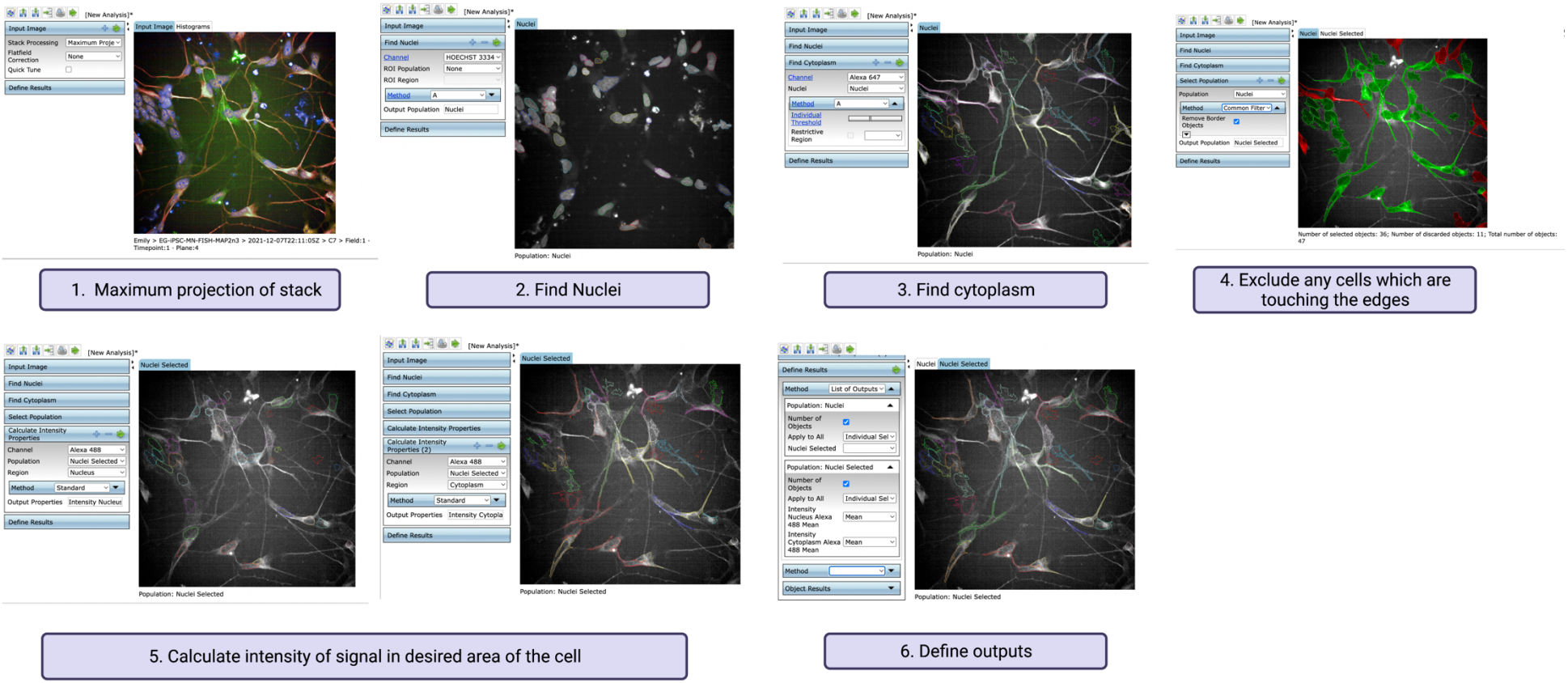


Figure 2.2: Image analysis pipeline for image analysis using PerkinElmer's Columbus software. *(1) The image stack is compressed to its maximum projection. (2) The nuclei are identified using the DAPI channel. (3) The cytoplasm is determined using an axonal/cytoplasmic marker (4) The cells touching the edges are excluded from the population (5) The intensity of the selected area (nuclei/cytoplasm/cell) can be calculated (6) The outputs are defined, in this case as the mean intensity of each area/channel.*

2.6 Statistical analysis

All data is presented as mean ± standard error of the mean (SEM). All experiments were performed at least in triplicate where possible. If less than three replicates were performed, no statistical analyses were performed. Data were analysed using GraphPad Prism Software (V9.2). The following statistical tests were applied: T-test, One way ANOVA, Two-way ANOVA. Statistical significance was considered to be $p < 0.05$ *, $p < 0.01$ **, $p < 0.005$ ***, $p < 0.001$ ****.

3. Developing and validating tools for the project

3.1 Introduction

Gle1 is a highly conserved, essential nuclear export factor (Kendirgi et al., 2003) which has also been implicated in multiple other pathways, such as stress granule formation (Aditi et al., 2015), translation initiation and translation termination (Bolger et al., 2008; Bolger and Wente, 2011). Mutations within this gene have been shown to be causative of a motor neuron disease called human lethal congenital contracture syndrome 1 (LCCS1) (Nousiainen et al., 2008) which is a severe foetal form of disease causing degeneration of motor neurons within the anterior horn of the spinal cord (Herva et al., 1988), but other mutations have also been implicated in amyotrophic lateral sclerosis (ALS) (Kaneb et al., 2015).

Spinal muscular atrophy (SMA) is a childhood motor neuron disease which is caused by loss of the *SMN1* gene (Lefebvre et al., 1995). This causes a progressive degeneration of lower motor neurons located in the anterior horn of the spinal cord, leading to atrophy of muscle (Prior et al., 1993). The most severe cases of SMA (SMA type 0) present in utero, at birth or very early in life as general muscle weakness and limb deformities (Dubowitz, 1999; Macleod et al., 1999).

Similarities between diseases resulting from the loss of either of these two proteins therefore raise questions about the parallels between their functions and what pathways they might be involved with within the cell, in particular motor neurons.

3.2 Aims

The aim of this project is to evaluate whether there is a functional link between Gle1 and SMN and whether this can be related to any aspects of SMA. To carry out this research, we focussed on the dominant isoform of human Gle1: Gle1b. Virally mediated protein overexpression was used to look at whether any observed SMA phenotypes could be rescued by increased levels of Gle1b. This chapter will describe the generation and validation of various lentiviral and AAV vectors designed to modulate Gle1 expression in cell models used in this project.

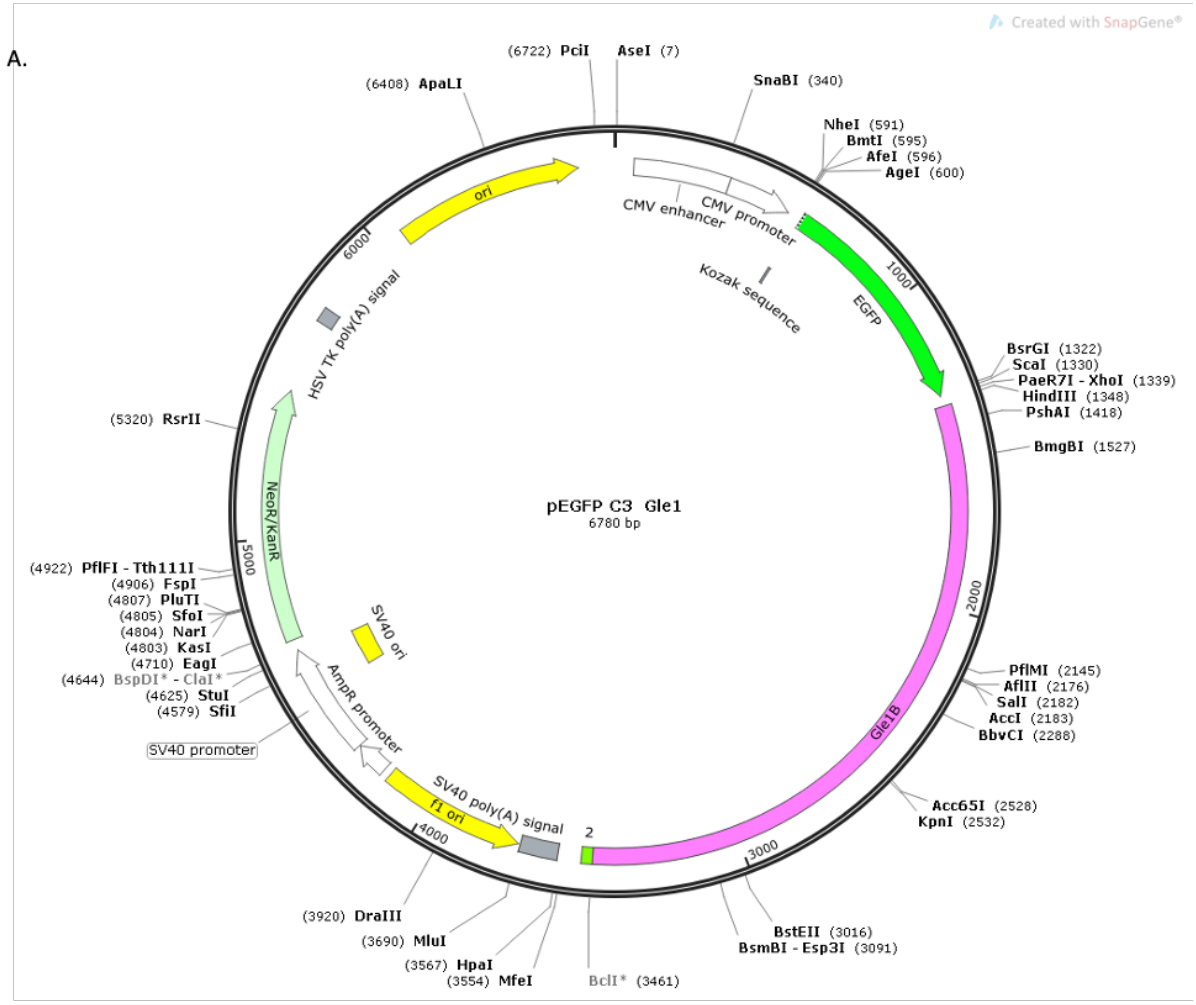
3.3 Results

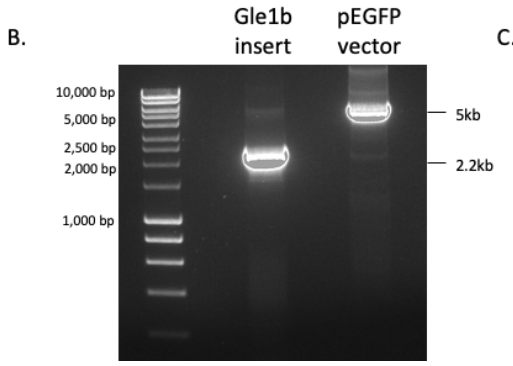
3.3.1 pEGFP-Gle1b Plasmid Cloning and Validation

To explore potential interactions between the Gle1 and SMN proteins we wanted to generate construct which would express a GFP-Gle1b fusion protein (**Fig 3.1 A**), this would then allow use of a GFP-trap immunoprecipitation kit to assess binding partners of Gle1.

This was carried out through a Gibson assembly using an in-house pEGFP C3 SMN N86 vector. This plasmid contained a short truncation of the SMN protein fused to GFP, previous work in the lab involved sequencing this backbone and so it was known to be a good starting point for further cloning. The sequence for the Gle1b cDNA was taken from the previously generated and validated AAV Gle1b plasmid (made in-house by Matthew Wyles).

Both the pEGFP backbone and the Gle1b cDNA was PCR amplified and gel extraction of the correctly sized products was carried out, prior to purification (**Fig 3.1 B**). Following ligation, the DNA was transformed into NEB stable-3 cells. Colonies were picked and grown in LB broth, successful growth was purified by MiniPrep (Qiagen). DNA was checked by digestion prior to sequencing those that looked successful (**Fig 3.1 C**).





C.

Enzyme	Expected Fragment Sizes (bp)
EcoRI	6226 554
NcoI	3922 1904 703 251

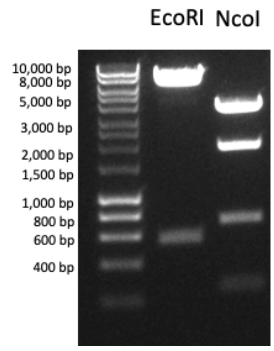


Figure 3.1: Gle1b cDNA was successfully cloned into a pEGFP vector. *(**A**) Plasmid map of the finalised pEGFP_C3_Gle1b plasmid (**B**) PCR amplification of the Gle1b insert and pEGFP vector backbone prior to gel extraction and purification. Ladder shown is 1kb NEB ladder. (**C**) Representative digest of a 'successful' colony with digest sizes matching those expected (as shown in table on the left). Ladder shown is 1kb NEB ladder.*

Following the return of correct sequencing data showing that the cDNA had correctly been inserted into the pEGFP backbone, we aimed to validate expression of our GFP-Gle1b fusion protein. HEK cells were transfected with the chosen clone and 48 hours post transfection cells were processed for both western blot and immunofluorescence.

Co-staining of Gle1 and GFP through ICC shows a clear colocalization of overexpression of Gle1 with GFP (**Fig 3.2 A**). Western blots probed with a Gle1 antibody show an endogenous band of 75 kDa, with a clear overexpression band at 100kDa in the transfected cells (**Fig 3.2 B**). This corresponds with the expected size of the GFP-Gle1b fusion protein. Probing of the protein lysate with a GFP antibody also shows a band at 100kDa in the transfected condition only (**Fig 3.2 C**), supporting the validation of the plasmid. GAPDH shows equal loading across the conditions.

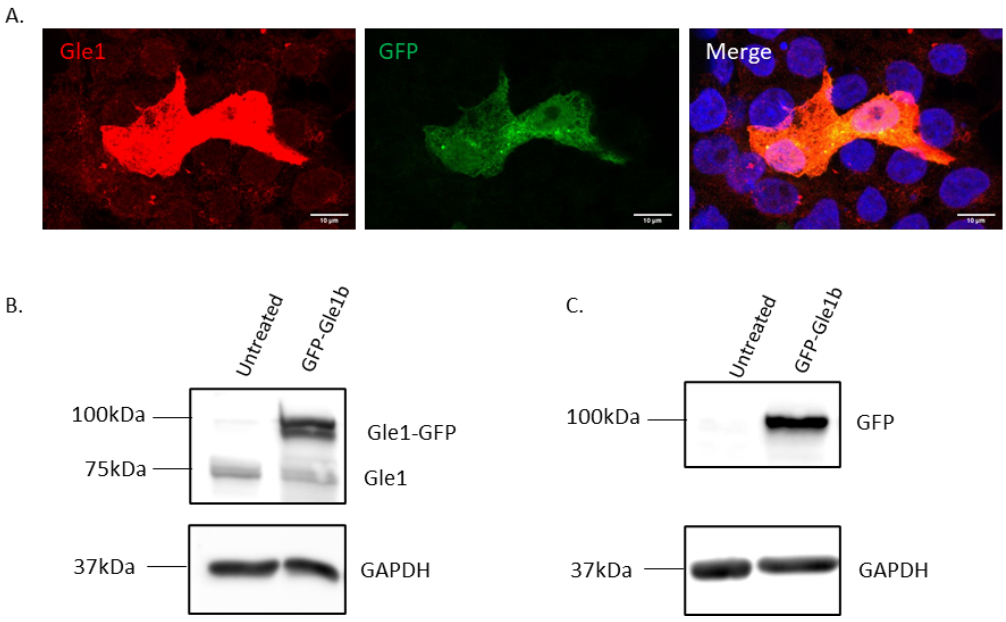


Figure 3.2: Validation of the GFP-Gle1b fusion protein overexpression plasmid. (A) *Immunocytochemistry of HEK cells transfected with 500ng of pEGFP-Gle1b plasmid. Co-localisation of Gle1 protein overexpression and GFP fluorescence confirmed with co-stained with Gle1 (red) and GFP (green) antibodies. Signal overlap seen in merge image. (**B**) Western blot probed with an anti-Gle1 antibody showing protein expression of both endogenous Gle1 at 75kDa and overexpressed GFP-tagged Gle1 at around 100kDa. GAPDH (at 37kDa) is used*

*to show equal loading (**C**) Western blot probed with an anti-GFP antibody showing protein expression of GFP-tagged Gle1 at around 100kDa. GAPDH (at around 36kDa) is used to show equal loading.*

3.2.2 LV FLAG-Gle1b Plasmid Cloning and Validation

Many experiments planned for this research project involve looking at overexpression of Gle1 in cells which are difficult to transfect, mainly being motor neurons and fibroblasts. As a result of this we needed to generate tagged lentiviral constructs. Using a previously generated 3xFLAG-tagged Gle1b pcDNA5_FRT plasmid (kindly cloned by Dr Guillaume Hautbergue at SITraN), and the in-house SIN-PGK-cPPT-WHV (pLenti-VOS) lentiviral backbone. Before starting the process of cloning the FLAG-tagged Gle1b plasmid was tested in HEK cells by western blot for overexpression of a FLAG-Gle1b isoform (**Fig 3.3**).

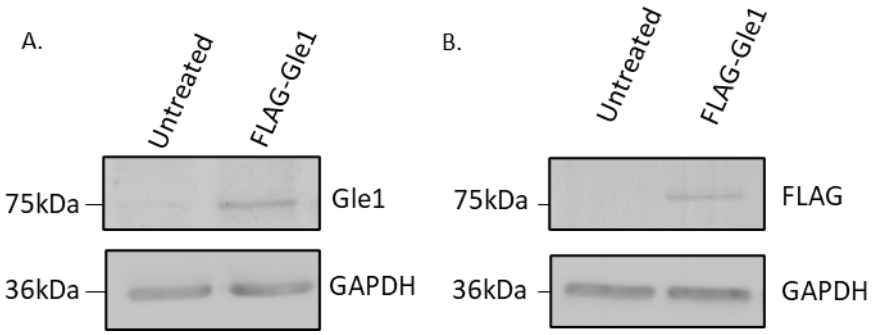


Figure 3.3: Western blot demonstration of overexpression of FLAG-Gle1b from 3xFLAG-tagged Gle1b pcDNA5_FRT plasmid. *(**A**) Western blot probed with anti-Gle1 antibody, showing protein expression of both endogenous Gle1 at 75kDa and overexpressed FLAG-tagged Gle1. GAPDH (at around 36kDa) is used to show equal loading (**B**) Western blot probed with an anti-FLAG antibody showing protein expression of FLAG-tagged Gle1 at around 75kDa. GAPDH (at around 36kDa) is used to show equal loading.*

Once this was validated, we moved to clone the insert from the pCDNA5_FRT plasmid into the multiple cloning site within the pLenti-VOS backbone using the MluI and NheI restriction

sites (**Fig 3.4 A**). Primers with complimentary cut sites were designed complimentary to the 3xFLAG-Gle1b insert from the pCDNA5_FRT plasmid. These were used to PCR amplify the insert and digest the vector, the correct sizes were verified by imaging the products on an agarose gel prior to gel extraction and purification (**Fig 3.4 B, C**). Ligation of the plasmid was carried out using the quick ligase kit and colonies were checked for the insert through digestions with MluI and NheI (**Fig 3.4 D**).

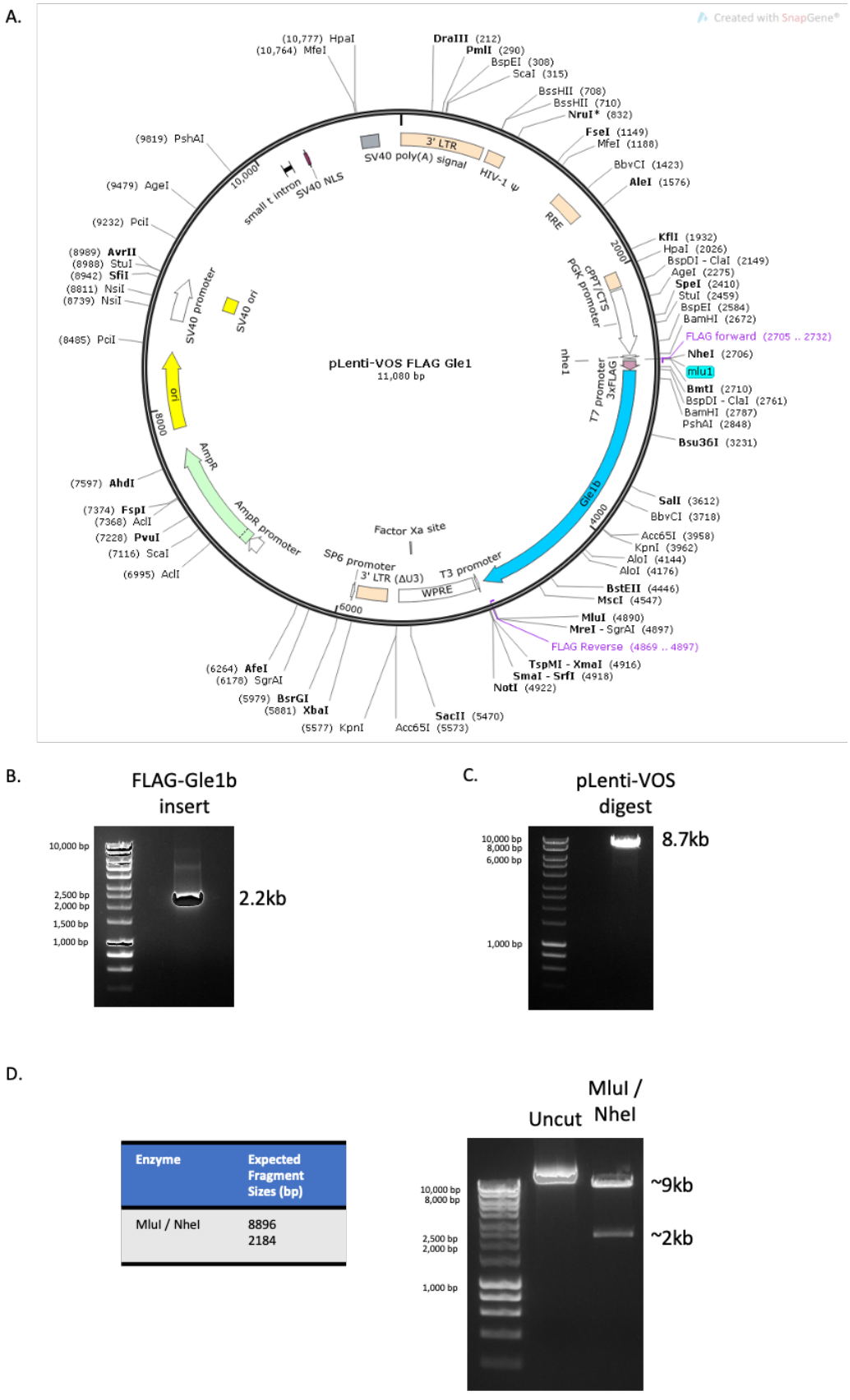


Figure 3.4: FLAG-Gle1b cDNA was successfully cloned into the pLenti-VOS vector. *(**A**) Plasmid map of the finalised pLenti-VOS FLAG-Gle1b plasmid (**B**) PCR amplification of the Gle1b insert*

*from the pCDNA5_FRT plasmid prior to purification. Ladder shown is 1kb NEB ladder. (**C**) Double digest of the pLenti-VOS plasmid prior to gel extraction and purification. Ladder shown is 1kb NEB ladder. (**D**) Representative digest of a 'successful' colony with digest sizes matching those expected from the maps (as shown in the schematics on the left). Ladder shown is 1kb NEB ladder.*

Validation of this plasmid was carried out through western blotting, as previously with the pCDNA5_FRT plasmid. Probing the blot with both a FLAG antibody and a Gle1 antibody showed a specific overexpression of the FLAG tagged Gle1b protein (**Fig 3.5**). GAPDH was used to indicate equal loading of total protein extracts.

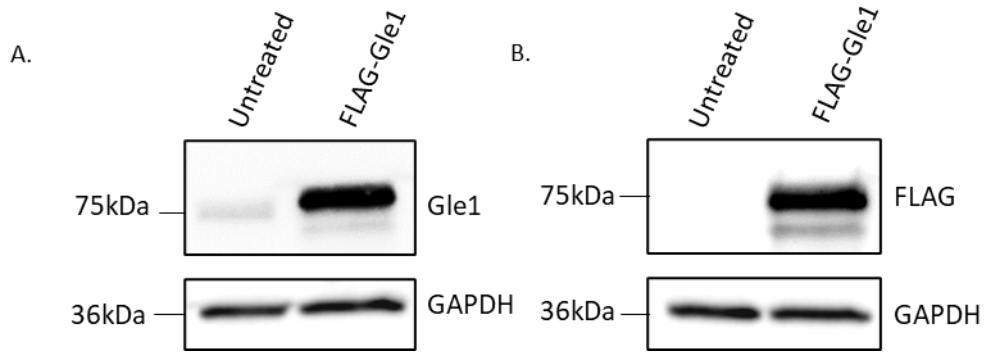


***Figure 3.5: Western blot demonstration of overexpression of FLAG-Gle1b LV_SIN-PGK-WPRE plasmid.** (**A**) Western blot probed with anti-Gle1 antibody, showing protein expression of both endogenous Gle1 at 75kDa and overexpressed FLAG-tagged Gle1b. GAPDH (at around 36kDa) is used to show equal loading. (**B**) Western blot probed with an anti-FLAG antibody showing protein expression of FLAG-tagged Gle1b at around 75kDa. GAPDH (at around 36kDa) is used to show equal loading.*

This plasmid was also validated by sequencing and so was confirmed to be ready for Lentiviral production.

3.2.3 LV FLAG-Gle1b Viral production and Validation

Lentiviral vector (LV FLAG Gle1b) was produced at a titre of 3.58 x 10^8 TU/ml. The produced virus was used for *in vitro* studies.

Validation of this virus was carried out through both western blotting and immunocytochemistry. A dose-response experiment was performed in human fibroblast cells to test for the optimal multiplicity of infection (MOI) to use for experiments. Western blotting demonstrated a clear dose dependant overexpression of a FLAG-tagged Gle1b protein (**Fig 3.6**).

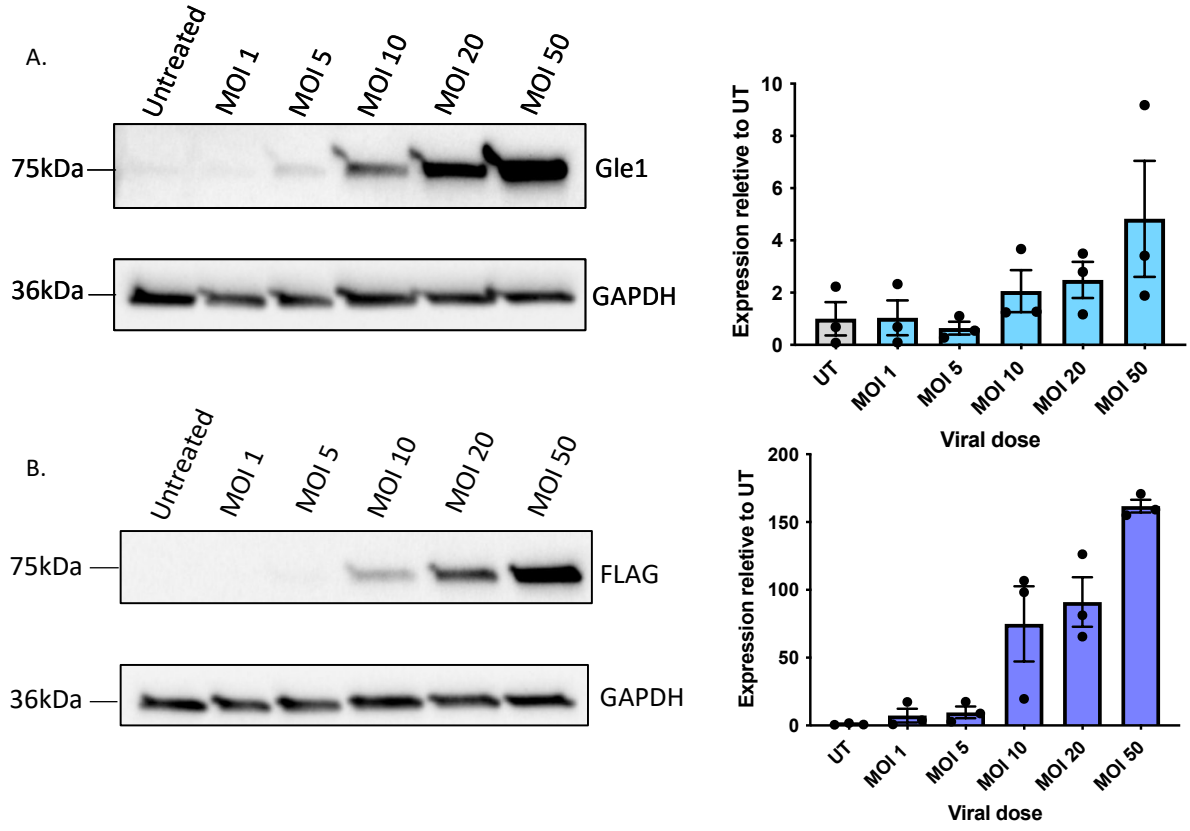


Figure 3.6: Western blot analysis of lentiviral FLAG-Gle1b overexpression in fibroblast cells (GM8680) 4 days after LV transduction. (A) *Western blot probed with anti-Gle1 antibody, showing protein expression of both endogenous Gle1 at 75kDa and overexpressed FLAG-tagged Gle1b. GAPDH (at around 36kDa) is used to show equal loading. Quantification by densitometric analysis of three biologically independent repeats (n=3), normalised to GAPDH levels. All values shown relative to Untransduced (UT). (**B**) Western blot probed with an anti-*

FLAG antibody showing protein expression of FLAG-tagged Gle1b at around 75kDa. GAPDH (at around 36kDa) is used to show equal loading. Quantification by densitometric analysis of three biologically independent repeats (n=3), normalised to GAPDH levels. All values shown relative to Untransduced (UT).

Transduction efficiency was also assessed through immunocytochemistry to determine the percentage of cells which were transduced with the virus at each dose (**Fig 3.7**). This data demonstrated that an MOI of 20 was the optimal dose of virus for both protein expression and numbers of cells transduced.

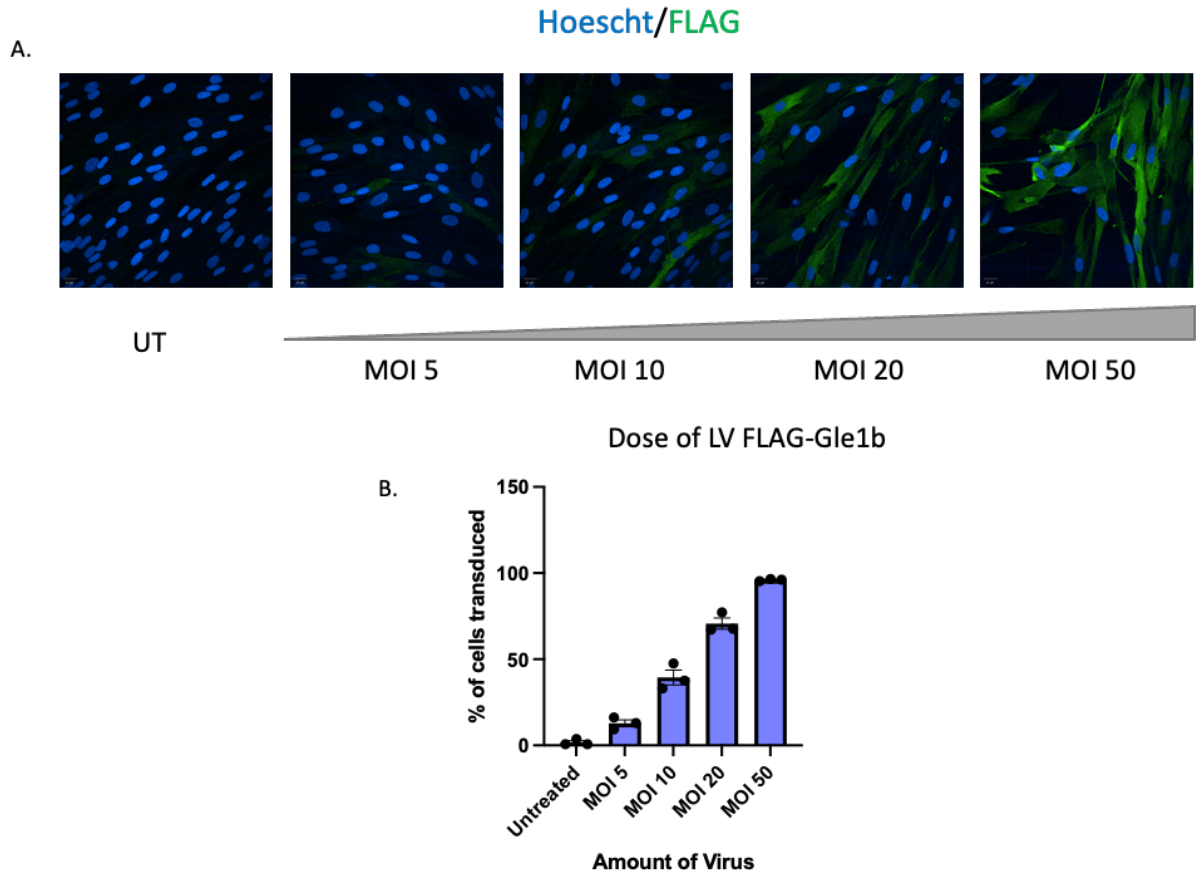
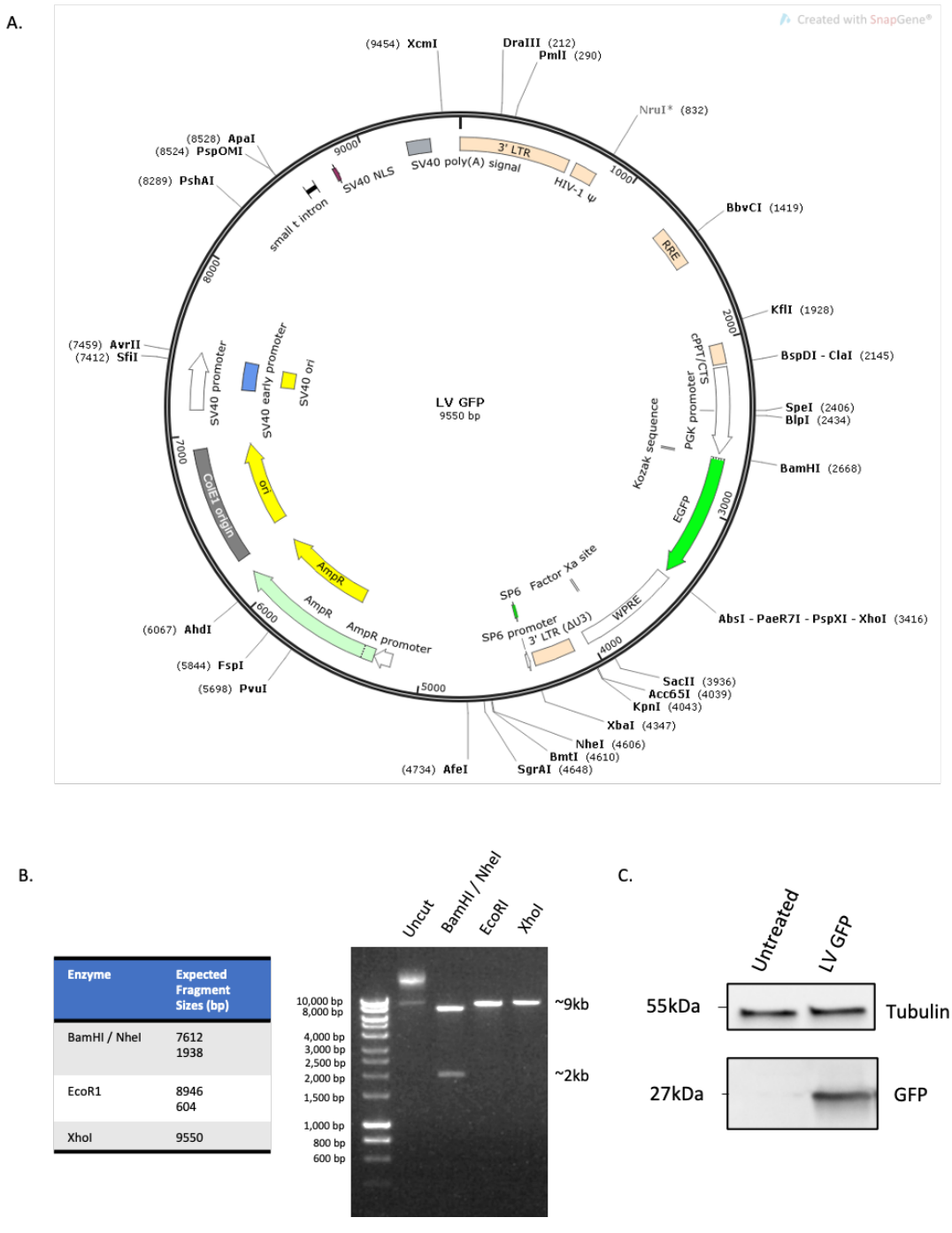


Figure 3.7: Immunocytochemistry analysis of lentiviral FLAG-Gle1b overexpression in fibroblast cell (GM8680) 4 days after LV transduction. *(**A**) Representative images of cells transduced with an increasing dose of LV FLAG-Gle1b. Transduced cells shown in green through staining with an anti-FLAG antibody, therefore showing overexpression of FLAG-*

*Gle1b. Hoechst staining of nuclei in blue. (**B**) Quantification of cells which were positive for FLAG staining indicating transduction, data from three technical repeats, 22 fields per repeat, at least 300 cells per condition.*

3.2.4 LV-GFP plasmid validation

A lentiviral vector containing a GFP reporter transgene (LV GFP) was used in experiments as a control virus. Prior to viral production this plasmid was validated following retransformation from an old plasmid stock. The plasmid was digested to confirm it was as expected (**Fig. 3.8b**). Western blots were performed to assess protein overexpression in HEK cell protein lysates which had been transfected with 500 ng plasmid DNA.



Enzyme	Expected Fragment Sizes (bp)
BamHI / NheI	7612 1938
EcoR1	8946 604
XhoI	9550

Figure 3.8: LV GFP Plasmid Validation. *(**A**) LV GFP plasmid map. (**B**) Digests to check retransformation of LV GFP plasmid from old plasmid stock. (**C**) Western blot of protein lysates from HEK cells transfected with LV GFP plasmid, probed with anti-GFP antibody, showing protein expression at 27kDa. Tubulin (at 55kDa) is used to show equal loading.*

3.2.5 LV GFP viral production and validation

Lentiviral production generated a stock of LV GFP virus to be used for in vitro studies. The final titre of this virus was 7.68 x 10^8 TU/ml.

Validation of LV GFP virus was carried out through both western blotting and immunocytochemistry. A dose-response experiment was carried out to test for the optimal MOI to use for experiments. Western blotting demonstrated a clear dose dependant overexpression of GFP protein (**Fig 3.9**). Immunocytochemistry demonstrated a good transduction efficiency from MOI 10 upwards (F**ig 3.10**). To match the MOI used of LV FLAG Gle1b, an MOI of 20 was chosen as the optimal dose for further experiments.

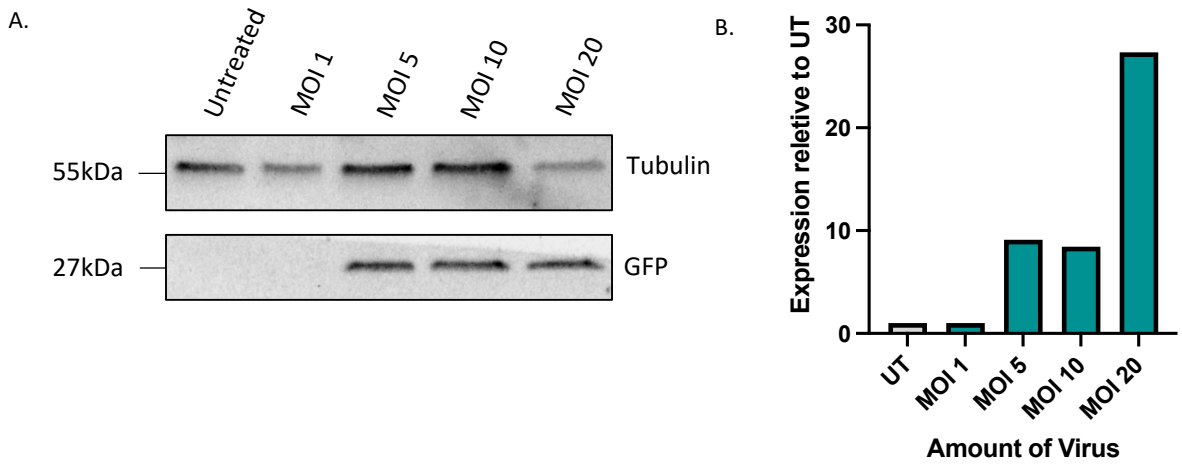


Figure 3.9: Western blot analysis of lentiviral GFP overexpression in fibroblast cell lines (GM8680) four days following LV transduction. (A) *Western blot probed with anti-GFP antibody, showing protein expression at 27kDa. Tubulin (at 55kDa) is used to show equal loading. (**B**) Quantification by densitometric analysis of one biological repeat (n=1), normalised to GAPDH levels. All values shown relative to Untransduced (UT).*

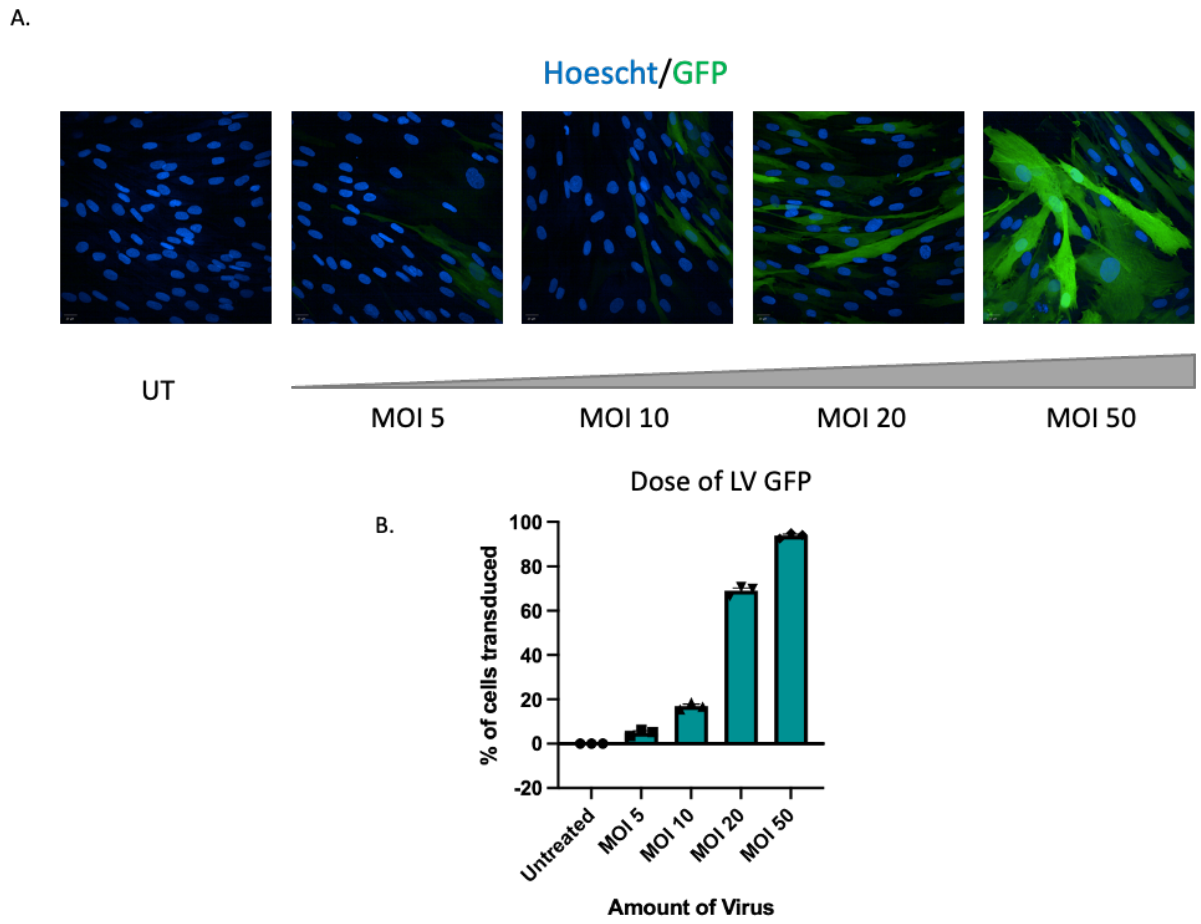


Figure 3.10: Immunocytochemistry analysis of lentiviral GFP overexpression in fibroblast cell lines (GM8680) four days following LV transduction. *(**A**) Representative images of cells transduced with an increasing dose of LV GFP. Transduced cells shown in green through staining with an anti-GFP antibody. Hoechst staining of nuclei in blue. (**B**) Quantification of cells which were positive for GFP staining indicating transduction, data from three technical repeats, 22 fields per repeat, at least 300 cells per condition.*

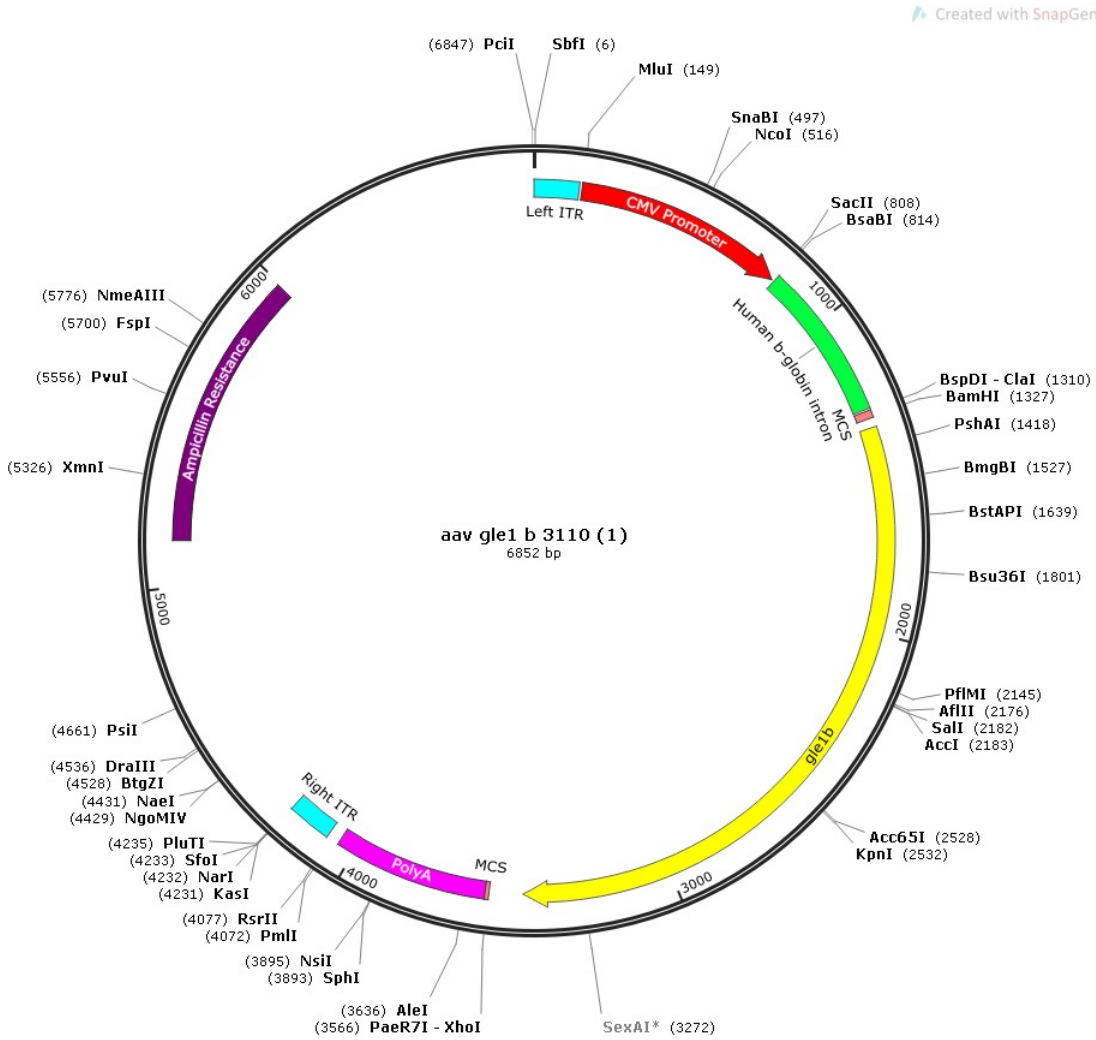
3.2.6 AAV Gle1b plasmid Validation

The work carried out in this thesis, combined with work done in the lab previously, led us to aim for investigation of the neuroprotective effect of Gle1 *in vivo*. To deliver the Gle1b transgene to the CNS *in vivo* an AAV vector was required. The human Gle1b isoform was previously cloned into an in-house pAAV_CMV by Matthew Wyles. As with the lentiviral

plasmids, prior to using this plasmid to produce large scale virus we wanted to ensure it was working as expected.

The plasmids were retransformed to create stocks, and these were checked by both digestion and sequencing. Digests of the plasmid and the inverted tandem repeats (ITRs) demonstrated that the plasmid contained the expected insert and the ITRs were importantly still intact, as these can easily be lost through retransformation (**Fig 3.11**).

A.



B.

Enzyme	Expected Fragment Sizes (bp)
BamHI / XhoI	4613 2239
SmaI	2854 2705 1271 11 11

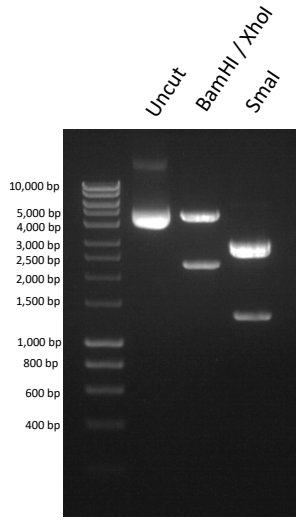


Figure 3.11: AAV Gle1b plasmid and digestions. *(**A**) Plasmid map of AAV Gle1b (**B**) Digestions to confirm retransformation of plasmid from old stock, in particular the ITR regions.*

Western blotting and immunocytochemistry of HEK cells transfected with the AAV Gle1b plasmid demonstrate a clear overexpression of the protein which is specific to the Gle1b isoform (**Fig 3.12 A, B**). Here, I also tested an AAV Gle1a plasmid (also generated by Matt Wyles previously). As Gle1a is not the dominant isoform, combined with data generated in later chapters of this thesis, I decided to focus specifically on Gle1b from this point onwards.

Immunocytochemistry also demonstrated an overexpression of Gle1b protein, this overexpression occurs in all areas of the cell rather than one region in particular (**Fig 3.12 C**).

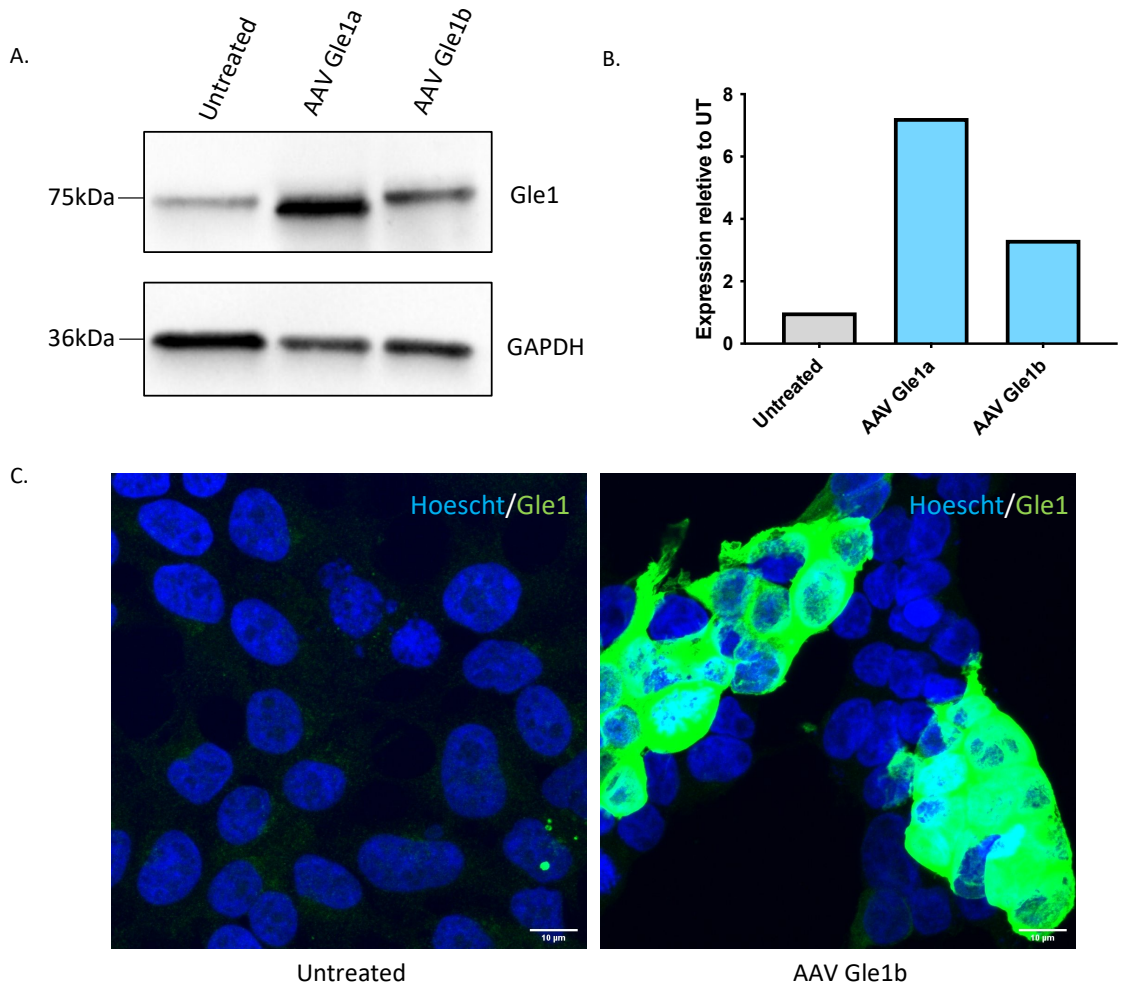


Figure 3.12: Validation of Gle1b overexpression from the pAAV_CMV_Gle1b plasmid. *(A) Western blot demonstrating protein overexpression of both Gle1 isoforms in HEK cells 48 hours after transfection. Probed with anti-Gle1 antibody, showing protein expression at 75kDa.*

*GAPDH (at 36kDa) is used to show equal loading. (**B**) Quantification by densitometric analysis of one biological repeats (n=1), normalised to GAPDH levels. All values shown relative to Untransfected (UT). (**C**) Immunocytochemistry showing overexpression of Gle1b protein (green) in HEK cells using a Gle1 antibody.*

3.2.7 AAV9 Gle1b viral production and validation

Following successful validation of the plasmid we proceeded with production of a pilot small scale AAV9 production (see materials and methods for more details). This would allow us to test whether the plasmid correctly produced a virus which would package and correctly express the transgene.

To test the outcome of the viral production, the capsid integrity and the virus's ability to overexpress the transgene were evaluated. Small volumes of virus were added to HEK cells alongside a transfected well, serving as a positive control. Protein lysate was run on a western blot and probed for Gle1 expression (**Fig 3.13 A**), quantification of Gle1 signal shows an increase in the transduced well compared to untreated cells, albeit lower than the transfected cells (**Fig 3.13 B**). Capsid integrity was evaluated by loading a small volume of virus mixed with laemelli buffer into a gel and probing for the VP1, VP2 and VP3 viral capsid proteins. These three proteins are visible at the correct sizes therefore confirming successful production of the viral capsid (**Fig 3.13 C**).

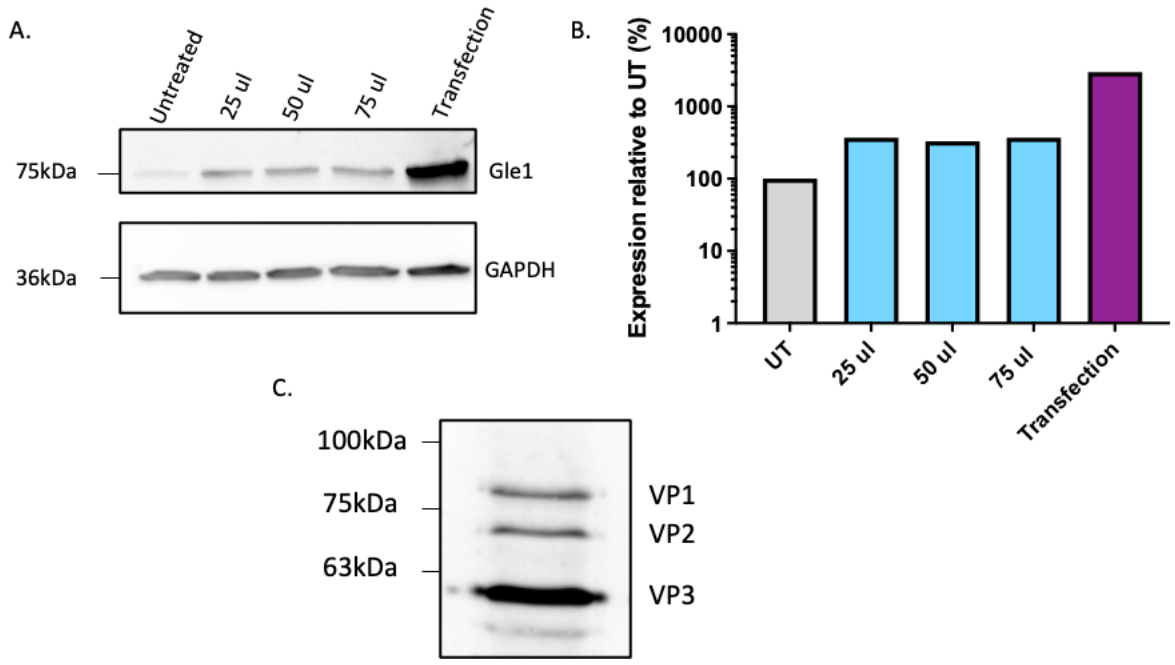


Figure 3.13: Validation of a small scale AAV9 Gle1b preparation. *(**A**) Western blot probed with anti-Gle1 antibody, showing protein expression at 75kDa. GAPDH (at 36kDa) is used to show equal loading. (**B**) Quantification by densitometric analysis (n=1), normalised to GAPDH levels. All values shown relative to Untransduced (UT). Scale is logarithmic. (**C**) Western blotting of viral capsid proteins (VP1, VP2 and VP3 at 87kDa, 73kDa and 61kDa respectively).*

As the plasmid appears to be capable of successfully produce functional and intact AAV9 virus, I proceeded with generation of a purified, larger sized AAV preparation, which would be suitable for *in vivo* work in the future.

The quality of the virus was assessed through SYPRO ruby staining. This allowed us to pool viral fractions into high- and low-quality groups (**Fig 3.14**), which could be mixed and further processed.

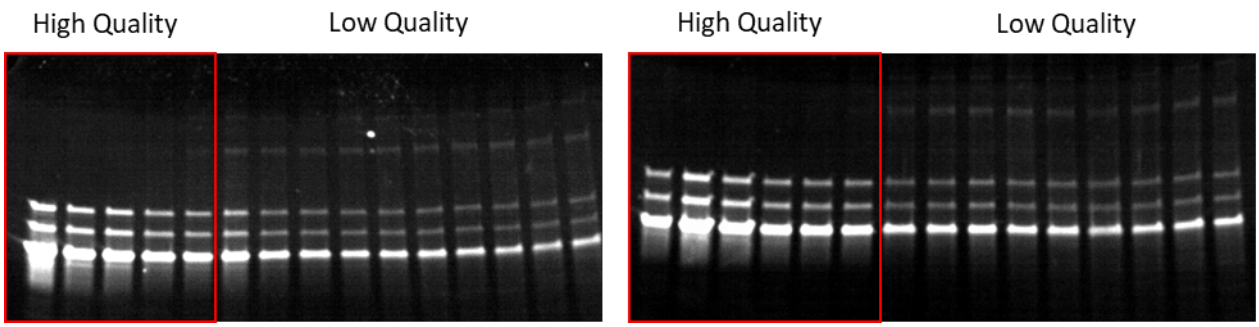


Figure 3.14: SYPRO ruby staining of viral fractions to determine high- and low- quality fractions. *Red boxes show fractions from each prep which were clean enough to be considered high-quality, the rest were pooled as low-quality fractions.*

Once the fractions were pooled, the virus was de-salted and aliquoted for long term storage at -80 ^{o}C. Following this process, the virus was assessed for integrity and transgene expression, as previously described in section **3.2.7**.

Firstly, the capsid integrity of the fractions was assessed through both western blotting (**Fig 3.15 A**) and SYPRO ruby staining (**Fig 3.15 B**). Both experiments demonstrated that the VP1, VP2 and VP3 proteins were present in all fractions, therefore demonstrating the presence of an intact capsid.

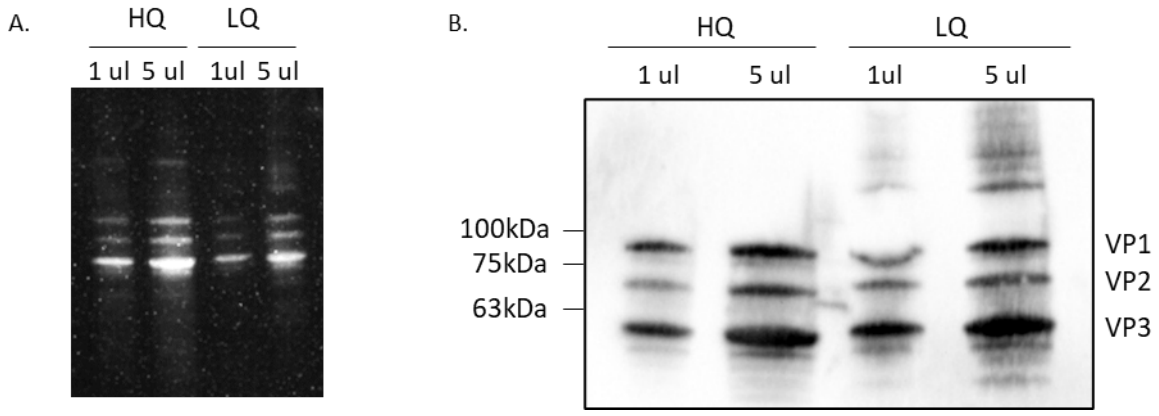


Figure 3.15: Validation of large-scale AAV9 Gle1b capsid. ***(A)*** *Sypro ruby staining of viral capsid in both high- and low- quality fractions.* ***(B)*** *Western blot probed with anti-VP1, VP2 and VP3 antibody at 87kDa, 73kDa and 61kDa respectively.*

After validating the presence of a correctly formed capsid, I performed a titration of both viral fractions, to assess how concentrated our fractions were for future experiments. Titres of both viral fractions can be seen in **Table 3.1**.

Table 3.1: Titre of large scale AAV9 Gle1b virus fractions. *Virus titred by qPCR in viral genomes copies per ml (vg/ml).*

	Titre of Virus
High Quality AAV9-Gle1b	1.93×10^{14}
Low Quality AAV9-Gle1b	1.22×10^{13}

To assess whether the virus is correctly expressing the Gle1b transgene, primary cortical neurons from mice were isolated and transduced with AAV9-Gle1b. AAV9 has a specific tropism to the CNS and so transduces these neurons with high efficiency. As we are planning to use this virus for *in vivo* study in mice, using primary neuronal cells from a wild-type mouse seemed the best cell model for initial testing.

Therefore, we performed both immunocytochemistry and protein extraction from cells which had been transduced with different MOI's of virus. Both the low-quality and high-quality fractions appear to demonstrate good protein expression at low and high doses (**Fig 3.16 A, B**).

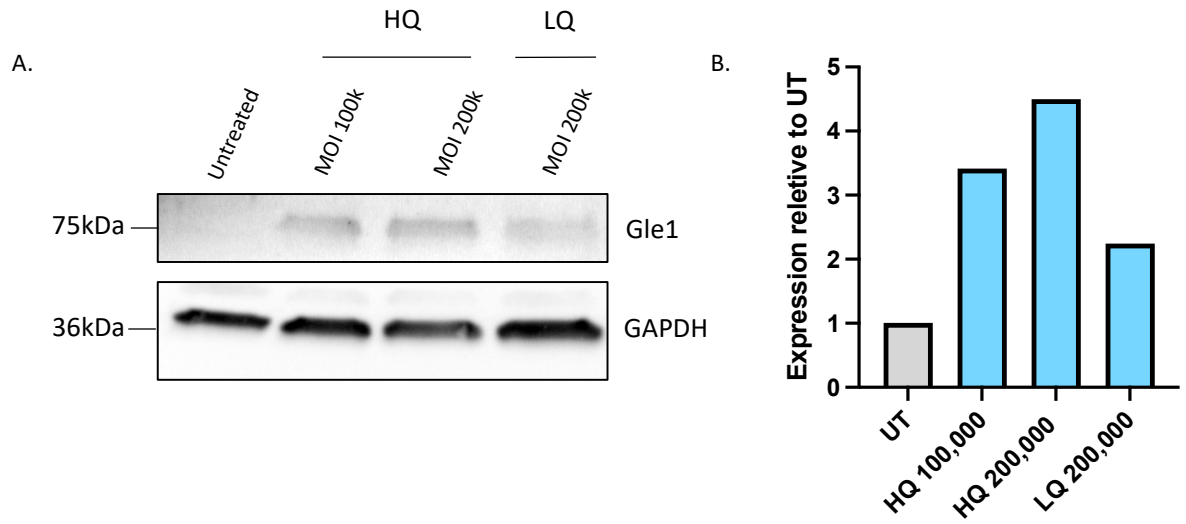


Figure 3.16: Western blot analysis of AAV9 Gle1b overexpression in wild-type mouse primary cortical neurons 7 days post transduction. (A) *Western blot probed with anti-Gle1 antibody, showing protein overexpression at 75kDa. GAPDH (at 36kDa) is used to show equal loading. (**B**) Quantification by densitometric analysis of one biological repeat (n=1), normalised to GAPDH levels. All values shown relative to Untransduced (UT).*

AAV9-Gle1b-mediated transduction efficiency was confirmed through immunofluorescence staining with a Gle1 antibody (**Fig 3.17 A, B**). Again, the overexpression of the transgene is expressed in both the nucleus and the cytoplasm.

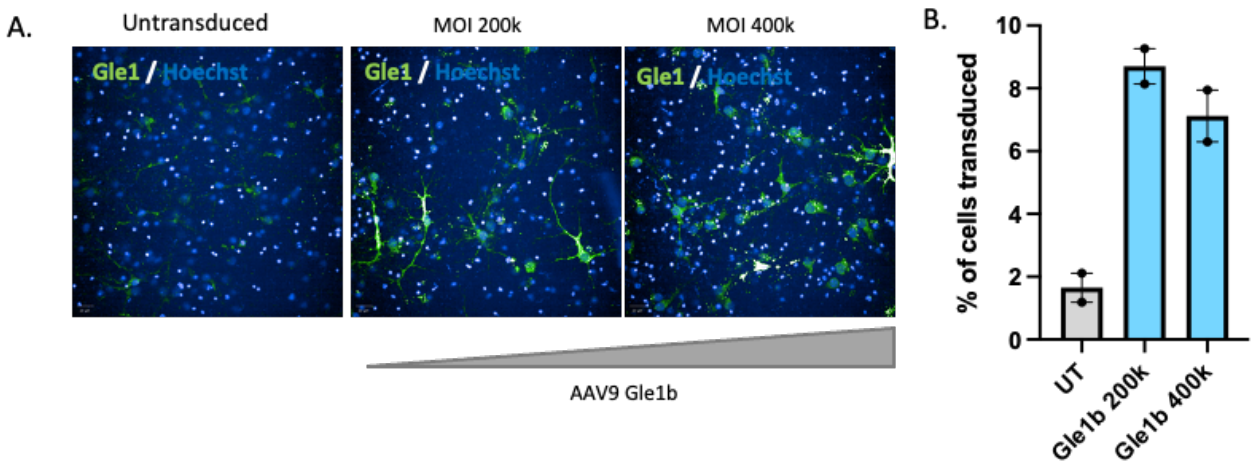


Figure 3.17: Validation of AAV9-Gle1b overexpression in primary cortical neurons 7 days following AAV transduction. (A) *Transduced cells shown in green through staining with an anti-Gle1 antibody. Hoechst staining of nuclei in blue. (**B**) Quantification of cells which were positive (above the level of the endogenous Gle1 staining seen in the untransduced condition) for Gle1 staining indicating transduction, data from two technical repeats, 32 fields per repeat, at least 500 cells per condition.*

3.3 Discussion

This chapter's focus was on describing the stages gone through to generate and optimise the plasmids and vectors which will be used throughout this project. This included design, generation and production of lentiviral and adeno-associated viral vectors aiming to overexpress the Gle1b protein, alongside corresponding controls. A plasmid overexpressing a GFP fusion protein was also developed with the aim of using this for immunoprecipitation experiments.

3.3.1 Choosing which Isoform of Gle1 to overexpress

Human Gle1 protein has two isoforms: hGle1a and hGle1b. These differ by a 43-amino acid domain at the C-terminus of the hGle1b isoform. qPCR analysis of HeLa cells demonstrated a 1000-fold prevalence of the hGle1b isoform in comparison to the levels of hGle1a, thereby suggesting that this is the dominant isoform (Kendirgi et al., 2003). Due to the similarity of these two isoforms, antibody recognition of specific isoforms is unfortunately not possible other than localisation within the cell. Therefore, for all overexpression constructs we decided to go forward with the more prevalent isoform, hGle1b, as this is the isoform responsible for the major roles of the *GLE1* gene (specifically nucleocytoplasmic shuttling of mRNA) (Kendirgi et al., 2005).

Diseases caused by mutations within the *GLE1* gene are spread across the different domains of the protein. However, many of the mutations which have been seen in LAAHD and ALS patients occur within the C-terminal domain, in particular the hCG1 binding domain which is specific to the hGle1b isoform (Kaneb et al., 2015; Nousiainen et al., 2008). This therefore could imply that it is loss of these isoform specific functions which are crucial to the survival of the motor neurons.

The Gle1b isoform is the predominant human isoform localised to the nuclear pore complex on the nuclear membrane. However, in this study to appears that overexpression of these Gle1b constructs is occurring predominantly in the cytoplasm. This could potentially result in the overexpressed Gle1b protein acting in a manner more comparable to Gle1a due to its localisation within the cell. A more thorough investigation of the impact of Gle1

overexpression would need to be carried out to fully establish which pathways the overexpressed protein could be implicated in as this may affect interpretation of later results with regards to isoform specific roles.

3.3.2 Production of Lentiviral vectors for *in vitro* study

Firstly, for the *in vitro* studies we aimed to produce a FLAG-tagged lentivirus overexpressing the hGle1b protein isoform. Lentiviral vectors are well characterised as delivery vectors to cells which are both dividing and nondividing (Azzouz et al., 2002). They have also been shown to have high transduction efficiency in neuronal cell populations (Naldini et al., 1996) and have been tested in clinical therapy's for neurological diseases such as Parkinson's disease (PD) (Palfi et al., 2018, 2014). More recent developments in the field have favoured different delivery vectors for clinical therapies as lentivirus has a predisposition to integrate randomly into genome. Multiple safety concerns have been raised due to the unpredictable nature of where these viruses integrate and the long term effects this may cause (Ciuffi, 2008; Sinn et al., 2005). In addition to this there is the substantial host immune response to both the vector and the transgene which needs to be considered when assessing the safety of these vectors *in vivo* (Annoni et al., 2019).

However, in our case for *in vitro* study, lentivirus provides a suitable method for efficient transduction of cell lines and primary neurons. These vectors can be produced in house at high titres. In this chapter, I have demonstrated the production and validation of two different lentiviruses, namely LV FLAG Gle1b and LV GFP. Both viruses were tested for the optimal dose which would give a good percentage of transduction in the desired cell populations, in this case fibroblast cells. This gave us a comparable percentage of transduced cells between the GFP and FLAG-tagged Gle1b overexpression, thereby allowing us to account for any variation in the observed conditions which may be being caused by the presence of the lentivirus itself rather than that of the transgene.

Nevertheless, there is still the expression of the GFP protein within the control virus which, in some cases, may cause increased immunogenicity, cytotoxicity or change the expression of specific markers which could be evaluated (Ansari et al., 2016). Therefore, this raises the question of how useful this is as a control in this situation. Ideally our lentiviral overexpression

vector would also contain a GFP tag as this would mean that any toxicity seen due to GFP would be equal across the two experimental conditions.

However, as we at this point in the project, did not have a vector which expressed a GFP-tagged Gle1 we decided to use the FLAG-tagged construct which we already had in our possession. Unfortunately, we were unable to produce a lentivirus expressing the 3xFLAG-tag only as this protein sequence is too short to be expressed alone. This would have been problematic as we would have had no way of detecting which cells were transduced within the population, therefore making comparison with the positively transduced FLAG-tagged cells difficult. For the experiments which are outlined in the reminder of this thesis we therefore use the LV-GFP virus as a control for our LV FLAG Gle1b virus.

It is also important to note that the optimal MOI for these viruses was tested in only one healthy line of fibroblasts (GM8680). Fibroblast cells can vary greatly in their characteristics depending on multiple factors (e.g., tissue type the sample was derived from, age of patient, genetic backgrounds). Therefore, it would have been optimal to have tested each virus in each of the fibroblast cell types which were to be used for experiments to ensure that the transduction efficiency was comparable across all groups.

3.3.3 Production of Adeno-associated viral vectors

Recombinant adeno-associated virus (rAAV) is the favoured vector for in vivo work, and therefore also clinical studies, as it functions episomally within the cytoplasm of the transduced cell using the cells machinery to produce the desired transgene (Wang et al., 2019). In the field of neuroscience, AAV9 is the preferred capsid as this preferentially transduces neuronal cell populations and has been shown to be able to cross the blood brain barrier *in vivo* (Aschauer et al., 2013; Foust et al., 2010; Valori et al., 2010).

In this chapter, I show the production of a functional AAV9-Gle1b virus which was produced to a high titre and quality for *in vivo* work in the future. This virus was validated in mouse primary cortical neurons as it was produced with the goal of injecting and transducing mouse brain and spinal cord tissue. Transduction with both low- and high-quality viruses demonstrates dose-dependent overexpression of Gle1b protein. Unfortunately, due to

COVID-19 related time restraints a mouse study was not anymore feasible, and so we did not use this virus for *in vivo* study during this PhD.

3.3.4 GFP-tagged Gle1b for immunoprecipitation

A GFP-based immunoprecipitation was chosen for the experiments performed in the following chapters as GFP-immunoprecipitations had already successfully been performed in the lab using GFP-tagged SMN constructs which we could use straight away to probe for Gle1 interaction. For ease of comparison, we therefore wanted to clone the Gle1b coding sequence into this same backbone. Our construct contains a N-terminal GFP tag, which according to the protein structure of the *GLE1* gene should mean it is located adjacent to the hNup155 binding domain (Nousiainen et al., 2008). This therefore may have caused disruption to Gle1b localisation and interaction with some of its known binding partners. It has also been demonstrated in the past that both N-terminal and C-terminal GFP tags can impact the native subcellular localisation of proteins (Hanson and Ziegler, 2004; Palmer and Freeman, 2004). I was able to successfully determine GFP expression by both immunofluorescence and through the eye piece of a fluorescent microscope with the cells in culture, therefore indicating that expression of the GFP tag itself was in no way impaired by the protein structure. Despite the limitations of using GFP-tagged constructs, this allowed us to produce a functional construct which we could easily perform immunoprecipitation experiments within the lab.

3.3.5 Summary

The experiments described in this chapter show successful design, generation and in vitro validation of tools which enable investigation into the dynamics between the Gle1 and SMN proteins in more depth and how they might function in a disease context. These constructs will be used throughout the rest of this thesis to test our hypotheses.

4. The Relationship between Gle1 and SMN

4.1. Introduction

Co-transcriptional processing of RNA is a complex, dynamic and multileveled process which is crucial for correct protein expression. RNA binding proteins (RBPs) are proteins which bind mRNA through highly conserved RNA-binding domains (Ravanidis et al., 2018). These proteins and their associated factors enable cells to regulate RNA transcription, post-transcriptional processing events (such as capping, splicing/alternative splicing and polyadenylation), mRNA nuclear export and translation prior to any post translational modifications at the protein level.

hGle1a and hGle1b have been implicated in modulation of stress granules through translation, and nucleocytoplasmic export of mRNA respectively (Aditi et al., 2015; Kendirgi et al., 2003). The role of SMN in the spliceosome and RNP biogenesis has been well characterised, with additional diverse functions in mRNA trafficking, stress granule assembly and local translation being more recent discoveries. Mutations within either of these proteins have been implicated in diseases, namely LCCS and SMA, which selectively affect motor neurons.

Two studies published during this PhD project begin to suggest that that knockdown of Gle1 may affect the expression levels of *SMN1* (Okamura et al., 2018), and that *GLE1* mRNA levels are reduced in SMA patient cells (Alrafiah et al., 2018a). However, beyond these initial findings the relationship between Gle1 and SMN has not been well studied. Due to the similarities in the biological function and the diseases which loss of either protein causes, the basis of my PhD was to try to understand whether there is any potential interaction between the two proteins.

4.1.2 Aims

The aims of this chapter are therefore to investigate whether there is any relationship between the Gle1 and SMN proteins. This will be investigated by asking the following questions:

1. Do Gle1 and SMN interact with each other?
2. Where abouts in the cell might this interaction be occurring?
3. Are the levels of Gle1 protein and mRNA affected in SMA disease models?
4. Can overexpression of one protein rescue loss of the other?

4.2 Results

4.2.1 Gle1 and SMN proteins interact

As there are similarities between the different functions of the Gle1 and SMN proteins, I was interested to investigate whether Gle1 and SMN play a common role in any shared pathways. Therefore, I looked at whether the two proteins interact.

By using a plasmid which overexpresses a GFP-SMN fusion protein, I was able to perform a GFP immunoprecipitation and probe for both known binding partners of SMN and our protein of interest, Gle1 (**Fig 4.1**). HEK293 cells were transfected with pEGFP and pEGFP-SMN plasmids for 48 hours prior to harvest and protein lysis. Total protein lysate was held back prior to incubation with the ChromoTek-GFP beads to run alongside the co-Immunoprecipitation (co-IP) samples as an input reference. This ensured that the inputs were equally expressing the proteins we were probing for and that they were loaded evenly. Protein lysates were also treated with RNase A to ensure that any interactions observed were due to protein-protein interactions rather than being bound through an RNA-intermediate.

Enrichment of the GFP-tagged SMN protein was confirmed through probing of the membrane with an anti-SMN antibody to show GFP-SMN overexpression at around 50 kDa, and an anti-GFP antibody to show GFP overexpression at around 27 kDa in the empty vector control samples. RNA helicase A was used as a positive control for a known binding partner of SMN (Pellizzoni et al., 2001b). Signal was seen for RNA helicase A (at around 155 kDa) both in the presence and absence of RNase A as expected. In the experimental condition of probing for endogenous Gle1 (at 75 kDa), there is a band in both the presence and absence of RNase A, indicative of interaction between the GFP-SMN and endogenous Gle1.

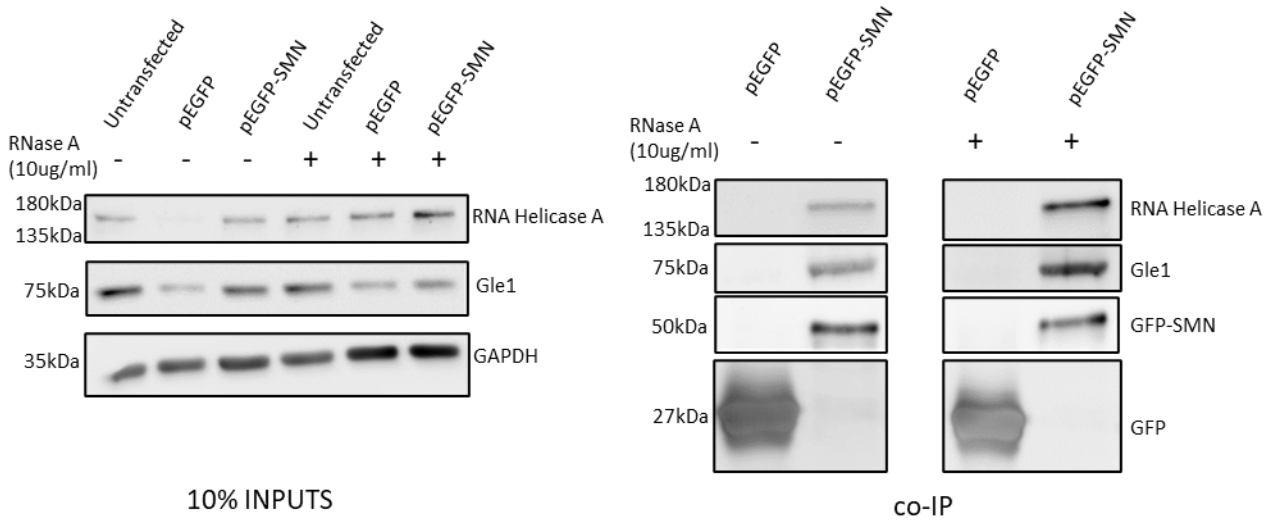


Figure 4.1: Immunoprecipitation of pEGFP-SMN plasmid shows presence of Gle1 protein. *GFP immunoprecipitation pulling down an overexpressed empty pEGFP construct or pEGFP-SMN. 10% protein inputs probed for RNA Helicase A, Gle1 and GAPDH to show even loading of IP protein. IP pellets probed for RNA helicase A as a published positive control with SMN. Interaction between Gle1 and SMN observed in pEGFP-SMN pellets both with and without addition of RNase A to the protein extracts prior to immunoprecipitation (n=3).*

To confirm the interaction seen by overexpression of GFP-tagged SMN protein, I decided to repeat this experiment in the reverse configuration. To do this I cloned the more abundant isoform of human Gle1 (hGle1b), into the same pEGFP vector as used for the SMN coding sequence (as described in **Chapter 3**). This construct was then used as previously with pEGFP-SMN. Again, protein lysate was taken aside prior to incubation with the GFP-beads to check the conditions were equally incubated and loaded on the gel. Samples were also treated with RNase A to ensure any interactions seen were protein-protein and not simply due to an RNA-intermediate. Probing of the membrane with an anti-Gle1 antibody demonstrated an enrichment within the IP of overexpressed GFP-tagged Gle1b protein at around 100 kDa, alongside endogenous Gle1 at 75 kDa. Probing this same lysate with an anti-SMN antibody showed signal for samples both untreated and treated with RNase A, indicating that GFP-Gle1b was interacting with endogenous SMN (**Fig 4.2**).

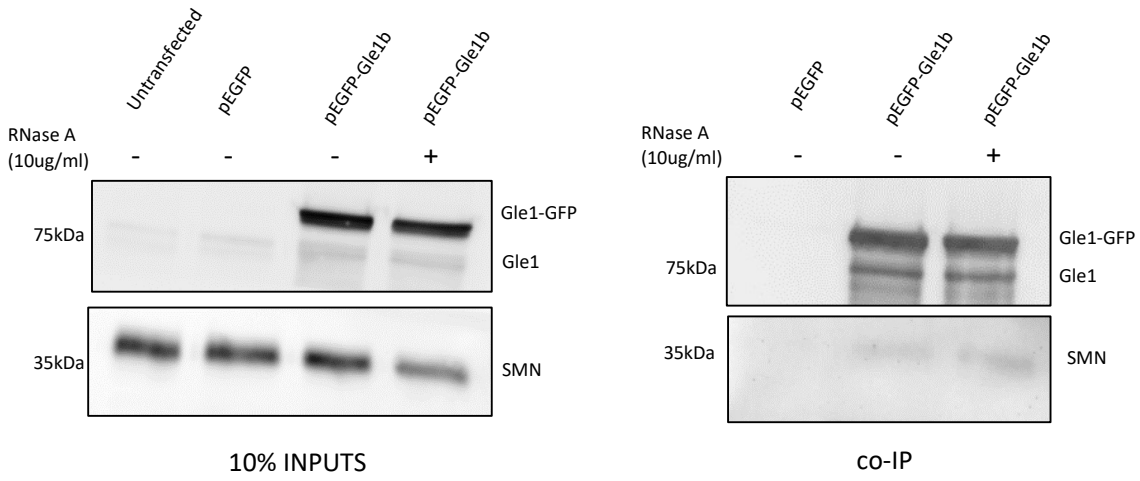


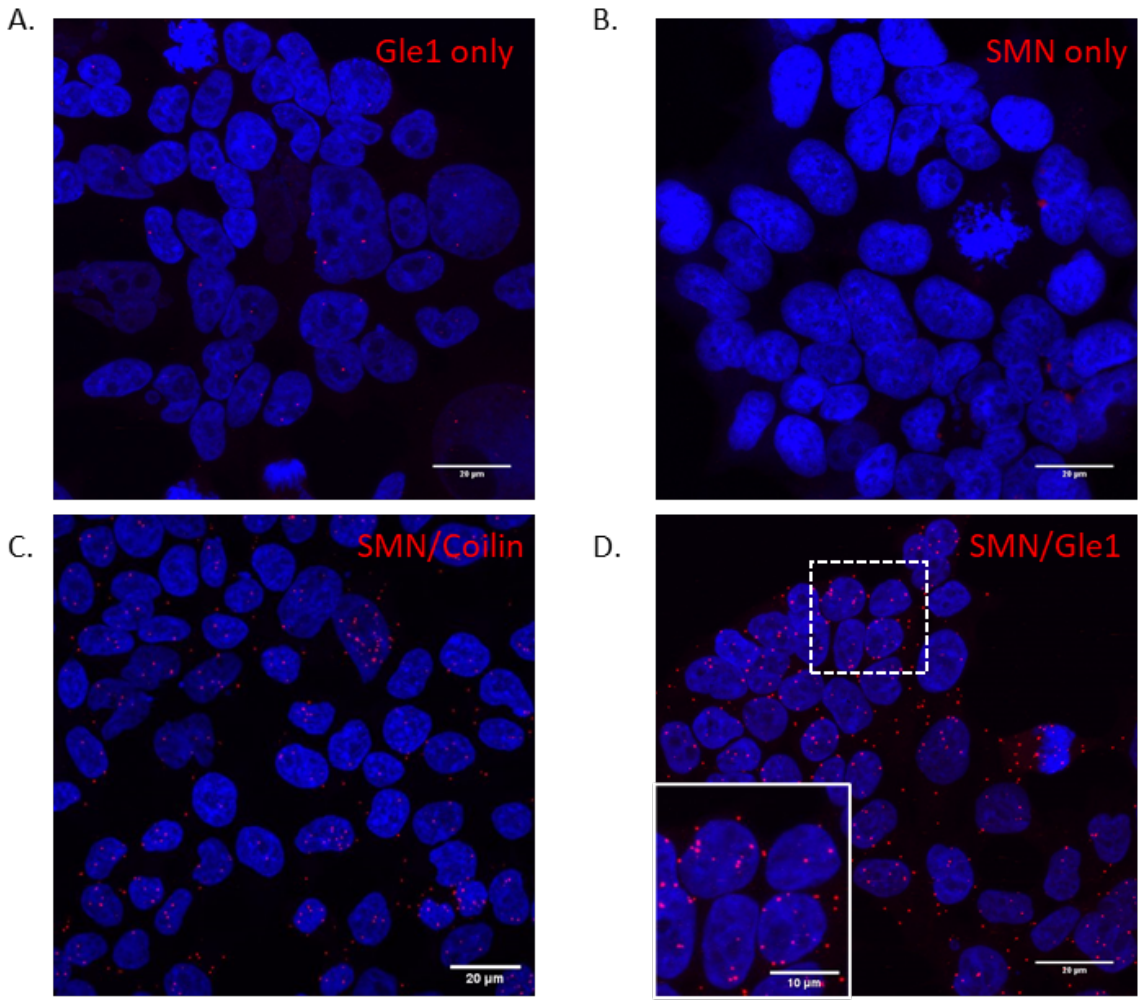
Figure 4.2: Immunoprecipitation of pEGFP-Gle1b plasmid shows presence of SMN protein. *GFP immunoprecipitation pulling down an overexpressed empty pEGFP construct or pEGFP-SMN. 10% protein inputs Gle1 and GAPDH to show even loading of IP protein. Interaction between Gle1b and SMN observed in pEGFP-SMN pellets both with and without addition of RNase A to the protein extracts prior to immunoprecipitation (n=3).*

Following these experiments looking at interactions based on overexpression of a tagged protein construct, I wanted to look at whether this interaction occurs on an endogenous level. To investigate this, I used a proximity ligation assay (PLA). This is an assay which allows visualisation proteins in situ which are in very close proximity (40 nm). Two protein targets can be labelled by using two primary antibodies of different species, PLA probes complimentary to these species then can bind the selected primary antibodies. A hybridisation reaction can then be carried out, hybridizing oligos can connect if the PLA probes are within a close enough proximity to one another. Incubation with a ligase can then form a circular DNA template necessary for rolling-circle amplification (RCA). Addition of a DNA polymerase generates concatemeric sequences during RCA, amplifying the signal from the PLA probe. Hybridisation of a labelled oligo to these sequences therefore allows localisation of the signal, which can then be quantified and evaluated. (Sigma, 2021).

To further investigate the interaction observed between Gle1 and SMN, a PLA was performed across a variety of different cell models to assess the localisation of the proteins and how it varies across cell types and conditions.

Firstly, I used HEK293 cells to assess whether the two proteins could be seen to localise and where abouts in the cell this occurred. For PLA, cells were treated with either anti-Gle1 antibody, anti-SMN antibody, or both anti-SMN and anti-Gle1 antibodies together. Anti-SMN and Anti-coilin antibodies were also used as a positive control to demonstrate the assay had worked in the absence of any signal in the experimental condition (Hebert et al., 2001). The PLA was performed as previously explained and signal was visualised through confocal microscopy. Due to the majority of the protein structure being the same between hGle1a and hGle1b, antibodies for this protein recognise both isoforms.

Minimal signal was seen in the negative control conditions of each antibody alone: a small number of foci could be observed in the anti-Gle1 antibody condition (**Fig 4.3 A**), with none seen in the anti-SMN antibody condition (**Fig 4.3 B**). Much stronger signal was seen in the conditions treated with both antibodies, both in the experimental condition and the positive control (**Fig 4.3 C, D)**, suggesting that SMN and Gle1 proteins are occurring in close proximity within the cell. Signal distribution was quantified looking at the distribution between the nucleus and cytoplasm, suggesting that the interaction occurs in equally in both the nucleus and the cytoplasm (**Fig 4.3 E**).



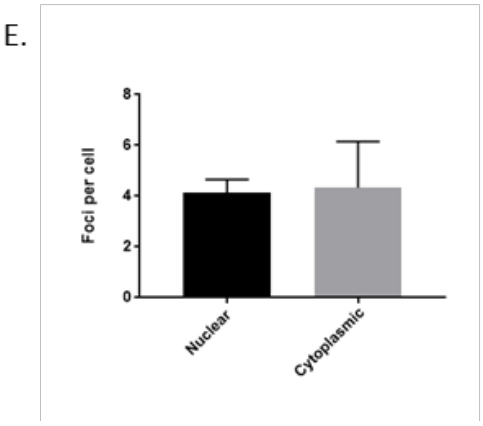
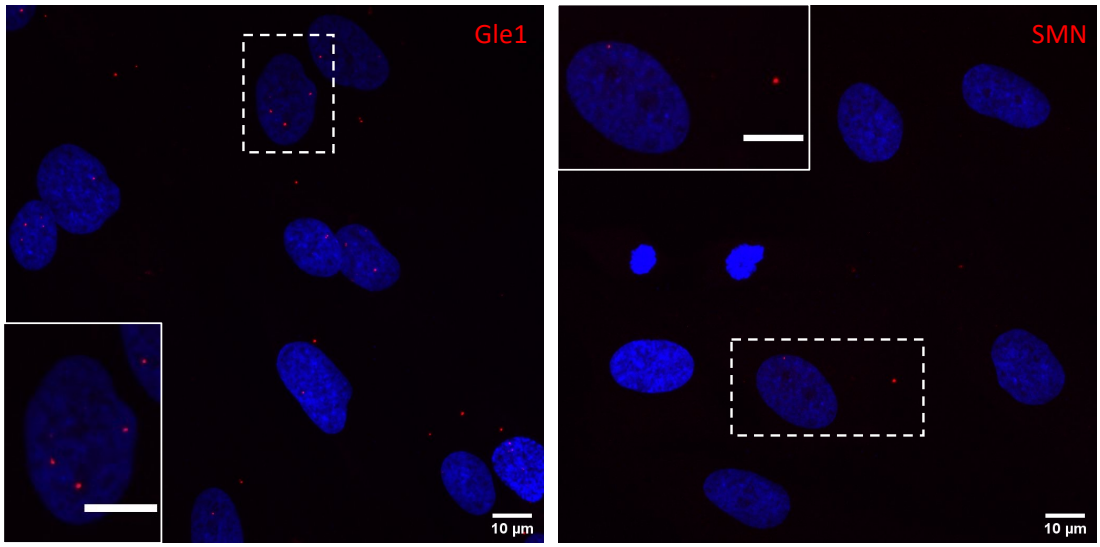


Figure 4.3: Proximity ligation assay showing co-localisation of Gle1 and SMN protein in HEK cells. *Negative controls with (**A**) Gle1 antibody only (**B**) SMN antibody only, and the positive control of (**C**) SMN and coilin, **(D)** experimental condition of Gle1 and SMN antibodies together. The red dots are indicative of co-localisation of Gle1 and SMN proteins. White box indicates which area of the image is shown to the right at a higher magnification. (**E**) Quantification of*

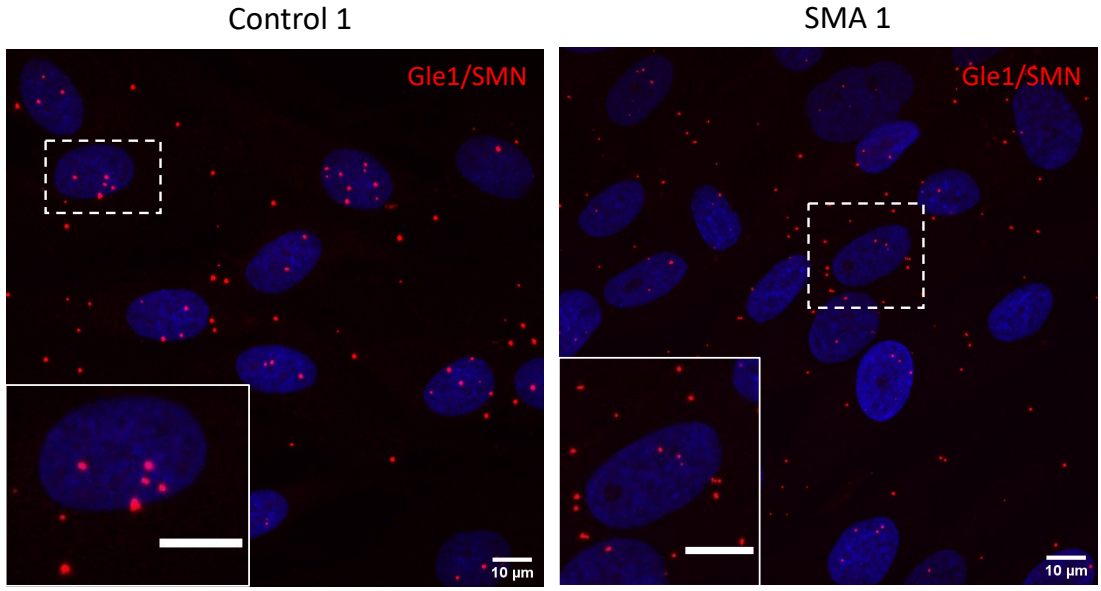
nuclear vs non-nuclear foci per cell (n=3 biological repeats, 50 cells per condition). Data displayed as mean +/- SEM.

Next, we were interested in whether the interaction between Gle1 and SMN was affected by the reduction of SMN in SMA patient cells. To investigate this, we used fibroblast cells taken from SMA patients. As seen in the HEK293 cells previously, we observed little to no signal in the negative control conditions (**Fig 4.4 A, B**). Signal in the experimental condition demonstrated that again Gle1 and SMN appeared to colocalise evenly across the nucleus and cytoplasm (**Fig 4.4 C-E**). The levels of signal between the control fibroblast cells and the Type I SMA patient cells did not appear to be different between the conditions. There is, however, a greater level of variability in these cells than the HEK293s. One repeat here had to be excluded due to problems with the kit/antibodies expiring/losing efficiency following the COVID-19 pandemic.

A.


Gle1
SMN
10 µm
10 µm

B.


Control 1
SMA 1
Gle1/SMN
Gle1/SMN
10 µm
10 µm

C.

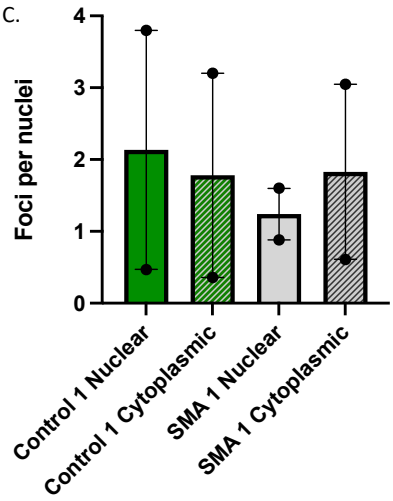
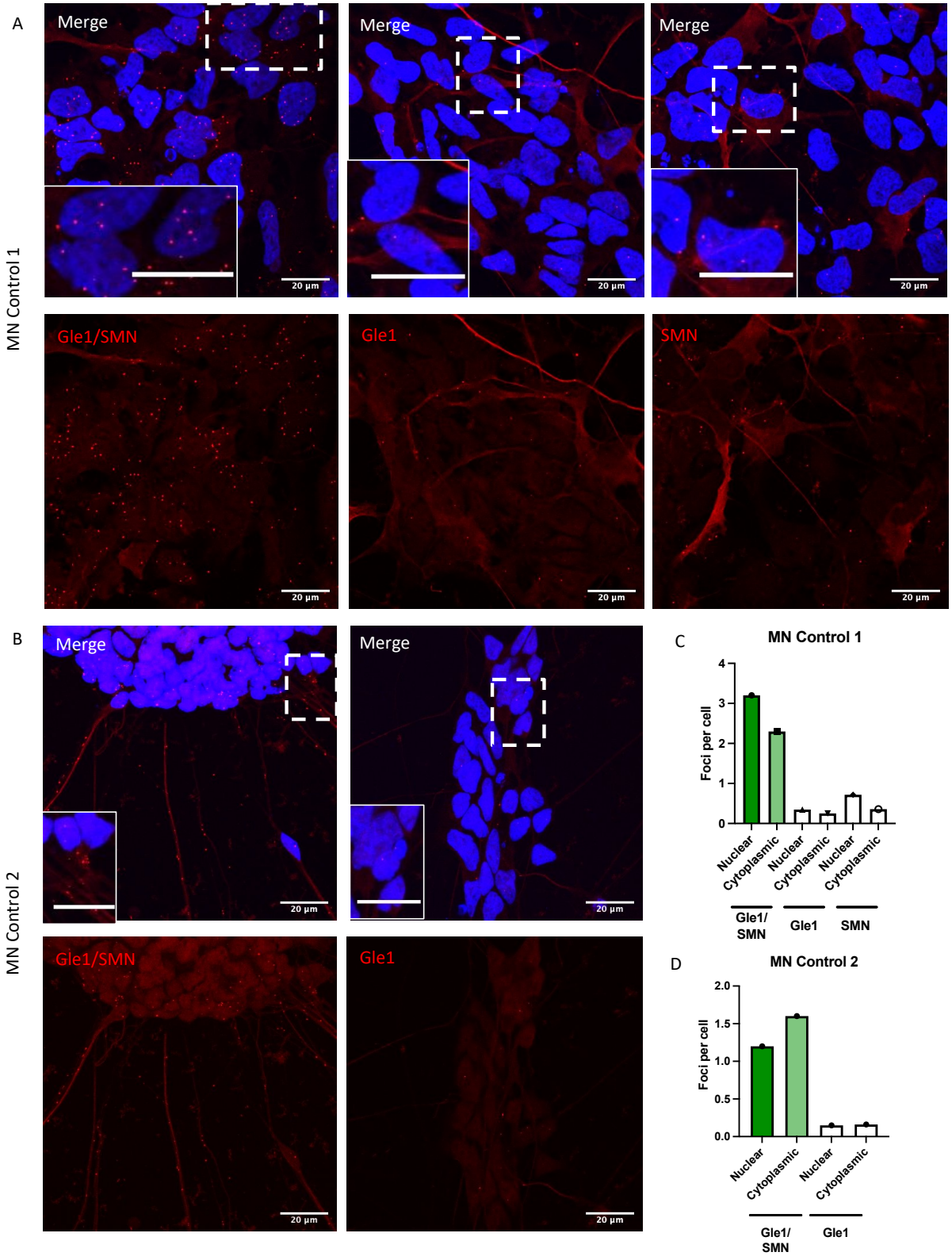
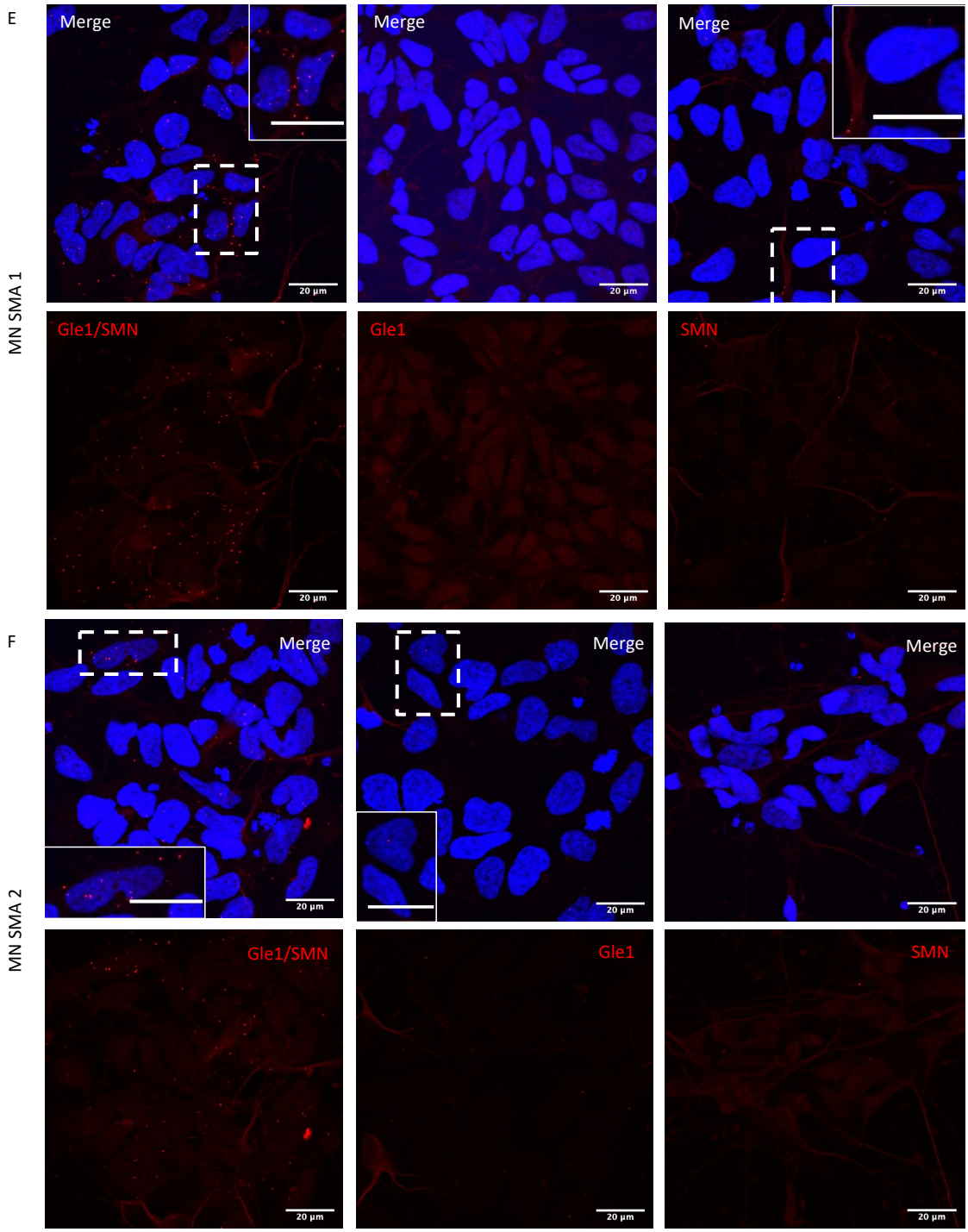

Foci per nuclei
0
1
2
3
4
Control 1 Nuclear
Control 1 Cytoplasmic
SMA 1 Nuclear
SMA 1 Cytoplasmic

Figure 4.4: Proximity ligation assay showing co-localisation of Gle1 and SMN protein in fibroblasts. (A) *Negative controls with Gle1 antibody only and SMN antibody only in control fibroblast cells (GM0489).* ***(B)*** *Experimental condition of Gle1 and SMN antibody together in control cells (GM0489) and SMA cells (GM09677). The red dots are indicative of co-localisation between Gle1 and SMN proteins. White box indicates which area of the image is shown to the right at a higher magnification. Scale bars shown as 10 µm. (**E**) Quantification of nuclear vs non-nuclear foci per cell (n=2 biological repeat, 50 cells per condition). Data displayed as mean +/- SEM.*

As the phenotypes often observed in SMA cases are specific to motor neurons, we also wanted to assess whether the localisation or levels of signal varied in control and SMA patient motor neurons. The location of the interaction in these cells could potentially point us towards a common pathway specific to motor neurons.

Here I looked at proximity of Gle1 and SMN protein in two control motor neuron populations: MN Control 1 / Cs14 and MN Control 2 / GM, alongside two SMA patient motor neuron populations: MN SMA 1 / SMA 77 and MN SMA 2 / SMA 84. As previously in the other cell types there is little to no signal observed in the single antibody only conditions which indicates our signal is showing true co-localisation of the two proteins (**Fig 4.5 A-H**). There is Gle1/SMN PLA signal within the nuclei, cytoplasm, and axons of the motor neurons of both control and SMA cells (**Fig 4.5**). However due to COVID-19 related time restraints and problems with delivery's, this data is only preliminary (n=1) and requires further replicates to draw any conclusions as to whether this interaction is disrupted in SMA cells.


A
MN Control 1
Merge
Merge
Merge
Gle1/SMN
Gle1
SMN
20 µm
B
MN Control 2
Merge
Merge
Gle1/SMN
Gle1
C
MN Control 1
Foci per cell
Nuclear
Cytoplasmic
Gle1/ SMN
Gle1
SMN
D
MN Control 2
Foci per cell
Nuclear
Cytoplasmic
Gle1/ SMN
Gle1


E
MN SMA 1
Merge
Merge
Merge
Gle1/SMN
Gle1
SMN
F
MN SMA 2
Merge
Merge
Merge
Gle1/SMN
Gle1
SMN
20 µm

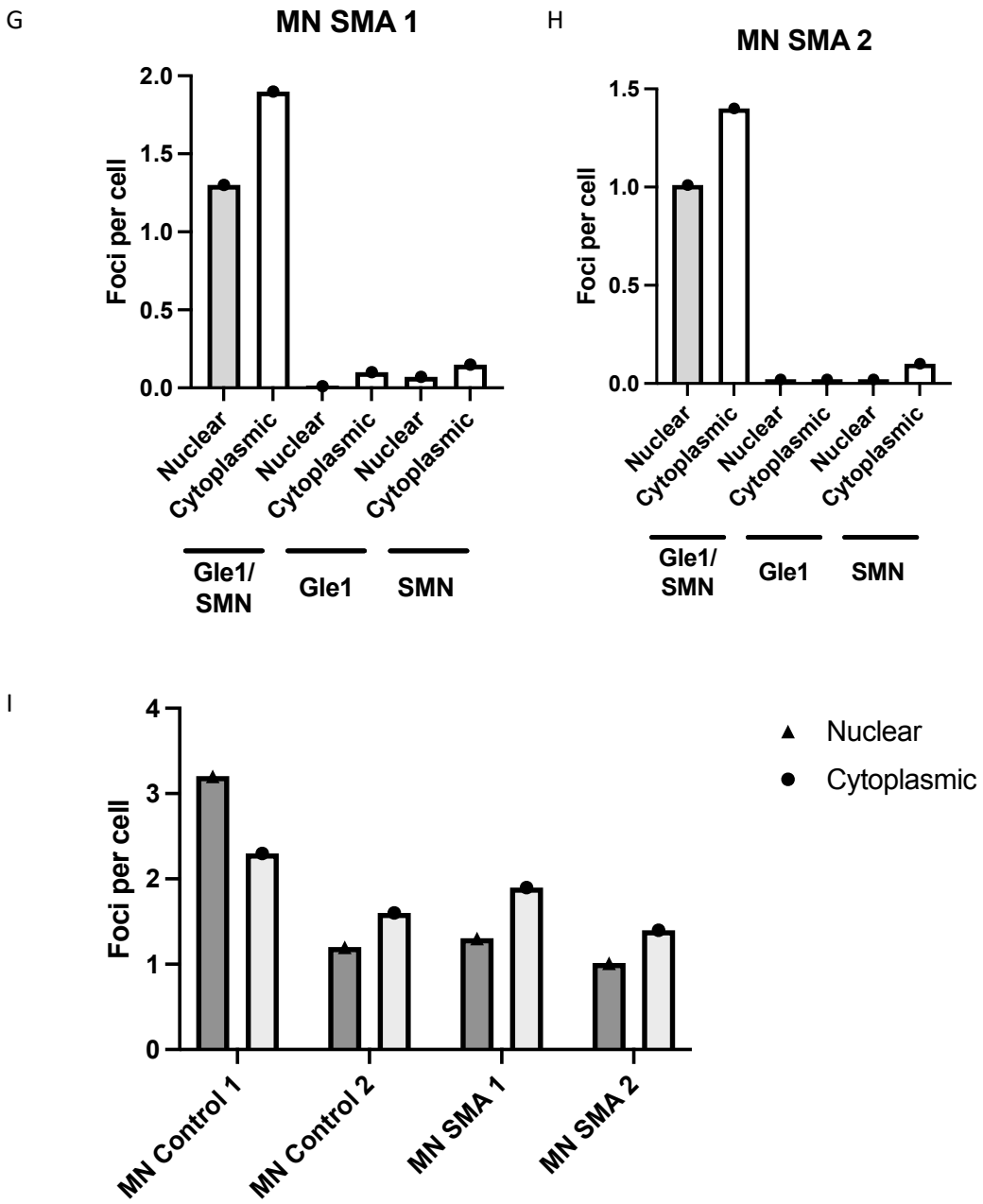


Figure 4.5: Proximity ligation assay showing co-localisation of Gle1 and SMN protein in motor neurons. *The red dots are indicative of interaction between Gle1 and SMN proteins. White box indicates which area of the image is shown to the right at a higher magnification. All scale bars 20 µm. All data displayed as mean. Background red staining from MAP2 axonal marker on far-red (647) channel. (**A**) MN Control 1 (Cs14) Gle1/SMN signal, Gle1 antibody alone negative control, SMN antibody alone negative control. (**B**) MN Control 2 (GM) Gle1/SMN signal, Gle1 antibody alone negative control. (**C**) Quantification of foci per cell in*

*MN Control 1 (Cs14) conditions (n=1, 50 cells per condition). (**D**) Quantification of foci per cell in MN Control 2 (Cs14) conditions (n=1). (**E**) MN SMA 1 (SMA 77) Gle1/SMN signal, Gle1 antibody alone negative control, SMN antibody alone negative control. (**F**) MN SMA 2 (SMA 84) Gle1/SMN signal, Gle1 antibody alone negative control. (**G**) Quantification of foci per cell in MN SMA 1 (SMA 77) conditions (n=1, 50 cells per condition). (**H**) Quantification of foci per cell in MN SMA 2 (SMA 84) conditions (n=1, 50 cells per condition). (**I**) Quantification of nuclear vs non-nuclear foci per cell.*

4.2.2 Gle1 levels are reduced in SMA cell models

As there appears to be a clear interaction between Gle1 and SMN proteins. I then wanted to look at whether the expression of *GLE1* was altered in SMA patients compared to control samples given that patients have reduced expression of SMN.

As these two genes are both ubiquitously expressed, I decided to assess whether the levels of Gle1 were altered in a cell type which was representative of a more general cell type rather than a post-mitotic specialised cell type. To test this, I looked at both endogenous protein expression and mRNA transcript levels in control and type I SMA patient fibroblasts. Where possible, each SMA fibroblast line was matched to an age and sex match control.

Protein lysates from each sample were extracted and run on a western blot. The membrane was then probed with an anti-Gle1 antibody to look at endogenous protein expression. Tubulin was used as a loading control to ensure equal loading across the samples, this was used for normalisation during the quantification of signal. In both pairs of age matched cells, SMA patients demonstrate a significant reduction of around 30-50% in Gle1 protein levels when compared to the control cells (Control 1 vs SMA 1 (paired t-test, p=0.0014, **, n=5), Control 2 vs SMA 2 (paired t-test, p=0.0312, *, n=4)). (**Fig 4.6 A-B**). The averages from each cell type were also grouped and compared; showing that although there is variability in expression across the controls, the expression levels in the SMA cells were both reduced comparatively (**Fig 4.6 C**).

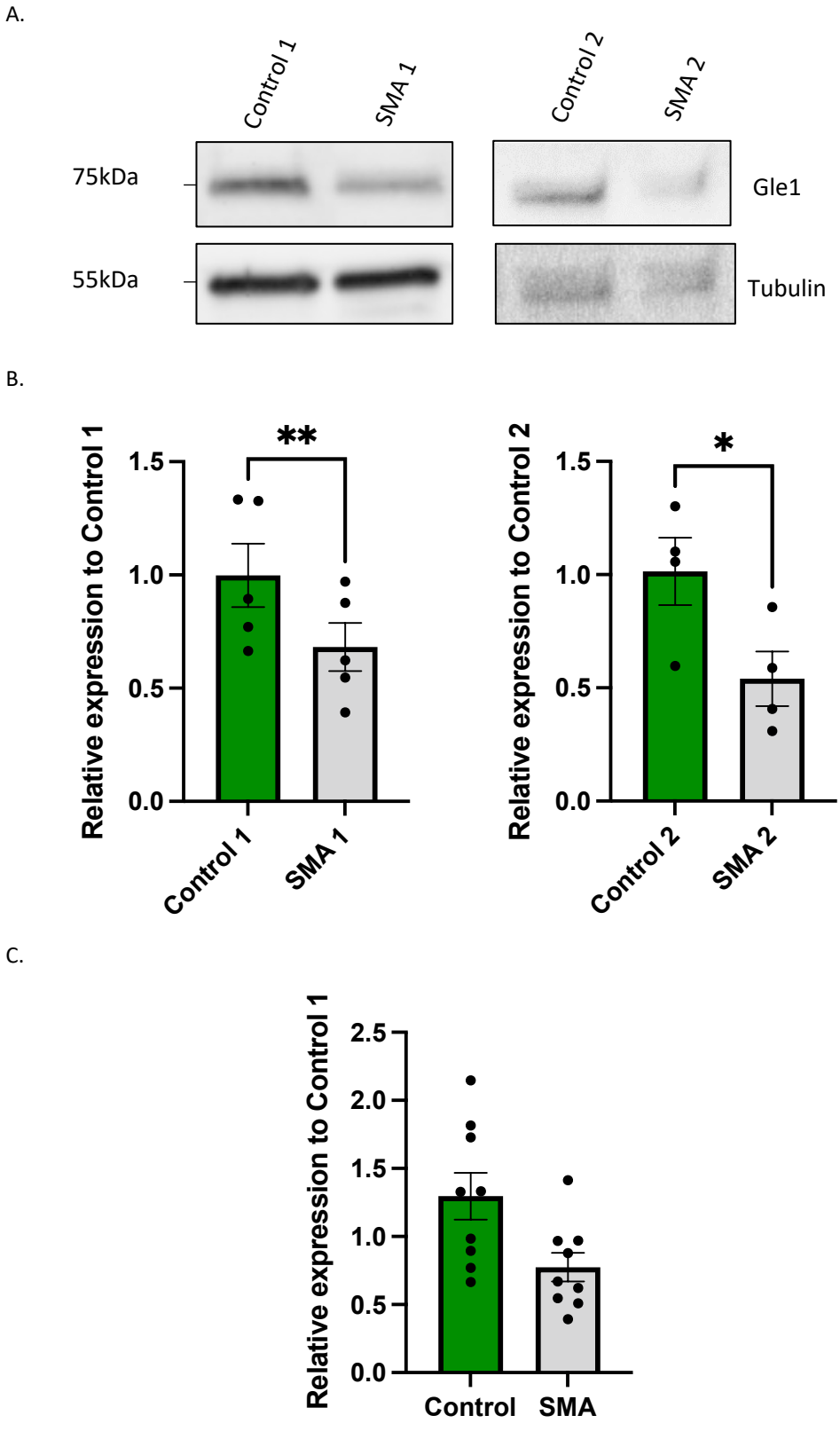


Figure 4.6: Gle1 protein levels are reduced in SMA fibroblast cells. (A) *Representative western blot showing protein expression in human fibroblast cells: Control 1 (GM0489), SMA 1 (GM09677), Control 2 (GM08680), SMA 2 (GM00232). Western blot probed with anti-Gle1*

antibody, showing protein expression of endogenous Gle1 at 75kDa. Tubulin (at 55kDa) is used to show equal loading of samples. ***(B)*** *Densitometric analysis shows a reduction in Gle1 protein levels across age-matched pairs. Control 1 vs SMA 1 (paired t-test, p=0.0014, **, n=5), Control 2 vs SMA 2 (paired t-test, p=0.0312, *, n=4). Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal.* ***(C)*** *Average values of densitometric analysis of each line shows consistency between the age matched pairs when all normalised to Control 1. Gle1 protein levels are reduced overall in SMA patient cells. Protein loading was normalised using Tubulin signal. Data displayed as mean +/- SEM.*

Complementary to western blots, cells were also processed in parallel for RNA extraction, *GLE1* mRNA transcript levels were measured through qPCR. By assessing both protein and mRNA expression in these cells we aimed to see where during the process of gene expression Gle1 levels are disrupted: at the transcript level, during RNA processing or during translation.

Again, cells were compared to their age and sex matched controls and data was normalised to a GAPDH loading control. Patient cells show a varied fold change expression of Gle1 mRNA. SMA 1 cells demonstrate 70% reduction compared to the Control 1 cells (unpaired t-test, p=0.0441, *; n=3), while the SMA 2 cells show an expression which is comparable to Control 2 cell (unpaired t-test, p=0.7353, ns; n=3) (**Fig 4.7 A, B**). The averages from each cell type were also grouped and compared; this shows that the levels of *GLE1* mRNA transcripts vary across both the control and SMA groups (**Fig 4.7 C**), in this case a third group of control and SMA cells would be useful to determine the average levels of transcripts across both groups of cells and whether this is reduced in patient cells.

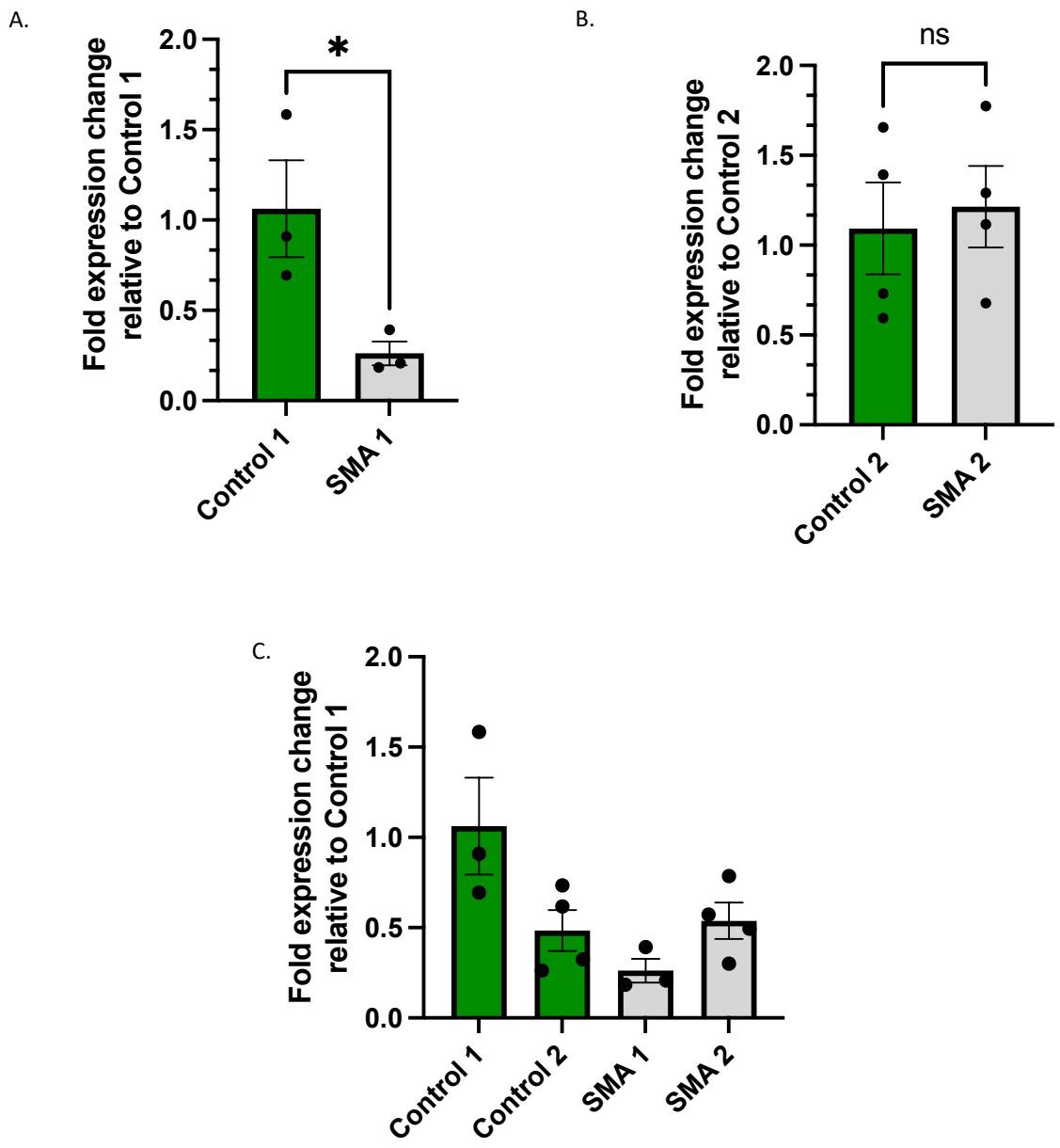


Figure 4.7: GLE1 mRNA levels are only reduced in one SMA patient sample. *Fold expression change worked out through delta delta CT analysis of qPCR data, data normalised to the Healthy cell condition in both cases.* ***(A)*** *SMA 1 (GM09677) show a 70% decrease in GLE1 mRNA levels compared against age-matched control GM0489 (unpaired t-test, p=0.0441, *, n=3). (**B**) SMA 2 (GM00232) however shows no difference in GLE1 mRNA expression levels compared to age matched control GM8680 (unpaired t-test, p=0.7353, ns, n=4). Data displayed as mean +/- SEM.* ***(C)*** *Average fold expression change of Control and SMA cells, normalised to Control 1.*

To assess whether this loss of Gle1 was specific to an isoform we performed a nuclear cytoplasmic fractionation of control and SMA patient fibroblast cells. As mentioned previously, there are two isoforms of human Gle1: hGle1b and hGle1a. The more abundant isoform, hGle1b, being located on the nuclear fibrils of the nuclear pore complex, with the hGle1a isoform found in the cytoplasm (Aditi et al., 2015; Kendirgi et al., 2005). I hypothesised that a nuclear loss would correspond to a loss of hGle1b, while a cytoplasmic loss could be attributed to a loss of the hGle1a isoform. This, in turn, could give us more clues as to which pathways may be affected in the SMA patient population by a reduction in Gle1 levels.

Successful fractionation of the nucleus and the cytoplasm of these cells was confirmed by western blotting (**Fig 4.8 A**) using an anti-SSRP1 antibody as a marker for the nucleus and an anti-Tuj1 antibody as a marker of cytoplasm. Unfortunately, due to the composition of fibroblast cells, I was unable to get a fully pure nuclear fraction as they are very heavily cytoplasmic. However, as the cytoplasmic fractions had no presence of nuclear protein it could still be used for RNA extraction and analysis by qPCR. As shown previously, there is a significant reduction in the total levels of *GLE1* mRNA (**Fig 4.8 B**). Additionally, there is a significant reduction in the levels of *GLE1* mRNA transcripts within the nucleus of the SMA patient cells (paired t-test, p=0.0192, *, n=3) (**Fig 4.8 C**). Cytoplasmic levels of *GLE1* mRNA transcripts are not significantly different from the controls (paired t-test, p>0.05, ns, n=3) (**Fig 4.8 D**).

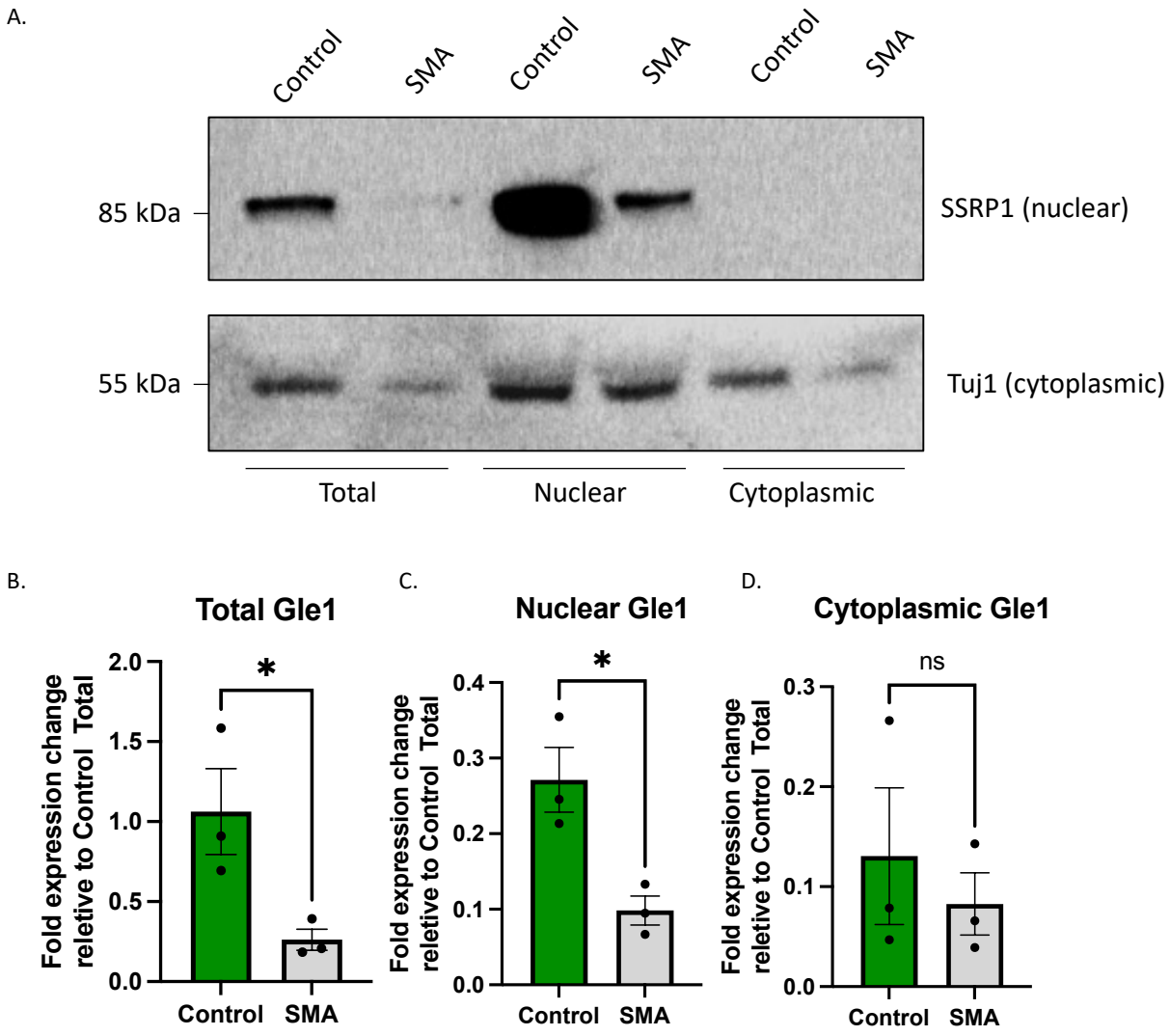


Figure 4.8: Gle1 mRNA levels are reduced specifically in the nucleus of SMA patient fibroblasts. (A) *Representative western blot showing successful fractionation of fibroblast samples: Control 1 (GM0489) and SMA 1 (GM09677). Western blot probed with anti-SSRP1 antibody to show nuclear protein expression at 85 kDa and anti-Tuj1 antibody to show cytoplasmic protein expression at 55 kDa.* ***(B)*** *Fold expression change worked out through delta delta CT analysis of qPCR data, normalised to the Control total condition in all cases. Data displayed as mean +/- SEM. GLE1 mRNA levels in the total fraction are significantly reduced in SMA cells compared to controls (unpaired t-test, p=0.0441, *, n=3).* ***(C)*** *GLE1 mRNA levels in the nuclear fraction are significantly reduced in SMA cells compared to controls (paired t-test, p=0.0192, *, n=3).* ***(D)*** *GLE1 mRNA levels in the cytoplasmic fraction are not significantly different in SMA cells compared to controls (paired t-test, p=0.6321, ns, n=3).*

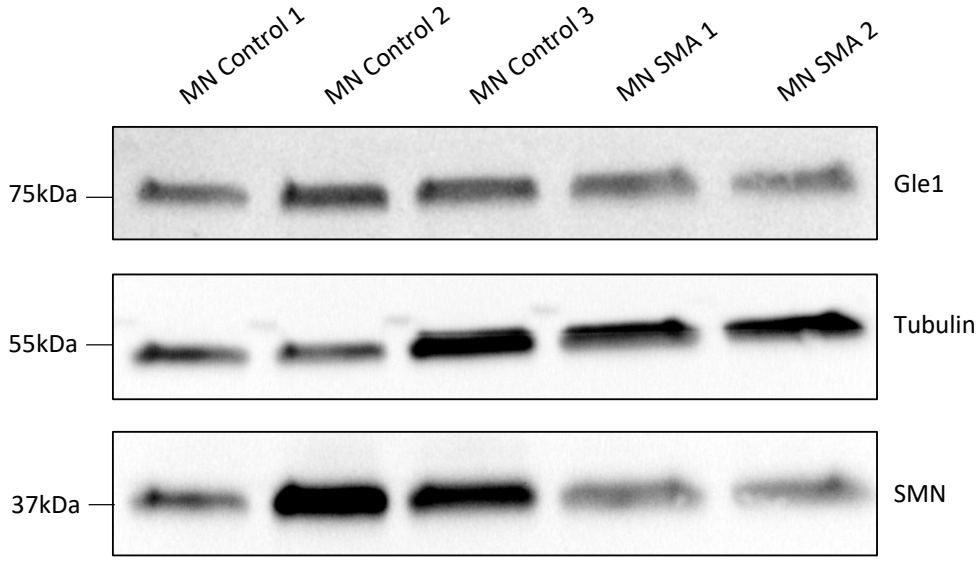
As loss of either Gle1 or SMN have been shown to give rise to diseases which specifically affect motor neurons, I wanted to assess whether Gle1 protein levels were also reduced in these cells.

Protein lysis was carried out of three control motor neuron populations: MIFF1 (MN Control 1), Cs14 (MN Control 2) and GM (MN Control 3), alongside two SMA patient motor neuron lines: SMA 77 (MN SMA 1) and SMA 84 (MN SMA 2). These experiments were performed on day 33 of culture, after motor neuron maturation, from three separate differentiations (n=3, biological replicates). As before, endogenous protein expression was calculated through densitometry analysis of western blots probed with an anti-Gle1 antibody. Here we also probed for endogenous SMN to confirm the phenotype of these cells (**Fig 4.9 A**).

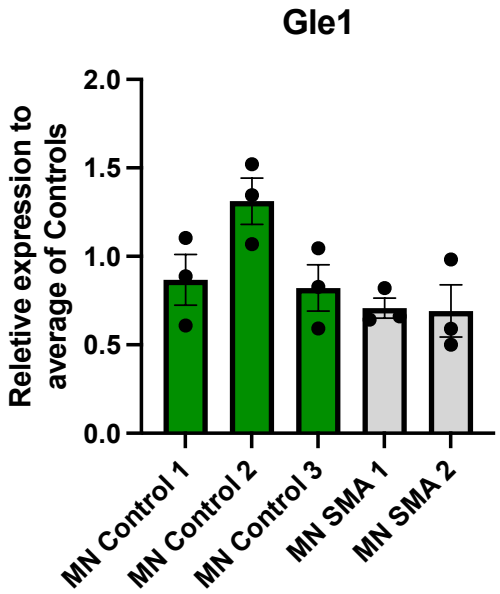
Quantification of Gle1 protein expression in these lines does not appear to give as striking a result as in fibroblasts with no significant difference being observed between the healthy and SMA motor neurons (RM one way ANOVA, $p>0.5$, ns, n=3) (**Fig 4.9 B**). Overall, when grouped together there is a trend towards there being around a 30% reduction in endogenous Gle1 protein expression compared to controls (**Fig 4.9 C**).

The same trend can be observed in the endogenous expression of SMN, with no significant difference being observed between the three control lines and the two SMA lines (RM one-way ANOVA, $p>0.05$, ns, n=3) (**Fig 4.9 D**). Overall there is a trend for reduction in SMN protein expression in SMA cases compared to controls (**Fig 4.9 E**).

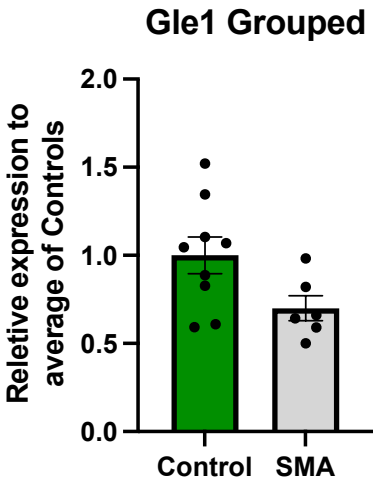
A.


MN Control 1
MN Control 2
MN Control 3
MN SMA 1
MN SMA 2
75kDa
Gle1
55kDa
Tubulin
37kDa
SMN

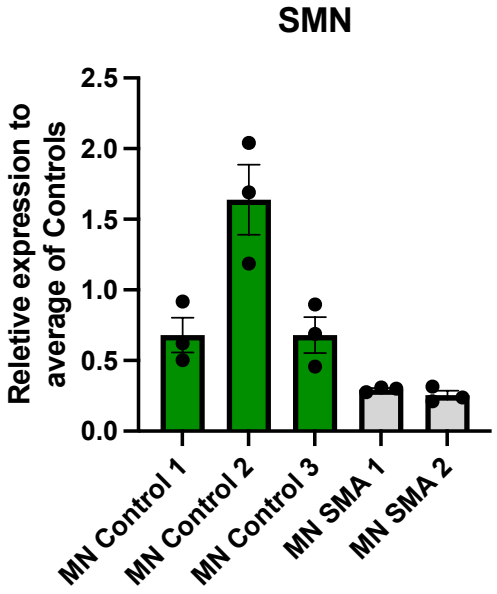
B.


Gle1
Relative expression to average of Controls
2.0
1.5
1.0
0.5
0.0
MN Control 1
MN Control 2
MN Control 3
MN SMA 1
MN SMA 2

C.


Gle1 Grouped
Relative expression to average of Controls
2.0
1.5
1.0
0.5
0.0
Control
SMA

D.


SMN
Relative expression to average of Controls
2.5
2.0
1.5
1.0
0.5
0.0
MN Control 1
MN Control 2
MN Control 3
MN SMA 1
MN SMA 2

E.

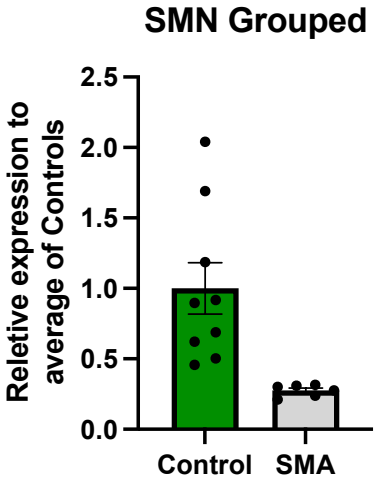

SMN Grouped
Relative expression to average of Controls
2.5
2.0
1.5
1.0
0.5
0.0
Control
SMA

Figure 4.9: Gle1 protein levels in SMA motor neurons are not significantly reduced but trend downwards from control cells. (A) *Representative western blot showing protein expression in human iPSC-derived motor neurons: MN Control 1 (Miff1), MN Control 2 (Cs14), MN Control 3 (GM), MN SMA 1 (SMA 77), MN SMA 2 (SMA 84). Western blot probed with anti-Gle1 antibody and anti-SMN antibody, showing protein expression of endogenous Gle1 at 75kDa and endogenous SMN at 36 kDa respectively. Tubulin (at 55kDa) is used to show equal loading of samples.* ***(B)*** *Densitometric analysis of Gle1 protein expression levels shows no significant difference across the different cell types (one-way ANOVA, p>0.05, ns, n=3). Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal.* ***(C)*** *Averages values from each cell type were grouped into Controls vs SMA. This shows a tendency for Gle1 protein expression levels to be reduced in SMA cells compared to controls. Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal.* ***(D)*** *Densitometric analysis of SMN protein expression levels shows no significant difference across the different cell types (one-way ANOVA, p>0.05, ns, n=3). Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal.* ***(E)*** *Averages values from each cell type were grouped into Controls vs SMA. This shows that SMA cells have reduced SMN protein expression compared to controls. Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal.*

4.2.3 Overexpression of Gle1b does not rescue loss of SMN and vice versa in fibroblast cells

Following on from previous experiments, to investigate whether reduction of Gle1 was linked to a reduction in SMN, I wanted to assess whether the loss of SMN protein in SMA patient fibroblasts could be rescued by restoration or elevation of Gle1 protein levels, and vice versa.

To perform these experiments, I used lentiviral vectors overexpressing either Gle1b or SMN, alongside a lentivirus containing GFP only as a control. As mentioned in Chapter 3, all MOI's for viral overexpression were previously optimised (LV SMN produced and optimised by Dr Evangelia Karyka in our team). We used control and SMA patient fibroblasts so that we could assess how overexpression of either protein could rescue the endogenous reduction seen in the patient cells previously.

Protein lysates were run on western blots to look at SMN and Gle1 protein levels. Tubulin was used to show even loading (**Fig 4.10 A**).

In this case, overexpression of SMN protein did not appear to increase the protein levels of Gle1 in either control cells or the SMA patient cells (paired t-Test, p=0.5783, ns, n=3). (**Fig 4.10 B**). Vice versa, overexpression of Gle1b protein did not act to restore SMN protein levels in SMA lines (paired t-Test, p=0.1727, ns, n=3) and had no effect on SMN protein expression in the control cells (**Fig 4.10 C**).

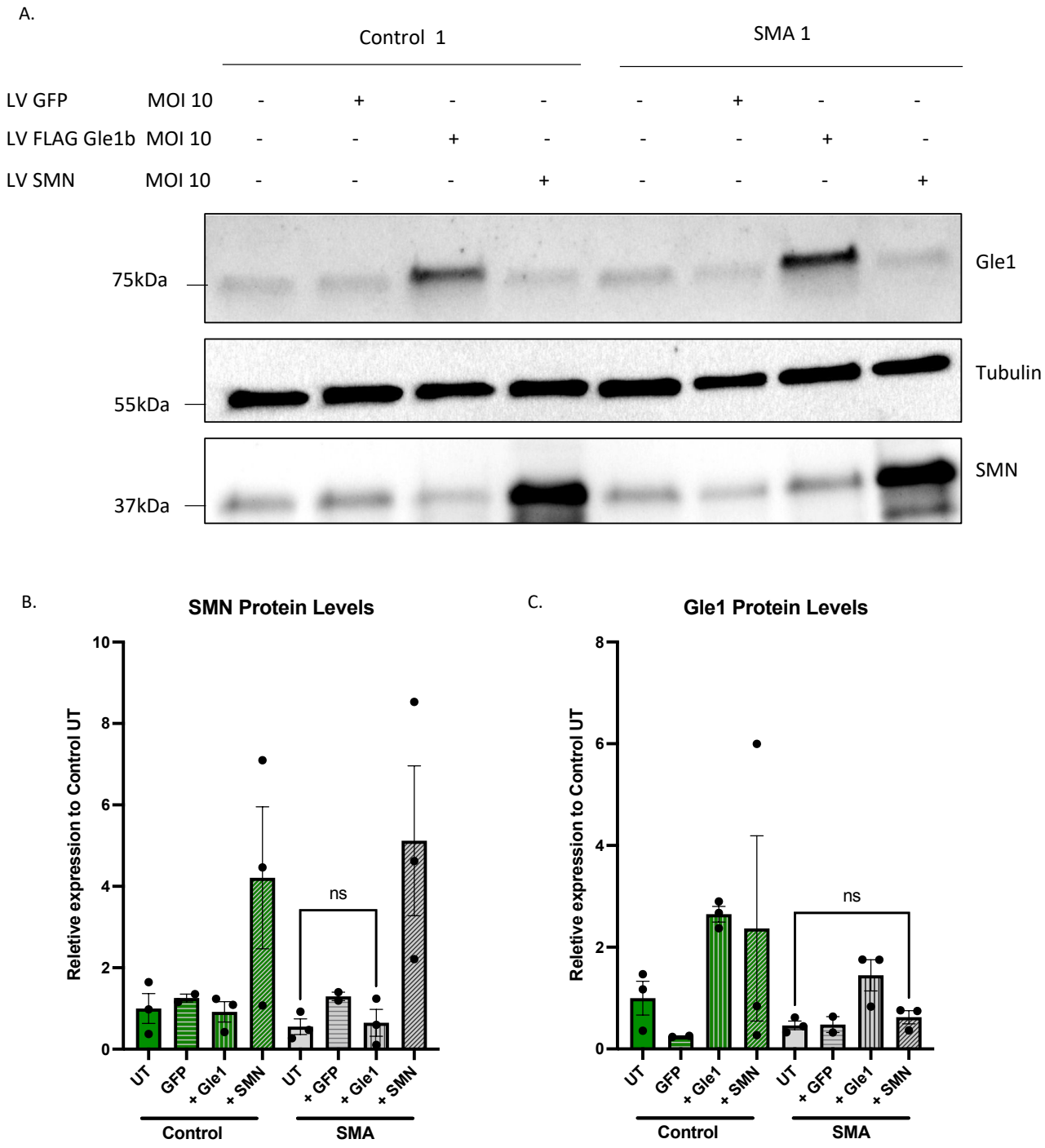


Figure 4.10: Gle1 and SMN protein levels do not influence each other by overexpression. (A) *Representative western blot showing protein expression in human fibroblast cells: Control (GM0489) and SMA (GM09677). Western blot probed with anti-Gle1 antibody or anti-SMN antibody, showing protein expression of Gle1 at 75kDa and SMN at 37kDa respectively. Tubulin (at 55kDa) is used to show equal loading of samples.* ***(B)*** *Densitometric analysis of SMN protein expression levels. There is no significant difference between SMN protein levels*

when LV FLAG Gle1b is overexpressed compared to the SMA untreated condition (paired t-Test, p=0.5783, ns, n=3). Protein loading was normalised using Tubulin signal. Data displayed as mean +/- SEM. ***(C)*** *Densitometric analysis of Gle1 protein expression levels. There is no significant difference between Gle1 protein levels when LV SMN is overexpressed compared to the SMA untreated condition (paired t-Test, p=0.1727, ns, n=3). Protein loading was normalised using Tubulin signal. Data displayed as mean +/- SEM.*

4.3 Discussion

The aim of the work which has been described in this chapter was to investigate whether there is a link between the Gle1 and SMN proteins. To assess this, I evaluated whether Gle1 and SMN proteins interact with each other on a protein-protein level through co-immunoprecipitations with anti-GFP coated beads. Proximity ligation assays in different cell models also provided a means to assess where in the cell interactions are likely to occur. I then set out to evaluate whether Gle1 expression was disrupted in an SMA disease context using two different cell models. Finally, I utilised the tools described in **Chapter 3** to test whether the reduction of either protein could be rescued through overexpression of the other.

4.3.1 There is protein-protein interaction between Gle1 and SMN proteins and the two proteins are found in close proximity in both the nucleus and the cytoplasm.

Due to both proteins' involvement in RNA processing pathways and we decided to investigate whether Gle1 and SMN could potentially interact with one another. Therefore, I performed GFP co-immunoprecipitations, using constructs previously generated in the lab (by Dr Eva Karyka) and those described in chapter 3, alongside proximity ligation assays in different cell lines.

These experiments provided novel data indicating that Gle1 and SMN proteins interact on a protein-protein level. Through assessment of interaction using different methods, we can have confidence that the interaction is likely to exist and therefore could indicate biological function. This is not an interaction which has previously been described in the literature, interestingly it also does not come up in an affinity purification mass spectrometry screen of overexpressed FLAG-Gle1b in HEK293 cells (Wolf et al., 2020), suggesting that it may involve another protein or occurs transiently within the cell.

The band seen where I overexpress and pulldown the GFP-Gle1b protein construct is much weaker than the band we observe when I overexpress and pulldown GFP-SMN. It would be interesting here to clone the hGle1A isoform into the same GFP construct and see whether this improves the signal. By western blot we cannot discriminate between the endogenous

isoforms as they are too similar in size and both are recognised by the same antibody. Through proximity ligation assay we see signal in both the nucleus and the cytoplasm of cells. As previously described, hGle1b is predominantly localised to the cytoplasmic fibrils of the nuclear pore complex and shuttles between this and the nucleus. The hGle1A isoform, however, is exclusively cytoplasmic and so may be the isoform responsible for the more distal association with SMN seen in this assay.

Of note is the location of the PLA signal observed in the iPSC-derived motor neuron model. The presence of Gle1 and SMN proteins colocalising within the axons of motor neurons could point towards involvement in SMN's role as a regulator of axonal mRNA transport, stability and translation (Akten et al., 2011; Fallini et al., 2012; Kye et al., 2014). This is also a novel finding and could help to explain the results seen previously in the lab where the axonal growth defect seen in SMN -/- motor neurons could be rescued by Gle1 overexpression (**Fig 1.5,** unpublished data). A potential experiment here could be to look at whether overexpression of Gle1 has any effect on axonal beta-actin mRNA expression within the growth cone, as this is disrupted specifically in SMA cases (Rathod et al., 2012; Rossoll et al., 2003).

It would also be interesting to see whether Gle1 also binds to any other well documented binding partners of SMN. Proximity ligation assay looking at Gle1 and Senataxin in HeLa cells demonstrates a close colocalization between the two proteins (Sharma and Wente, 2020). The AP-MS screen mentioned earlier suggests that FLAG-Gle1b also pulls down MYBBP1A and DHX9 (Wolf et al., 2020), which are known binding partners of SMN (Fuller et al., 2009; Pellizzoni et al., 2001b). Although proximity ligation assay between Gle1 and DHX9 in HeLa cells failed to produce a signal (Sharma and Wente, 2020).

As the signal was much weaker in the Gle1 pulldown, it would perhaps also be beneficial to see whether the GFP-SMN co-IP gives us any information about whether it binds to any of Gle1's known interacting partners, for example Nup42 or DDX19.

4.3.2 Gle1 protein levels are reduced in SMA patient fibroblasts

The genetic cause of SMA is a mutation within the *SMN1* gene which causes a 80-90% loss of SMN protein (Lefebvre et al., 1995). Other proteins have also been identified as having reduced levels within SMA models, for example Plastin3 (Hao et al., 2012), and Zinc finger Protein ZPR1 (Gangwani et al., 2001). Following our link of Gle1 and SMN proteins, I aimed to investigate whether Gle1 expression is altered in SMA cases.

In this chapter I have demonstrated a reduction (between 30-50%) in the protein levels of Gle1 within two different pairs of age matched fibroblasts.

A lack of Gle1 protein has previously been observed in ALS patients who possess mutations within the *GLE1* gene indicative of a role in neuronal health (Kaneb et al., 2015). Furthermore, a zebrafish model deficient in Gle1 protein displays reduced spinal motor neurons, specifically reducing the secondary and tertiary motor neuronal branches (Jao et al., 2012; Seytanoglu et al., 2016). This phenotype can also be observed in SMN-/- zebrafish embryos, which can then be rescued by overexpression of SMN protein (Hao et al., 2015). Therefore, potentially linking the reduction in Gle1 protein levels seen in SMA patient cells to the severe neurodegenerative phenotype seen in this disease.

Next, I evaluated the quantities of mRNA within these same cell lines. The fibroblast line GM09677 demonstrated a clear reduction in the levels of Gle1 mRNA, specifically in the nucleus. Decreases in nuclear Gle1 protein have been previously observed in the *GLE1* p.I684T mutation responsible for LAAHD (Paakkola et al., 2017). This could potentially point towards a loss of hGle1b, as this is the isoform which is predominantly localised to the nucleus and the NPC (Kendirgi et al., 2003). Therefore suggesting that there may be disruption in the mRNA export pathway which this isoform is heavily implicated in (Adams et al., 2017; Kendirgi et al., 2003).

However, the patient line GM00232 showed no difference in mRNA levels despite having a significant reduction in the levels of Gle1 protein. Both of these cell lines are classified as Type I SMA cases, and are therefore homozygous for deletions in exons 7 and 8 of the SMN1 gene, with GM09677 having 3 copies of the SMN2 gene, and GM00232 having 2 copies of the SMN2

gene (Coriell, 2021a, 2021b; Stabley et al., 2015). Due to the significant reduction we see on a protein level in these cells, we could speculate that this could be due to a different problem in the gene expression pathway, this cell line may produce mRNA which can be detected by these specific *GLE1* primers, but it may not lead to fully functional protein. It would be interesting to investigate whether splicing is occurring correctly in these cells as SMN is a major component of the spliceosome, similar experiments were performed when assessing how Plastin3 expression is effected by a loss of *SMN1* (Hao et al., 2012). Potentially this could also be due to variation in the control fibroblasts too. The GM08680 control cells which we use here have been reported to have a 10% loss or gain of chromosomes (Coriell, 2021c), this could be causing problems when comparing transcript levels to these SMA cells. Comparing between the two control cell types, Control 2 cells (GM08680) show only 50% of *GLE1* mRNA transcript levels as compared to Control 1 cells (GM0489). This therefore could be masking a drop in the mRNA levels in the SMA cells.

Due to this inconsistency, it is difficult to fully assess whether this change in expression is occurring at a transcriptional or translational level. This would be best evaluated through qPCR analysis of an additional age matched control and Type I SMA fibroblast sample, alongside analysis of other components of the RNA processing pathway as mentioned previously. Additional fibroblast cells would also help to assess the levels of variability within control fibroblast cell types themselves. Therefore, strengthening the differences in Gle1 levels between control and SMA through ensuring they are not occurring due to natural variability within fibroblasts cells. However, this was not possible due to time constraints.

A study done in 2018 suggests there is a reduction in *Gle1* mRNA in type I SMA fibroblasts (Alrafiah et al., 2018a), however, there are certain limitations to this paper. Despite grouping two parental fibroblast samples with one Type I SMA fibroblast line (GM03813), there is still a significant reduction in the *GLE1* mRNA transcripts compared to controls. Indicating that potentially the carrier parental cells (GM03814, GM03815) may also have reduced levels of *GLE1* transcripts compared to controls. More research would have to be conducted with these cells to draw any conclusions from this data.

Protein expression was also investigated in iPSC-derived mature motor neuron protein lysates, here the difference was less striking than in the fibroblast cell lines with no significant

difference being seen between the three control lines (MIFF1, Cs14 and GM) and the two SMA lines. When grouped together as control vs SMA, there does appear to be a trend for decreased Gle1 protein expression. More replicates looking at the expression of these two proteins may help to decrease the variability which we see in these samples and give a clearer picture of whether there is a decrease in the protein levels of Gle1 in these motor neurons.

Differences in expression levels have been observed between different cell types previously. Studies have demonstrated that there is significant difference in the gene expression between iPSC-motor neurons and the fibroblasts from which they were derived (Fuller et al., 2015), indicating that there are certain cell type specific changes which are important to assess within both populations.

4.3.3 Reduction of Gle1 and SMN expression levels are not rescued by overexpression of the other protein

As there is a reduction in the protein levels of Gle1 in SMA fibroblasts, I wanted to investigate whether this was directly because of the lack of SMN in this population, or whether it was independent of SMN levels. To evaluate this hypothesis, I used the GM09677 fibroblast cells and its age matched control GM0498. Using these cells would allow me to see whether overexpression of Gle1b or SMN could rescue the reduced levels of the other protein. For these experiments I used the lentiviral vectors which were described in chapter 3; LV_FLAG-Gle1b and LV-GFP. I also used a full-length LV-SMN which was previously produced in the lab by Dr Eva Karyka.

There was no observable increase in the levels of Gle1 when SMN levels were restored and vice versa, meaning that restoration of one protein's expression level will not rescue the other. ZPR1 has been shown to be a positive transcriptional modifier of both *SMN1* and *SMN2* and when overexpressed raises the levels of SMN protein both *in vitro* and *in vivo*, thereby improving known SMA disease phenotypes in a SMN-dependant manner (Kannan et al., 2020). Although known to be involved in transcription and translation processes, Gle1 does not show this same affect and so may be involved with SMN in an alternative manner.

Importantly, to answer this question thoroughly I would have liked to also perform knockdown experiments of the two proteins, using siRNA to knockdown both Gle1 and SMN expression levels in healthy control fibroblasts or motor neurons. This would give us more information about whether the reduction of Gle1 we see in SMA cases is directly because of a loss of SMN or is indeed occurring independently of SMN loss.

4.3.4 Gle1 and SMN proximity ligation assay signal is not altered between control and SMA patient samples.

Surprisingly, although both the expression levels of SMN and Gle1 are reduced in SMA cases, in both SMA fibroblast cells and iPSC-derived motor neuron cells there does not appear to be a significant reduction in the PLA signal. Although there is substantial variability in the fibroblast conditions which may be masking any potential reductions, more repeats would need to be performed in these cells to fully assess whether there is any difference between signal in controls vs patient cells. Especially since in this patient line we see a specific nuclear reduction in *GLE1* mRNA transcripts through nuclear/cytoplasmic fractionation, it would be interesting to see whether there is a reduction in nuclear PLA signal which corresponds to this.

Despite this variation, there does appear to be a difference in the foci size between the conditions, with the control cell foci appearing slightly larger than that of SMA. This could potentially indicate increased protein co-localisation in the control cells but would possibly require more investigation before any conclusions could be made.

4.3.5 Cell models used in this chapter

In this chapter I use different cell lines and systems to investigate my hypotheses. For initial experiments to assess whether there is a potential link between the Gle1 and SMN proteins, we used HEK293 cells as our model cell line. These are an immortalised cell line derived from an aborted foetus in 1973. These cells were optimal for initial investigation due to their easily transfectable nature and fast growth rate.

For experiments involving patient samples I began by looking into fibroblasts cells from patients and age-matched controls. Fibroblasts are primary cells derived from skin. Multiple

studies looking at neurological diseases have used fibroblasts as a model as they display disease phenotypes which can be studied as biomarkers (Bell et al., 2020; Macdonald et al., 2018; Mahadik and Mukherjee, 1996). There are various advantages and disadvantages to using this as a model of disease. Firstly, as they can be derived from patient skin biopsies, it is a good method of generating cell models from patients and age matched control samples. They are also robust cells which can be grown easily in the lab, although they do have a much slower growth rate than immortalised cell lines. The nature of these cells can also mean that they are much more difficult to treat with drugs or virus's than other cell lines, therefore making certain experiments less effective without sufficient optimisation (Auburger et al., 2012). However, as they are primarily skin cells, they may not be fully representative of pathways which are particularly suspectable in more refined cell populations – in the case of this study: motor neurons.

Therefore, I wanted to also study whether the results seen in the fibroblast populations were conserved in motor neurons. This would help us to understand more about how a reduction of either Gle1 or SMN specifically causes degeneration of motor neurons. An iPSC-derived population of Neural Progenitor Cells were consequently generated (with the kind help of Dr Cleide De Souza Santos at SITraN) using previously described protocols (Du et al., 2015). Culturing of motor neurons requires the addition of multiple neurotrophic factors over a period of up to 40 days. These neurons are highly sensitive to external stress, such as lack of factors, changes in temperature/CO_2 and other variables which may occur during handling. This makes culture of healthy motor neuron populations challenging protocols which stretch over many months.

Human iPSC models have been extensively used in SMA research, with groups demonstrating that these cell populations display a lack of *SMN1* and motor neuron specific cell death progressively over time (Barrett et al., 2014; Ebert and Svendsen, 2010; Sareen et al., 2012).

4.3.6 Summary

In summary, in this chapter I have demonstrated that there is a reduction in Gle1 expression levels in SMA patient fibroblast cells, this may be consistent in iPSC-derived motor neurons, but more work is required to confirm this hypothesis. I have also seen that reduction at an

mRNA level is sample specific with only one cell type showing a reduction, specifically in the nucleus of the cells.

I have also demonstrated that Gle1 and SMN are binding partners in close proximity in both the nucleus and the cytoplasm. This interaction does not appear to be altered by the decreased levels of the SMN and Gle1 proteins in SMA cases, this may indicate that these two proteins are functioning in a pathway which is crucial for the survival of the motor neurons, but more research is needed into how these two proteins may be functioning together.

Therefore, in chapter 5 of this thesis I will look at potential pathways which may involve these two proteins and whether overexpression of Gle1b can act to ameliorate SMA disease phenotypes.

5. Can overexpression of Gle1b rescue disease phenotypes of Spinal Muscular Atrophy in *in vitro* cell models?

5.1 Introduction

Therapies to treat Spinal Muscular Atrophy have largely focussed on replacing mutated *SMN1*. During the course of this PhD, Zolgensma® (a novel gene therapy) was approved for clinical use in Type I SMA patients (Agency, 2020). This is an adeno-associated viral vector which can specifically deliver functional *SMN1* to the nervous system in order to improve motor function and prevent degeneration of motor neurons (Mendell et al., 2017).

In addition to these SMN-dependant therapies, researchers have also worked for many years on looking at different genetic modifiers which can act to rescue SMA pathology despite of the lack of SMN protein, for example *Plastin3* (Oprea et al., 2008), *ZPR1* (Kannan et al., 2020) or *Neurocalcin* (Riessland et al., 2017). It has been suggested that these therapies can also be used in conjunction with treatments to increase SMN expression in order to ameliorate the disease phenotypes (Kaifer et al., 2017).

In Chapter 4, I demonstrated that Gle1 expression levels could potentially be downregulated in SMA patient models and that Gle1 and SMN likely function together in a shared pathway. Combined with previous studies in our lab which demonstrated that overexpression of Gle1 protein in a SMA disease context can act to rescue axonal growth defects (**Fig 1.5,** unpublished data), we wanted to evaluate whether addition of Gle1b to SMA patient cells could reduce the severity of disease phenotypes.

5.1.2 Aims

In this chapter I therefore aim to investigate whether overexpression of hGle1b can rescue any disease phenotypes which we expect to see in SMA cell models. This includes looking at published phenotypes of SMA pathology such as increased DNA damage (Jangi et al., 2017). I will also investigate phenotypes which appear to be relevant to Gle1 function that have been shown to be dysfunctional in SMA cases, such as localisation of poly(A)+ mRNA (Narcís et al., 2018) and Golgi morphology (Custer et al., 2019; Ting et al., 2012).

5.2 Results

5.2.1 DNA Damage

DNA damage is a well characterised disease phenotype within SMA cases (Fayzullina and Martin, 2016; Jangi et al., 2017; Yanling Zhao et al., 2016). Two papers published during this PhD implicated components of the mRNA export pathway with DNA damage. The first studied the role of DDX19 in resolving R-loops, where the protein transiently moves from its cytoplasmic localisation to the NPC to the nucleus upon induction of DNA damage (Hodroj et al., 2017a). As part of this study, Hodroj and colleagues, observed increased γH2AX (a marker of DNA double strand breaks) upon induced DNA damage and replication stress when DDX19 was depleted by RNAi. Interestingly, they also observed an increase in γH2AX levels when Gle1 is knocked down (Hodroj et al., 2017a). The second paper of interest (Okamura et al., 2018), also studied the effects that knockdown of Gle1 had on components of the DNA damage pathway. This study indicated that knockdown of Gle1 alters the DNA damage response, showing a significant increase in γH2AX after siRNA-mediated Gle1 knockdown. DNA damage detecting and repairing proteins were also downregulated by Gle1 knockdown (Okamura et al., 2018).

Therefore, I was interested to see whether Gle1 is in any way involved in the DNA damage response and whether this could help to reduce the levels of DNA damage we see in SMA patients. I hypothesised that as DDX19 is translocating from the nuclear pore to the nucleus, specifically to sites of R-loops, then potentially Gle1b (as its activator) may also be moving alongside it and having a potential role in DDX19-mediated R-loop resolution (**Fig 5.1**).

To evaluate this hypothesis, I began by staining control and patient SMA fibroblasts with a marker for R-loops, this would allow us to assess whether Gle1 does indeed translocate from its position on the NPC to sites of DNA damage within the nucleus. These patient cells have been used previously in the lab to demonstrate an increase in R-loops in SMA patient cells compared to controls. Staining taken from three separate plating's of cells demonstrated a consistent co-localisation of Gle1 with the S9.6 marker of R-loops. Staining appears to demonstrate a co-localisation of Gle1 protein to both R-loops (visualised using the S9.6 antibody) and the nucleolus (visualised using a nucleolin antibody) (**Fig 5.2 A, B**). The majority

of the colocalisation of Gle1 appeared to be within these nucleolar structures. The specificity of the S9.6 antibody and the presence of R-loops in nucleolar structures has been previously validated in our group (**Fig 5.3**) (Karyka et al., 2022; Walker et al., 2017).

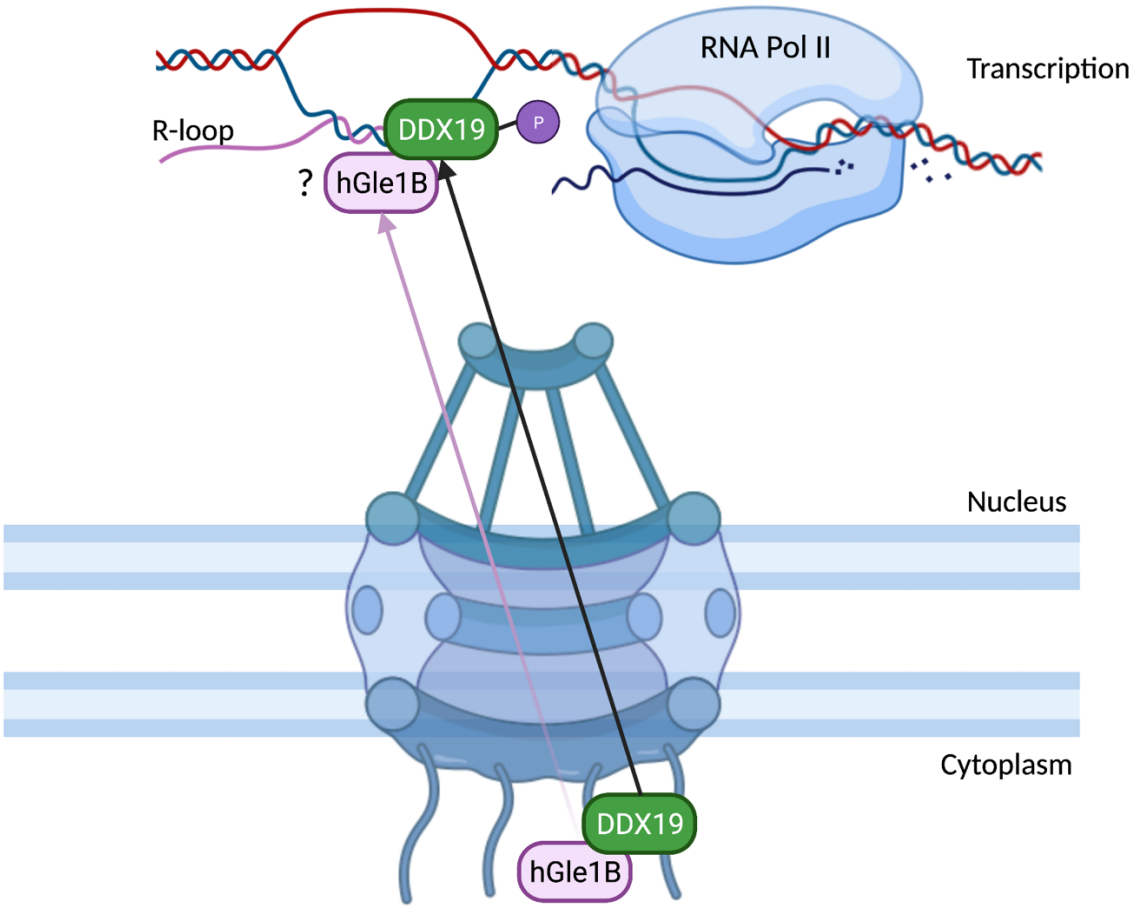


Figure 5.1: Schematic of hypothesis focussing on DNA damage phenotype. *As described in (Hodroj et al., 2017) DDX19 has a role in resolving nuclear R-loops which form during replication stress or DNA damage. DDX19, upon phosphorylation by Chk1, relocates from its cytoplasmic position on the nuclear pore complex to the nucleus, therefore resolving R-loops through its activity as a helicase. This occurs in a senataxin independent manner. As GLE1 is a known activator of DDX19 activity, I hypothesise that Gle1b may also transiently move from the cytoplasmic fibrils of the nuclear pore complex inside the nucleus and may have some potential link to this pathway. Figure created with biorender.com.*

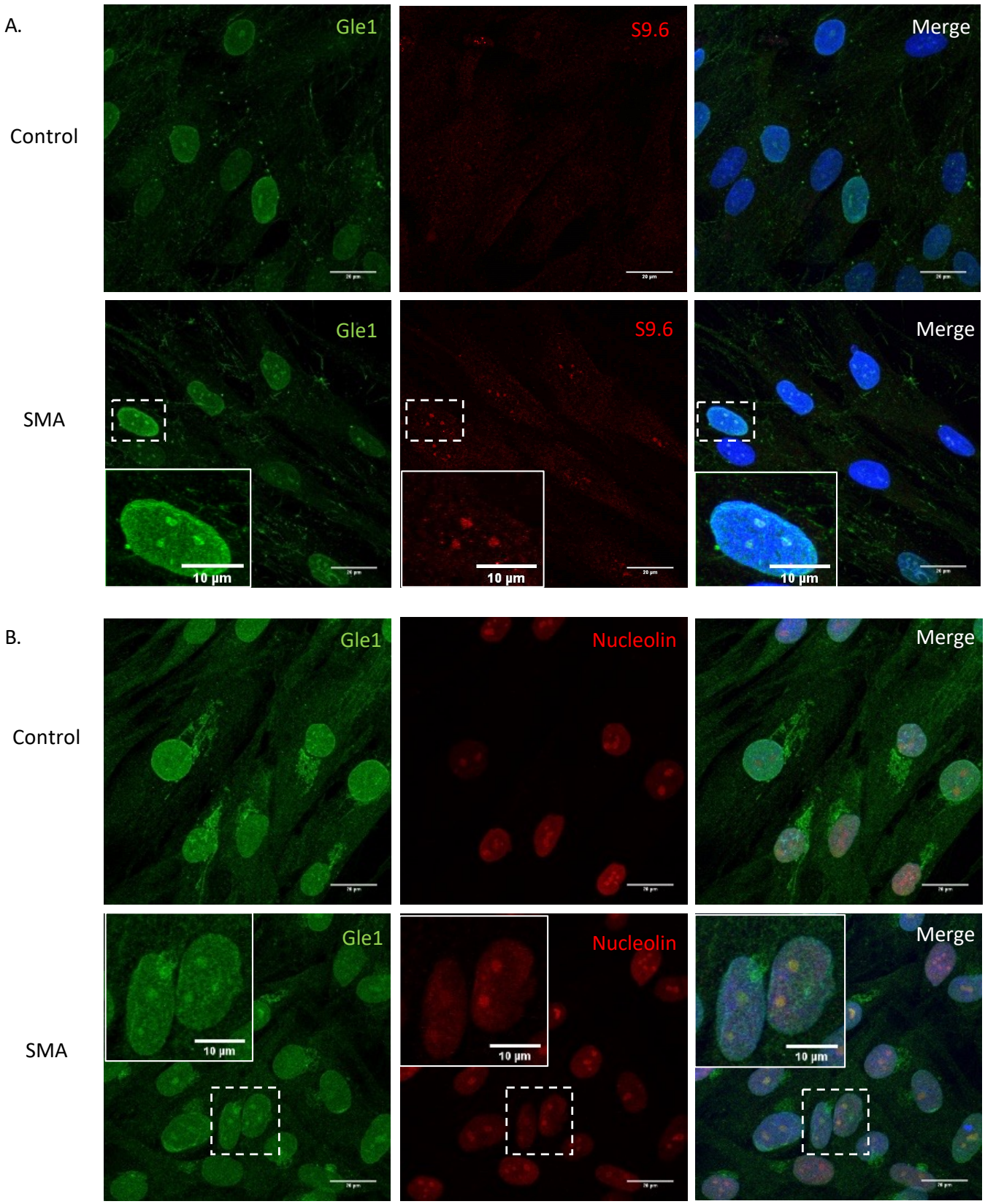


Figure 5.2: GLE1 staining by immunocytochemistry in SMA patient fibroblast cells co-localises with markers of the nucleolus and R-loops (S6.9). (A) *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677) stained with S9.6 antibody for visualising R-*

loop staining in red. GLE1 endogenous protein can be seen in green. Scale bar is 20 µm. For high magnification images scale bar is 10 µm, area taken as high magnification is marked in white. Data collected from three biologically independent replicates (n=3). ***(B)*** *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677) stained with nucleolin for visualising nucleoli stained in red. GLE1 endogenous protein can be seen in green. Scale bar is 20 µm. For high magnification images scale bar is 10 µm, area taken as high magnification is marked in white. Data collected from three biologically independent replicates (n=3).*

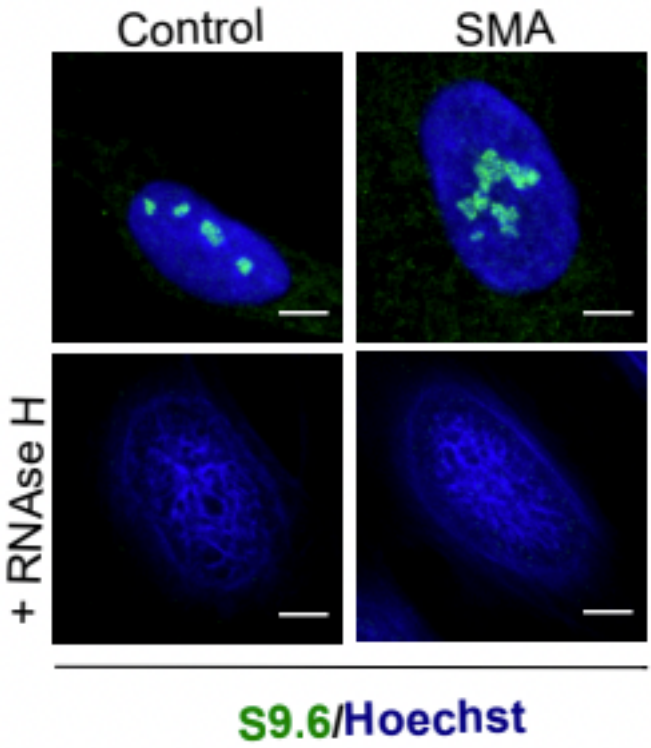


Figure 5.3: RNase H treatment of Control and SMA fibroblasts showing specificity of S9.6 antibody. *Fixed control and SMA fibroblasts were pre-treated with RNaseH and stained for R-loops (S9.6).* ***Data generated and figure made by Dr Evangelia Karyka*** (Karyka et al., 2022)***.***

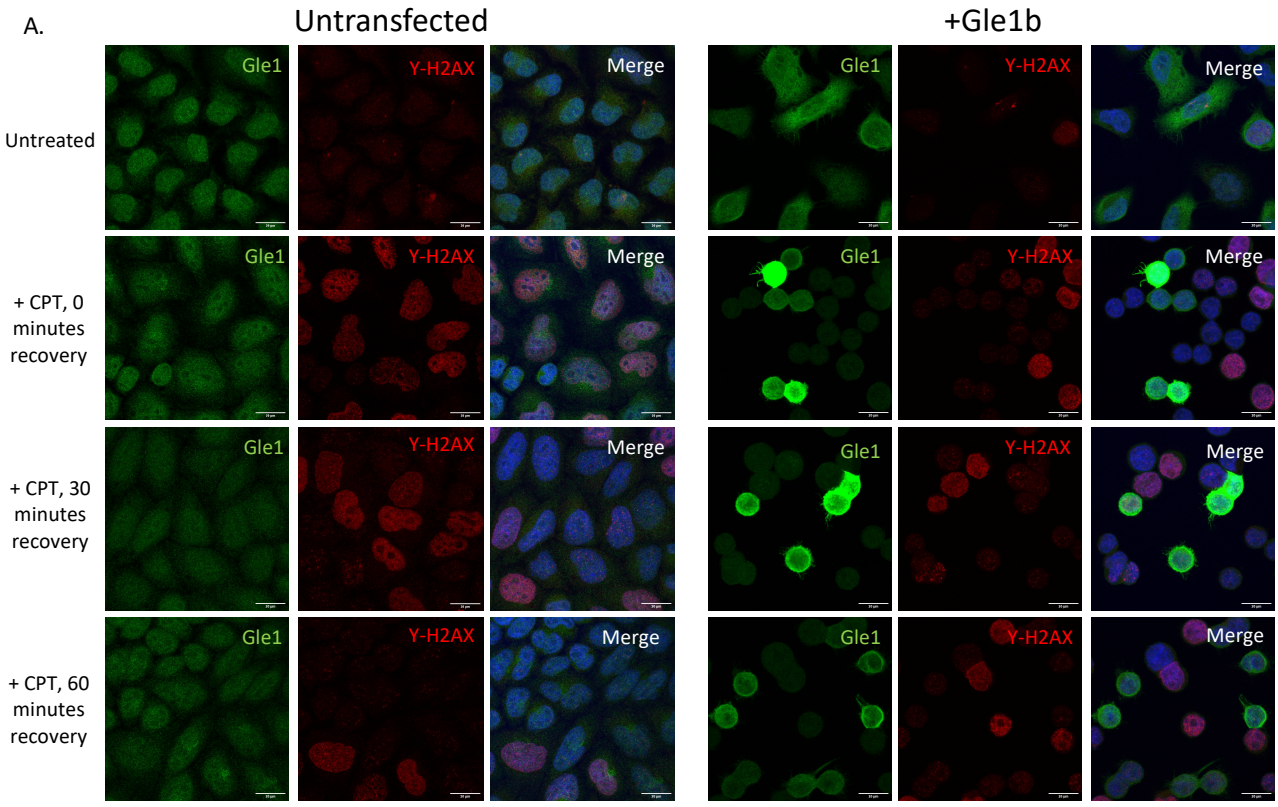
Due to the localisation of endogenous Gle1 to these markers, I moved into looking into whether Gle1b might play a functional role in the reduction of DNA damage. Here we use the

marker of double strand breaks: γ-H2AX. Double strand breaks are formed frequently during DNA replication and transcription alongside cellular stress. DNA damage observed in SMA cases is thought to be transcription-associated (unpublished data from our lab). The marker for R-loops was not used for this and following experiments due to the variability and unspecific nature of the staining it produces, although a better marker for SMA associated DNA damage (Ed et al., 2021).

For initial experiments I used HeLa cells as a cellular model in which to induce transcriptionally associated DNA damage through treatment with campthothecin (CPT). This is a drug which inhibits topoisomerase I. Cells were transfected with an overexpression plasmid for Gle1b, 48 hours following transfection cells were treated with CPT for 1 hour at 37 ^{O}C and left for a range of recovery times prior to fixation (**Fig 5.4 A**). Intensity of the γH2AX staining was carried out through Corrected Total Cell Fluorescence (CTCF) method, in the cells overexpressing Gle1b, only cells showing increased Gle1 staining were counted.

Here I saw that there was an increase in the intensity of the γH2AX staining immediately following CPT treatment in both the un-transfected condition and the cells transfected with the Gle1b overexpression plasmid. This increase was only statistically significant in the cells overexpressing Gle1b (RM one way ANOVA; +Gle1b: UT vs CPT R0, p=0.0296, *, n=3), however was close to significance in the untransfected cells (RM one way ANOVA, untransfected; UT vs CPT R0, p=0.0549, ns, n=3) (**Fig 5.4 B**). Addition of Gle1b to the cells did not have any effect on the levels of γH2AX staining compared to the untransfected controls across each of the time points (**Fig 5.4 B**). This data therefore indicates that addition of Gle1b does not have any effect on the induction of double strand breaks caused by CPT and so is likely not to be involved in any DNA damage response pathways.

A.


Untransfected
+Gle1b
Untreated
+ CPT, 0 minutes recovery
+ CPT, 30 minutes recovery
+ CPT, 60 minutes recovery
Gle1
Y-H2AX
Merge

B.

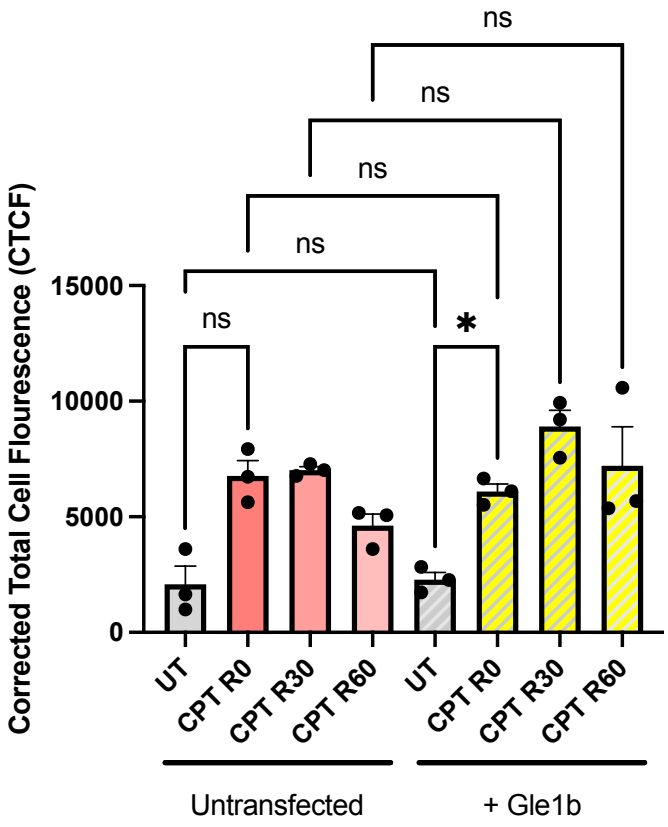

Corrected Total Cell Flourescence (CTCF)
15000
10000
5000
0
ns
ns
ns
ns
ns
*
UT
CPT R0
CPT R30
CPT R60
UT
CPT R0
CPT R30
CPT R60
Untransfected
+ Gle1b

Figure 5.4: Overexpression of Gle1b doesn't protect human cell lines against induced DNA damage. (A) *Immunocytochemistry of untreated and AAV Gle1b transfected HeLa cells. Endogenous and overexpressed Gle1 protein stained in green and γH2AX levels stained in red. Cells were then treated for 1 hour with campthothecin (CPT) and fixed 0 (R0), 30 (R30) or 60 (R60) minutes after treatment. Scale bar is 20 µm. Data collected from three biologically independent replicates (n=3, 50 cells per condition).* ***(B)*** *Quantification of γH2AX staining by corrected total cell fluorescence (CTCF) method. R0: Zero minutes recovery time after CPT treatment, R30: 30 minutes recovery time following CPT treatment, R60: 60 minutes recovery time following CPT treatment. There was no significant difference in the γH2AX signal in the untransfected cells which had been treated with CPT (zero recovery time) (RM one way ANOVA, untransfected; UT vs CPT R0, p=0.0549, ns, n=3). Cells overexpressing Gle1b showed a significant increase between untreated and treated with CPT (zero recovery time) (RM one way ANOVA; +Gle1b: UT vs CPT R0, p=0.0296, *, n=3). There was no significant difference between each of the time points when comparing untransfected cells to those overexpressing Gle1b (RM one way ANOVA; CPT R0, p=0.8925, ns, n=3; CPT R30, p=0.4750, ns, n=3; CPT R60, p=0.6997, ns, n=3).*

As previously demonstrated (**Chapter 4**), within the SMA patient fibroblasts Gle1 protein expression appears to be reduced. Therefore, I wanted to investigate whether restoration of the levels of Gle1 could potentially act to ameliorate some of SMA's disease phenotypes, using DNA damage as our initial example.

Using lentiviral constructs (described in **Chapter 3**), I aimed to investigate whether overexpression of Gle1b in SMA patient fibroblasts could reduce the endogenously increased levels of γH2AX which these cells display. For this experiment, I stained healthy control and SMA fibroblasts with γH2AX and Flag antibodies (**Fig 5.5 A**). For the transduced conditions, only cells which were positive for Flag staining were quantified. As expected, we observed a significant increase in the corrected total cell fluorescence (CTCF) values of the γH2AX staining between the Control and SMA cells (RM one way ANOVA; Control vs SMA p=0. 0291, *, n=3). Control cells which were treated with LV Flag Gle1b showed a significant increase in the levels of γH2AX staining (RM one way ANOVA; Control vs. Control +Gle1 p=0.0145, *, n=3). SMA cells transduced with LV Flag Gle1b do not show a significant change in γH2AX levels

compared to untreated SMA cells (RM one way ANOVA; SMA vs SMA +Gle1 p=0.9908, ns, n=3*)* (**Fig 5.5 A, B**).

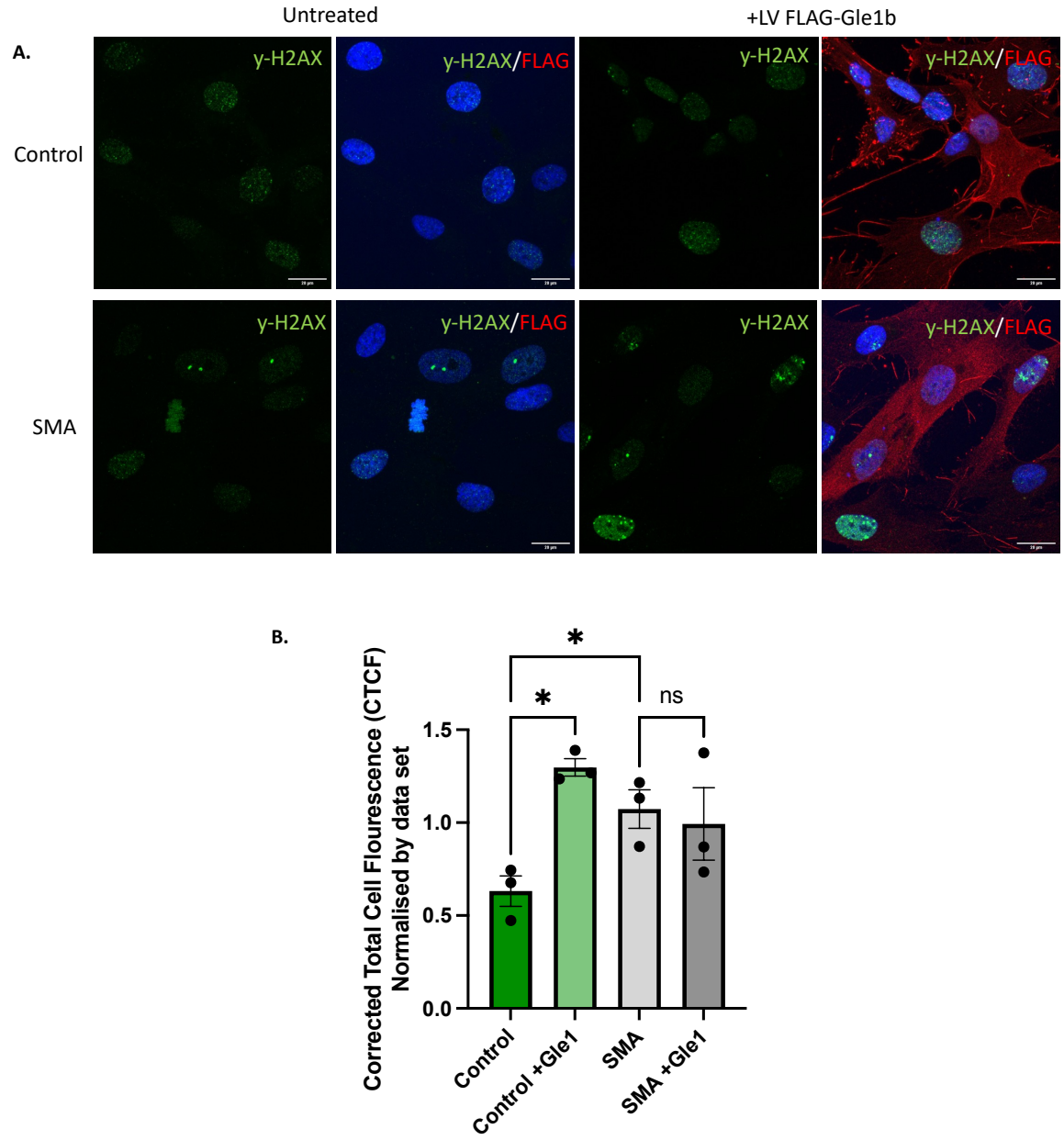


Figure 5.5: Overexpression of Gle1b in SMA patient fibroblast cells doesn't reduce endogenous DNA damage. (A) *Immunocytochemistry of Healthy (GM8680) and SMA (GM9677) patient fibroblasts. FLAG tagged Gle1b protein stained in red and γH2AX levels stained in green. Scale bar is 20 µm. Data collected from three biologically independent replicates (n=3, 50 cells per condition).* ***(B)*** *Quantification of γH2AX staining by corrected total*

*cell fluorescence (CTCF) method. Each data set is normalised to its average value. There is a significant increase in γH2AX signal between the control cells and control cells treated with LV Flag Gle1b. There is a significant increase in γH2AX signal between the control cells and SMA cells. There is no significant difference in γH2AX signal between SMA cells and SMA cells treated with LV Flag Gle1b (RM one way ANOVA; Control vs. Control +Gle1 p=0.0145, *; Control vs SMA p=0.0291, *; SMA vs SMA +Gle1 p=0.9908, ns; n=3).*

Fibroblasts are primary cells but still undergo cell replication. Therefore, double strand breaks occur much more frequently as the cell goes through the cell cycle, specifically S phase where the DNA itself is replicated during transcription and double strand breaks are exposed. During this phase cells could potentially have an increased level of γH2AX staining which may not be directly related to the phenotype we are looking at.

Consequently, as a final readout we wanted to look at the DNA damage levels within the iPSC-derived motor neuron SMA cells (**Fig 5.6 A**). Unfortunately, in this case over three biologically independent repeats no difference in the endogenous levels of γH2AX staining could be detected between control cells and SMA cases (**Fig 5.6 B**). Therefore, we did not continue with this line of experiments to assess impact of Gle1b on DNA damage.

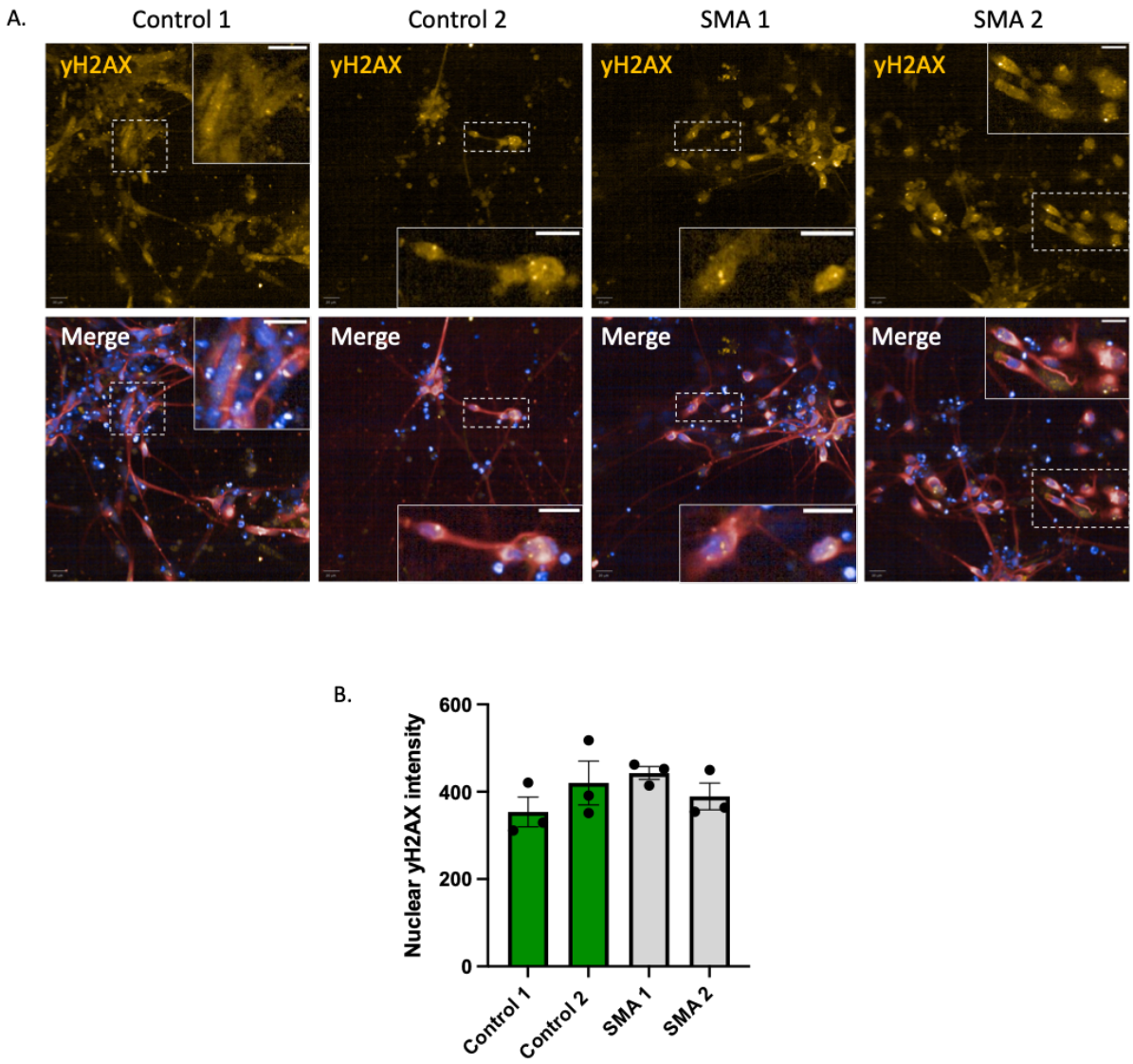
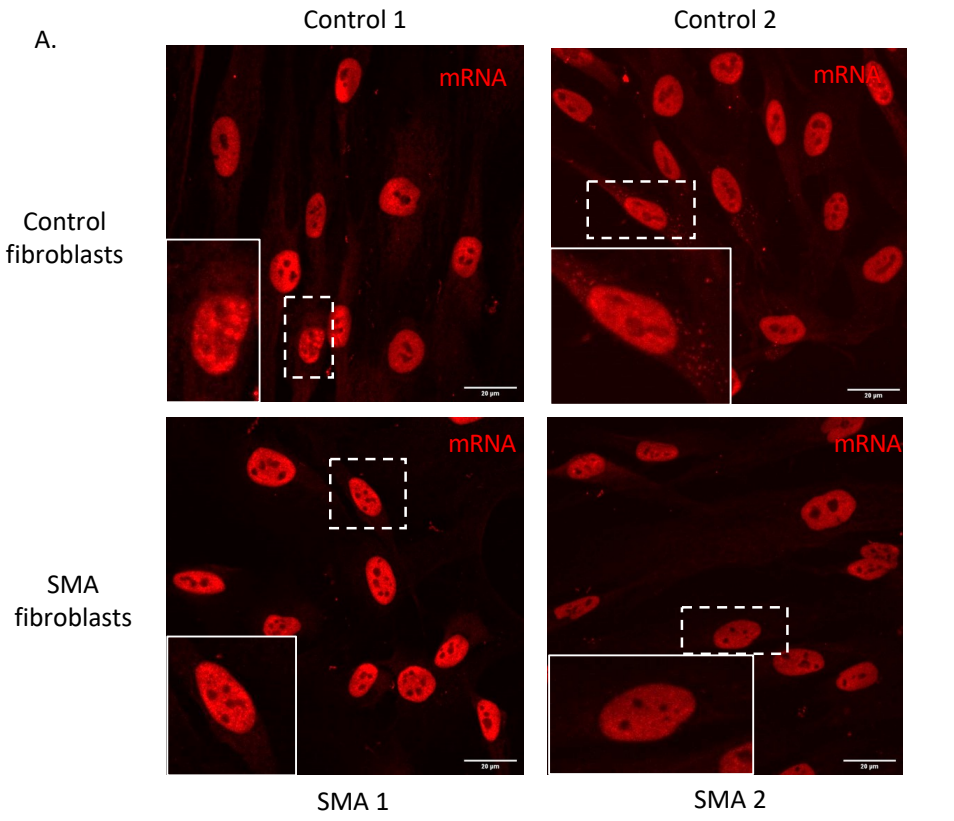


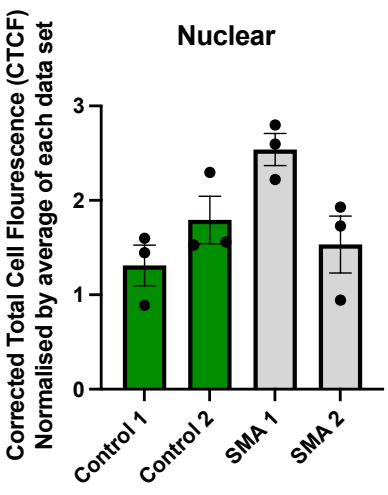
Figure 5.6: There was no observable increase in γH2AX staining in iPSC derived SMA motor neuron populations. (A) *Immunocytochemistry of Healthy 1 (Cs14, Control 2 (GM), SMA 1 (SMA 77) and SMA 2 (SMA 84) iPSC derived motor neurons. γH2AX visualised in orange, motor neuron axons visualised by MAP2 in red with nuclei in blue. Scale bar is 20 µm. Data collected from three biologically independent replicates, (n=3, data average from 24 fields, 2 technical replicates per biological repeat, at least 200 cells per condition).* ***(B)*** *Quantification of γH2AX staining from average intensity of 568 channel signal. There is no significant difference between any of the cell types (Repeated measures one-way ANOVA, P=0.1735, ns, n=3).*

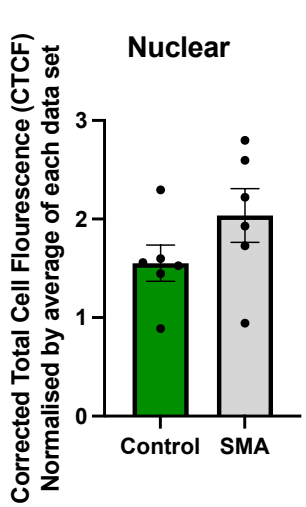
5.2.1 RNA nuclear export defects

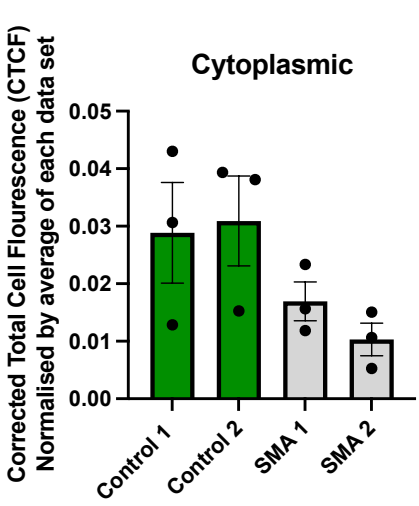
As Gle1 is primarily a protein which functions in the process of mRNA nuclear export. I thought it was important to look at this pathway in SMA cases as a reduction of Gle1 protein (specifically a nuclear reduction) may point to dysregulation of mRNA export from the nucleus. There have been reports in the literature that mRNA nuclear export is disrupted in the SMN Δ7 mouse model, with accumulation of poly(A) RNA nuclear granules (Narcís et al., 2018), which supports our hypothesis. I wanted first to assess whether this was a phenotype which we could also see in the SMA cells which we are using. The plan is to assess whether overexpression of Gle1b may act to reduce nuclear accumulation of poly(A) RNA, including mRNAs and long intergenic non-coding RNAs (lincRNAs).

To investigate whether this phenotype is present in our cells, I first looked at the two sets of fibroblast cells available in our lab. Fluorescent in situ hybridisation (FISH) within these cells using a probe for poly(A)+ RNA demonstrated strong nuclear staining, with diffuse staining visible in the cytoplasm (**Fig 5.7 A**). This cytoplasmic staining appeared to trend towards being stronger in the Control cell lines (in particular Control 2, where cytoplasmic foci can be seen) (**Fig 5.7 A**). There was no significant difference between the nuclear staining intensity between controls and SMA cells, although there was a trend for there to be increased signal in the SMA cases (**Fig 5.7 B, C**). Across all four of the cell types there was no significant difference between any of the groups cytoplasmic staining intensity (RM one way ANOVA, ns, n=3) (**Fig 5.7 D**). However, when grouped together as Control and SMA, there is a trend towards a decrease in the level of cytoplasmic staining (**Fig 5.6 E**).


A.
Control 1
Control 2
Control fibroblasts
SMA fibroblasts
SMA 1
SMA 2
mRNA


B.
Nuclear
Corrected Total Cell Flourescence (CTCF) Normalised by average of each data set
0
1
2
3
Control 1
Control 2
SMA 1
SMA 2


C.
Nuclear
Corrected Total Cell Flourescence (CTCF) Normalised by average of each data set
0
1
2
3
Control
SMA


D.
Cytoplasmic
Corrected Total Cell Flourescence (CTCF) Normalised by average of each data set
0.00
0.01
0.02
0.03
0.04
0.05
Control 1
Control 2
SMA 1
SMA 2

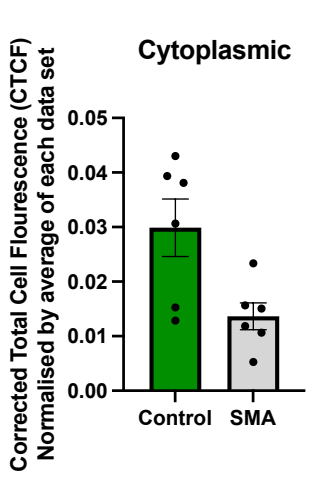
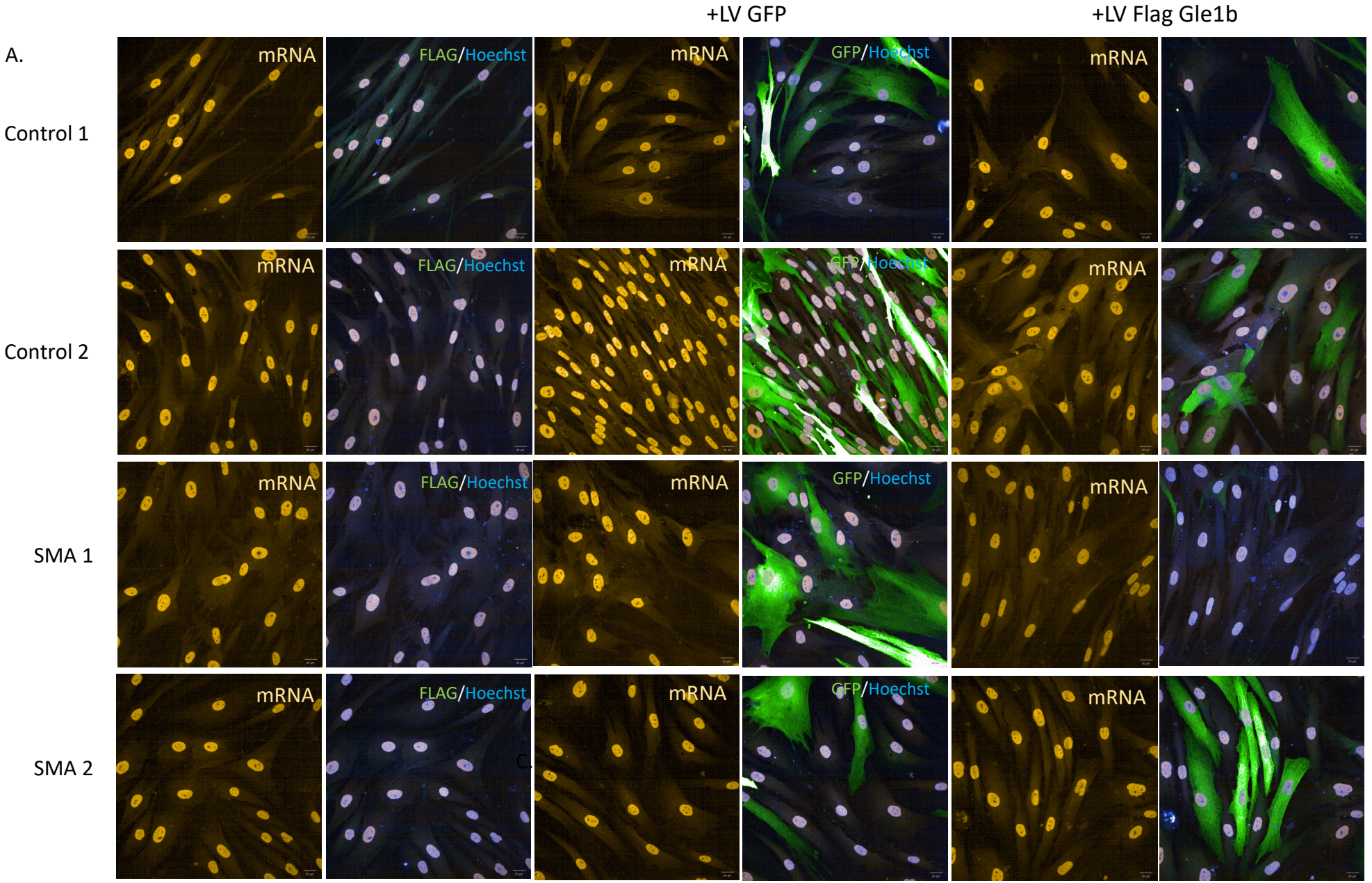

E.
Cytoplasmic
Corrected Total Cell Flourescence (CTCF) Normalised by average of each data set
0.00
0.01
0.02
0.03
0.04
0.05
Control
SMA

Figure 5.7: There is a reduction in cytoplasmic poly(A)+ RNA staining in SMA patient fibroblasts compared to control cells. (A) *Fluorescent in situ hybridisation (FISH) staining of poly(A)+ mRNA within the nucleus and cytoplasm. Control 1 (GM4989), Control 2 (GM680), SMA 1 (GM09677), and SMA 2 (GM00232) fibroblasts. Data collected from 3 biologically independent replicates, 50 cells per condition per repeat (n=3). Scale bar is 20 µM. Area of higher magnification images shown in white. (**B**) Quantification of poly (A) probe staining within the nucleus by corrected total cell fluorescence (CTCF) method, data is normalised to the average of each dataset. There is no statistical significance between any of the cell types (RM one way ANOVA, ns, n=3) (**C**) Average nuclear value from each data set grouped as control vs SMA. (**D**) Quantification of poly (A) probe staining within the nucleus by corrected total cell fluorescence (CTCF) method, data is normalised to the average of each dataset. A perinuclear ROI was taken for cytoplasmic quantification as visualised by a cytoplasmic marker. There is no statistical significance between any of the cell types (RM one way ANOVA, ns, n=3). **(E)** Average cytoplasmic value from each data set grouped as control vs SMA. There is a trend toward decreased cytoplasmic signal in SMA cells than in Controls*

Due to decreased levels of cytoplasmic poly(A)+ RNA staining in these SMA patient fibroblast cells, I decided to test whether lentiviral overexpression of FLAG tagged Gle1b protein would have any impact on the distribution of this staining. As a control for the lentivirus itself I used LV GFP. Unfortunately, due to time constraints caused by Covid-19 and problems with our confocal microscope, for these experiments I used a different method of imaging; an automated Opera Phenix high throughput imaging system. The trend we saw through using confocal imaging was not conserved through this method as there was little difference across any of the cells, and so I therefore did not continue with this experiment (**Fig 5.8 A-C**).


+LV GFP
+LV Flag Gle1b
A.
Control 1
Control 2
SMA 1
SMA 2
mRNA
FLAG/Hoechst
GFP/Hoechst
C.

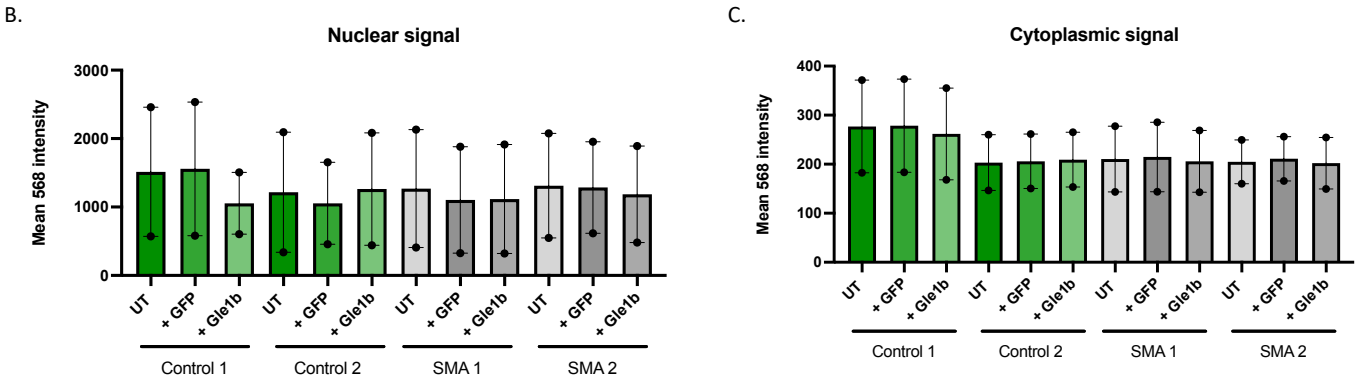


Figure 5.8: Overexpression of Gle1b in SMA patient fibroblasts doesn't affect poly-(A)+ RNA distribution. (A) *Fluorescent in situ hybridisation (FISH) staining of poly(A)+ mRNA within the nucleus and cytoplasm. Control 1 (GM4989), Control 2 (GM680), SMA 1 (GM09677), and SMA 2 (GM00232) fibroblasts. Data collected from 2 biologically independent replicates (n=2), data average from 24 fields, 2 technical replicates per biological repeat, at least 200 cells per condition. Scale bar is 20 µm. (**B**) Quantification of poly (A) probe staining within the nucleus by taking the mean value of the 568 channel per condition. (**C**) Quantification of poly (A) probe staining within the cytoplasm by taking the mean value of the 568 channel per condition.*

As the previous study demonstrating a nuclear accumulation of poly(A)+ mRNA granules within the nucleus focussed on motor neurons derived from the SMN Δ7 mouse model (Narcís et al., 2018), I wanted to investigate whether the iPSC-derived motor neurons which I was working with also displayed a similar phenotype which may suggest that there is impaired mRNA nuclear export in this model. Therefore, I repeated the FISH staining in this model. Unfortunately, as in the SMA fibroblast lines, there was no detectable difference between the healthy and control cells in either the nuclear or cytoplasmic intensity of the poly(A)+ RNA probe (**Fig 5.9 A, B**).

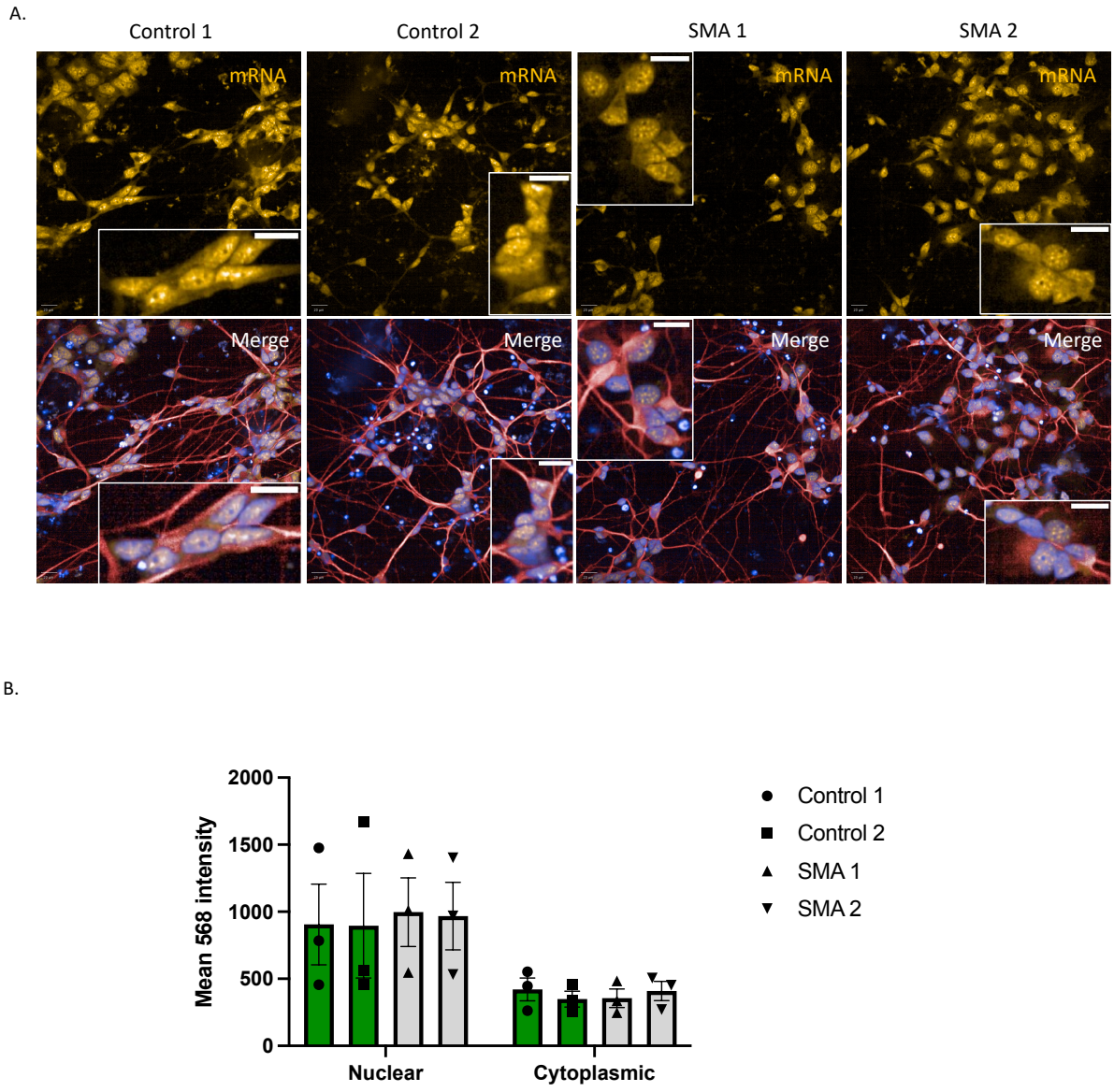


Figure 5.9: There is no detectable difference between RNA distribution in Control iPSC derived motor neurons compared to SMA patient iPSC derived motor neurons. (A) *Immunocytochemistry of Healthy 1 (Cs14, Control 2 (GM), SMA 1 (SMA 77) and SMA 2 (SMA 84) iPSC derived motor neurons. γH2AX visualised in orange, motor neuron axons visualised by MAP2 in red with nuclei in blue. Scale bar is 20 µm. Data collected from three biologically independent replicates (n=3), data average from 24 fields, 2 technical replicates per biological repeat, at least 200 cells per condition.* ***(B)*** *Quantification of γH2AX staining from average intensity of 568 channel signal. There is no significant difference between any of the cell types.*

As there appeared to be no detectable global defect in the nuclear export of poly(A)+ RNA in the SMA cells which I am using here, I hypothesised that the nuclear loss of Gle1 on both a protein and mRNA level in one set of the SMA fibroblast cells (as demonstrated in **chapter 4**) may be only affecting specific transcripts which Gle1 is specifically responsible for the nuclear export of. A list of these transcripts was recently published (Sharma and Wente, 2020). In this paper, Sharma and colleagues used a Gle1-SD peptide which binds to the 39 amino acid domain which is responsible for Gle1b's nuclear-cytoplasmic shuttling activity, therefore blocking Gle1 shuttling. Consequently, there is significant nuclear accumulation of poly(A)+ mRNA, with a specific increase in the transcripts of 70 genes. Therefore we can infer that these transcripts specifically require Gle1b for their nuclear export (Sharma and Wente, 2020) (**Fig 5.10**).

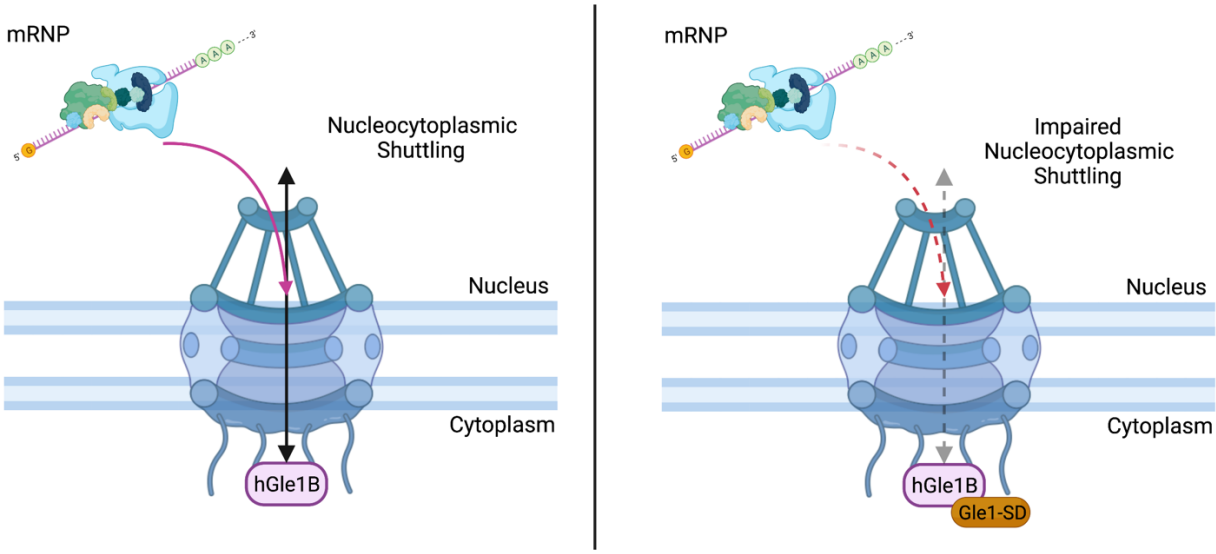


Figure 5.10: Schematic of certain transcripts which are selectively exported from the nucleus when Gle1b shuttling blocked. *(**left**) Under normal conditions, hGle1B shuttles between the nucleus and cytoplasm to help export the mRNP complex. (**right**) Nucleocytoplasmic shuttling of Gle1b is impaired when the Gle1 shuttling domain is blocked by a synthetic peptide (Gle1-SD) which is complimentary for this 39 amino acid sequence. This causes nuclear retention of a specific set of mRNA transcripts. Figure created with biorender.com.*

To investigate my hypothesis, I went through the list of mRNA transcripts which are reported in this study (Sharma and Wente, 2020) and grouped them by pathways. Of particular interest

were the pathways which are known to be disrupted in SMA cases already in the literature. My focus therefore came down to a list of 9 transcripts which covered the pathways of growth factors, cytoskeletal proteins, the MAPK pathway, ribosomal proteins, the Golgi and zinc finger proteins (**Table 5.1**). qPCR primers for all these transcripts were designed, ordered, and optimised. Three of the transcript pairs were suitably optimised within the time frame and so these transcripts were used for preliminary investigation. These genes were *Fos*; an immediate early gene involved in gene expression modulation, differentiation, and apoptosis, *ZFP91-CNTF*; a read-through transcript of the *ZFP91* and *CNTF* genes which is thought to be non-coding and *RPL13ap7*; a ribosomal pseudogene.

Table 5.1: A list of mRNA transcripts which are selectively retained in the nucleus when Gle1b's shuttling activity is blocked (Sharma and Wente, 2020)***.*** *Transcripts highlighted in yellow were optimised and used for subsequent qPCR analysis. All information about the roles of each transcript was found on uniprot.*

Pathway	*Transcript*	*Fold Change*	*Role (all taken from uniport.com)*
Growth Factor	EGR1	19.73	Transcriptional regulator
Cytoskeletal	TPP3	88.73	Regulator of microtubule dynamic
	XIRP1	9.79	Protect actin filaments during depolymerization
	RND1	8.68	Regulate the organization of the actin cytoskeleton in response to extracellular growth factors
MAPK Pathway	Fos	72.87	Regulation of cell proliferation, differentiation, apoptotic cell death
Ribosomal Proteins	RPL13Ap7	16.02	Pseudogene (non-coding)
	RPS17	7.87	Component of 40S subunit

Golgi	GOLGA6L18	20.74	Golgi-associated protein
Zinc Finger Proteins	ZFP91-CNTF	128	Readthrough transcript of nucleic acid binding protein and a nervous system hormone

For these experiments age and sex matched pair of fibroblasts - Control 1 (GM00489) and SMA 1 (GM09677) - were used for nuclear and cytoplasmic fractionation, followed by qPCR analysis of the transcript levels. Fractionation of the nucleus and cytoplasm was confirmed by western blotting of the protein lysates and probing for a nuclear marker (SSRP1) and cytoplasmic marker (Tuj1) (**Fig 5.11 A**). As mentioned previously in Chapter 4, there is some cytoplasmic contamination of the nuclear fractions. However, due to the nature of these fibroblast cells, this was the optimal fraction which we could obtain.

qPCR analysis shows that in the cases of *ZFP91-CNTF* and *RPL13a7* there was no difference in total levels between control and SMA cells (ZFP91-CNTF: one way ANOVA, p=0.8340, ns, n=4. RPL13a7: one way ANOVA, p=0.8776, ns, n=4). There was also no significant difference between the mRNA transcript levels in the nucleus and cytoplasm in control and SMA cells for each (*ZFP91-CNTF*: Control nuclear vs SMA nuclear: one way ANOVA, p=0.9647, ns, n=4; Control Cytoplasmic vs SMA Cytoplasmic: one way ANOVA, p=0.9973, ns, n=4. *RPL13a7*: Control nuclear vs SMA nuclear: one way ANOVA, p=0.9368, ns, n=3; Control Cytoplasmic vs SMA Cytoplasmic: one way ANOVA, p=0.9024, ns, n=3.) (**Fig 5.11 C, D**). *Fos* mRNA transcripts are significantly downregulated in the nucleus of SMA cells compared to controls (One way ANOVA; Control Nuclear vs Cytoplasmic Nuclear, p=0.0005, ***, n=4). However, when looking at the cytoplasmic fractions, there does not seem to be an increase in the mRNA transcripts retained in the nucleus as the cytoplasmic fractions are not significantly different from one another (One way ANOVA, Control cytoplasmic vs SMA cytoplasmic, p=8163, ns, n=4) despite the reduction seen in the total/nuclear levels (**Fig 5.11 B**).

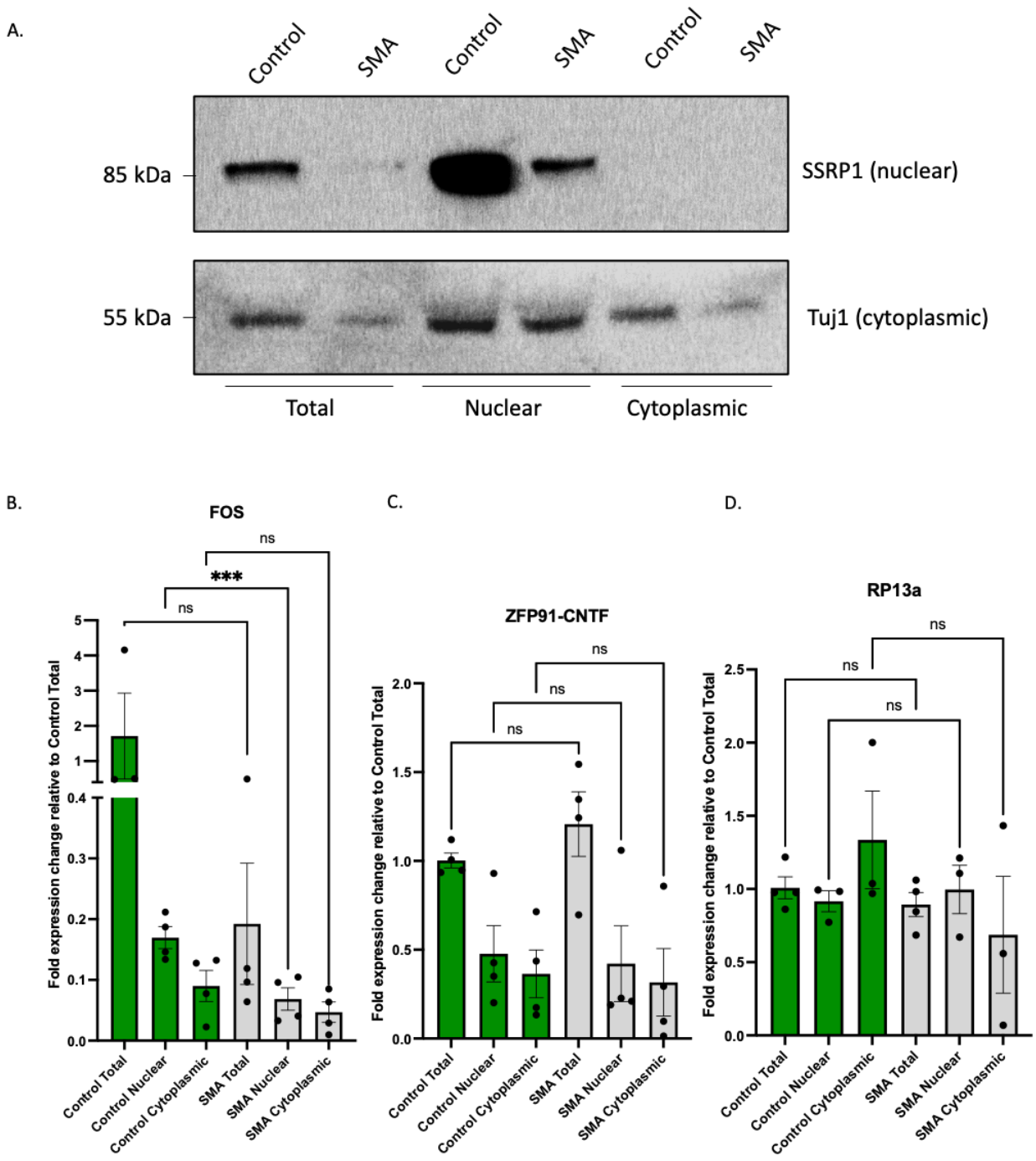


Figure 5.11: Nuclear/cytoplasmic distribution of mRNA transcripts selectively exported by Gle1b are not altered in SMA patient fibroblasts compared to control cells. (A) *Representative western blot showing successful fractionation of fibroblast samples: Control (GM00489) and SMA (GM09677). Western blot probed with anti-SSRP1 antibody to show nuclear protein expression at 85 kDa and anti-Tuj1 antibody to show cytoplasmic protein expression at 55 kDa.* ***(B)*** *Fold expressions change of FOS mRNA transcripts, worked out through delta delta CT analysis of qPCR data, shown relative to the Control total condition in all cases. Data displayed as mean +/- SEM, from four biologically independent replicates (N=4).*

*Data normalised to GAPDH loading control. FOS levels are reduced in SMA patient cells, but there is no difference in Cytoplasmic levels of mRNA transcripts (Control Total vs SMA Total; one way ANOVA, p=7081, ns, n=3. Control Nuclear vs SMA Nuclear; one way ANOVA, p=0.0005, ***, n=4. Control Cytoplasmic vs SMA Cytoplasmic; one way ANOVA, p=8163, ns, n=4)* ***(C)*** *Fold expressions change of ZFP91-CNTF mRNA transcripts, worked out through delta delta CT analysis of qPCR data, shown relative to the Control total condition in all cases. Data displayed as mean +/- SEM, from four biologically independent replicates (N=4). Data normalised to GAPDH loading control. There is no difference in mRNA transcript expression between any of the conditions (Control Total vs SMA Total; one way ANOVA, p=0.8340, ns, n=4. Control Nuclear vs SMA Nuclear; one way ANOVA, p=0.9647, ns, n=4. Control Cytoplasmic vs SMA Cytoplasmic; one way ANOVA, p=9973, ns, n=4).* ***(D)*** *Fold expressions change of RPL13a7 mRNA transcripts, worked out through delta delta CT analysis of qPCR data, shown relative to the Control total condition in all cases. Data displayed as mean +/- SEM, from three biologically independent replicates (N=3). Data normalised to GAPDH loading control. There is no difference in mRNA transcript expression between any of the conditions (Control Total vs SMA Total; one way ANOVA, p=8776, ns, n=4. Control Nuclear vs SMA Nuclear; one way ANOVA, p=0.9368, ns, n=3. Control Cytoplasmic vs SMA Cytoplasmic; one way ANOVA, p=9024, ns, n=3).*

5.2.3 Golgi dysregulation

During preliminary experiments staining for endogenous levels of Gle1 protein in the control and SMA fibroblast cells, it was noticed that the staining was picking up signal in a perinuclear structure. Investigative staining established that this staining colocalised closely with markers of both the cis- and trans-Golgi (GM130 and TGN46 respectively) (**Fig 5.12 A, B**).

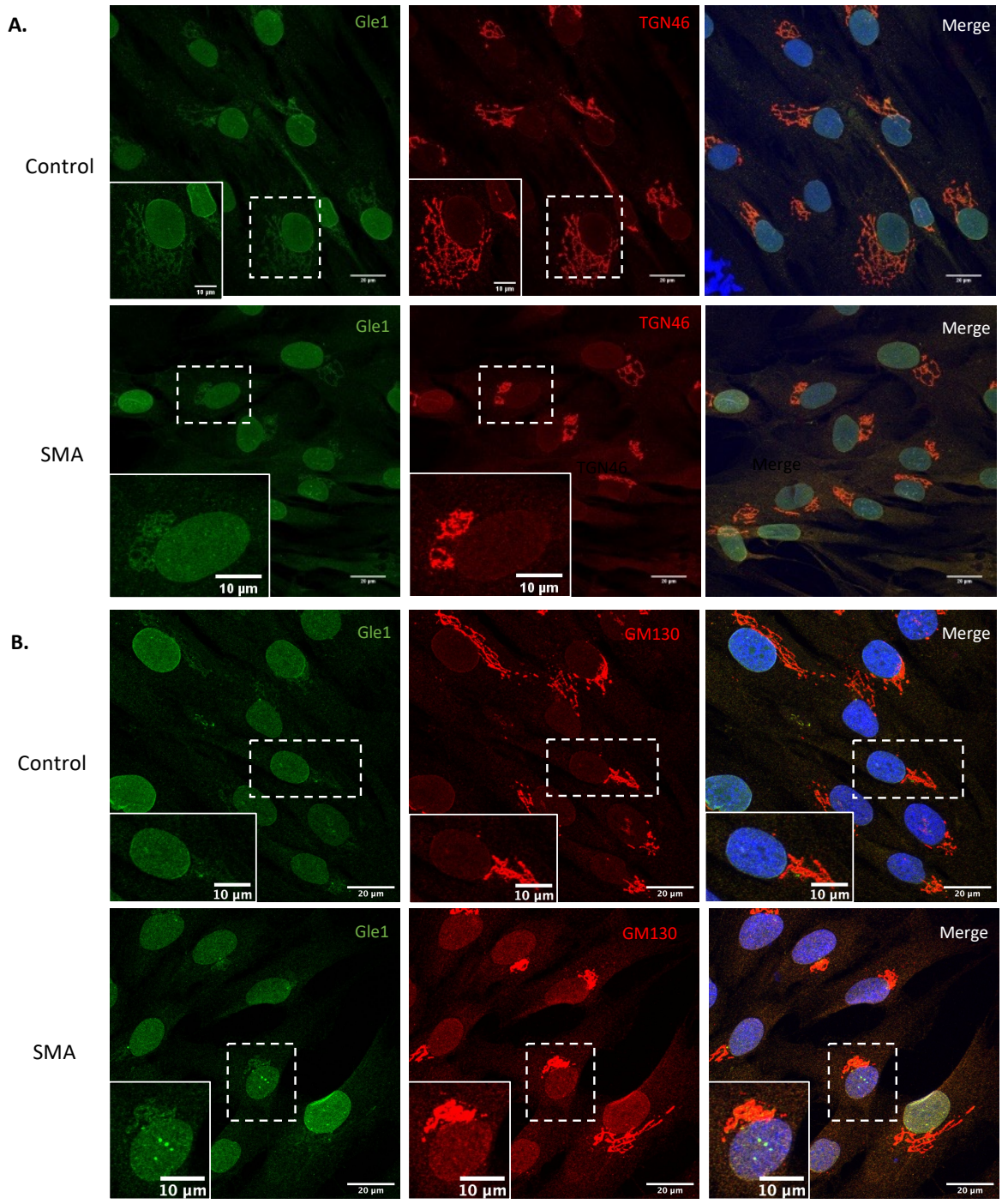
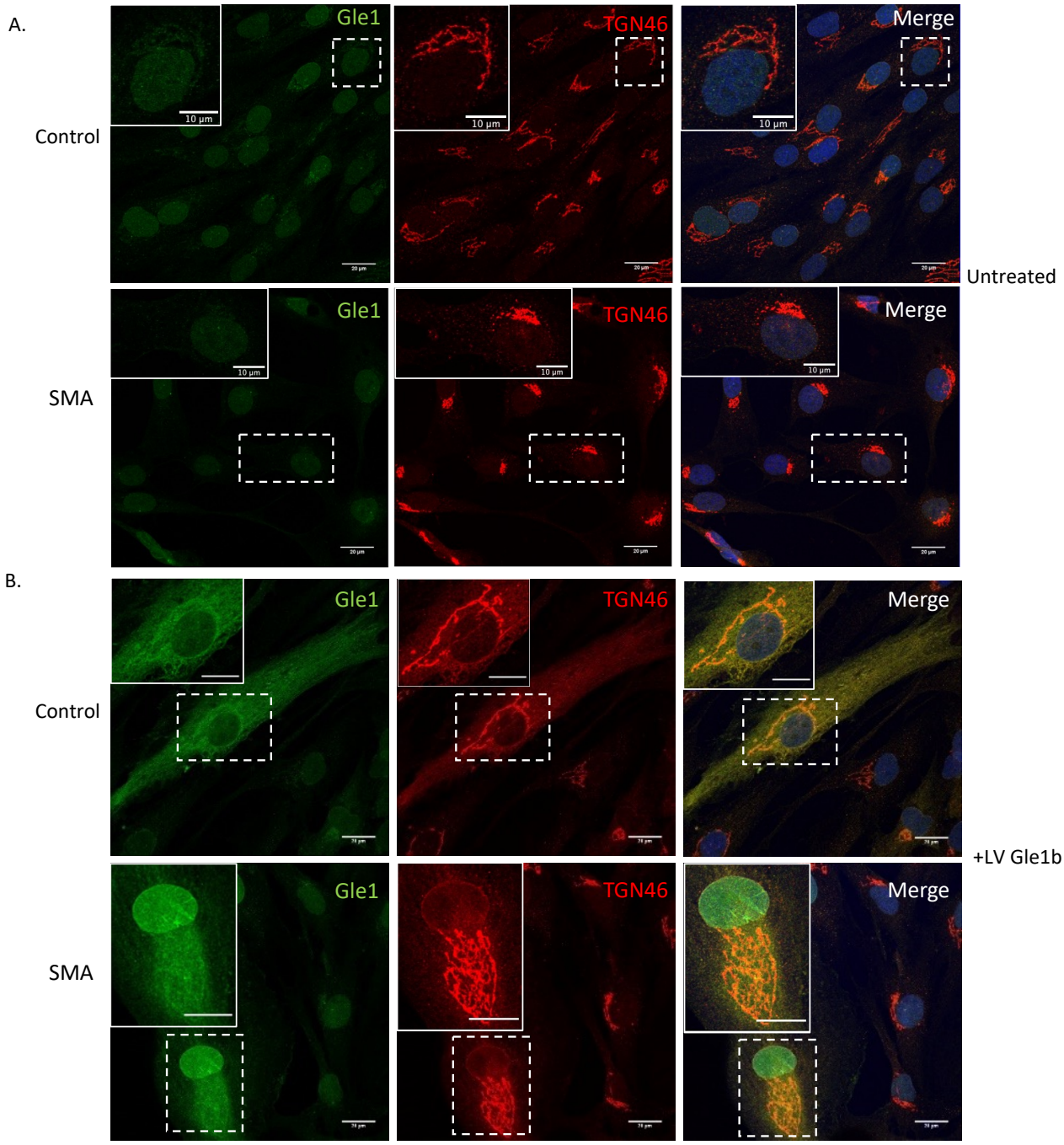


Figure 5.12: Localisation of Gle1b in trans and cis golgi in both control and SMA fibroblast cells (A) *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677) stained with TGN46 for visualising trans-golgi in red. Gle1 endogenous protein can be seen in green. Scale bar is 20 µM. For high magnification images scale bar is 10 µM, area taken as high magnification is marked in white. Data collected from three biologically independent replicates (n=3).* ***(B)*** *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677) stained with GM130 antibody for visualising cis-golgi in red. Gle1 endogenous protein can be seen in green. Scale bar is 20 µM. For high magnification images scale bar is 10 µM, area taken as*

high magnification is marked in white. Data collected from three biologically independent replicates (n=3).

Initial impressions from this staining demonstrated that there appeared to be a difference in the Golgi morphology between the control cells and the SMA patient cell lines. There is evidence in the literature that SMN is linked to the Golgi network (Ting et al., 2012), and that there has been abnormal Golgi morphology reported with decreased levels of COPI in SMA cells (Custer et al., 2019). Therefore, I was interested to see whether this difference in morphology was consistent with that reported in the literature, as well as assessing whether overexpression of Gle1b could act to restore the morphology to that of the control cells.

Control and SMA patient cells were therefore stained with the marker of the trans Golgi; TGN 46 as well as a Gle1 antibody to assess both endogenous Gle1 levels alongside overexpressed Gle1b protein (**Fig 5.13 A, B**). Golgi morphology was categorised using the method demonstrated in (Custer et al., 2019). Cells were categorised according to the scale seen in **Fig 5.13 C**, with 1 being a more clustered morphology and 3 being a more diffuse, 'ribbon'-like structure. Quantification of three biologically independent replicates was performed blinded by two individuals, as the outcome was consistent between the two quantifications only one has been displayed here (**Fig 5.13 D, Appendix 2**). This analysis demonstrated that the control cells had a ribbon like structure (typically associated with Golgi morphology), while the SMA cells displayed a much more compacted structure. When Gle1b was overexpressed, using a previously generated LV Gle1b virus, the morphology of the Golgi in cells showing increased Gle1 staining appeared to be more consistent with that of the control cells than the untreated SMA cells, therefore resulting in many more of the cells being categorised as category 3 rather than 1 prior to treatment. (**Fig 5.13 D**).


A.
Control
SMA
Gle1
TGN46
Merge
Untreated
B.
Control
SMA
+LV Gle1b

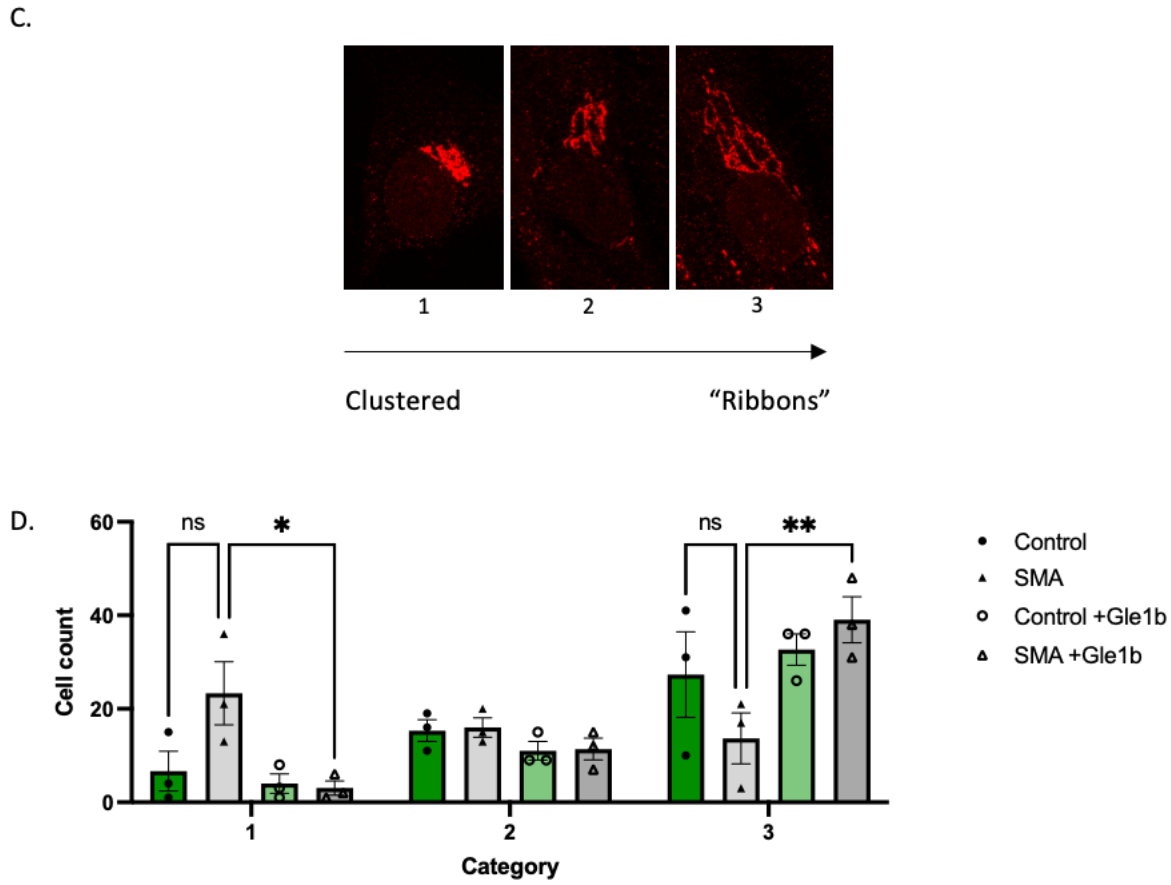


Figure 5.13: Golgi has a more 'clustered' phenotype in SMA fibroblast cell lines. Overexpression of Gle1b led to the rescue of the disrupted Golgi phenotype in SMA fibroblast cell lines. (A) *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677) stained with TGN46 antibody for visualising trans-Golgi in red. Gle1 protein can be seen in green. Scale bar is 20 µm. Data collected from three biologically independent replicates (n=3). Zoomed scale bar is 10 µm.* ***(B)*** *Control fibroblast cells (GM8680) and SMA patient fibroblasts (GM9677), treated with LV Gle1b (MOI 20) for four days, stained with TGN46 antibody for visualising trans-Golgi in red. Gle1 protein can be seen in green. Scale bar is 20 µm. Data collected from three biologically independent replicates (n=3). Zoomed scale bar is 10 µm.* ***(C)*** *Method used for categorisation of Golgi staining. Adapted from* (Custer et al., 2019). ***(D)*** *Quantification of cells categorised by method shown in C. Data displayed as mean +/- SEM. Data collected from three biologically independent replicates (n=3), 60 cells per condition. SMA cells tend to show a more clustered phenotype compared to controls (Two-way ANOVA, Control vs SMA, Category 1; p=0.0634, ns, n=3. Control vs SMA, Category 3; p=1606, ns, n=3). When Gle1 protein is overexpressed, there is a significant increase from Type*

*1 to Type 3 in SMA cells (Two-way ANOVA, SMA vs SMA+Gle1b, Category 1; p=0.0.0276, *, n=3. SMA vs SMA+Gle1b, Category 3; p=0.0026, **, n=3).*

As explained previously, as SMA is a disease which specifically effects motor neurons, I was interested to see whether this observed phenotype was conserved in iPSC-derived motor neurons. In the interest of time, these experiments were imaged on the Opera Phenix high throughput imaging system, meaning that the images are a lower magnification than previously. Here, two sets of control and two sets of SMA patient motor neurons were stained with trans-Golgi marker TGN46 (in yellow) to visualise the Golgi Apparatus (**Fig 5.14 A**), neurons were stained with neuronal marker MAP2 (red) and nuclei with Hoechst (Blue). Categorisation of the Golgi morphology was done as previously described by (Custer et al., 2019) (**Fig 5.14 B**). As seen in the fibroblast cells, the SMA patient cells tended to display a more clustered Golgi morphology than the control cell lines (**Fig 5.14 C**). Due to time limitations and Covid-19 related constraints, rescue experiments including Gle1b overexpression in motor neurons are still on going.

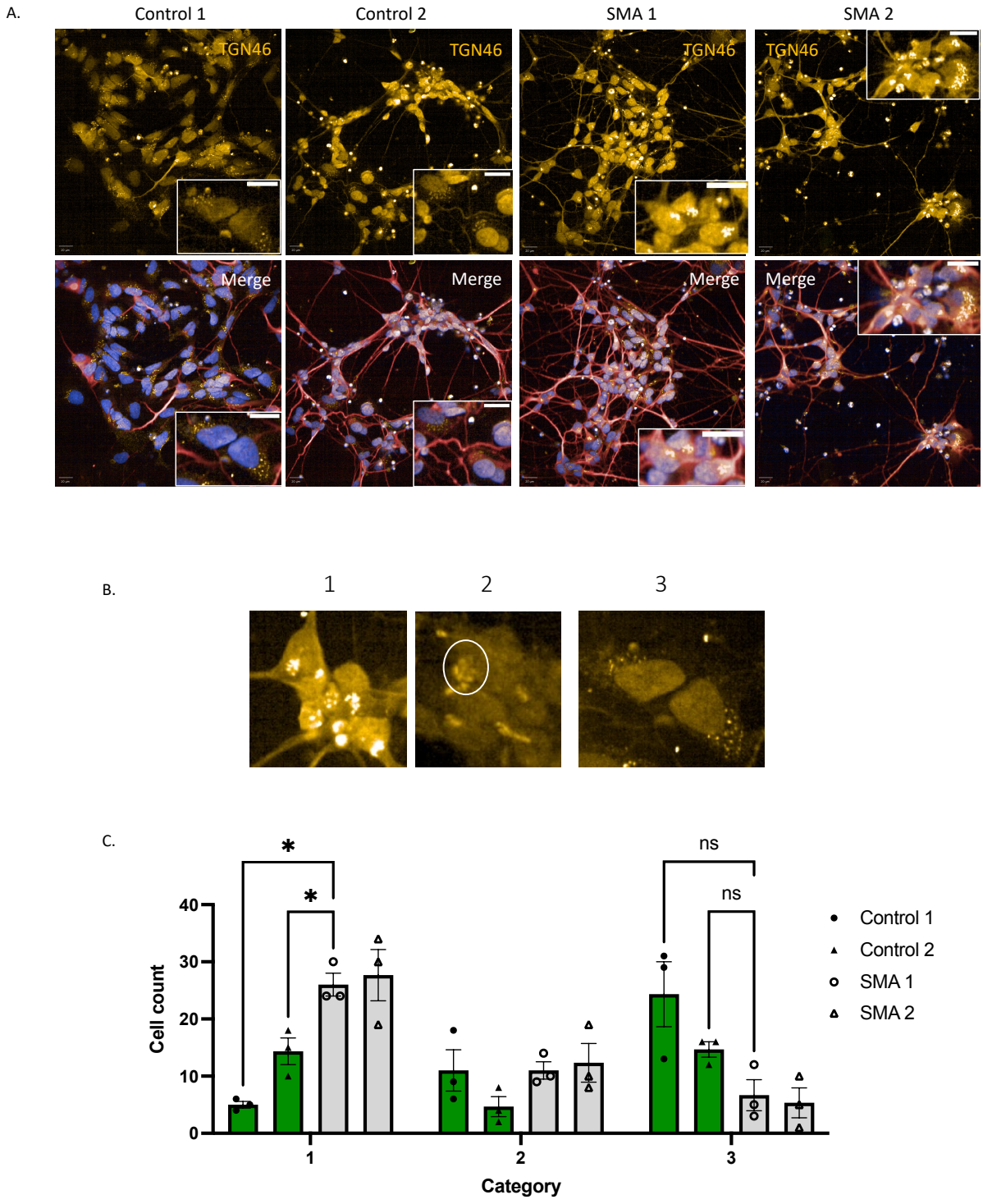


Figure 5.14: Golgi phenotype observed in SMA fibroblasts is conserved in iPSC-derived motor neurons. (A) *Immunocytochemistry of Control 1 (Cs14), Control 2 (GM), SMA 1 (SMA 77) and SMA 2 (SMA 84) iPSC derived motor neurons. TGN46 visualised in yellow, motor neuron axons visualised by MAP2 in red with nuclei in blue. Scale bar is 20 µM. Data collected from three biologically independent replicates (n=3).* ***(B)*** *Method used for categorisation of golgi staining.*

Adapted from (Custer et al., 2019). ***(C)*** *Quantification of cells categorised by method shown in C. Data displayed as mean +/- SEM. Data collected from three biologically independent replicates (n=3), 40 cells per condition. SMA cells tend to show a more clustered phenotype compared to controls (Two-way ANOVA, Control 1 vs SMA 1, Category 1; p=0.0133, *, n=3. Control 2 vs SMA 1, Category 1; p=0.0376, *, n=3. Control 1 vs SMA 1, Category 3; p=0.1855, ns, n=3. Control 2 vs SMA 1, Category 3; p=0.4115, ns, n=3).*

To explore potential mechanisms of the observed changes in SMA cases, I wanted to investigate the pathways involved in Golgi morphology. GOLPH3 is a Golgi membrane tether which when downregulated promotes a condensed Golgi morphology (Dippold et al., 2009) (**Fig 5.15**). Through interaction with Phosphatidylinositol-4-phosphate (PtdIns(4)P) at the Golgi membrane, GOLPH3 binds MYO18A to link the Golgi to F-actin filaments. This interaction causes tensile force and extension of the Golgi network thus forming the ribbon structure. GOLPH3 can be phosphorylated by the DNA damage responder; DNA-Pk. Therefore, activation of GOLPH3 by upregulation of DNA-Pk in response to increased DNA damage cause increased tensile strength to be placed on the Golgi network by F-actin, therefore leading to Golgi fragmentation. In SMA cases, it has been reported that despite increased levels of DNA damage, there is a downregulation of DNA-Pk specifically (Kannan et al., 2020). Hence, I hypothesise that within these SMA patient cells which are displaying a compacted Golgi phenotype that there could potentially be a reduction in the levels of GOLPH3, alongside a reduction in DNA-Pk which may be resulting in a downregulation of this GOLPH3 based pathway.

To assess this, protein lysates taken from iPSC-derived motor neurons were run on western blots and probed with antibodies against DNA-Pk and GOLPH3 (**Fig 5.16 A**). Three control motor neuron cell types (MIFF1 (MN Control 1), Cs14 (MN Control 2) and GM (MN Control 3)) were used alongside two SMA patient motor neuron cell types (SMA 77 / MN SMA 1 and SMA 84 / MN SMA 2). These experiments were performed on day 33 of culture, after motor neuron maturation, from three separate differentiations (n=3, biological replicates). As before, endogenous protein expression was calculated through densitometric analysis (**Fig 5.16 B-E**).

There was no difference between the protein expression levels of DNA-Pk (**Fig 5.16 B, C**) or GOLPH3 (**Fig 5.16 D, E**) in control motor neurons compared to SMA patient motor neurons.

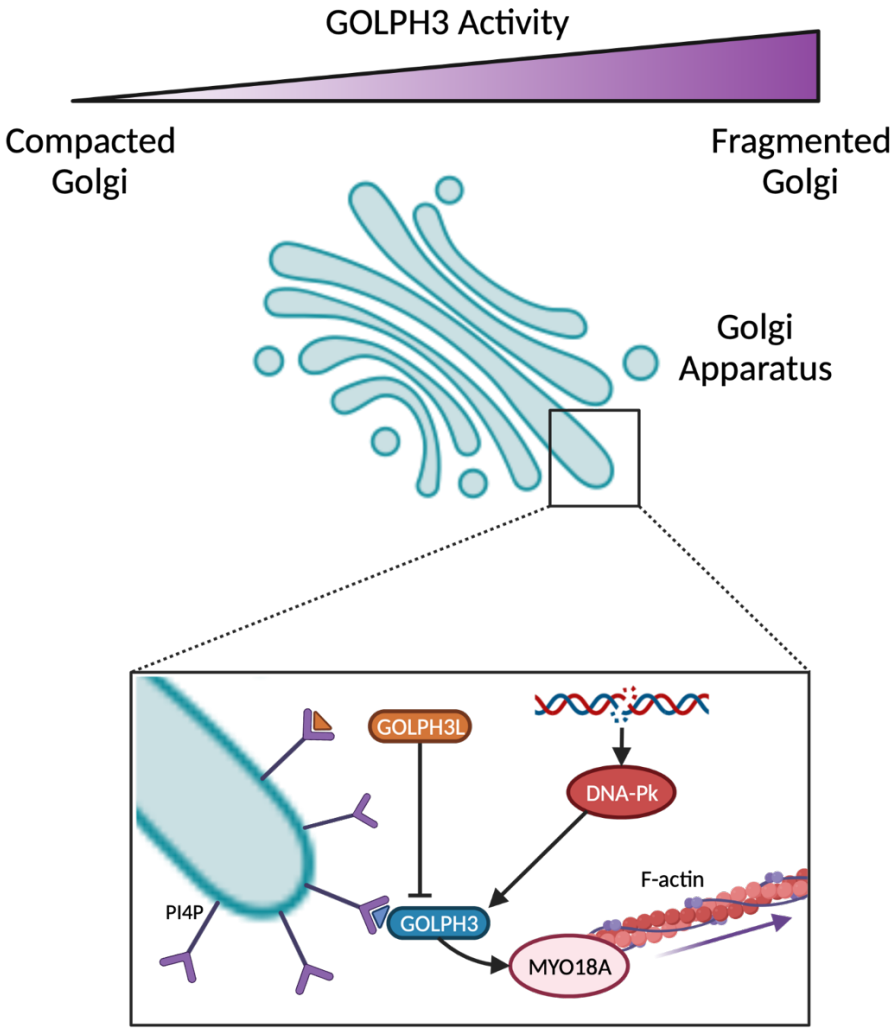
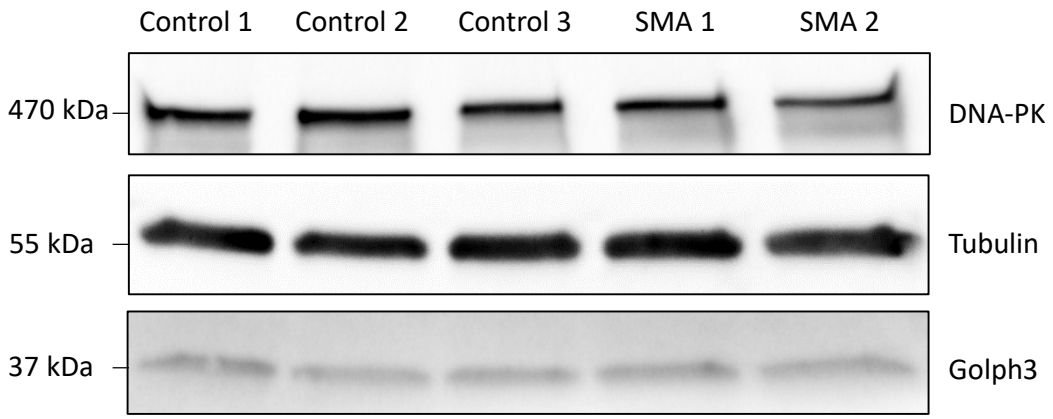
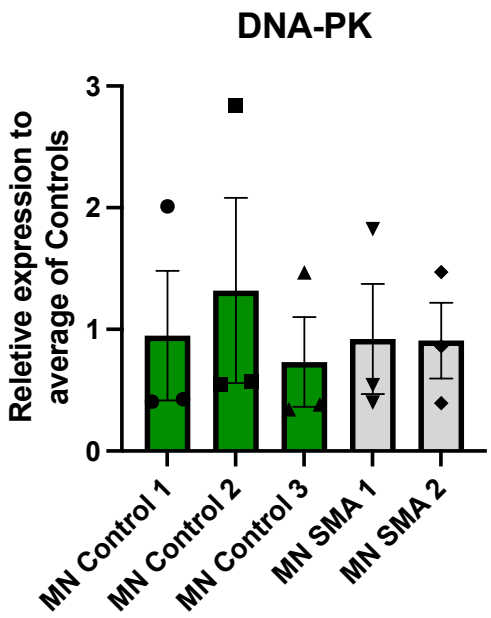


Figure 5.15: Potential pathways related to Golgi dysfunction. *The levels of GOLPH3 activity can influence the morphology of the Golgi. Potential low levels of GOLPH3, can mean that the Golgi appears compacted due to actin depolymerisation. High levels of GOLPH3 activity stimulates MYO18A, which in turn causes F-actin to exert tensile strength onto the Golgi, causing it to form its classic ribbon-like structure. Pathways such as the DNA damage response cause activation of DNA damage responder DNA-Pk, this increases the activity of GOLPH3, therefore stimulating Golgi extension by F-Actin through MYO18A, eventually resulting in Golgi fragmentation* (Dippold et al., 2009). *Figure created with biorender.com.*

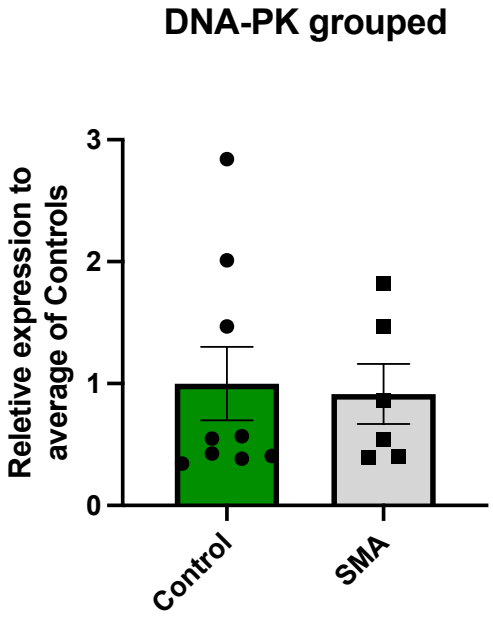
A.


Control 1
Control 2
Control 3
SMA 1
SMA 2
470 kDa
DNA-PK
55 kDa
Tubulin
37 kDa
Golph3

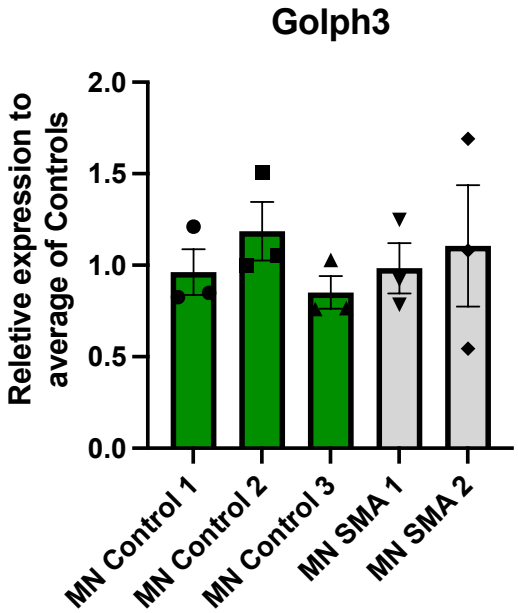
B.


DNA-PK
Relative expression to average of Controls
0
1
2
3
MN Control 1
MN Control 2
MN Control 3
MN SMA 1
MN SMA 2

C.


DNA-PK grouped
Relative expression to average of Controls
0
1
2
3
Control
SMA

D.


Golph3
Relative expression to average of Controls
0.0
0.5
1.0
1.5
2.0
MN Control 1
MN Control 2
MN Control 3
MN SMA 1
MN SMA 2

E.

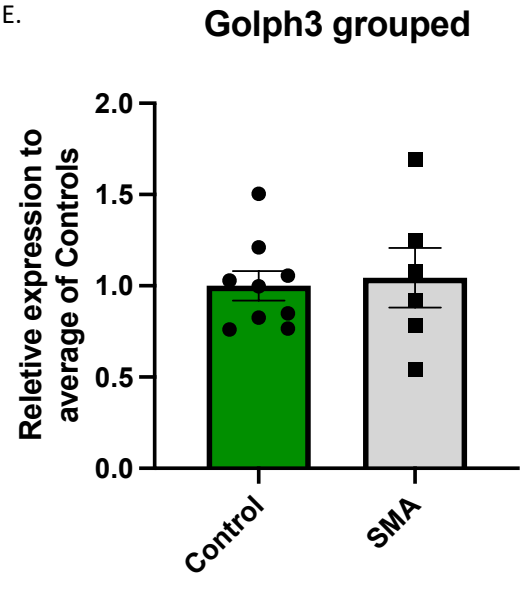

Golph3 grouped
Relative expression to average of Controls
0.0
0.5
1.0
1.5
2.0
Control
SMA

Figure 5.16: There is no difference in protein expression levels of DNA-Pk and Golph3 between control and SMA iPSC-derived Motor neurons. ***(A)*** *Representative western blot showing protein expression in human iPSC-derived motor neurons: MN Control 1 (Miff1), MN Control 2 (Cs14), MN Control 3 (GM), MN SMA 1 (SMA 77), MN SMA 2 (SMA 84). Western blot probed with anti-DNA-Pk antibody and anti-Golph3 antibody, showing protein expression of DNA-Pk at 490kDa and Golph3 at 36 kDa respectively. Tubulin (at 55kDa) is used to show equal loading of samples.* ***(B)*** *Densitometric analysis of DNA-Pk protein expression levels shows no significant difference across the different cell types (one-way ANOVA, p>0.05, ns, n=3). Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal. Data normalised to the average of the controls.* ***(C)*** *Values from each cell type were grouped into Controls vs SMA. There is no difference between DNA-Pk levels in controls vs SMA cells. Data displayed as mean +/- SEM.* ***(D)*** *Densitometric analysis of SMN protein expression levels shows no significant difference across the different cell types (one-way ANOVA, p>0.05, ns, n=3). Data displayed as mean +/- SEM. Protein loading was normalised using Tubulin signal. Data normalised to the average of the controls.* ***(E)*** *Values from each cell type were grouped into Controls vs SMA. There is no difference between Golph3 levels in controls vs SMA cells. Data displayed as mean +/- SEM.*

5.3 Discussion

This chapter aimed to see whether overexpression of Gle1b protein could act to rescue any of the SMA disease phenotypes in our experimental models. I looked at established disease phenotypes of SMA, starting with DNA damage. Overexpression of Gle1b did not have any protective effect when DNA double strand breaks were induced by CPT treatment. There was also no significant rescue of endogenous DNA double strand breaks in SMA fibroblast cells which were treated with a lentivirus overexpressing Gle1b.

I then looked at whether there were any observable RNA export defects in our SMA cell models. After some promising preliminary data, I did not observe any significant nuclear accumulation of poly (A)+ RNA in SMA patient fibroblasts or motor neurons. A subset of mRNA transcripts which were reported previously to specifically require Gle1b for nuclear export were assessed in SMA patient fibroblasts (Sharma and Wente, 2020). This allowed me to see whether nuclear export of these transcripts is disrupted in SMA cells. There appeared to be no obvious defects in nuclear export of a number of these transcripts between control and SMA cases, despite a novel global reduction in *Fos* transcript levels between control and SMA cases.

Finally, I evaluated the morphology of the Golgi apparatus in SMA cells compared to controls. The Golgi morphology in SMA fibroblasts and motor neurons appears to display a more clustered phenotype than that of the control cells, a finding which has not previously been identified. In fibroblast cells, overexpression of Gle1b led to the rescue of the abnormal morphology of the Golgi. GOLPH3 was identified as a potential target for this pathway, with downregulation of GOLPH3 being shown to result in a more clustered Golgi phenotype. I therefore evaluated the levels of GOLPH3 in these cells alongside DNA-Pk, a DNA damage responder which evidence suggests is also downregulated in SMA cases. However, in this case, there was no difference in the protein levels of either of these proteins in our cell models.

5.3.1 Gle1b plays no part in the DNA damage response and endogenous DNA damage levels in SMA cases are not rescued by overexpression of Gle1b

Gle1b acts as a regulator of the DEAD-box helicase DDX19B; the activation of DDX19B ATPase activity by Gle1b and IP_6 facilitates export of mRNA from the nucleus into the cytoplasm (Alcázar-Román et al., 2006; Montpetit et al., 2011; Weirich et al., 2006). Knockdown of Gle1, DDX19 and IP_6 all resulted in increased γH2AX levels, with Gle1 knockdown showing the most significant increase. A reduction in mRNA for DNA repair factors (BRCA1 and FANCD2) were also observed when any of these three factors were knocked down, suggesting a delay in the DNA damage response (Okamura et al., 2018). Other studies have also identified knockdown of DDX19B to result in increased γH2AX signal (Paulsen et al., 2009) and as a result, selective activation of the ATM kinase pathway (Hodroj et al., 2017a).

R-loops are formed from displaced single stranded DNA molecule and a RNA:DNA hybrid. They form frequently during transcription as a result of replicative stress-induced genomic instability (Aguilera and García-Muse, 2012). These R-loops can expose the coding ssDNA and result in DNA breaks through exposure to DNA deaminases, which can give rise to DNA double strand breaks (Rinaldi et al., 2021). Within mammalian cells, SETX is a nuclear protein which can help to facilitate resolution of R loops during the termination of transcription (Alzu et al., 2012; Skourti-Stathaki et al., 2011). More recent papers have linked SETX and DNA repair factor BRCA1 as a complex which specifically repairs R-loop based DNA damage located within transcriptional pause sites (Hatchi et al., 2015; Hill et al., 2014). DDX19B has been proposed as a novel resolver of these R-loop structures, acting independently of SETX and functioning through an ATR-Chk1 dependant mechanism, which therefore acts in response to replication stress (Hodroj et al., 2017a).

SMA cases possess widespread RNA processing defects, largely intron retention, due to the major role of SMN in the spliceosome. This further results in induction of DNA damage through the formation of R-loops during transcription (Jangi et al., 2017). There are also reports of a reduction in the levels of SETX within SMA fibroblast cells alongside increased γH2AX and R-loop signal. Despite there being increased levels of DNA damage, there was also a significant downregulation of DNA-PK, therefore causing defects within the non-homologous end joining (NHEJ)-mediated pathway of DNA repair (Kannan et al., 2018).

Therefore, due to the links between the Gle1b/DDX19/IP_6 complex and R-loop resolution I was interested to look at Gle1 in the context of SMA cases to see if there was any link to the increased formation of R-loops. Immunocytochemistry of Gle1 and R-loops (visualised by the S6.9 antibody) demonstrated co-localisation between the two staining's. This mainly did occur in the nucleolus, as demonstrated by co-localisation with nucleolin antibody. This corresponds to R-loops which are forming during transcription of ribosomal RNA by RNA polymerase I (RNAP I).

Antibody staining with S9.6 has previously been optimised in our group with RNase H treatment to show resolution of the R-loops (Walker et al., 2017). This would be interesting in this situation to confirm that the staining we are seeing from the Gle1 antibody is also specific to R-loops themselves and not just a coincidence or interaction with another protein within the nucleolus. A study by Sharma and colleagues published in 2020, also confirmed the co-localisation of Gle1 with R-loops through proximity ligation assay, therefore supporting our observations in the absence of this control condition (Sharma and Wente, 2020).

To evaluate whether Gle1b plays a role alongside DDX19B in the resolution of R-loops (and therefore DNA double strand break resolution), we first set out to assess whether exogenously induced DNA damage caused by Campthothecin (CPT). This is a topioisomerase I inhibitor and so induces DNA double strand breaks through blocking the super coil relaxing ability of Top I which occurs during transcription and replication. Here I stained cells with γH2AX, a widely used marker of double strand breaks (DSB). I decided to use this marker due to it being an early-stage marker of DNA damage as it recognises double stand breaks and forms foci around them, these foci represent the DSB's quantitatively and so can be used as a direct readout of DNA damage levels within the cell. Ideally we would have also used a marker for R-loops, however due to widespread issues with the reliability of this antibody and the staining it produces we decided to use γH2AX as the main marker (Ed et al., 2021; Smolka et al., 2021). In this assay, I hypothesised that if Gle1b was in some way involved in recruitment of DNA repair proteins then overexpression may help either to be in some way protective to the cells and prevent a significant rise in γH2AX staining or improve the speed of the cell's recovery from the induced DSBs. However, there was no reduction in the amount of induced DNA damage in HeLa cells when Gle1b was overexpressed both upon immediate

fixation of the cells, and after recovery periods of 30 minutes and an hour. This therefore indicated that there was no involvement of Gle1b with the DNA damage response.

It would also be interesting to assess whether DDX19B protein expression or localisation is affected in SMA cases, as disruption to an additional R-loop resolving helicase may also be contributing to the increases seen in R-loops/DNA damage levels in SMA cases (Kannan et al., 2020, 2018).

As previously mentioned, a study published later into this PhD also demonstrated clearly that Gle1 co-localises with R-loops through positive signal in a proximity ligation assay (Sharma and Wente, 2020). This study also demonstrated that blockage of Gle1 shuttling from the nucleus to the cytoplasm increased the levels of S9.6 staining within the nucleus (excluding the nucleolar staining). They link this rise in R-loops to the mRNA's which Gle1 selectively exports from the nucleus being rich in R-loops formed during their transcription rather than due to other means of DNA damage. Which they confirm by seeing no increase in γH2AX signal in the cells where Gle1 shuttling is blocked. This is verified by also inducing DNA damage with hydroxyurea and again blocking Gle1 nuclear shuttling, here they see no changes in the levels of γH2AX intensity. Therefore linking Gle1b to R-loops formed during transcription rather than the DNA damage response (Sharma and Wente, 2020).

I was interested to see whether overexpression of Gle1b would restore the reduction in Gle1 levels which we saw in SMA cells in **Chapter 4** and may therefore act to ameliorate some of the disease phenotypes. Cells transduced with lentivirus expressing Gle1b did not show any reduction in the levels of γH2AX staining intensity compared to untreated SMA cells. Therefore, indicating that overexpression of Gle1b does not help to rescue disrupted R-loop formation which results in increased levels of DNA damage.

However, it appears as if overexpression of Gle1 within healthy cells may predispose them to cellular stress. The HeLa cells in **Figure 5.4** overexpressing Gle1b appear much more rounded when exposed to CPT and when treating healthy fibroblasts with LV FLAG Gle1b, there is a significant increase in γH2AX staining (**Fig 5.5**). This may indicate that if Gle1 is to be used to ameliorate disease phenotypes then careful consideration needs to be paid to the dosage.

Unfortunately, there was no difference in the levels of γH2AX signal between control and SMA patient motor neurons. This could have been due to external stress on these cells such as CO_2 levels or cell culture conditions as during the time that these experiments were conducted there were multiple issues in the tissue culture lab. Due to time restraints these experiments could not be repeated.

5.3.2 Discrepancies in global polyadenylated RNA export defects in SMA patient cells

As Gle1b plays a major role in the export of mRNA from the nucleus to the cytoplasm, I next wanted to investigate whether there were any RNA export defects in the SMA patient cells which could be related to a reduction in Gle1 protein levels. Reports in the literature indicated that in motor neurons isolated from the SMN Δ7 mouse model there was the presence of nuclear aggregates of polyadenylated mRNA, alongside a cytoplasmic depletion of polyadenylated mRNA (Casafont et al., 2010; Narcís et al., 2018).

I therefore stained cells for poly(A)+ RNA using an oligo-dT Cy3 probe. Initial confocal imaging of staining in fibroblasts showed a promising trend with there being a significant reduction in the poly(A)+ mRNA within the cytoplasm of SMA fibroblasts compared to the control cells. Unfortunately, due to time restraints a different method of imaging was used for overexpression experiments. For these experiments I used the automated Opera Phenix high throughput imaging system, however, there was no observable difference between patient and control cells and so despite there being clear overexpression of Gle1b in the treated cells, I did not continue with these experiments in this case. This difference in imaging utilised a different magnification (40x lens vs 63x lens on the confocal), however, this should not be a large enough difference to account for a difference in overall signal. Therefore, this could indicate that there could have potentially been issues with the probe. Further investigation would be required to check the experiment is working as required, for example including a control condition treated with a drug such as actinomycin A. This would block transcription and so show a reduction in mRNA in the nucleus of treated cells, therefore validating the specificity of the probe.

5.3.3 Transcripts which are specifically exported through Gle1-mediated pathways do not appear to be disrupted in SMA patient fibroblasts

As there did not appear to be a global RNA export defect seen in these cells, I hypothesised that there might be a more specific set of transcripts which were disrupted due to a reduction in Gle1 protein levels (Sharma and Wente, 2020). The original list of mRNA transcripts disrupted when Gle1 nuclear shuttling is blocked is quite extensive, containing 70 candidates. Therefore, I wanted to narrow it down before ordering primers for optimisation. By grouping the transcripts into common pathways, it became easier to see which ones could potentially have some relation to defects seen in SMA cases.

Optimisation of primers for many of the selected transcripts proved difficult due to the low expression levels within the non-neuronal cell type which we were using as our model system. However, the three primers which I did manage to optimise did give us a good representation across the pathways which we wished to look at.

Overall, there was no difference in any of the mRNA transcripts in the mRNA distribution between the control and SMA cell types. Therefore, it appears as if Gle1b mediated transport of mRNA from the nucleus to the cytoplasm isn't disrupted in these cells despite the decrease in Gle1 protein levels.

Interestingly, although proportionally there was no difference in the nuclear/cytoplasmic distribution of mRNA transcript levels for *FOS* in control and SMA cells, there is a significant decrease in *FOS* mRNA transcripts in SMA cells compared to controls. *Fos* is an immediate early gene which forms heterodimeric complexes with members of the Jun family creating the AP-1 transcription factor. c-Fos has been shown to be responsible for regulating cellular mechanisms which mediate neuronal excitably and survival (Zhang et al., 2002), neuronal activity and plasticity (Joo et al., 2015). Previously in the literature, c-Jun NH2-terminal kinase 3 (JNK3) has been implicated in neuronal degeneration which arises from SMN deficiency (Genabai et al., 2015). The involvement of JNK activation in neuronal apoptosis has been also been characterised in other neurodegenerative disorders previously (Eilers et al., 1998; Kuan et al., 2003; Le-Niculescu et al., 1999; Morishima et al., 2001). c-Fos has previously been shown to downregulate the interaction between c-Jun and ATF2 and so reduce the activation

of apoptotic pathways (Yuan et al., 2009), therefore downregulation in SMA cases of c-Fos may explain the activation of JNK3 pathways which lead to neurodegeneration.

5.3.4 Gle1b overexpression rescues the abnormal Golgi morphology in SMA fibroblast and motor neurons.

The Golgi Apparatus is a dynamic organelle and can be affected by multiple different cellular processes: such as stress, DNA repair, trafficking and replication (Kulkarni-Gosavi et al., 2019; Makhoul et al., 2019; Stalder and Gershlick, 2020). A range of neurodegenerative disorders have been demonstrated to show morphological defects within the Golgi, including amyotrophic lateral sclerosis (Statland et al., 2015; Sundaramoorthy et al., 2015) and Alzheimer's disease (Joshi et al., 2015).

Links between the Golgi and SMA have been investigated previously in the field. Mutations within *bicaudal D homolog 2 (BICD2)*, a golgin and motor adaptor protein, have been identified as pathological in three families affected with autosomal dominant SMA (Neveling et al., 2013). The Coatomer complex is involved in mediating Golgi trafficking pathways. A major subunit of this complex, alpha-COP, has been identified as an SMN binding protein which is important for SMN localisation to axonal growth cones and is in turn downregulated in SMN depleted cells (Custer et al., 2019; Peter et al., 2011).

In this chapter, I describe an observed morphological difference between SMA patient and control cells. This was blindly categorised by two independent operators, which showed that the SMA patient cells displays a clustered Golgi phenotype than the control cells. This appeared to be conserved across both fibroblast cells and iPSC-derived motor neurons. Interestingly this is the opposite phenotype of what is observed by Custer and colleagues (2019), who see a more fragmented Golgi in their fibroblast populations compared to their control cells. This difference could be due to them using an antibody for the cis-Golgi (GM130) as opposed to the trans-Golgi marker used in my experiment (TGN46). However, colocalization of the two antibodies in these fibroblast lines display a very similar morphology in our hands.

Staining of Gle1 by immunocytochemistry was observed within Golgi structures in the fibroblast cells, whether this is an artefact of the staining or has any functional relevance remains an unanswered question which would need further investigation to assess. Golgi fractionation would allow us to assess more accurately whether Gle1 protein is localising to the Golgi apparatus (Taguchi et al., 2003). There are studies in yeast which demonstrate that Gle1 interacts with SEC15, which is a protein which is responsible for vesicle trafficking from the Golgi to the cell surface (Bowser and Novick, 1991; Davierwala et al., 2005). However, this is yet to be confirmed in mammalian systems. The human homolog of SEC15 is Exocyst Complex component 6B (EXOC6B) (NCBI, 2021). There is no current published research looking at whether this relationship between Gle1 and EXOC6B is conserved in mammalian systems.

Upon overexpression of Gle1b, the morphology of the Golgi Apparatus within SMA patient cells shows restoration to a more diffuse morphology as seen in the control cells. However, as mentioned previously due to the highly dynamic nature of this organelle, morphology may not be the most accurate indication of dysfunction. Therefore, we wanted to also look at markers which may change in response to pathogenic or beneficial changes within the cell. Interestingly, one of the transcripts whose nuclear export was specifically mediated by Gle1 was a Golgi related protein (GOLGA6L18). Unfortunately, due to low transcript levels I was unable to assess whether this was disrupted in SMA cases.

GOLPH3 is a trans-Golgi membrane bound protein which binds to actin filaments through MYO18A, thus generating acto-myosin tensile force. Depending on whether GOLPH3 is activated or downregulated, this can either cause the Golgi to extend or compact respectively. Therefore, in our case, we hypothesised that the levels of GOLPH3 may be being downregulated in SMA patient cell models due to the morphological phenotype which was being observed. However, expression of GOLPH3 protein appeared to uniform across all cell types. It may perhaps be more beneficial to look for ways to observe GOLPH3 phosphorylation to measure its activity levels rather than looking at protein levels.

As DNA-Pk is also implicated in activation of the GOLPH3 pathway, I hypothesised that the lack of DNA-Pk which has been reported in SMN deficient cells may be causing the Golgi to be displaying a more compacted morphology rather than the extended one seen in the controls.

In the cell models used in this study there does not appear to be any difference in the DNA-Pk levels in motor neurons between control and SMA patient cells. However, the DNA-Pk protein is very large (470 kDa), so perhaps further optimisation of this antibody is required to accurately represent the protein levels within these cells.

5.3.5 Conclusions

In this chapter I have assessed multiple pathways in SMA patient fibroblasts and iPSC-derived motor neurons. I have described novel findings demonstrating a disrupted Golgi morphology in SMA patient fibroblasts and iPSC-derived motor neurons. I have also shown a significant reduction in the expression of *Fos* mRNA transcripts in SMA patient fibroblasts, this is not a finding which has been described in the literature and so is an interesting avenue for future research.

Overall, it is unclear whether overexpression of Gle1b protein is beneficial to these cells and further research is needed to reach a conclusion either way. There does appear to be rescue of certain SMA disease phenotypes, Golgi morphology and axonal growth defects (as shown previously in the lab). However whether this improves the survival of the cells has not been assessed. Preliminary evidence seems to suggest that a reduction in the levels of Gle1 protein in SMA cells does not have any impact on the export of RNA from the nucleus. It would be useful here to look at other pathways in which Gle1 has been shown to be involved to see if overexpression can have any effect, for example stress granule formation.

6. General Discussion

SMA is a devastating childhood motor neuron disease resulting from a loss of full length SMN protein, characterised by death of anterior horn motor neurons within in the spinal cord. SMN itself has roles in multiple aspects of RNA processing, from splicing to local translation within axons. In recent years, major advances have been made in SMA treatment, with approval of AAV9 mediated SMN expression (Zolgensma®) and antisense oligonucleotide (Spinraza®)-based therapies. However, these treatments have limitations for older SMA patients (SMA Types II-IV) as they are most effective if administered prior to disease onset or in the very early stages of disease. Therefore, there is still potential scope for improved therapeutics in SMA with combinatorial therapies offering greater benefit in these cases. To develop therapies which may help to ameliorate disease, it is of great importance to try to understand the disease pathogenesis as thoroughly as possible, allowing targeted treatment of disrupted pathways across the lifespan of affected individuals.

Gle1 is an essential, well conserved protein involved in the modulation of DEAD-box proteins which facilitate multiple pathways within the cell, mainly mRNA export, transcription termination and stress granule dynamics. Mutations within this gene can cause severe forms of early-onset motor neuron diseases and have also been found in ALS patients.

Considering the similarities between the pathways in which SMN and Gle1 proteins function and the diseases that they cause when they are mutated, this thesis was focussed on establishing: **i)** whether there is any functional interaction between these two proteins and **ii)** the potential role of *GLE1* as a disease-modifying gene.

6.1 Key findings from this PhD project

To achieve my project objectives, I first set out to design and generate constructs and vectors essential for the planned studies. This was done through cloning of the Gle1b transgene into various viral vectors and tagged constructs. Successful large-scale productions of viral vectors (lentivirus and adeno-associated virus) encoding Gle1b were validated and optimised for both *in vitro* and *in vivo* use.

Following successful generation of GFP-tagged Gle1b constructs, co-immunoprecipitation of Gle1 and SMN indicated interaction between the two proteins which was not RNA-dependant. This is a novel interaction which has not been reported in the literature previously. Proximity ligation assays demonstrated co-localisation of the two proteins in a variety of cell types, with signal being conserved in number between controls and SMA patient cells. The distribution of the co-localised Gle1 and SMN signal was split roughly equally between the nucleus and the cytoplasm, with signal being visible in the axons of motor neurons potentially indicating a role in the local translation of mRNA.

A reduction in the protein expression of Gle1 can also be seen in SMA patient cells. This difference is more prominent in the fibroblast cells than in iPSC-derived motor neurons, potentially pointing towards a neuronal specific role of the proteins which is important for survival. A paper published during the course of this PhD suggested that mRNA levels of *GLE1* are reduced in patient fibroblasts (Alrafiah et al., 2018a), in our hands there was variability in mRNA levels between these cell lines which may be masking any potential difference. Nonetheless, the finding that Gle1 protein expression is reduced in SMA patient cells is a novel and interesting finding which may be important in understanding more about the pathways affected in this disease.

To assess whether Gle1 and SMN can restore the reduction in levels seen in both proteins in SMA cells, I also looked to see whether overexpression of one protein could rescue the other. In this case, there was no change in SMN levels when Gle1b was overexpressed and vice versa. Therefore, indicating that if Gle1b has neuroprotective effects in SMA cases, it is not through SMN-dependant means.

Following these findings, I wanted to investigate whether restoration of Gle1 levels within SMA cells could help to ameliorate any disease phenotypes which occur because of SMN deficiency. Previous data from our lab (**Figure 1.5, Appendix 1, Unpublished data**), demonstrated that overexpression of Gle1 in primary mouse motor neurons from SMNΔ7 SMA mouse model (Le et al., 2005) could rescue the axonal growth defect seen in these cells (Alrafiah et al., 2018b). I therefore wanted to look at whether overexpression of Gle1b could result in improvements in several other disrupted pathways: DNA damage, mRNA export and Golgi morphological changes.

Data presented in this thesis, alongside research published during the course of this PhD, indicates that Gle1 co-localises with R-loops and blockage of Gle1b shuttling between the nucleus and cytoplasm results in an increase in R-loops within the nucleus (Sharma and Wente, 2020). This was subsequently proved to be due to Gle1b's role in modulation of DDX1 during termination of transcription rather than due to any role of Gle1 in the DNA damage response (Sharma and Wente, 2020). Within this thesis, I show that overexpression of Gle1b in HeLa cells exposed to CPT, widely used to induce DNA double strand breaks, has no protective effect. In addition to this, SMA cells treated with a lentivirus overexpressing Gle1b do not show a reduction in the overall levels of DNA damage, suggesting that restoration of Gle1 levels does not influence the transcriptionally associated increased DNA damage seen in SMA patient populations.

Defects within the mRNA export pathway in SMA cases is not an area which is very widely studied. Recent evidence suggests that there is an accumulation of mRNA granules within the nucleus of SMA primary motor neurons from the SMNΔ7 mouse model (Narcís et al., 2018), however, research beyond this study is limited. In this thesis I demonstrate that confocal imaging of Oligo dT FISH staining of poly(A)+ RNA shows a trend towards an increased nuclear signal and reduction in cytoplasmic signal in the SMA fibroblasts, indicating that there is an mRNA export defect in these cells. Unfortunately, due to time restrictions Gle1-mediated rescue experiments were imaged using a different imaging system and this result could not be replicated. As mutations in both *GLE1* and *SMN1* affect motor neurons specifically, together with the initial finding in primary motor neurons I wanted to see whether this mRNA export defect was more obvious in motor neurons. However, again, no difference across the cells could be seen.

When Gle1b's shuttling ability is impaired, there is a subset of mRNA transcripts which selectively are retained in the nucleus. To assess whether Gle1 specific mRNA transport is disrupted in the SMA fibroblast cells where we saw the largest difference in Oligo dT FISH poly(A)+ mRNA staining, I performed nuclear/cytoplasmic fractionation and RNA extraction from these cells. This allowed me to probe for these transcripts to see whether the reduction in Gle1 levels seen in these cells is affecting nuclear export of mRNA. In each of the three transcripts I looked at: *c-Fos*, *ZFP91-CNTF* and *RPL13ap7*, there was no difference between

mRNA distribution in the nucleus and the cytoplasm. However, these results did highlight the novel finding that Fos levels are significantly reduced in SMA patient fibroblasts compared to controls.

The final phenotype assessed was Golgi morphology. A disrupted Golgi has been implicated in multiple neurodegenerative diseases including amyotrophic lateral sclerosis (ALS) and Alzheimer's disease (Joshi et al., 2015; Sundaramoorthy et al., 2015). Fragmented Golgi has previously been observed in SMA patient fibroblasts (Custer et al., 2019). However, in this thesis I demonstrate that SMA cells display a more clustered Golgi phenotype than that of the control cells. Interestingly, overexpression of Gle1b in SMA fibroblasts restored the morphology of the Golgi to a more ribbon like structure as seen in control cells. This, however, does not give us any readout on Golgi function so this would be important to consider in future. Custer and colleagues did not observe an increased sensitivity of SMA cells to thapsigarin (a means of measuring ER-golgi stress), so the disrupted Golgi phenotype may be occurring due to other means such as loss of F-actin tensile strength.

I also investigated a potential pathway which may be causing a compacted Golgi phenotype: the GOLPH3 pathway. Decreases in GOLPH3 and its activity can result in a compacted Golgi as we have observed in the SMA patient cells. Therefore, I looked at protein levels of GOLPH3 and DNA-Pk in iPSC-derived patient cells. Unfortunately, this showed no indication that protein levels in controls and SMA patients were different. A method of looking at GOLPH3 activity maybe a more useful readout in this situation.

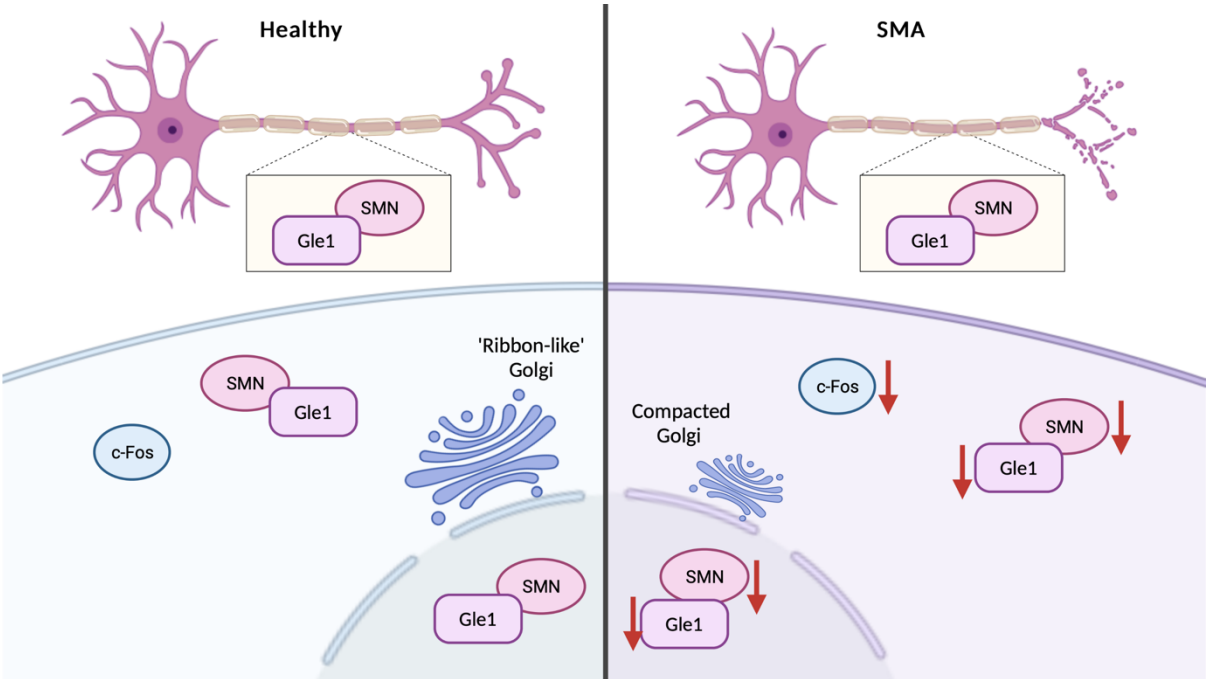


Figure 6.1: Summary of main conclusions from PhD project. *Total protein levels of Gle1, c-Fos and SMN are reduced in SMA patient cells compared to healthy controls. Gle1 and SMN colocalise in both the nucleus and the cytoplasm of control and SMA cells, with colocalisation also occurring in the axons. The Golgi morphology appears more compacted in SMA patient cells compared to healthy controls. Figure created with biorender.com.*

6.2 Potential Pathways for Gle1-SMN interaction

Here, we demonstrate a novel interaction between Gle1 and SMN. There are multiple pathways that both proteins function in which could potentially involve some sort of interaction between them. Interestingly, PLA data demonstrates that the interaction is occurring in across different regions of the cell potentially pointing towards interaction of both Gle1 isoforms with SMN.

The potential interaction between Gle1 and SMN is still yet to be fully explored and much work is needed to understand how these two proteins may be involved with each other. It would be of benefit to perform *in vitro* binding assays and mass spectrometry to establish whether the interaction between the two proteins is direct or whether they might be forming part of a scaffold or protein complex.

A mass spectrometry screen using FLAG-Gle1b was performed by Eric Wolf and colleagues (Wolf et al., 2020). As mentioned previously, results from this screen do not show SMN as a potential binding factor of Gle1. This could suggest that either this interaction is occurring transiently within the cell, there is another scaffolding protein required for interaction/complex formation or the interaction maybe more relevant to the Gle1a isoform. These hypotheses would have to be investigated before a conclusion can be reached.

As Gle1 has been established as a modulator of DEAD-box proteins, it could also be interesting to investigate known DEAD-box binding partners of SMN. DDX20 (also known as GEMIN 3), is a major component of the SMN complex. Gemin3 is also a known binding partner of ALS-associated proteins (Cacciottolo et al., 2019). Due to Gle1's link to ALS and our findings here, it may be worth investigating whether Gle1 is involved in the modulation of this DEAD-box protein in addition to its activity with DDX19, DDX3 and DDX1.

In the nucleus, Gle1b has recently been implicated in transcription termination through modulation of DDX1 (Sharma and Wente, 2020). In this paper they show through PLA that Gle1 also is in close proximity to Senataxin (SETX), a known binding partner of SMN, and propose that Gle1 may also be involved in modulating the role of SETX in R-loop resolution (Sharma and Wente, 2020). This is interesting as SMN is responsible for the recruitment of SETX to the R-loops at sites of transcription termination (Yanling Zhao et al., 2016). Therefore, potentially linking the nuclear roles of SMN and Gle1 through the pathway of R-loop resolution at sites of transcription termination.

As explained in **Chapter 1**, the role of SMN in trafficking and local translation of mRNA in the axons of motor neurons could help to explain the increased sensitivity of motor neurons within SMA. Mutations within the *GLE1* gene also result in neuronal specific deficits, although the role of Gle1 in these cell types is yet to be explored. Therefore, the presence of Gle1/SMN PLA signal within the axons of motor neurons could prove to be an interesting and exciting finding which needs further investigation, especially coupled with the previous data from our lab that overexpression of Gle1 in primary motor neurons from the SMNΔ7 mouse model rescued the axonal growth defect. Investigation into whether Gle1 associates with any of the neuronal binding partners of SMN, for example HuD or α-COP, may help to understand more

about the mechanisms by which SMN and Gle1 are functioning within motor neurons and why their loss is so detrimental.

Another important cytoplasmic pathway that could prove interesting to study is stress granule formation. Both proteins are implicated in the assembly of stress granules (Aditi et al., 2015; Zou et al., 2011) and points towards another pathway worth investigating.

DDX21 is a nucleolar DEAD-box helicase which is involved in rRNA transcription and processing, including the resolution of R-loops (Song et al., 2017). Recent research in our lab group has shown a motor neuron specific deficit of DDX21 within SMN-deficient models (Karyka et al., unpublished data). DDX21 is a high hit on the Gle1b mass spectrometry screen mentioned previously (Wolf et al., 2020). Coupled with our data here, indicating that Gle1 ICC staining colocalises with nucleolin, this could be an interesting mechanism to look further into with regards to Gle1 and SMN interaction.

6.3 *GLE1* as a potential a genetic disease modifier for SMA

Loss of SMN is the primary cause of SMA in patients. SMN protein depletion is likely to affect its binding partners and downstream pathways. Therefore, it is interesting to observe that the levels of Gle1 are downregulated in SMA fibroblasts and motor neurons. Another disease modifying gene, *Plastin3*, has also been shown to be downregulated in SMA cases. Hao and colleagues demonstrate that SMN post-transcriptionally regulates the protein levels of Pls3 (Hao et al., 2012). In this thesis, analysis of *GLE1* mRNA in SMA cell models was not conclusive with large amounts of variation in transcript levels between control cells. Therefore, more work would be needed here to fully investigate whether the transcription and RNA processing of *GLE1* is affected by a loss of SMN. Alternatively, it could also be investigated whether loss of SMN affects Gle1 translation or protein stability (Hao et al., 2012). This would give us more information about where in the disease progression depletion of Gle1 could be occurring and what this might mean for disease pathogenesis.

As discussed previously, Gle1 knockout models in zebrafish display reduced numbers of spinal cord motor neurons, abnormal arborisation and defective Schwann cell development (Jao et al., 2012; Seytanoglu et al., 2016). SMN knockout zebrafish display very similar characteristics

(Hao et al., 2015). Schwann cells in SMA mouse models have been shown to require SMN for their development, with a lack of SMN resulting in myelination defects hypothesised to be due to SMN's role in translational control of myelin proteins (Hunter et al., 2014). These defects were seen to be rescued when SMN was overexpressed specifically in the Schwann cells, thereby improving neuromuscular function but interestingly not survival of the mice (Hunter et al., 2016). It would therefore be interesting to investigate the relationship of Gle1 and SMN in the context of glial cells, as prior studies looking at SMA related therapies have focussed on motor neurons specifically. This paper suggests that other SMA-related disease pathologies can be targeted using a therapy which targets a range of CNS cell types (Hunter et al., 2016).

Other known disease modifiers of SMA include *ZPR1*. Overexpression of ZPR1 in SMA disease models can be shown to transcriptionally upregulate *SMN2* expression, therefore ameliorating SMA disease phenotypes through SMN-dependant means. (Kannan et al., 2020). The work shown in this thesis revealed that overexpression of Gle1b does not appear to have any impact on the protein levels of SMN, indicating that if there is any improvement in SMA disease phenotypes (e.g. axonal growth or rescue of Golgi morphology), that it would be occurring in an SMN-independent manner.

Assessment of the effect of Gle1b overexpression in SMA cell models does not clearly indicate that there is a therapeutic benefit of Gle1 in SMA cases. There are promising results in certain aspects, for example Golgi morphology and axonal growth defects. However, more research is needed to confidently establish whether these effects have an impact on survival or overall disease severity.

6.4 Novel SMA disease features

The work carried out as part of this PhD has uncovered some phenotypes within SMA cell models which have, to my knowledge, not been observed in SMA populations previously.

The first readout is a compacted Golgi morphology. A recent study which focusses on the morphology of the Golgi in SMA fibroblast cells (Custer et al., 2019), reported that the Golgi in SMA patient fibroblasts display a more fragmented Golgi morphology when compared to

healthy control cells. There are a few differences between our study and the work carried in this paper, firstly they use a marker of the cis-Golgi (GM130), while we look at a marker of the trans-Golgi (TGN46). This could potentially influence the morphology observed in these cells as these are two distinct compartments of the Golgi Apparatus with different interacting partners. A variety of different Golgi morphology's is reviewed in depth in (Makhoul et al., 2019), where they suggest different morphology's of each Golgi compartment. Despite the differences in morphology seen between us and this group, I have shown that the compacted Golgi phenotype is conserved between fibroblast cells and iPS-derived motor neurons.

Disruption of the Golgi Apparatus has now been linked to multiple diseases and can result in functional changes such as disrupted glycosylation and membrane trafficking. Morphological changes within the Golgi can often precede other pathological phenotypes in neurodegenerative diseases and therefore could be an important disease marker to be aware of (Liu et al., 2021).

Another interesting finding within this project was the downregulation of *FOS* mRNA transcripts in SMA patient cells. This finding was not investigated further due to the significant delay caused by Covid-19 restrictions, but it still an interesting finding with implications for SMA disease pathology. c-Fos protein expression is heavily involved in learning and memory pathways, with knockout models displaying defects in long-term memory formation (Gallo et al., 2018). Although primarily a disease affecting motor neurons, cognitive impairment has been seen in SMA Type I patients (Polido et al., 2019). This could be a novel pathway which explains some of these symptoms displayed in patients, although much more research is needed into this hypothesis.

6.5 Future Directions

It would be interesting to assess whether the expression levels of Gle1 are also downregulated in SMA mouse models. Using CNS and muscle tissue from SMN delta7 mice at a range of time points, from birth to end stage of disease, would demonstrate whether there was a downregulation of Gle1 levels and at what stage of disease this might occur.

In vivo proof-of-concept experiments would also be interesting to carry out. In **Chapter 3** I describe the generation and validation of a high quality ssAAV9-Gle1b virus suitable for in vivo experiments. In our lab previously it has been demonstrated that AAV9-mediated delivery of transgenes via cisterna magna injections at post-natal day 1 leads to significant transduction of spinal motor neurons and can result in significant increases of survival in the SMN delta 7 mouse model (Alrafiah et al., 2018a). Therefore, it would be interesting to see whether in this model, overexpression of Gle1 would have an effect on disease progression. This would provide a clearer answer in relation to the therapeutic benefit of Gle1 in SMA.

The role of Gle1 within neurons is something which has not been well studied. Due to the neuronal specific diseases which arise from mutations within Gle1, it would be very interesting to further investigate what these neuronal specific roles could potentially be and whether mutant forms of Gle1 disrupt neuronal pathways.

6.6 Final Conclusions

In conclusion, the data presented in this thesis implicates Gle1 as a novel protein of interest in relation to SMA and its causative *SMN* gene. Gle1 and SMN colocalise within the nucleus, cytoplasm, and axons of motor neurons, with SMA patient cells displaying reduction of Gle1 protein. Overexpression of Gle1 can act to rescue axonal growth defects and disrupted Golgi morphology although does not appear to improve endogenous increases in DNA damage seen in SMA patient cells. More research is needed to fully understand the pathways in which Gle1 and SMN are both involved in together and how this may be influencing SMA disease pathogenesis. Understanding more about these pathways can help the development of therapies that can be used in conjunction with pre-approved SMN-dependant treatments to create a comprehensive cure for all SMA patients.

References

Ackermann, B., Kröber, S., Torres-benito, L., Borgmann, A., Peters, M., Hosseini barkooie, S.M., Tejero, R., Jakubik, M., Schreml, J., Milbradt, J., Wunderlich, T.F., Riessland, M., Tabares, L., Wirth, B., 2013. Plastin 3 ameliorates spinal muscular atrophy via delayed axon pruning and improves neuromuscular junction functionality. Hum. Mol. Genet. 22, 1328–1347.

Adams, R.L., Mason, A.C., Glass, L., Aditi, Wente, S.R., 2017. Nup42 and IP 6 coordinate Gle1 stimulation of Dbp5/DDX19B for mRNA export in yeast and human cells. Traffic 18, 776–790.

Adams, R.L., Terry, L.J., Wente, S.R., 2014. Nucleoporin FG domains facilitate mRNP remodeling at the cytoplasmic face of the nuclear pore complex. Genetics 197, 1213–1224.

Aditi, Folkmann, A.W., Wente, S.R., 2015. Cytoplasmic hGle1A regulates stress granules by modulation of translation. Mol. Biol. Cell 26, 1476–1490.

Aditi, Glass, L., Dawson, T.R., Wente, S.R., 2016. An amyotrophic lateral sclerosis-linked mutation in GLE1 alters the cellular pool of human Gle1 functional isoforms. Adv. Biol. Regul. 62, 25–36.

Aditi, Mason, A.C., Sharma, M., Renee Dawson, T., Wente, S.R., 2019. MAPK- and glycogen synthase kinase 3–mediated phosphorylation regulates the DEAD-box protein modulator Gle1 for control of stress granule dynamics. J. Biol. Chem. 294, 559.

Agency, E.M., 2020. New gene therapy to treat spinal muscular atrophy (corrected) | European Medicines Agency [WWW Document]. Viewed 2021-11-25. <https://www.ema.europa.eu/en/news/new-gene-therapy-treat-spinal-muscular-atrophy-corrected>

Aguilera, A., García-Muse, T., 2012. R Loops: From Transcription Byproducts to Threats to Genome Stability. Mol. Cell 46, 115–124.

Ahmad, S., Wang, Y., Shaik, G.M., Burghes, A.H., Gangwani, L., 2012. The zinc finger protein ZPR1 is a potential modifier of spinal muscular atrophy. Hum. Mol. Genet. 21, 2745–2758.

Akey, C.W., Radermacher, M., 1993. Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J. Cell Biol. 122, 1–19.

Akten, B., Kye, M.J., Hao, L.T., Wertz, M.H., Singh, S., Nie, D., Huang, J., Merianda, T.T., Twiss, J.L., Beattie, C.E., Steen, J.A.J., Sahin, M., 2011. Interaction of survival of motor neuron (SMN) and HuD proteins with mRNA cpg15 rescues motor neuron axonal deficits. Proc. Natl. Acad. Sci. U. S. A. 108, 10337–10342.

Al-Zaidy, S.A., Kolb, S.J., Lowes, L., Alfano, L.N., Shell, R., Church, K.R., Nagendran, S., Sproule, D.M., Feltner, D.E., Wells, C., Ogrinc, F., Menier, M., L'Italien, J., Arnold, W.D., Kissel, J.T., Kaspar, B.K., Mendell, J.R., 2019. AVXS-101 (Onasemnogene Abeparvovec) for SMA1: Comparative Study with a Prospective Natural History Cohort. J. Neuromuscul. Dis. 6, 307–317.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Sali, A., Rout, M.P., 2007. The molecular architecture of the nuclear pore complex. Nature 450, 695–701.

Alcazar-Roman, A.R., Bolger, T.A., Wente, S.R., 2010. Control of mRNA export and translation termination by inositol hexakisphosphate requires specific interaction with Gle1. J. Biol. Chem. 285, 16683–16692.

Alcázar-Román, A.R., Tran, E.J., Guo, S., Wente, S.R., 2006. Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. Nat. Cell Biol. 8, 711–716.

Alrafiah, A., Alghanmi, M., Almashhadi, S., Aqeel, A., Awaji, A., 2018a. The expression of SMN1, MART3, GLE1 and FUS genes in spinal muscular atrophy. Folia Histochem. Cytobiol. 56, 215–

221.

Alrafiah, A., Karyka, E., Coldicott, I., Iremonger, K., Lewis, K.E., Ning, K., Azzouz, M., 2018b. Plastin 3 Promotes Motor Neuron Axonal Growth and Extends Survival in a Mouse Model of Spinal Muscular Atrophy. Mol. Ther. Methods Clin. Dev. 9, 81–89.

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., Liberi, G., 2012. Senataxin Associates with Replication Forks to Protect Fork Integrity across RNA-Polymerase-II-Transcribed Genes. Cell 151, 835–846.

Andrews, J.A., Miller, T.M., Vijayakumar, V., Stoltz, R., James, J.K., Meng, L., Wolff, A.A., Malik, F.I., 2018. CK-2127107 amplifies skeletal muscle response to nerve activation in humans. Muscle Nerve 57, 729–734.

Annoni, A., Gregori, S., Naldini, L., Cantore, A., 2019. Modulation of immune responses in lentiviral vector-mediated gene transfer. Cell. Immunol. 342.

Ansari, A.M., Ahmed, A.K., Matsangos, A.E., Lay, F., Born, L.J., Marti, G., Harmon, J.W., Sun, Z., 2016. Cellular GFP Toxicity and Immunogenicity: Potential Confounders in in Vivo Cell Tracking Experiments. Stem Cell Rev. 12, 553.

Aryanpur, P.P., Regan, C.A., Collins, J.M., Mittelmeier, T.M., Renner, D.M., Vergara, A.M., Brown, N.P., Bolger, T.A., 2017. Gle1 regulates RNA binding of the DEAD-box helicase Ded1 in its complex role in translation initiation. Mol. Cell. Biol.

Aschauer, D.F., Kreuz, S., Rumpel, S., 2013. Analysis of transduction efficiency, tropism and axonal transport of AAV serotypes 1, 2, 5, 6, 8 and 9 in the mouse brain. PLoS One 8.

Auburger, G., Klinkenberg, M., Drost, J., Marcus, K., Morales-Gordo, B., Kunz, W.S., Brandt, U., Broccoli, V., Reichmann, H., Gispert, S., Jendrach, M., 2012. Primary Skin Fibroblasts as a Model of Parkinson's Disease. Mol. Neurobiol. 46, 20.

Azzouz, M., Martin-Rendon, E., Barber, R.D., Mitrophanous, K.A., Carter, E.E., Rohll, J.B., Kingsman, S.M., Kingsman, A.J., Mazarakis, N.D., 2002. Multicistronic Lentiviral Vector-Mediated Striatal Gene Transfer of Aromatic l-Amino Acid Decarboxylase, Tyrosine Hydroxylase, and GTP Cyclohydrolase I Induces Sustained Transgene Expression, Dopamine Production, and Functional Improvement in a Rat Model of Parkinson's Disease. J. Neurosci. 22, 10302–10312.

Barrett, D., Bilic, S., Chyung, Y., Cote, S.M., Iarrobino, R., Kacena, K., Kalra, A., Long, K., Nomikos, G., Place, A., Still, J.G., Vrishabhendra, L., 2021. A Randomized Phase 1 Safety, Pharmacokinetic and Pharmacodynamic Study of the Novel Myostatin Inhibitor Apitegromab (SRK-015): A Potential Treatment for Spinal Muscular Atrophy. Adv. Ther. 38, 3203–3222.

Barrett, R., Ornelas, L., Yeager, N., Mandefro, B., Sahabian, A., Lenaeus, L., Targan, S.R., Svendsen, C.N., Sareen, D., 2014. Reliable generation of induced pluripotent stem cells from human lymphoblastoid cell lines. Stem Cells Transl. Med. 3, 1429–1434.

Bäumer, D., Lee, S., Nicholson, G., Davies, J.L., Parkinson, N.J., Murray, L.M., Gillingwater, T.H., Ansorge, O., Davies, K.E., Talbot, K., 2009. Alternative Splicing Events Are a Late Feature of Pathology in a Mouse Model of Spinal Muscular Atrophy. PLoS Genet. 5, e1000773.

Beck, M., Förster, F., Ecke, M., Plitzko, J.M., Melchior, F., Gerisch, G., Baumeister, W., Medalia, O., 2004. Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. Science 306, 1387–90.

Bell, S.M., De Marco, M., Barnes, K., Shaw, P.J., Ferraiuolo, L., Blackburn, D.J., Mortiboys, H., Venneri, A., 2020. Deficits in Mitochondrial Spare Respiratory Capacity Contribute to the Neuropsychological Changes of Alzheimer's Disease. J. Pers. Med. 2020, Vol. 10, Page 32 10, 32.

Benkhelifa-Ziyyat, S., Besse, A., Roda, M., Duque, S., Astord, S., Carcenac, R., Marais, T., Barkats, M.,

2013. Intramuscular scAAV9-SMN injection mediates widespread gene delivery to the spinal cord and decreases disease severity in SMA mice. Mol. Ther. 21, 282–290.

Bernabò, P., Tebaldi, T., Groen, E.J.N., Lane, F.M., Perenthaler, E., Mattedi, F., Newbery, H.J., Zhou, H., Zuccotti, P., Potrich, V., Shorrock, H.K., Muntoni, F., Quattrone, A., Gillingwater, T.H., Viero, G., 2017. In Vivo Translatome Profiling in Spinal Muscular Atrophy Reveals a Role for SMN Protein in Ribosome Biology. Cell Rep. 21, 953–965.

Bertini, E., Dessaud, E., Mercuri, E., Muntoni, F., Kirschner, J., Reid, C., Lusakowska, A., Comi, G.P., Cuisset, J.M., Abitbol, J.L., Scherrer, B., Ducray, P.S., Buchbjerg, J., Vianna, E., van der Pol, W.L., Vuillerot, C., Blaettler, T., Fontoura, P., André, C., Bruno, C., Chabrol, B., Deconinck, N., Estournet, B., Fontaine-Carbonnel, S., Goemans, N., Gorni, K., Govoni, A., Guglieri, M., Lochmuller, H., Magri, F., Mayer, M., Müller-Felber, W., Rivier, F., Roper, H., Schara, U., Scoto, M., van den Berg, L., Vita, G., Walter, M.C., 2017. Safety and efficacy of olesoxime in patients with type 2 or non-ambulatory type 3 spinal muscular atrophy: a randomised, double-blind, placebo-controlled phase 2 trial. Lancet. Neurol. 16, 513–522.

Bléoo, S., Sun, X., Hendzel, M.J., Rowe, J.M., Packer, M., Godbout, R., 2001. Association of Human DEAD Box Protein DDX1 with a Cleavage Stimulation Factor Involved in 3'-End Processing of Pre-mRNA. Mol. Biol. Cell 12, 3046.

Bolger, T.A., Folkmann, A.W., Tran, E.J., Wente, S.R., 2008. The mRNA export factor Gle1 and inositol hexakisphosphate regulate distinct stages of translation. Cell 134, 624–633.

Bolger, T.A., Wente, S.R., 2011. Gle1 is a multifunctional DEAD-box protein regulator that modulates Ded1 in translation initiation. J. Biol. Chem. 286, 39750–39759.

Bowerman, M., Anderson, C.L., Beauvais, A., Boyl, P.P., Witke, W., Kothary, R., 2009. SMN, profilin IIa and plastin 3: a link between the deregulation of actin dynamics and SMA pathogenesis. Mol. Cell. Neurosci. 42, 66–74.

Bowerman, M., Murray, L.M., Beauvais, A., Pinheiro, B., Kothary, R., 2012. A critical smn threshold in mice dictates onset of an intermediate spinal muscular atrophy phenotype associated with a distinct neuromuscular junction pathology. Neuromuscul. Disord. 22, 263–276.

Bowser, R., Novick, P., 1991. Sec15 protein, an essential component of the exocytotic apparatus, is associated with the plasma membrane and with a soluble 19.5S particle. J. Cell Biol. 112, 1117–1131.

Brzustowicz, L.M., Lehner, T., Castilla, L.H., Penchaszadeh, G.K., Wilhelmsen, K.C., Daniels, R., Davies, K.E., Leppert, M., Ziter, F., Wood, D., Dubowitz, V., Zerres, K., Hausmanowa-Petrusewicz, I., Ott, J., Munsat, T.L., Gilliam, T.C., 1990. Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2-13.3. Nature 344, 540–541.

Bui, K.H., von Appen, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.-T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., Beck, M., 2013. Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. Cell 155, 1233–1243.

Cacciottolo, R., Ciantar, J., Lanfranco, M., Borg, R.M., Vassallo, N., Bordonné, R., Cauchi, R.J., 2019. SMN complex member Gemin3 self-interacts and has a functional relationship with ALS-linked proteins TDP-43, FUS and Sod1. Sci. Reports 2019 91 9, 1–19.

Casafont, I., Berciano, M.T., Lafarga, M., 2010. Bortezomib induces the formation of nuclear poly(A) RNA granules enriched in Sam68 and PABPN1 in sensory ganglia neurons. Neurotox. Res. 17, 167–178.

Cerino, M., Di Meglio, C., Albertini, F., Audic, F., Riccardi, F., Boulay, C., Philip, N., Bartoli, M., Lévy, N., Krahn, M., Chabrol, B., 2020. Extension of the phenotypic spectrum of GLE1-related disorders to a mild congenital form resembling congenital myopathy. Mol. Genet. genomic

Med. 8.

Chakraborty, P., Huang, J.T.J., Hiom, K., 2018. DHX9 helicase promotes R-loop formation in cells with impaired RNA splicing. Nat. Commun. 9.

Chiriboga, C.A., Swoboda, K.J., Darras, B.T., Iannaccone, S.T., Montes, J., De Vivo, D.C., Norris, D.A., Bennett, C.F., Bishop, K.M., 2016. Results from a phase 1 study of nusinersen (ISIS-SMNRx) in children with spinal muscular atrophy. Neurology 86, 890.

Ciuffi, A., 2008. Mechanisms governing lentivirus integration site selection. Curr. Gene Ther. 8, 419–429.

Cobben, J.M., Van der Steege, G., Grootscholten, P., De Visser, M., Scheffer, H., Buys, C.H.C.M., 1995. Deletions of the survival motor neuron gene in unaffected siblings of patients with spinal muscular atrophy. Am. J. Hum. Genet. 57, 805.

Coriell, 2021a. GM09677 [WWW Document]. Viewed 2021-11-24. <https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM09677>

Coriell, 2021b. GM00232 [WWW Document]. Viewed 2021-11-24. <https://catalog.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM00232&Product=CC>

Coriell, 2021c. GM08680 [WWW Document]. Viewed 2021-11-24. <https://www.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM08680>

Crawford, T.O., Paushkin, S. V., Kobayashi, D.T., Forrest, S.J., Joyce, C.L., Finkel, R.S., Kaufmann, P., Swoboda, K.J., Tiziano, D., Lomastro, R., Li, R.H., Trachtenberg, F.L., Plasterer, T., Chen, K.S., Bell, M., Jacoby, D., McBurney, R., Chung, W., Simard, L., Sahin, M., 2012. Evaluation of SMN protein, transcript, and copy number in the biomarkers for spinal muscular atrophy (BforSMA) clinical study. PLoS One 7.

Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., Matunis, M.J., 2002. Proteomic analysis of the mammalian nuclear pore complex. J. Cell Biol. 158, 915–27.

Custer, S.K., Foster, J.N., Astroski, J.W., Androphy, E.J., 2019. Abnormal Golgi morphology and decreased COPI function in cells with low levels of SMN. Brain Res. 1706, 135–146.

Custer, S.K., Todd, A.G., Singh, N.N., Androphy, E.J., 2013. Dilysine motifs in exon 2b of SMN protein mediate binding to the COPI vesicle protein α-COP and neurite outgrowth in a cell culture model of spinal muscular atrophy. Hum. Mol. Genet. 22, 4043.

D'Amico, A., Mercuri, E., Tiziano, F.D., Bertini, E., 2011. Spinal muscular atrophy. Orphanet J. Rare Dis. 6, 1–10.

Davierwala, A.P., Haynes, J., Li, Z., Brost, R.L., Robinson, M.D., Yu, L., Mnaimneh, S., Ding, H., Zhu, H., Chen, Y., Cheng, X., Brown, G.W., Boone, C., Andrews, B.J., Hughes, T.R., 2005. The synthetic genetic interaction spectrum of essential genes. Nat. Genet. 37, 1147–1152.

De Vivo, D.C., Bertini, E., Swoboda, K.J., Hwu, W.L., Crawford, T.O., Finkel, R.S., Kirschner, J., Kuntz, N.L., Parsons, J.A., Ryan, M.M., Butterfield, R.J., Topaloglu, H., Ben-Omran, T., Sansone, V.A., Jong, Y.J., Shu, F., Staropoli, J.F., Kerr, D., Sandrock, A.W., Stebbins, C., Petrillo, M., Braley, G., Johnson, K., Foster, R., Gheuens, S., Bhan, I., Reyna, S.P., Fradette, S., Farwell, W., 2019. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. Neuromuscul. Disord. 29, 842.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G.-Y.R., Karydas, A., Seeley, W.W., Josephs, K.A., Coppola, G., Geschwind, D.H., Wszolek, Z.K.,

Feldman, H., Knopman, D.S., Petersen, R.C., Miller, B.L., Dickson, D.W., Boylan, K.B., Graff-Radford, N.R., Rademakers, R., 2011. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron 72, 245–256.

Dhillon, S., 2020. Risdiplam: First Approval. Drugs 2020 8017 80, 1853–1858.

Dippold, H.C., Ng, M.M., Farber-Katz, S.E., Lee, S.K., Kerr, M.L., Peterman, M.C., Sim, R., Wiharto, P.A., Galbraith, K.A., Madhavarapu, S., Fuchs, G.J., Meerloo, T., Farquhar, M.G., Zhou, H., Field, S.J., 2009. GOLPH3 Bridges Phosphatidylinositol-4- Phosphate and Actomyosin to Stretch and Shape the Golgi to Promote Budding. Cell 139, 337–351.

Doran, B., Gherbesi, N., Hendricks, G., Flavell, R.A., Davis, R.J., Gangwani, L., 2006. Deficiency of the zinc finger protein ZPR1 causes neurodegeneration. Proc. Natl. Acad. Sci. U. S. A. 103, 7471.

Du, Z.W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C.T.L., Zhong, X., Fan, F., Zhang, S.C., 2015. Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. Nat. Commun. 2015 61 6, 1–9.

Dubowitz, V., 1999. Very severe spinal muscular atrophy (SMA type 0): an expanding clinical phenotype. Eur. J. Paediatr. Neurol. 3, 49–51.

Ebert, A.D., Svendsen, C.N., 2010. Stem cell model of spinal muscular atrophy. Arch. Neurol. 67, 665–669.

Ed, F., Ch Edin, E., Hartono, S.R., Sanz, L.A., Vanoosthuyse, V., 2021. Best practices for the visualization, mapping, and manipulation of R-loops. EMBO J. 40, e106394.

Edens, B.M., Ajroud-Driss, S., Ma, L., Ma, Y.C., 2015. Molecular mechanisms and animal models of spinal muscular atrophy. Biochim. Biophys. Acta - Mol. Basis Dis. 1852, 685–692.

Eilers, A., Whitfield, J., Babij, C., Rubin, L.L., Ham, J., 1998. Role of the Jun kinase pathway in the regulation of c-Jun expression and apoptosis in sympathetic neurons. J. Neurosci. 18, 1713–1724.

Ellard, S., Kivuva, E., Turnpenny, P., Stals, K., Johnson, M., Xie, W., Caswell, R., Lango Allen, H., 2015. An exome sequencing strategy to diagnose lethal autosomal recessive disorders. Eur. J. Hum. Genet. 23, 401–404.

Estruch, F., Cole, C.N., 2003. An early function during transcription for the yeast mRNA export factor Dbp5p/Rat8p suggested by its genetic and physical interactions with transcription factor IIH components. Mol. Biol. Cell 14, 1664–76.

Fallini, C., Bassell, G.J., Rossoll, W., 2012. Spinal muscular atrophy: The role of SMN in axonal mRNA regulation. Brain Res. 1462, 81–92.

Fallini, C., Donlin-Asp, P.G., Rouanet, J.P., Bassell, G.J., Rossoll, W., 2016. Deficiency of the Survival of Motor Neuron Protein Impairs mRNA Localization and Local Translation in the Growth Cone of Motor Neurons. J. Neurosci. 36, 3811–3820.

Fallini, C., Rouanet, J.P., Donlin-Asp, P.G., Guo, P., Zhang, H., Singer, R.H., Rossoll, W., Bassell, G.J., 2014. Dynamics of survival of motor neuron (SMN) protein interaction with the mRNA-binding protein IMP1 facilitates its trafficking into motor neuron axons. Dev. Neurobiol. 74, 319–332.

Fallini, C., Zhang, H., Su, Y., Silani, V., Singer, R.H., Rossoll, W., Bassell, G.J., 2011. The Survival of Motor Neuron (SMN) Protein Interacts with the mRNA-Binding Protein HuD and Regulates Localization of Poly(A) mRNA in Primary Motor Neuron Axons. J. Neurosci. 31, 3914.

Farrar, M.A., Kiernan, M.C., 2015. The Genetics of Spinal Muscular Atrophy: Progress and Challenges. Neurotherapeutics 12, 290–302.

Fayzullina, S., Martin, L.J., 2016. DNA Damage Response and DNA Repair in Skeletal Myocytes From a

Mouse Model of Spinal Muscular Atrophy. J. Neuropathol. Exp. Neurol. 75, 889–902.

Fayzullina, S., Martin, L.J., 2014. Skeletal muscle DNA damage precedes spinal motor neuron DNA damage in a mouse model of Spinal Muscular Atrophy (SMA). PLoS One 9.

Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T.F., Wirth, B., 2002. Quantitative Analyses of SMN1 and SMN2 Based on Real-Time LightCycler PCR: Fast and Highly Reliable Carrier Testing and Prediction of Severity of Spinal Muscular Atrophy. Am. J. Hum. Genet. 70, 358–368.

Finkel, R.S., Chiriboga, C.A., Vajsar, J., Day, J.W., Montes, J., De Vivo, D.C., Yamashita, M., Rigo, F., Hung, G., Schneider, E., Norris, D.A., Xia, S., Bennett, C.F., Bishop, K.M., 2016. Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. Lancet (London, England) 388, 3017–3026.

Finkel, R.S., McDermott, M.P., Kaufmann, P., Darras, B.T., Chung, W.K., Sproule, D.M., Kang, P.B., Reghan Foley, A., Yang, M.L., Martens, W.B., Oskoui, M., Glanzman, A.M., Flickinger, J., Montes, J., Dunaway, S., O'Hagen, J., Quigley, J., Riley, S., Benton, M., Ryan, P.A., Montgomery, M., Marra, J., Gooch, C., De Vivo, D.C., 2014. Observational study of spinal muscular atrophy type I and implications for clinical trials. Neurology 83, 810–817.

Finkel, R.S., Mercuri, E., Darras, B.T., Connolly, A.M., Kuntz, N.L., Kirschner, J., Chiriboga, C.A., Saito, K., Servais, L., Tizzano, E., Topaloglu, H., Tulinius, M., Montes, J., Glanzman, A.M., Bishop, K., Zhong, Z.J., Gheuens, S., Bennett, C.F., Schneider, E., Farwell, W., De Vivo, D.C., 2017. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. N. Engl. J. Med. 377, 1723–1732.

Folkmann, A.W., Collier, S.E., Zhan, X., Aditi, Ohi, M.D., Wente, S.R., 2013. Gle1 functions during mRNA export in an oligomeric complex that is altered in human disease. Cell 155, 582–593.

Folkmann, A.W., Dawson, T.R., Wente, S.R., 2014. Insights into mRNA export-linked molecular mechanisms of human disease through a Gle1 structure-function analysis. Adv. Biol. Regul. 54, 74–91.

Folkmann, A.W., Noble, K.N., Cole, C.N., Wente, S.R., 2011. Dbp5, Gle1-IP 6 and Nup159. Nucleus 2, 540–548.

Foust, K.D., Wang, X., McGovern, V.L., Braun, L., Bevan, A.K., Haidet, A.M., Le, T.T., Morales, P.R., Rich, M.M., Burghes, A.H.M., Kaspar, B.K., 2010. Rescue of the spinal muscular atrophy phenotype in a mouse model by early postnatal delivery of SMN. Nat. Biotechnol. 28, 271–274.

Fuller, H.R., Man, N.T., Lam, L.T., Thanh, L.T., Keough, R.A., Asperger, A., Gonda, T.J., Morris, G.E., 2009. The SMN Interactome Includes Myb-Binding Protein 1a. J. Proteome Res. 9, 556–563.

Fuller, H.R., Mandefro, B., Shirran, S.L., Gross, A.R., Kaus, A.S., Botting, C.H., Morris, G.E., Sareen, D., 2015. Spinal Muscular Atrophy Patient iPSC-Derived Motor Neurons Have Reduced Expression of Proteins Important in Neuronal Development. Front. Cell. Neurosci. 9, 506.

Gabanella, F., Butchbach, M.E.R., Saieva, L., Carissimi, C., Burghes, A.H.M., Pellizzoni, L., 2007. Ribonucleoprotein Assembly Defects Correlate with Spinal Muscular Atrophy Severity and Preferentially Affect a Subset of Spliceosomal snRNPs. PLoS One 2, e921.

Gabanella, F., Pisani, C., Borreca, A., Farioli-Vecchioli, S., Ciotti, M.T., Ingegnere, T., Onori, A., Ammassari-Teule, M., Corbi, N., Canu, N., Monaco, L., Passananti, C., Di Certo, M.G., 2016. SMN affects membrane remodelling and anchoring of the protein synthesis machinery. J. Cell Sci. 129, 804–816.

Gallo, F.T., Katche, C., Morici, J.F., Medina, J.H., Weisstaub, N. V., 2018. Immediate early genes, memory and psychiatric disorders: Focus on c-Fos, Egr1 and Arc. Front. Behav. Neurosci. 12, 79.

Gangwani, L., Flavell, R.A., Davis, R.J., 2005. ZPR1 Is Essential for Survival and Is Required for

Localization of the Survival Motor Neurons (SMN) Protein to Cajal Bodies. Mol. Cell. Biol. 25, 2744.

Gangwani, L., Mikrut, M., Theroux, S., Sharma, M., Davis, R.J., 2001. Spinal muscular atrophy disrupts the interaction of ZPR1 with the SMN protein. Nat. Cell Biol. 3, 376–383.

Genabai, N.K., Ahmad, S., Zhang, Z., Jiang, X., Gabaldon, C.A., Gangwani, L., 2015. Genetic inhibition of JNK3 ameliorates spinal muscular atrophy. Hum. Mol. Genet. 24, 6986.

Gilman, B., Tijerina, P., Russell, R., 2017. Distinct RNA unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. Biochem. Soc. Trans. 45, 1313.

Glass, L., Wente, S.R., 2019. Gle1 mediates stress granule-dependent survival during chemotoxic stress. Adv. Biol. Regul. 71, 156.

Gross, T., Siepmann, A., Sturm, D., Windgassen, M., Scarcelli, J.J., Seedorf, M., Cole, C.N., Krebber, H., 2007. The DEAD-Box RNA Helicase Dbp5 Functions in Translation Termination. Science (80-.). 315, 646–649.

Hahnen, E., Forkert, R., Marke, C., Rudnik-schöneborn, S., Schönling, J., Zerres, K., Creavin, T., Wirth, B., 1995. Molecular analysis of candidate genes on chromosome 5q13 in autosomal recessive spinal muscular atrophy: evidence of homozygous deletions of the SMN gene in unaffected individuals. Hum. Mol. Genet. 4, 1927–1933.

Hanson, D.A., Ziegler, S.F., 2004. Fusion of green fluorescent protein to the C-terminus of granulysin alters its intracellular localization in comparison to the native molecule. J. Negat. Results Biomed. 3.

Hao, L.T., Burghes, A.H.M., Beattie, C.E., 2011. Generation and Characterization of a genetic zebrafish model of SMA carrying the human SMN2 gene. Mol. Neurodegener. 6, 24.

Hao, L.T., Duy, P.Q., Jontes, J.D., Beattie, C.E., 2015. Motoneuron development influences dorsal root ganglia survival and Schwann cell development in a vertebrate model of spinal muscular atrophy. Hum. Mol. Genet. 24, 346–360.

Hao, L.T., Wolman, M., Granato, M., Beattie, C.E., 2012. Survival Motor Neuron Affects Plastin 3 Protein Levels Leading to Motor Defects. J. Neurosci. 32, 5074.

Hao Le, T., Duy, P.Q., An, M., Talbot, J., Iyer, C.C., Wolman, M., Beattie, C.E., 2017. HuD and the Survival Motor Neuron Protein Interact in Motoneurons and Are Essential for Motoneuron Development, Function, and mRNA Regulation. J. Neurosci. 37, 11559.

Hatchi, E., Skourti-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., Dimitrov, S., Pathania, S., McKinney, K.M., Eaton, M.L., Kellis, M., Hill, S.J., Parmigiani, G., Proudfoot, N.J., Livingston, D.M., 2015. BRCA1 Recruitment to Transcriptional Pause Sites Is Required for R-Loop-Driven DNA Damage Repair. Mol. Cell 57, 636–647.

Hautbergue, G.M., 2017. Personalised Medicine 1007.

Hebert, M.D., Szymczyk, P.W., Shpargel, K.B., Gregory Matera, A., 2001. Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. Genes Dev. 15, 2720–2729.

Heesen, L., Peitz, M., Torres-Benito, L., Hölker, I., Hupperich, K., Dobrindt, K., Jungverdorben, J., Ritzenhofen, S., Weykopf, B., Eckert, D., Hosseini-Barkooie, S.M., Storbeck, M., Fusaki, N., Lonigro, R., Heller, R., Kye, M.J., Brüstle, O., Wirth, B., 2016. Plastin 3 is upregulated in iPSC-derived motoneurons from asymptomatic SMN1-deleted individuals. Cell. Mol. Life Sci. 73, 2089–2104.

Helmken, C., Hofmann, Y., Schoenen, F., Oprea, G., Raschke, H., Rudnik-Schöneborn, S., Zerres, K., Wirth, B., 2003. Evidence for a modifying pathway in SMA discordant families: Reduced SMN level decreases the amount of its interacting partners and Htra2-beta1. Hum. Genet. 114, 11–21.

Hensel, N., Kubinski, S., Claus, P., 2020. The Need for SMN-Independent Treatments of Spinal Muscular Atrophy (SMA) to Complement SMN-Enhancing Drugs. Front. Neurol. 11, 45.

Herva, R., Conradi, N.G., Kalimo, H., Leisti, J., Sourander, P., 1988. A syndrome of multiple congenital contractures: Neuropathlogical analysis on five fetal cases. Am. J. Med. Genet. 29, 67–76.

Hill, S.J., Rolland, T., Adelmant, G., Xia, X., Owen, M.S., Dricot, A., Zack, T.I., Sahni, N., Jacob, Y., Hao, T., McKinney, K.M., Clark, A.P., Reyon, D., Tsai, S.Q., Joung, J.K., Beroukhim, R., Marto, J.A., Vidal, M., Gaudet, S., Hill, D.E., Livingston, D.M., 2014. Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. Genes Dev. 28, 1957–1975.

Hinshaw, J.E., Carragher, B.O., Milligan, R.A., 1992. Architecture and design of the nuclear pore complex. Cell 69, 1133–1141.

Hodge, C.A., Colot, H. V, Stafford, P., Cole, C.N., 1999. Rat8p/Dbp5p is a shuttling transport factor that interacts with Rat7p/Nup159p and Gle1p and suppresses the mRNA export defect of xpo1-1 cells. EMBO J. 18, 5778–5788.

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., Maiorano, D., 2017a. An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. EMBO J. 36, 1182–1198.

Hodroj, D., Serhal, K., Maiorano, D., 2017b. Ddx19 links mRNA nuclear export with progression of transcription and replication and suppresses genomic instability upon DNA damage in proliferating cells. Nucleus 8, 1–7.

Hosseinibarkooie, S., Peters, M., Torres-Benito, L., Rastetter, R.H.H., Hupperich, K., Hoffmann, A., Mendoza-Ferreira, N., Kaczmarek, A., Janzen, E., Milbradt, J., Lamkemeyer, T., Rigo, F., Bennett, C.F., Guschlbauer, C., Büschges, A., Hammerschmidt, M., Riessland, M., Kye, M.J.J., Clemen, C.S.S., Wirth, B., 2016. The Power of Human Protective Modifiers: PLS3 and CORO1C Unravel Impaired Endocytosis in Spinal Muscular Atrophy and Rescue SMA Phenotype. Am. J. Hum. Genet. 99, 647–665.

Hoy, S.M., 2019. Onasemnogene Abeparvovec: First Global Approval. Drugs 79, 1255–1262.

Hsieh-Li, H.M., Chang, J.G., Jong, Y.J., Wu, M.H., Wang, N.M., Tsai, C.H., Li, H., 2000. A mouse model for spinal muscular atrophy. Nat. Genet. 24, 66–70.

Hua, Y., Zhou, J., 2004. Survival motor neuron protein facilitates assembly of stress granules. FEBS Lett. 572, 69–74.

Hunter, G., Powis, R.A., Jones, R.A., Groen, E.J.N., Shorrock, H.K., Lane, F.M., Zheng, Y., Sherman, D.L., Brophy, P.J., Gillingwater, T.H., 2016. Restoration of SMN in Schwann cells reverses myelination defects and improves neuromuscular function in spinal muscular atrophy. Hum. Mol. Genet. 25, 2853.

Hunter, G., Sarvestany, A.A., Roche, S.L., Symes, R.C., Gillingwater, T.H., 2014. SMN-dependent intrinsic defects in Schwann cells in mouse models of spinal muscular atrophy. Hum. Mol. Genet. 23, 2235–2250.

Iino, S., Kobayashi, S., Hidaka, H., 1998. Neurocalcin-immunopositive nerve terminals in the muscle spindle, Golgi tendon organ and motor endplate. Brain Res. 808, 294–299.

Ivings, L., Pennington, S.R., Jenkins, R., Weiss, J.L., Burgoyne, R.D., 2002. Identification of Ca2+-dependent binding partners for the neuronal calcium sensor protein neurocalcin δ: interaction

with actin, clathrin and tubulin. Biochem. J. 363, 599–608.

Jablonka, S., Schrank, B., Kralewski, M., Rossoll, W., Sendtner, M., 2000. Reduced survival motor neuron (Smn) gene dose in mice leads to motor neuron degeneration: an animal model for spinal muscular atrophy type III. Hum. Mol. Genet. 9, 341–346.

Jangi, M., Fleet, C., Cullen, P., Gupta, S. V., Mekhoubad, S., Chiao, E., Allaire, N., Bennett, C.F., Rigo, F., Krainer, A.R., Hurt, J.A., Carulli, J.P., Staropoli, J.F., 2017. SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. Proc. Natl. Acad. Sci. 114, E2347–E2356.

Jao, L.-E., Akef, A., Wente, S.R., 2017. A role for Gle1, a regulator of DEAD-box RNA helicases, at centrosomes and basal bodies. Mol. Biol. Cell 28, 120–127.

Jao, L.-E., Appel, B., Wente, S.R., 2012. A zebrafish model of lethal congenital contracture syndrome 1 reveals Gle1 function in spinal neural precursor survival and motor axon arborization. Development 139, 1316–1326.

Jarmoskaite, I., Russell, R., 2011. DEAD-box proteins as RNA helicases and chaperones. Wiley Interdiscip. Rev. RNA 2, 135.

Joo, J.Y., Schaukowitch, K., Farbiak, L., Kilaru, G., Kim, T.K., 2015. Stimulus-specific combinatorial functionality of neuronal c-fos enhancers. Nat. Neurosci. 2016 191 19, 75–83.

Joshi, G., Bekier, M.E., Wang, Y., 2015. Golgi fragmentation in Alzheimer's disease. Front. Neurosci. 9.

Kaifer, K.A., Villalón, E., Osman, E.Y., Glascock, J.J., Arnold, L.L., Cornelison, D.D.W., Lorson, C.L., 2017. Plastin-3 extends survival and reduces severity in mouse models of spinal muscular atrophy. JCI insight 2.

Kaneb, H.M., Folkmann, A.W., Belzil, V. V, Jao, L.-E., Leblond, C.S., Girard, S.L., Daoud, H., Noreau, A., Rochefort, D., Hince, P., Szuto, A., Levert, A., Vidal, S., Andre-Guimont, C., Camu, W., Bouchard, J.-P., Dupre, N., Rouleau, G.A., Wente, S.R., Dion, P.A., 2015. Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. Hum. Mol. Genet. 24, 1363–1373.

Kannan, A., Bhatia, K., Branzei, D., Gangwani, L., 2018. Combined deficiency of Senataxin and DNA-PKcs causes DNA damage accumulation and neurodegeneration in spinal muscular atrophy. Nucleic Acids Res. 46, 8326–8346.

Kannan, A., Jiang, X., He, L., Ahmad, S., Gangwani, L., 2020. ZPR1 prevents R-loop accumulation, upregulates SMN2 expression and rescues spinal muscular atrophy. Brain 143, 69–93.

Karyka, E., Ramirez, N.B., Webster, C.P., Marchi, P.M., Graves, E.J., Godena, V.K., Marrone, L., Bhargava, A., Ray, S., Ning, K., Crane, H., Hautbergue, G.M., El-Khamisy, S.F., Azzouz, M., 2022. SMN-deficient cells exhibit increased ribosomal DNA damage. Life Sci. Alliance 5.

Kendirgi, F., Barry, D.M., Griffis, E.R., Powers, M.A., Wente, S.R., 2003. An essential role for hGle1 nucleocytoplasmic shuttling in mRNA export. J. Cell Biol. 160, 1029–1040.

Kendirgi, F., Rexer, D.J., Alcazar-Roman, A.R., Onishko, H.M., Wente, S.R., 2005. Interaction between the shuttling mRNA export factor Gle1 and the nucleoporin hCG1: a conserved mechanism in the export of Hsp70 mRNA. Mol. Biol. Cell 16, 4304–4315.

Kim, H.H., Lee, S.J., Gardiner, A.S., Perrone-Bizzozero, N.I., Yoo, S., 2015. Different motif requirements for the localization zipcode element of β-actin mRNA binding by HuD and ZBP1. Nucleic Acids Res. 43, 7432–7446.

Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter,

B.D., Hogan, J.A., Upla, P., Chemmama, I.E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A.S., Wang, J., Williams, R., Unruh, J.R., Greenberg, C.H., Jacobs, E.Y., Yu, Z., de la Cruz, M.J., Mironska, R., Stokes, D.L., Aitchison, J.D., Jarrold, M.F., Gerton, J.L., Ludtke, S.J., Akey, C.W., Chait, B.T., Sali, A., Rout, M.P., 2018. Integrative structure and functional anatomy of a nuclear pore complex. Nature 555, 475–482.

Köhler, A., Hurt, E., 2007. Exporting RNA from the nucleus to the cytoplasm. Nat. Rev. Mol. Cell Biol. 8, 761–773.

Kolb, S.J., Kissel, J.T., 2015. Spinal Muscular Atrophy. Neurol. Clin. 33, 831.

Kuan, C.Y., Whitmarsh, A.J., Yang, D.D., Liao, G., Schloemer, A.J., Dong, C., Bao, J., Banasiak, K.J., Haddad, G.G., Flavell, R.A., Davis, R.J., Rakic, P., 2003. A critical role of neural-specific JNK3 for ischemic apoptosis. Proc. Natl. Acad. Sci. U. S. A. 100, 15184–15189.

Kulkarni-Gosavi, P., Makhoul, C., Gleeson, P.A., 2019. Form and function of the Golgi apparatus: scaffolds, cytoskeleton and signalling. FEBS Lett. 593, 2289–2305.

Kye, M.J. eon., Niederst, E.D., Wertz, M.H., Gonçalves, I. do C.G., Akten, B., Dover, K.Z., Peters, M., Riessland, M., Neveu, P., Wirth, B., Kosik, K.S., Sardi, S.P., Monani, U.R., Passini, M.A., Sahin, M., 2014. SMN regulates axonal local translation via miR-183/mTOR pathway. Hum. Mol. Genet. 23, 6318–6331.

Lamond, A.I., Spector, D.L., 2003. Nuclear speckles: a model for nuclear organelles. Nat. Rev. Mol. Cell Biol. 2003 48 4, 605–612.

Lauria, F., Bernabò, P., Tebaldi, T., Groen, E.J.N., Perenthaler, E., Maniscalco, F., Rossi, A., Donzel, D., Clamer, M., Marchioretto, M., Omersa, N., Orri, J., Dalla Serra, M., Anderluh, G., Quattrone, A., Inga, A., Gillingwater, T.H., Viero, G., 2020. SMN-primed ribosomes modulate the translation of transcripts related to spinal muscular atrophy. Nat. Cell Biol. 2020 2210 22, 1239–1251.

Le-Niculescu, H., Bonfoco, E., Kasuya, Y., Claret, F.-X., Green, D.R., Karin, M., 1999. Withdrawal of survival factors results in activation of the JNK pathway in neuronal cells leading to Fas ligand induction and cell death. Mol. Cell. Biol. 19, 751–763.

Le, T.T., Pham, L.T., Butchbach, M.E.R., Zhang, H.L., Monani, U.R., Coovert, D.D., Gavrilina, T.O., Xing, L., Bassell, G.J., Burghes, A.H.M., 2005. SMNΔ7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. Hum. Mol. Genet. 14, 845–857.

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., 1995. Identification and characterization of a spinal muscular atrophy-determining gene. Cell 80, 155–65.

Lefebvre, S., Burlet, P., Liu, Q., Bertrandy, S., Clermont, O., Munnich, A., Dreyfuss, G., Melki, J., 1997. Correlation between severity and SMN protein level in spinal muscular atrophy. Nat. Genet. 16, 265–269.

Li, D.K., Tisdale, S., Lotti, F., Pellizzoni, L., 2014. SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. Semin. Cell Dev. Biol. 32, 22–29.

Li, H., Custer, S.K., Gilson, T., Hao, L.T., Beattie, C.E., Androphy, E.J., 2015. α-COP binding to the survival motor neuron protein SMN is required for neuronal process outgrowth. Hum. Mol. Genet. 24, 7295–7307.

Li, Y., Sun, B., Wang, Z., He, Z., Yang, F., Wang, H., Cui, F., Chen, Z., Ling, L., Wang, C., Huang, X., 2021. Mutation Screening of the GLE1 Gene in a Large Chinese Cohort of Amyotrophic Lateral Sclerosis Patients. Front. Neurosci. 15.

Lin, D.H., Correia, A.R., Cai, S.W., Huber, F.M., Jette, C.A., Hoelz, A., 2018. Structural and functional

analysis of mRNA export regulation by the nuclear pore complex. Nat. Commun. 9, 2319.

Linder, B., Plöttner, O., Kroiss, M., Hartmann, E., Laggerbauer, B., Meister, G., Keidel, E., Fischer, U., 2008. Tdrd3 is a novel stress granule-associated protein interacting with the Fragile-X syndrome protein FMRP. Hum. Mol. Genet. 17, 3236–3246.

Ling, S.H.M., Song, H., 2010. Mechanistic insights into mRNA export through structures of Dbp5. RNA Biol. 7, 23–27.

Liu, J., Huang, Y., Li, T., Jiang, Z., Zeng, L., Hu, Z., 2021. The role of the Golgi apparatus in disease (Review). Int. J. Mol. Med. 47.

Liu, Q., Dreyfuss, G., 1996. A novel nuclear structure containing the survival of motor neurons protein. EMBO J. 15, 3555–65.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408.

Long, K.K., O'Shea, K.M., Khairallah, R.J., Howell, K., Paushkin, S., Chen, K.S., Cote, S.M., Webster, M.T., Stains, J.P., Treece, E., Buckler, A., Donovan, A., 2019. Specific inhibition of myostatin activation is beneficial in mouse models of SMA therapy. Hum. Mol. Genet. 28, 1076.

Lorson, C.L., Hahnen, E., Androphy, E.J., Wirth, B., 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc. Natl. Acad. Sci. U. S. A. 96, 6307–11.

Lyons, D.A., Pogoda, H.-M., Voas, M.G., Woods, I.G., Diamond, B., Nix, R., Arana, N., Jacobs, J., Talbot, W.S., 2005. erbb3 and erbb2 Are Essential for Schwann Cell Migration and Myelination in Zebrafish. Curr. Biol. 15, 513–524.

Macdonald, R., Barnes, K., Hastings, C., Mortiboys, H., 2018. Mitochondrial abnormalities in Parkinson's disease and Alzheimer's disease: can mitochondria be targeted therapeutically? Biochem. Soc. Trans. 46, 891–909.

Macleod, M.J., Taylor, J.E., Lunt, P.W., Mathew, C.G., Robb, S.A., 1999. Prenatal onset spinal muscular atrophy. Eur. J. Paediatr. Neurol. 3, 65–72.

Mahadik, S.P., Mukherjee, S., 1996. CULTURED SKIN FIBROBLASTS AS A CELL MODEL FOR INVESTIGATING SCHIZOPHRENIA. J. psy~hiat. Res 30, 421–439.

Mahmoudi, S., Henriksson, S., Weibrecht, I., Smith, S., Söderberg, O., Strömblad, S., Wiman, K.G., Farnebo, M., 2010. WRAP53 is essential for Cajal body formation and for targeting the survival of motor neuron complex to Cajal bodies. PLoS Biol. 8.

Mailman, M.D., Heinz, J.W., Papp, A.C., Snyder, P.J., Sedra, M.S., Wirth, B., Burghes, A.H.M., Prior, T.W., 2002. Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. Genet. Med. 4, 20–26.

Mäkelä-Bengs, P., Järvinen, N., Vuopala, K., Suomalainen, A., Ignatius, J., Sipilä, M., Herva, R., Palotie, A., Peltonen, L., 1998. Assignment of the Disease Locus for Lethal Congenital Contracture Syndrome to a Restricted Region of Chromosome 9q34, by Genome Scan Using Five Affected Individuals. Am. J. Hum. Genet. 63, 506–516.

Makhoul, C., Gosavi, P., Gleeson, P.A., 2019. Golgi Dynamics: The Morphology of the Mammalian Golgi Apparatus in Health and Disease. Front. Cell Dev. Biol. 7, 112.

Mason, A.C., Wente, S.R., 2020. Functions of Gle1 are governed by two distinct modes of self-association. J. Biol. Chem. 295, 16813.

McWhorter, M.L., Monani, U.R., Burghes, A.H.M., Beattie, C.E., 2003. Knockdown of the survival motor neuron (Smn) protein in zebrafish causes defects in motor axon outgrowth and

pathfinding. J. Cell Biol. 162, 919–932.

Melki, J., Abdelhak, S., Sheth, P., Bachelot, M.F., Burlet, P., Marcadet, A., Aicardi, J., Barois, A., Carriere, J.P., Fardeau, M., Fontan, D., Ponsot, G., Billette, T., Angelini, C., Barbosa, C., Ferriere, G., Lanzi, G., Ottolini, A., Babron, M.C., Cohen, D., Hanauer, A., Clerget-Darpoux, F., Lathrop, M., Munnich, A., Frezal, J., 1990. Gene for chronic proximal spinal muscular atrophies maps to chromosome 5q. Nature 344, 767–768.

Melki, J., Lefebvre, S., Burglen, L., Burlet, P., Clermont, O., Millasseau, P., Reboullet, S., Bénichou, B., Zeviani, M., Le Paslier, D., Cohen, D., Weissenbach, J., Munnich, A., 1994. De novo and inherited deletions of the 5q13 region in spinal muscular atrophies. Science 264, 1474–1477.

Mendell, J.R., Al-Zaidy, S., Shell, R., Arnold, W.D., Rodino-Klapac, L.R., Prior, T.W., Lowes, L., Alfano, L., Berry, K., Church, K., Kissel, J.T., Nagendran, S., L'Italien, J., Sproule, D.M., Wells, C., Cardenas, J.A., Heitzer, M.D., Kaspar, A., Corcoran, S., Braun, L., Likhite, S., Miranda, C., Meyer, K., Foust, K.D., Burghes, A.H.M., Kaspar, B.K., 2017. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. N. Engl. J. Med. 377, 1713–1722.

Mercuri, E., Darras, B.T., Chiriboga, C.A., Day, J.W., Campbell, C., Connolly, A.M., Iannaccone, S.T., Kirschner, J., Kuntz, N.L., Saito, K., Shieh, P.B., Tulinius, M., Mazzone, E.S., Montes, J., Bishop, K.M., Yang, Q., Foster, R., Gheuens, S., Bennett, C.F., Farwell, W., Schneider, E., De Vivo, D.C., Finkel, R.S., 2018. Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. N. Engl. J. Med. 378, 625–635.

Meyer, K., Ferraiuolo, L., Schmelzer, L., Braun, L., McGovern, V., Likhite, S., Michels, O., Govoni, A., Fitzgerald, J., Morales, P., Foust, K.D., Mendell, J.R., Burghes, A.H.M., Kaspar, B.K., 2015. Improving single injection CSF delivery of AAV9-mediated gene therapy for SMA: a dose-response study in mice and nonhuman primates. Mol. Ther. 23, 477–487.

Miller, A.L., Suntharalingam, M., Johnson, S.L., Audhya, A., Emr, S.D., Wente, S.R., 2004. Cytoplasmic inositol hexakisphosphate production is sufficient for mediating the Gle1-mRNA export pathway. J. Biol. Chem. 279, 51022–51032.

Monani, U.R., Lorson, C.L., Parsons, D.W., Prior, T.W., Androphy, E.J., Burghes, A.H., McPherson, J.D., 1999. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. Hum. Mol. Genet. 8, 1177–83.

Monani, U.R., Sendtner, M., Coovert, D.D., Parsons, D.W., Andreassi, C., Le, T.T., Jablonka, S., Schrank, B., Rossol, W., Prior, T.W., Morris, G.E., Burghes, A.H.M., 2000. The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn–/– mice and results in a mouse with spinal muscular atrophy. Hum. Mol. Genet. 9, 333–339.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., Weis, K., 2011. A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature 472, 238–242.

Morishima, Y., Gotoh, Y., Zieg, J., Barrett, T., Takano, H., Flavell, R., Davis, R.J., Shirasaki, Y., Greenberg, M.E., 2001. Beta-amyloid induces neuronal apoptosis via a mechanism that involves the c-Jun N-terminal kinase pathway and the induction of Fas ligand. J. Neurosci. 21, 7551–7560.

Muntoni, F., Bertini, E., Comi, G., Kirschner, J., Lusakowska, A., Mercuri, E., Scoto, M., van der Pol, W.L., Vuillerot, C., Burdeska, A., El-Khairi, M., Fontoura, P., Ives, J., Gorni, K., Reid, C., Fuerst-Recktenwald, S., 2020. Long-term follow-up of patients with type 2 and non-ambulant type 3 spinal muscular atrophy (SMA) treated with olesoxime in the OLEOS trial. Neuromuscul. Disord. 30, 959–969.

Murphy, R., Wente, S.R., 1996. An RNA-export mediator with an essential nuclear export signal.

Nature 383, 357–360.

Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., Trono, D., 1996. In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. Science (80-.). 272, 263–267.

Narcís, J.O., Tapia, O., Tarabal, O., Piedrafita, L., Calderó, J., Berciano, M.T., Lafarga, M., 2018. Accumulation of poly(A) RNA in nuclear granules enriched in Sam68 in motor neurons from the SMNΔ7 mouse model of SMA. Sci. Reports 2018 81 8, 1–16.

Narkis, G., Ofir, R., Manor, E., Landau, D., Elbedour, K., Birk, O.S., 2007. Lethal Congenital Contractural Syndrome Type 2 (LCCS2) Is Caused by a Mutation in ERBB3 (Her3), a Modulator of the Phosphatidylinositol-3-Kinase/Akt Pathway. Am. J. Hum. Genet. 81, 589–595.

NCBI, 2021. HomoloGene - NCBI [WWW Document]. Viewed 2021-12-21. <https://www.ncbi.nlm.nih.gov/homologene/44781>

Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., McCluskey, L.F., Miller, B.L., Masliah, E., Mackenzie, I.R., Feldman, H., Feiden, W., Kretzschmar, H.A., Trojanowski, J.Q., Lee, V.M.-Y., 2006. Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. Science (80-.). 314, 130–133.

Neveling, K., Martinez-Carrera, L.A., Hölker, I., Heister, A., Verrips, A., Hosseini-Barkooie, S.M., Gilissen, C., Vermeer, S., Pennings, M., Meijer, R., Te Riele, M., Frijns, C.J.M., Suchowersky, O., Maclaren, L., Rudnik-Schöneborn, S., Sinke, R.J., Zerres, K., Lowry, R.B., Lemmink, H.H., Garbes, L., Veltman, J.A., Schelhaas, H.J., Scheffer, H., Wirth, B., 2013. Mutations in BICD2, which encodes a golgin and important motor adaptor, cause congenital autosomal-dominant spinal muscular atrophy. Am. J. Hum. Genet. 92, 946–954.

Noble, K.N., Tran, E.J., Alcazar-Roman, A.R., Hodge, C.A., Cole, C.N., Wente, S.R., 2011. The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. Genes Dev. 25, 1065–1077.

Nousiainen, H.O., Kestilä, M., Pakkasjärvi, N., Honkala, H., Kuure, S., Tallila, J., Vuopala, K., Ignatius, J., Herva, R., Peltonen, L., 2008. Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease. Nat. Genet. 40, 155–157.

Obado, S.O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N.E., Chait, B.T., Field, M.C., Rout, M.P., 2016. Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. PLOS Biol. 14, e1002365.

Okamura, M., Yamanaka, Y., Shigemoto, M., Kitadani, Y., Kobayashi, Y., Kambe, T., Nagao, M., Kobayashi, I., Okumura, K., Masuda, S., 2018. Depletion of mRNA export regulator DBP5/DDX19, GLE1 or IPPK that is a key enzyme for the production of IP6, resulting in differentially altered cytoplasmic mRNA expression and specific cell defect. PLoS One 13, e0197165.

Oprea, G.E., Kröber, S., McWhorter, M.L., Rossoll, W., Müller, S., Krawczak, M., Bassell, G.J., Beattie, C.E., Wirth, B., 2008. Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. Science (80-.). 320, 524–527.

Paakkola, T., Vuopala, K., Kokkonen, H., Ignatius, J., Valkama, M., Moilanen, J.S., Fahiminiya, S., Majewski, J., Hinttala, R., Uusimaa, J., 2017. A homozygous I684T in GLE1 as a novel cause of arthrogryposis and motor neuron loss. Clin. Genet.

Pakkasjärvi, N., Ritvanen, A., Herva, R., Peltonen, L., Kestilä, M., Ignatius, J., 2006. Lethal congenital contracture syndrome (LCCS) and other lethal arthrogryposes in Finland—An epidemiological study. Am. J. Med. Genet. Part A 140A, 1834–1839.

Palfi, S., Gurruchaga, J.M., Lepetit, H., Howard, K., Ralph, G.S., Mason, S., Gouello, G., Domenech, P., Buttery, P.C., Hantraye, P., Tuckwell, N.J., Barker, R.A., Mitrophanous, K.A., 2018. Long-Term Follow-Up of a Phase I/II Study of ProSavin, a Lentiviral Vector Gene Therapy for Parkinson's Disease. Hum. Gene Ther. Clin. Dev. 29, 148.

Palfi, S., Gurruchaga, J.M., Scott Ralph, G., Lepetit, H., Lavisse, S., Buttery, P.C., Watts, C., Miskin, J., Kelleher, M., Deeley, S., Iwamuro, H., Lefaucheur, J.P., Thiriez, C., Fenelon, G., Lucas, C., Brugières, P., Gabriel, I., Abhay, K., Drouot, X., Tani, N., Kas, A., Ghaleh, B., Le Corvoisier, P., Dolphin, P., Breen, D.P., Mason, S., Guzman, N.V., Mazarakis, N.D., Radcliffe, P.A., Harrop, R., Kingsman, S.M., Rascol, O., Naylor, S., Barker, R.A., Hantraye, P., Remy, P., Cesaro, P., Mitrophanous, K.A., 2014. Long-term safety and tolerability of ProSavin, a lentiviral vector-based gene therapy for Parkinson's disease: a dose escalation, open-label, phase 1/2 trial. Lancet (London, England) 383, 1138–1146.

Palmer, E., Freeman, T., 2004. Investigation Into the use of C- and N-terminal GFP Fusion Proteins for Subcellular Localization Studies Using Reverse Transfection Microarrays. Comp. Funct. Genomics 5, 342.

Paulsen, R.D., Soni, D. V., Wollman, R., Hahn, A.T., Yee, M.C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., Meyer, T., Cimprich, K.A., 2009. A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. Mol. Cell 35, 228–239.

Pellizzoni, L., Baccon, J., Charroux, B., Dreyfuss, G., 2001a. The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. Curr. Biol. 11, 1079–1088.

Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., Dreyfuss, G., 2001b. A functional interaction between the survival motor neuron complex and RNA polymerase II. J. Cell Biol. 152, 75–85.

Pellizzoni, L., Yong, J., Dreyfuss, G., 2002. Essential role for the SMN complex in the specificity of snRNP assembly. Science 298, 1775–1779.

Peter, C.J., Evans, M., Thayanithy, V., Taniguchi-Ishigaki, N., Bach, I., Kolpak, A., Bassell, G.J., Rossoll, W., Lorson, C.L., Bao, Z.Z., Androphy, E.J., 2011. The COPI vesicle complex binds and moves with survival motor neuron within axons. Hum. Mol. Genet. 20, 1701.

Poirier, A., Weetall, M., Heinig, K., Bucheli, F., Schoenlein, K., Alsenz, J., Bassett, S., Ullah, M., Senn, C., Ratni, H., Naryshkin, N., Paushkin, S., Mueller, L., 2018. Risdiplam distributes and increases SMN protein in both the central nervous system and peripheral organs. Pharmacol. Res. Perspect. 6.

Polido, G.J., Miranda, M.M.V. De, Junior, N.C., Mendonça, R.D.H., Caromano, F.A., Reed, U.C., Zanoteli, E., Voos, M.C., 2019. Cognitive performance of children with spinal muscular atrophy: Asystematic review. Dement. Neuropsychol. 13, 436.

Prior, T.W., Leach, M.E., Finanger, E., 1993. Spinal Muscular Atrophy. GeneReviews®.

Prior, T.W., Swoboda, K.J., Scott, H.D., Hejmanowski, A.Q., 2004. Homozygous SMN1 Deletions in Unaffected Family Members and Modification of the Phenotype by SMN2. Am. J. Med. Genet. A 0, 307.

Protter, D.S.W., Parker, R., 2016. Principles and Properties of Stress granules. Trends Cell Biol. 26, 668.

Rage, F., Boulisfane, N., Rihan, K., Neel, H., Gostan, T., Bertrand, E., Bordonné, R., Soret, A., 2013. Genome-wide identification of mRNAs associated with the protein SMN whose depletion decreases their axonal localization. RNA 19, 1755.

Rathod, R., Havlicek, S., Frank, N., Blum, R., Sendtner, M., 2012. Laminin induced local axonal

translation of β-actin mRNA is impaired in SMN-deficient motoneurons. Histochem. Cell Biol. 138, 737–748.

Ravanidis, S., Kattan, F.G., Doxakis, E., 2018. Unraveling the Pathways to Neuronal Homeostasis and Disease: Mechanistic Insights into the Role of RNA-Binding Proteins and Associated Factors. Int. J. Mol. Sci. 2018, Vol. 19, Page 2280 19, 2280.

Rayala, H.J., Kendirgi, F., Barry, D.M., Majerus, P.W., Wente, S.R., 2004. The mRNA export factor human Gle1 interacts with the nuclear pore complex protein Nup155. Mol. Cell. Proteomics 3, 145–155.

Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., Aebi, U., 1990. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. J. Cell Biol. 110, 883–94.

Renton, A.E., Chiò, A., Traynor, B.J., 2013. State of play in amyotrophic lateral sclerosis genetics. Nat. Neurosci. 17, 17–23.

Renvoisé, B., Colasse, S., Burlet, P., Viollet, L., Meier, U.T., Lefebvre, S., 2009. The loss of the snoRNP chaperone Nopp140 from Cajal bodies of patient fibroblasts correlates with the severity of spinal muscular atrophy. Hum. Mol. Genet. 18, 1181–1189.

Riessland, M., Kaczmarek, A., Schneider, S., Swoboda, K.J., Löhr, H., Bradler, C., Grysko, V., Dimitriadi, M., Hosseinibarkooie, S., Torres-Benito, L., Peters, M., Upadhyay, A., Biglari, N., Kröber, S., Hölker, I., Garbes, L., Gilissen, C., Hoischen, A., Nürnberg, G., Nürnberg, P., Walter, M., Rigo, F., Bennett, C.F., Kye, M.J., Hart, A.C., Hammerschmidt, M., Kloppenburg, P., Wirth, B., 2017. Neurocalcin Delta Suppression Protects against Spinal Muscular Atrophy in Humans and across Species by Restoring Impaired Endocytosis. Am. J. Hum. Genet. 100, 297–315.

Riethmacher, D., Sonnenberg-Riethmacher, E., Brinkmann, V., Yamaai, T., Lewin, G.R., Birchmeier, C., 1997. Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. Nature 389, 725–730.

Rigo, F., Hua, Y., Krainer, A.R., Frank Bennett, C., 2012. Antisense-based therapy for the treatment of spinal muscular atrophy. J. Cell Biol. 199, 21–25.

Rinaldi, C., Pizzul, P., Longhese, M.P., Bonetti, D., 2021. Sensing R-Loop-Associated DNA Damage to Safeguard Genome Stability. Front. Cell Dev. Biol. 8, 1657.

Rochette, C.F., Gilbert, N., Simard, L.R., 2001. SMN gene duplication and the emergence of the SMN2 gene occurred in distinct hominids: SMN2 is unique to Homo sapiens. Hum. Genet. 108, 255–266.

Rossoll, W., Jablonka, S., Andreassi, C., Kröning, A.K., Karle, K., Monani, U.R., Sendtner, M., 2003. Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. J. Cell Biol. 163, 801–812.

Said, E., Chong, J.X., Hempel, M., Denecke, J., Soler, P., Strom, T., Nickerson, D.A., Kubisch, C., Bamshad, M.J., Lessel, D., 2017. Survival beyond the perinatal period expands the phenotypes caused by mutations in GLE1. Am. J. Med. Genet. A 173, 3098–3103.

Sanchez, G., Dury, A.Y., Murray, L.M., Biondi, O., Tadesse, H., El fatimy, R., Kothary, R., Charbonnier, F., Khandjian, E.W., Côté, J., 2013. A novel function for the survival motoneuron protein as a translational regulator. Hum. Mol. Genet. 22, 668–684.

Sareen, D., Ebert, A.D., Heins, B.M., McGivern, J. V., Ornelas, L., Svendsen, C.N., 2012. Inhibition of apoptosis blocks human motor neuron cell death in a stem cell model of spinal muscular atrophy. PLoS One 7.

Schmitt, C., von Kobbe, C., Bachi, A., Panté, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M.,

Séraphin, B., Carmo-Fonseca, M., Izaurralde, E., 1999. Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO J. 18, 4332–4347.

Schmutz, J., Martin, J., Terry, A., Couronne, O., Grimwood, J., Lowry, S., Gordon, L.A., Scott, D., Xie, G., Huang, W., Hellsten, U., Tran-Gyamfi, M., She, X., Prabhakar, S., Aerts, A., Altherr, M., Bajorek, E., Black, S., Branscomb, E., Caoile, C., Challacombe, J.F., Chan, Y.M., Denys, M., Detter, J.C., Escobar, J., Flowers, D., Fotopulos, D., Glavina, T., Gomez, M., Gonzales, E., Goodstein, D., Grigoriev, I., Groza, M., Hammon, N., Hawkins, T., Haydu, L., Israni, S., Jett, J., Kadner, K., Kimball, H., Kobayashi, A., Lopez, F., Lou, Y., Martinez, D., Medina, C., Morgan, J., Nandkeshwar, R., Noonan, J.P., Pitluck, S., Pollard, M., Predki, P., Priest, J., Ramirez, L., Retterer, J., Rodriguez, A., Rogers, S., Salamov, A., Salazar, A., Thayer, N., Tice, H., Tsai, M., Ustaszewska, A., Vo, N., Wheeler, J., Wu, K., Yang, J., Dickson, M., Cheng, J.F., Eichler, E.E., Olsen, A., Pennacchio, L.A., Rokhsar, D.S., Richardson, P., Lucas, S.M., Myers, R.M., Rubin, E.M., 2004. The DNA sequence and comparative analysis of human chromosome 5. Nat. 2004 4317006 431, 268–274.

Schorling, D.C., Pechmann, A., Kirschner, J., 2020. Advances in Treatment of Spinal Muscular Atrophy – New Phenotypes, New Challenges, New Implications for Care. J. Neuromuscul. Dis. 7, 1.

Schrank, B., Götz, R., Gunnersen, J.M., Ure, J.M., Toyka, K. V, Smith, A.G., Sendtner, M., 1997. Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. Proc. Natl. Acad. Sci. U. S. A. 94, 9920–5.

Setola, V., Terao, M., Locatelli, D., Bassanini, S., Garattiini, E., Battaglia, G., 2007. Axonal-SMN (a-SMN), a protein isoform of the survival motor neuron gene, is specifically involved in axonogenesis. Proc. Natl. Acad. Sci. U. S. A. 104, 1959.

Seytanoglu, A., Alsomali, N.I., Valori, C.F., McGown, A., Kim, H.R., Ning, K., Ramesh, T., Sharrack, B., Wood, J.D., Azzouz, M., 2016. Deficiency in the mRNA export mediator Gle1 impairs Schwann cell development in the zebrafish embryo. Neuroscience 322, 287–297.

Sharma, M., Wente, S.R., 2020. Nucleocytoplasmic shuttling of Gle1 impacts DDX1 at transcription termination sites. Mol. Biol. Cell 31, 2398.

Shinomiya, H., 2012. Plastin family of actin-bundling proteins: its functions in leukocytes, neurons, intestines, and cancer. Int. J. Cell Biol. 2012.

Sigma, 2021. How Proximity Ligation Assays (PLA) Work [WWW Document]. Viewed 2021-11-16. <https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/protein-biology/protein-and-nucleic-acid-interactions/how-pla-works>

Simard, L.R., Bélanger, M.C., Morissette, S., Wride, M., Prior, T.W., Swoboda, K.J., 2007. Preclinical validation of a multiplex real-time assay to quantify SMN mRNA in patients with SMA. Neurology 68, 451–456.

Singh, R.N., Howell, M.D., Ottesen, E.W., Singh, N.N., 2017. Diverse role of Survival Motor Neuron Protein. Biochim. Biophys. Acta 1860, 299.

Singh, R.N., Ottesen, E.W., Singh, N.N., 2020. The First Orally Deliverable Small Molecule for the Treatment of Spinal Muscular Atrophy. Neurosci. insights 15.

Sinn, P.L., Sauter, S.L., McCray, P.B., 2005. Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. Gene Ther. 12, 1089–1098.

Skourti-Stathaki, K., Proudfoot, N.J., Gromak, N., 2011. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. Mol. Cell 42, 794–805.

Smith, C., Parboosingh, J.S., Boycott, K.M., Bonnemann, C.G., Mah, J.K., Lamont, R.E., Micheil Innes, A., Bernier, F.P., 2017. Expansion of the GLE1-associated arthrogryposis multiplex congenita clinical spectrum. Clin. Genet. 91, 426–430.

Smolka, J.A., Sanz, L.A., Hartono, S.R., Chédin, F., 2021. Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. J. Cell Biol. 220.

Song, C., Hotz-Wagenblatt, A., Voit, R., Grummt, I., 2017. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. Genes Dev. 31, 1370–1381.

Stabley, D.L., Harris, A.W., Holbrook, J., Chubbs, N.J., Lozo, K.W., Crawford, T.O., Swoboda, K.J., Funanage, V.L., Wang, W., Mackenzie, W., Scavina, M., Sol-Church, K., Butchbach, M.E.R., 2015. SMN1 and SMN2 copy numbers in cell lines derived from patients with spinal muscular atrophy as measured by array digital PCR. Mol. Genet. genomic Med. 3, 248–257.

Stalder, D., Gershlick, D.C., 2020. Direct trafficking pathways from the Golgi apparatus to the plasma membrane. Semin. Cell Dev. Biol. 107, 112–125.

Stanley, G.J., Fassati, A., Hoogenboom, B.W., 2017. Biomechanics of the transport barrier in the nuclear pore complex. Semin. Cell Dev. Biol. 68, 42–51.

Statland, J.M., Barohn, R.J., McVey, A.L., Katz, J.S., Dimachkie, M.M., 2015. Patterns of Weakness, Classification of Motor Neuron Disease, and Clinical Diagnosis of Sporadic Amyotrophic Lateral Sclerosis. Neurol. Clin. 33, 735–48.

Strahm, Y., Fahrenkrog, B., Zenklusen, D., Rychner, E., Kantor, J., Rosbach, M., Stutz, F., 1999. The RNA export factor Gle1p is located on the cytoplasmic fibrils of the NPC and physically interacts with the FG-nucleoporin Rip1p, the DEAD-box protein Rat8p/Dbp5p and a new protein Ymr 255p. EMBO J. 18, 5761–5777.

Sumner, C.J., Kolb, S.J., Harmison, G.G., Jeffries, N.O., Schadt, K., Finkel, R.S., Dreyfuss, G., Fischbeck, K.H., 2006. SMN mRNA and protein levels in peripheral blood. Neurology 66, 1067–1073.

Sundaramoorthy, V., Sultana, J.M., Atkin, J.D., 2015. Golgi fragmentation in amyotrophic lateral sclerosis, an overview of possible triggers and consequences. Front. Neurosci. 9.

Suraweera, A., Lim, Y.C., Woods, R., Birrell, G.W., Nasim, T., Becherel, O.J., Lavin, M.F., 2009. Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. Hum. Mol. Genet. 18, 3384–3396.

Taguchi, T., Pypaert, M., Warren, G., 2003. Biochemical Sub-Fractionation of the Mammalian Golgi Apparatus. Traffic 4, 344–352.

Tan, Q.K.-G., McConkie-Rosell, A., Juusola, J., Gustafson, K.E., Pizoli, C.E., Buckley, A.F., Jiang, Y.-H., 2017. The importance of managing the patient and not the gene: expanded phenotype of GLE1-associated arthrogryposis. Cold Spring Harb. Mol. case Stud.

Ting, C.H., Wen, H.L., Liu, H.C., Hsieh-Li, H.M., Li, H., Lin-Chao, S., 2012. The spinal muscular atrophy disease protein SMN is linked to the Golgi network. PLoS One 7.

Tiziano, F.D., Pinto, A.M., Fiori, S., Lomastro, R., Messina, S., Bruno, C., Pini, A., Pane, M., D'Amico, A., Ghezzo, A., Bertini, E., Mercuri, E., Neri, G., Brahe, C., 2009. SMN transcript levels in leukocytes of SMA patients determined by absolute real-time PCR. Eur. J. Hum. Genet. 2010 181 18, 52–58.

Todd, A.G., Lin, H., Ebert, A.D., Liu, Y., Androphy, E.J., 2013. COPI transport complexes bind to specific RNAs in neuronal cells. Hum. Mol. Genet. 22, 729–736.

Tseng, S.S.-I., Weaver, P.L., Liu, Y., Hitomi, M., Tartakoff, A.M., Chang, T.H., 1998. Dbp5p, a cytosolic

RNA helicase, is required for poly(A)+ RNA export. EMBO J. 17, 2651–2662.

Upadhyay, A., Hosseinibarkooie, S., Schneider, S., Kaczmarek, A., Torres-Benito, L., Mendoza-Ferreira, N., Overhoff, M., Rombo, R., Grysko, V., Kye, M.J., Kononenko, N.L., Wirth, B., 2019. Neurocalcin delta knockout impairs adult neurogenesis whereas half reduction is not pathological. Front. Mol. Neurosci. 12, 19.

Valori, C.F., Ning, K., Wyles, M., Mead, R.J., Grierson, A.J., Shaw, P.J., Azzouz, M., 2010. Systemic delivery of scAAV9 expressing SMN prolongs survival in a model of spinal muscular atrophy. Sci. Transl. Med. 2.

Vuopala, K., Ignatius, J., Herva, R., 1995. Lethal arthrogryposis with anterior horn cell disease. Hum. Pathol. 26, 12–9.

Wadman, R.I., Jansen, M.D., Curial, C.A.D., Groen, E.J.N., Stam, M., Wijngaarde, C.A., Medic, J., Sodaar, P., Van Eijk, K.R., Huibers, M.M.H., Van Kuik, J., Lemmink, H.H., Van Rheenen, W., Veldink, J.H., Van Den Berg, L.H., Van Der Pol, W.L., 2019. Analysis of FUS, PFN2, TDP-43, and PLS3 as potential disease severity modifiers in spinal muscular atrophy. Neurol. Genet. 6.

Wadman, R.I., Jansen, M.D., Stam, M., Wijngaarde, C.A., Curial, C.A.D., Medic, J., Sodaar, P., Schouten, J., Vijzelaar, R., Lemmink, H.H., van den Berg, L.H., Groen, E.J.N., van der Pol, W.L., 2020. Intragenic and structural variation in the SMN locus and clinical variability in spinal muscular atrophy. Brain Commun. 2.

Wadman, R.I., Stam, M., Gijzen, M., Lemmink, H.H., Snoeck, I.N., Wijngaarde, C.A., Braun, K.P.J., Schoenmakers, M.A.G.C., Van Den Berg, L.H., Dooijes, D., Van Der Pol, W.L., 2017. Association of motor milestones, SMN2 copy and outcome in spinal muscular atrophy types 0–4. J. Neurol. Neurosurg. Psychiatry 88, 365–367.

Wadman, R.I., Wijngaarde, C.A., Stam, M., Bartels, B., Otto, L.A.M., Lemmink, H.H., Schoenmakers, M.A.G.C., Cuppen, I., van den Berg, L.H., van der Pol, W.L., 2018. Muscle strength and motor function throughout life in a cross-sectional cohort of 180 patients with spinal muscular atrophy types 1c-4. Eur. J. Neurol. 25, 512–518.

Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W., Lukashchuk, V., Chiang, S.-C., Ray, S., Mulcahy, P.J., Jurga, M., Tsagakis, I., Iannitti, T., Chandran, J., Coldicott, I., De Vos, K.J., Hassan, M.K., Higginbottom, A., Shaw, P.J., Hautbergue, G.M., Azzouz, M., El-Khamisy, S.F., 2017. C9orf72 expansion disrupts ATM-mediated chromosomal break repair. Nat. Neurosci. 20, 1225–1235.

Walsh, M.B., Janzen, E., Wingrove, E., Hosseinibarkooie, S., Muela, N.R., Davidow, L., Dimitriadi, M., Norabuena, E.M., Rubin, L.L., Wirth, B., Hart, A.C., 2020. Genetic modifiers ameliorate endocytic and neuromuscular defects in a model of spinal muscular atrophy. BMC Biol. 18.

Wang, D., Tai, P.W.L., Gao, G., 2019. Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. 2019 185 18, 358–378.

Wang, W., Zhou, Z., Zhao, W., Huang, Y., Tang, R., Ying, K., Xie, Y., Mao, Y., 2001. Molecular cloning, mapping and characterization of the human neurocalcin delta gene (NCALD). Biochim. Biophys. Acta - Gene Struct. Expr. 1518, 162–167.

Watkins, J.L., Murphy, R., Emtage, J.L., Wente, S.R., 1998. The human homologue of Saccharomyces cerevisiae Gle1p is required for poly(A)+ RNA export. Proc. Natl. Acad. Sci. U. S. A. 95, 6779–6784.

Weirich, C.S., Erzberger, J.P., Flick, J.S., Berger, J.M., Thorner, J., Weis, K., 2006. Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. Nat. Cell Biol. 8, 668–676.

Wolf, E.J., Miles, A., Lee, E.S., Nabeel-Shah, S., Greenblatt, J.F., Palazzo, A.F., Tropepe, V., Emili, A., 2020. MKRN2 Physically Interacts with GLE1 to Regulate mRNA Export and Zebrafish Retinal Development. Cell Rep. 31.

Wolff, L., Strathmann, E.A., Müller, I., Mählich, D., Veltman, C., Niehoff, A., Wirth, B., 2021. Plastin 3 in health and disease: a matter of balance. Cell. Mol. Life Sci. 78, 5275.

Workman, E., Kolb, S.J., Battle, D.J., 2012. Spliceosomal small nuclear ribonucleoprotein biogenesis defects and motor neuron selectivity in spinal muscular atrophy. Brain Res. 1462, 93–9.

Yamatani, H., Kawasaki, T., Mita, S., Inagaki, N., Hirata, T., 2010. Proteomics analysis of the temporal changes in axonal proteins during maturation. Dev. Neurobiol. 70, 523–537.

Yanling Zhao, D., Gish, G., Braunschweig, U., Li, Y., Ni, Z., Schmitges, F.W., Zhong, G., Liu, K., Li, W., Moffat, J., Vedadi, M., Min, J., Pawson, T.J., Blencowe, B.J., Greenblatt, J.F., 2016. SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. Nature 529, 48–53.

Yates, T.M., Campeau, P.M., Ghoumid, J., Kibaek, M., Larsen, M.J., Smol, T., Albaba, S., Hertz, J.M., Balasubramanian, M., 2020. Biallelic variants in GLE1 with survival beyond neonatal period. Clin. Genet. 98, 622–625.

York, J.D., Odom, A.R., Murphy, R., Ives, E.B., Wente, S.R., 1999. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. Science 285, 96–100.

Young, P.J., Day, P.M., Zhou, J., Androphy, E.J., Morris, G.E., Lorson, C.L., 2002. A direct interaction between the survival motor neuron protein and p53 and its relationship to spinal muscular atrophy. J. Biol. Chem. 277, 2852–2859.

Yuan, Z., Gong, S., Luo, J., Zheng, Z., Song, B., Ma, S., Guo, J., Hu, C., Thiel, G., Vinson, C., Hu, C.-D., Wang, Y., Li, M., 2009. Opposing Roles for ATF2 and c-Fos in c-Jun-Mediated Neuronal Apoptosis. Mol. Cell. Biol. 29, 2431–2442.

Zerres, K., Davies, K.E., 1999. 59th ENMC International Workshop: Spinal Muscular Atrophies: recent progress and revised diagnostic criteria 17–19 April 1998, Soestduinen, The Netherlands. Neuromuscul. Disord. 9, 272–278.

Zhang, J., Zhang, D., McQuade, J.S., Behbehani, M., Tsien, J.Z., Xu, M., 2002. c-fos regulates neuronal excitability and survival. Nat. Genet. 2002 304 30, 416–420.

Zhang, Q.J., Li, J.J., Lin, X., Lu, Y.Q., Guo, X.X., Dong, E.L., Zhao, M., He, J., Wang, N., Chen, W.J., 2017. Modeling the phenotype of spinal muscular atrophy by the direct conversion of human fibroblasts to motor neurons. Oncotarget 8, 10945.

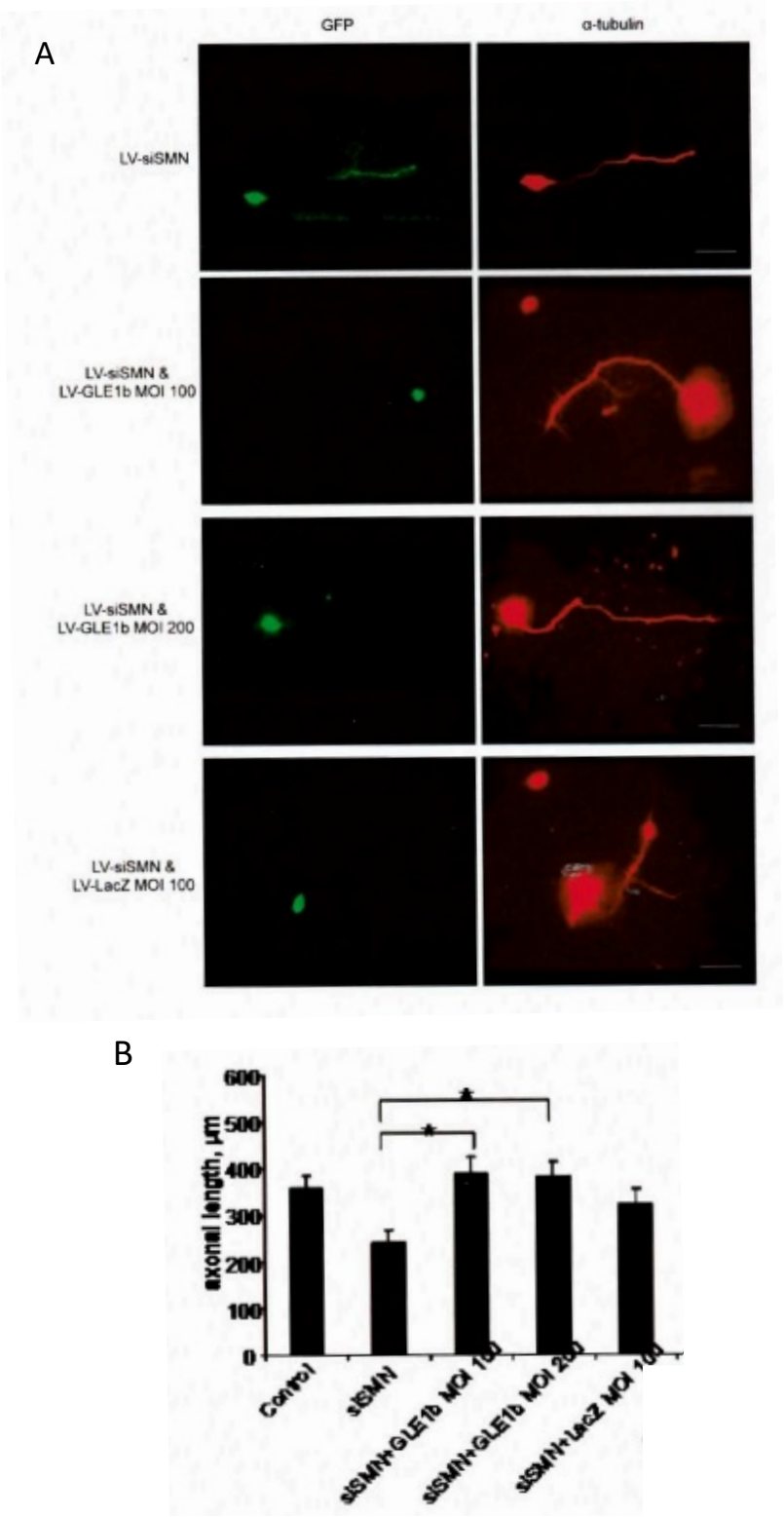
Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M., Dreyfuss, G., 2008. SMN Deficiency Causes Tissue-Specific Perturbations in the Repertoire of snRNAs and Widespread Defects in Splicing. Cell 133, 585–600.

Zhao, J., Jin, S.-B., Björkroth, B., Wieslander, L., Daneholt, B., 2002. The mRNA export factor Dbp5 is associated with Balbiani ring mRNP from gene to cytoplasm. EMBO J. 21, 1177–87.

Zou, J., Barahmand-Pour, F., Blackburn, M.L., Matsui, Y., Chansky, H.A., Yang, L., 2004. Survival motor neuron (SMN) protein interacts with transcription corepressor mSin3A. J. Biol. Chem. 279, 14922–14928.

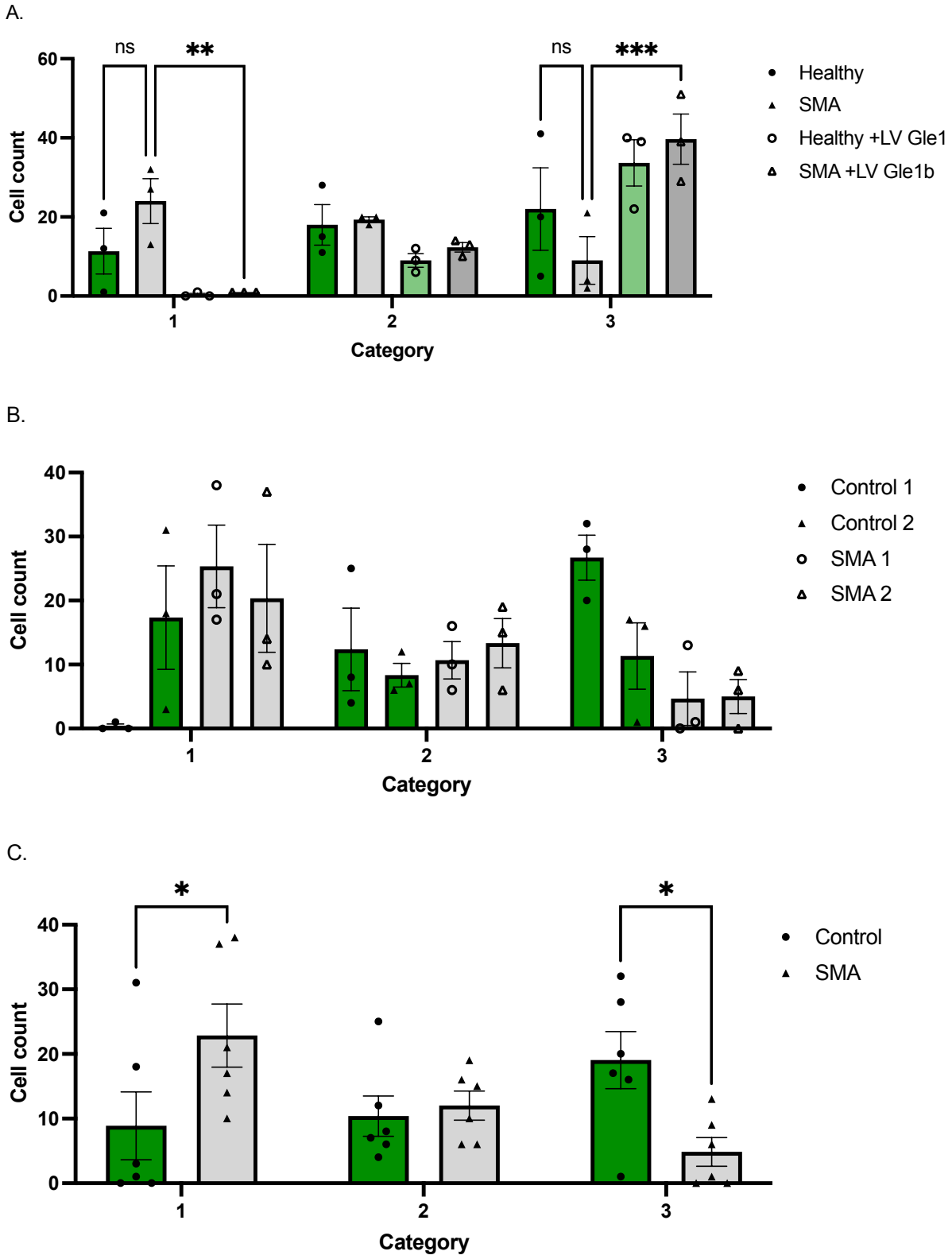
Zou, T., Yang, X., Pan, D., Huang, J., Sahin, M., Zhou, J., 2011. SMN deficiency reduces cellular ability to form stress granules, sensitizing cells to stress. Cell. Mol. Neurobiol. 31, 541–550.

Appendix



Appendix 1: Lentiviral mediated overexpression of Gle1 in WT primary mouse motor neurons transduced with LV-shSMN rescues axonal growth defect. (*A) Immunocytochemistry staining of alpha-tubulin of WT mouse primary motor neurons. GFP immunoflourescenece shows transduced cells. (B) Quantification of axonal length (One way*

*ANOVA+Bonferroni post hoc test, p<0.05, *).* ***Data taken from Masters Thesis by Regina Marketou, 2009***



Appendix 2: Additional categorisation graphs for Golgi quantification in fibroblasts and motor neurons, to show conservation of trends across different individual quantifiers. (A)

*Quantification of fibroblast cells. Data displayed as mean +/- SEM. Data collected from three biologically independent replicates (n=3), 60 cells per condition. SMA cells tend to show a more clustered phenotype compared to controls (Two-way ANOVA, Control vs SMA, Category 1; p=0.1303, ns, n=3. Control vs SMA, Category 3; p=1167, ns, n=3). When Gle1 protein is overexpressed, there is a significant increase from Type 1 to Type 3 in SMA cells (Two-way ANOVA, SMA vs SMA+Gle1b, Category 1; p=0.0.0027, **, n=3. SMA vs SMA+Gle1b, Category 3; p=0.0001, ***, n=3).* ***(B)*** *Quantification of iPSC-derived motor neurons. Data displayed as mean +/- SEM. Data collected from three biologically independent replicates (n=3), 40 cells per condition. SMA cells tend to show a more clustered phenotype compared to controls (**C**) Motor neuron data grouped into control and SMA. SMA cells show a significantly higher tendency to be more category 1 (clustered) than control cells, which are more likely to be category 3 (Two-way ANOVA: Category 1: p=0.477, *, n=6. Category 3: p=0.445, *, n=6).*

Appendix 3: Plasmid Sequences

pEGFP Gle1b

ACTGGGTATTCGTCATGTCGATACCGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGA
AACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACTGCGGTTGCGA
TTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTAT
GTTGTTGATATCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATCTCCGG
TCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAA
GTATTCTGGAGGCTGCATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACATCCTGAA
ACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTGTGCAGTCGGCCCTTGATGGTAAAACCATCCC
TCACTGGTATCGCATGATTAACCGTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAG
GCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACGGTGATTGGCTACCGTGGCGGCA
ACTGGATTTATGAGTGGGCCCCGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACC
TACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTG
TGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTT
GCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGA
AAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTT
GCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAAC
CTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATT
AATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGT
GCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGCTTTTACTTGCTTTAAAAAACCTCCCACACCTCCC
CCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA
TCTTATCATGTGTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAACTTCCCACCCCATACCCTATTACC
ACTGCCAAATTACCTGTGGTTTCATTTACTCTAAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTT
GTTAAATATGTACTACAAACTTAGTAGTTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTG
GATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCT
TTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCT
TGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCC
TAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGATATCGAGCTTGCTACAAGGG
ACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCAT
ATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTA
GGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTC
TGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGA
AAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGG
GGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTC
AGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAA
ATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATAC
AGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGA
AGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGG
GACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAA
AGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAG
CAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCC
AGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAAC
TCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCT
GGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCA
AGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGG
CTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATA
GTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGC
CCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGG
TATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCA
ACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTATCGATGGTCGAGTACCGGGTA

GGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTG
GCCTCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCGCCA
CCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGT
AGCACGTCTCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGG
CCAATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGG
CGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTC
TGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCTCTAGGATCCATGGACTACAAAGACCATGACGGTG
ATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGGGATCCATGCCGTCTGAGGGTCGCTGCTGGGA
GACCTTGAAGGCCCTACGCAGTTCCGACAAAGGTCGCCTTTGCTACTACCGCGACTGGCTGCTGCGGCGCGAGGAT
GTTTTAGAAGAATGTATGTCTCTTCCCAAGCTATCTTCTTATTCTGGATGGGTGGTAGAGCACGTCCTACCCCATATG
CAGGAGAACCAACCTCTGTCTGAGACTTCGCCATCCTCTACGTCAGCTTCAGCCCTAGATCAACCCTCATTTGTTCCC
AAATCTCCTGACGCAAGCTCTGCCTTTTCCCCAGCCTCCCCTGCAACACCAAATGGAACCAAGGGCAAAGATGAGTC
CCAGCACACAGAATCTATGGTACTTCAGTCCTCACGGGGGATCAAAGTGGAAGGCTGCGTCCGAATGTACGAACTG
GTACACAGAATGAAAGGAACAGAGGGCCTGAGGCTATGGCAGGAGGAGCAGGAGAGGAAGGTGCAAGCCCTCTC
GGAGATGGCATCTGAACAACTGAAGCGGTTTGATGAATGGAAGGAACTGAAGCAGCATAAAGAATTCCAGGACTT
GCGGGAAGTAATGGAGAAGAGCTCCAGAGAAGCCTTGGGACACCAAGAGAAGCTAAAAGCTGAGCACCGTCACA
GAGCAAAGATTCTCAACCTGAAGCTGCGGGAAGCAGAGCAGCAGCGCGTGAAGCAAGCAGAACAGGAGCGGCTT
CGGAAGGAAGAAGGCCAGATCCGCCTGCGGGCCCTCTATGCTCTGCAGGAGGAGATGCTGCAGCTCAGCCAGCAG
CTGGATGCCTCTGAGCAGCACAAAGCCCTGCTTAAGGTCGACCTGGCTGCCTTCCAGACCCGAGGCAACCAGCTGT
GCAGCCTCATCTCAGGGATCATCCGGGCCTCTTCAGAGAGCAGCTATCCCACAGCAGAGAGTCAAGCTGAGGCTGA
GCGAGCTCTGCGGGAAATGCGGGACCTCCTGATGAACTTGGGGCAGGAGATCACCAGAGCCTGCGAAGACAAGA
GGAGGCAGGATGAAGAAGAGGCCCAGGTAAAGCTGCAAGAGGCACAGATGCAGCAGGGACCAGAGGCCCACAA
AGAGCCCCCAGCTCCCAGCCAGGGCCCAGGAGGGAAACAGAATGAAGACCTCCAGGTGAAGGTACAAGACATTAC
AATGCAGTGGTACCAGCAGCTGCAGGATGCTTCCATGCAGTGTGTGTTGACCTTTGAGGGCCTGACCAACAGCAAG
GACAGTCAGGCCAAAAAGATAAAGATGGACCTCCAGAAGGCTGCTACCATCCCAGTGAGCCAAATCTCTACCATTG
CAGGCTCAAAACTGAAGGAGATCTTTGACAAGATCCACAGCCTGCTCTCTGGAAAACCTGTTCAATCTGGTGGGCG
CTCTGTGTCTGTCACACTTAACCCACAGGGGCTGGACTTTGTTCAATACAAACTGGCAGAGAAATTTGTGAAACAAG
GCGAGGAGGAAGTGGCCTCTCACCATGAAGCAGCATTCCCCATTGCAGTTGTGGCATCCGGGATCTGGGAGCTCCA
CCCCAGAGTGGGGGACCTCATTCTTGCTCATCTACATAAGAAGTGTCCTTACTCTGTTCCTTTCTATCCCACTTTCAAG
GAGGGAATGGCTTTGGAAGACTATCAGAGGATGCTTGGTTACCAAGTAAAGGATTCCAAAGTGGAGCAGCAAGAC
AACTTTCTAAAACGCATGTCAGGGATGATCCGTCTCTACGCTGCTATCATCCAGCTCCGGTGGCCATATGGAAACCG
ACAGGAGATTCACCCTCATGGCTTAAATCATGGATGGCGCTGGTTGGCACAGATCTTAAACATGGAGCCCTTGTCA
GATGTGACAGCCACCCTCCTCTTTGACTTCCTGGAGGTGTGTGGGAATGCCCTCATGAAGCAATACCAGGTTCAGTT
CTGGAAGATGCTAATTCTCATCAAAGAGGACTACTTTCCCAGAATTGAAGCTATCACAAGCTCAGGACAGATGGGC
TCCTTCATACGCCTCAAGCAGTTCTTGGAGAAATGTTTGCAACACAAGGACATTCCTGTCCCCAAGGGCTTTCTGACT
TCCTCCTTCTGGCGCTCCTGACTCGAGGGAATTCCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACT
GGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCC
GTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGC
AACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTT
TCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGG
GGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGT
GTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGC
GGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCATCGGGAATTCGAGCTCGGTACCTTTAAAACCAATGACTTACAAGGCAGCTGTAAATCTTAGCCACTTTT
TAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAAACAAAATCTGCTTTTTGCTTGTACTGGGTC
TCTCTGGTTAGACCAAATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCT
TGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAAACCCTTTTAGT
CAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCCGTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGAT
ACATAAGGTTATGTATTAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTGGCTTACT
GAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTTTGGTTGGACGAACCTTCTGAGTTTCTGGTA
ACGCCGTCCCGCACCCGGAAATGGTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACG
CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGA
TCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACT
GTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCT

TCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
TTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAG
ACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAG
GGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCC
TGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTT
GCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGC
ACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTC
GCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA
GTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG
GACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTA
ACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCAC
TTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATC
ATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGG
ATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTC
ATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATG
ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGA
TCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGC
CGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGG
GCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTG
AGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA
GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC
TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACG
GTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACC
GCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA
AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGTGTGTCAGTT
AGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC
CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTT
TTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
AGGCTTTTGCAAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCACTTTATCCCGCGTCAGG
GAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGAAATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCG
CAACCTATTTTCCCCTCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAAAAATACAT
CGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGC
ATTATTGCCGTAAGCCGTGGCGGTCTGTACCGGGTGCGTTACTGGCGCGTGA

LV GFP Plasmid Sequence:

TGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCT
GATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAG
TTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGA
TGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGC
TGCATCCGGAGTACTTCAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAG
GCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGG
GTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAA
GCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTT
AGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCG
ACGCAGGACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAAT
TTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGC
GATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCAGG

GAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTAC
AACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAA
AGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGC
ACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAAT
ATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAA
AGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATG
ACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGG
CGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATA
CCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATG
CTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAAT
TACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAAT
TAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATA
GTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACC
ATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGA
GAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGGTTAACTTTTAAAAGAAAAGGGGG
GATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAA
ACAAATTACAAAAATTCAAAATTTTATCGATGGTCGAGTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGG
AGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCTGGCCTCGCACACATTCCACATCCA
CCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCC
CCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTCTCACTAGTCTCGTGCAGATGGA
CAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCTTTCTGG
GCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCGGGGCGGGCGCCCGA
AGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCG
GGCCTTTCGACCTCTAGGATCCACCGGTCGCCACCatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgag
ctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccac
cggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgactt
cttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagg
gcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagcca
caacgtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgac
cactaccagcagaacaccccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccctgagcaaagaccccaacg
agaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtaaagcggccgcCTCGAGG
GAATTCCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTAC
GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATA
AATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCT
GACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATT
GCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCG
TGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCC
TTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCG
CGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCGGGAATTCGAGCTCGGTAC
CTTTAAAACCAATGACTTACAAGGCAGCTGTAAATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCT
AATTCACTCCCAACGAAAACAAAATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAAATCTGAGCCTGGGA
GCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCC
GTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAAACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAG
AATTAATTCCGTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTATTAATTGTAGCCGC
GTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTGGCTTACTGAAGCAGACCCTATCATCTCTCTCGTAAA
CTGCCGTCAGAGTCGGTTTGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATGGTCAG
CGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACA
GGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGC
GCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCAT
TCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGA
GCGTCGATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACC
CGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC
ATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGT
TAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTT

TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA
CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCT
CAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTAT
GTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTT
GGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATA
ACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGC
ACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC
GTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCC
CGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT
GGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAG
ATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTT
CATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCG
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGC
TTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGT
AACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACT
CTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTA
CCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG
AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGG
GGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTC
AGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCT
CACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCT
CCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCC
CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGC
AGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCC
TAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCT
CGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGGACACAA
GACAGGCTTGCGAGATATGTTTGAGAATACCACTTTATCCCGCGTCAGGGAGAGGCAGTGCGTAAAAAGACGCGG
ACTCATGTGAAATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCCTCGAACACTTTTTA
AGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAAAAATACATCGTCACCTGGGACATGTTGCAGATCCAT
GCACGTAAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGTCTGT
ACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTTTGTATTTCCAGCTACGATCACGACA
ACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGG
ATACCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAACCGGCT
GGTCGTCCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCG
TCGTATTCGTCCCGCCAATCTCCGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAACTTCAG
GCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGCATCCATGACACAGGCAAACCTGAGCGAAACCCTGTT
CAAACCCCGCTTTAAACATCCTGAAACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTGTGCAGT
CGGCCCTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACCGTCTGATGTGGATCTGGCGCGGCATTGAC
CCACGCGAAATCCTCGACGTCCAGGCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATA
CGGTGATTGGCTACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTGTGAAGGAACCTTACTTCTGTG
GTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGT
GTTAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAA
TGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAAC
ATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGT
CATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAG
AAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTC
CACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAAGGGG
TTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGCTTTTACTT
GCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATT
GCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGT

TGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTGTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCC
AAACTTCCCACCCCATACCCTATTACCACTGCCAAATTACCTGTGGTTTCATTTACTCTAAACCTGTGATTCCTCTGAA
TTATTTTCATTTTAAAGAAATTGTATTTGTTAAATATGTACTACAAACTTAGTAGT

LV FLAG Gle1b Plasmid Sequence

ACTGGGTATTCGTCATGTCGATACCGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGA
AACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACTGCGGTTGCGA
TTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTAT
GTTGTTGATATCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATCTCCGG
TCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAA
GTATTCTGGAGGCTGCATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACATCCTGAA
ACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTGTGCAGTCGGCCCTTGATGGTAAAACCATCCC
TCACTGGTATCGCATGATTAACCGTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAG
GCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACGGTGATTGGCTACCGTGGCGGCA
ACTGGATTTATGAGTGGGCCCCGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACC
TACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGTTTG
TGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTT
GCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGA
AAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTT
GCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAAC
CTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATT
AATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGT
GCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGCTTTTACTTGCTTTAAAAAACCTCCCACACCTCCC
CCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA
TCTTATCATGTGTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAACTTCCCACCCCATACCCTATTACC
ACTGCCAAATTACCTGTGGTTTCATTTACTCTAAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTT
GTTAAATATGTACTACAAACTTAGTAGTTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTG
GATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCT
TTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCT
TGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCC
TAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGATATCGAGCTTGCTACAAGGG
ACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCAT
ATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTA
GGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTC
TGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGA
AAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGG
GGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTC
AGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAA
ATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATAC
AGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGA
AGAGCAAAACAAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGG
GACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAA
AGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAG
CAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCC
AGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAAC
TCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGATTGGAATCACACGACCT
GGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCA
AGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGG
CTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATA
GTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGC

CCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGG
TATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCA
ACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTATCGATGGTCGAGTACCGGGTA
GGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTG
GCCTCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGGCCCCTTCGCGCCA
CCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGT
AGCACGTCTCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCAGCGG
CCAATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGG
CGGGCTCAGGGGCGGGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAAGCGCACGTC
TGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCTCTAGGATCCATGGACTACAAAGACCATGACGGTG
ATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGGGATCCATGCCGTCTGAGGGTCGCTGCTGGGA
GACCTTGAAGGCCCTACGCAGTTCCGACAAAGGTCGCCTTTGCTACTACCGCGACTGGCTGCTGCGGCGCGAGGAT
GTTTTAGAAGAATGTATGTCTCTTCCCAAGCTATCTTCTTATTCTGGATGGGTGGTAGAGCACGTCCTACCCCATATG
CAGGAGAACCAACCTCTGTCTGAGACTTCGCCATCCTCTACGTCAGCTTCAGCCCTAGATCAACCCTCATTTGTTCCC
AAATCTCCTGACGCAAGCTCTGCCTTTTCCCCAGCCTCCCCTGCAACACCAAATGGAACCAAGGGCAAAGATGAGTC
CCAGCACACAGAATCTATGGTACTTCAGTCCTCACGGGGGATCAAAGTGGAAGGCTGCGTCCGAATGTACGAACTG
GTACACAGAATGAAAGGAACAGAGGGCCTGAGGCTATGGCAGGAGGAGCAGGAGAGGAAGGTGCAAGCCCTCTC
GGAGATGGCATCTGAACAACTGAAGCGGTTTGATGAATGGAAGGAACTGAAGCAGCATAAAGAATTCCAGGACTT
GCGGGAAGTAATGGAGAAGAGCTCCAGAGAAGCCTTGGGACACCAAGAGAAGCTAAAAGCTGAGCACCGTCACA
GAGCAAAGATTCTCAACCTGAAGCTGCGGGAAGCAGAGCAGCAGCGCGTGAAGCAAGCAGAACAGGAGCGGCTT
CGGAAGGAAGAAGGCCAGATCCGCCTGCGGGCCCTCTATGCTCTGCAGGAGGAGATGCTGCAGCTCAGCCAGCAG
CTGGATGCCTCTGAGCAGCACAAAGCCCTGCTTAAGGTCGACCTGGCTGCCTTCCAGACCCGAGGCAACCAGCTGT
GCAGCCTCATCTCAGGGATCATCCGGGCCTCTTCAGAGAGCAGCTATCCCACAGCAGAGAGTCAAGCTGAGGCTGA
GCGAGCTCTGCGGGAAATGCGGGACCTCCTGATGAACTTGGGGCAGGAGATCACCAGAGCCTGCGAAGACAAGA
GGAGGCAGGATGAAGAAGAGGCCCAGGTAAAGCTGCAAGAGGCACAGATGCAGCAGGGACCAGAGGCCCACAA
AGAGCCCCCAGCTCCCAGCCAGGGCCCAGGAGGGAAACAGAATGAAGACCTCCAGGTGAAGGTACAAGACATTAC
AATGCAGTGGTACCAGCAGCTGCAGGATGCTTCCATGCAGTGTGTGTTGACCTTTGAGGGCCTGACCAACAGCAAG
GACAGTCAGGCCAAAAAGATAAAGATGGACCTCCAGAAGGCTGCTACCATCCCAGTGAGCCAAATCTCTACCATTG
CAGGCTCAAAACTGAAGGAGATCTTTGACAAGATCCACAGCCTGCTCTCTGGAAAACCTGTTCAATCTGGTGGGCG
CTCTGTGTCTGTCACACTTAACCCACAGGGGCTGGACTTTGTTCAATACAAACTGGCAGAGAAATTTGTGAAACAAG
GCGAGGAGGAAGTGGCCTCTCACCATGAAGCAGCATTCCCCATTGCAGTTGTGGCATCCGGGATCTGGGAGCTCCA
CCCCAGAGTGGGGGACCTCATTCTTGCTCATCTACATAAGAAGTGTCCTTACTCTGTTCCTTTCTATCCCACTTTCAAG
GAGGGAATGGCTTTGGAAGACTATCAGAGGATGCTTGGTTACCAAGTAAAGGATTCCAAAGTGGAGCAGCAAGAC
AACTTTCTAAAACGCATGTCAGGGATGATCCGTCTCTACGCTGCTATCATCCAGCTCCGGTGGCCATATGGAAACCG
ACAGGAGATTCACCCTCATGGCTTAAATCATGGATGGCGCTGGTTGGCACAGATCTTAAACATGGAGCCCTTGTCA
GATGTGACAGCCACCCTCCTCTTTGACTTCCTGGAGGTGTGTGGGAATGCCCTCATGAAGCAATACCAGGTTCAGTT
CTGGAAGATGCTAATTCTCATCAAAGAGGACTACTTTCCCAGAATTGAAGCTATCACAAGCTCAGGACAGATGGGC
TCCTTCATACGCCTCAAGCAGTTCTTGGAGAAATGTTTGCAACACAAGGACATTCCTGTCCCCAAGGGCTTTCTGACT
TCCTCCTTCTGGCGCTCCTGACTCGAGGGAATTCCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACT
GGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCC
GTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGC
AACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTT
TCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGG
GGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGT
GTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGC
GGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCATCGGGAATTCGAGCTCGGTACCTTTAAAACCAATGACTTACAAGGCAGCTGTAAATCTTAGCCACTTTT
TAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAAACAAAATCTGCTTTTTGCTTGTACTGGGTC
TCTCTGGTTAGACCAAATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCT
TGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAAACCCTTTTAGT
CAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCCGTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGAT
ACATAAGGTTATGTATTAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTGGCTTACT
GAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTTTGGTTGGACGAACCTTCTGAGTTTCTGGTA
ACGCCGTCCCGCACCCGGAAATGGTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACG

CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGA
TCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACT
GTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCT
TCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAG
TTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAG
ACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAG
GGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCC
TGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTT
GCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGC
ACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTC
GCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACA
GTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG
GACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTA
ACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCAC
TTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATC
ATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGG
ATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTC
ATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATG
ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGA
TCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGC
CGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTG
CTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGG
GCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTG
AGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA
GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC
TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACG
GTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACC
GCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA
AGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGTGTGTCAGTT
AGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAG
GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC
CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTT
TTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT
AGGCTTTTGCAAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCACTTTATCCCGCGTCAGG
GAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGAAATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCG
CAACCTATTTTCCCCTCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAAAAATACAT
CGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGC
ATTATTGCCGTAAGCCGTGGCGGTCTGTACCGGGTGCGTTACTGGCGCGTGA

AAV Gle1b Plasmid Sequence

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTC
GCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATACTAGGGGTTCCTGCGGCCGCAC
GCGTGGAGCTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACAT
AACTTACGGTAAATGGCCCGCCTGGCTACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCA
TAGTAACGTCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT
AAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC
ATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATGCTATTACCATGGTGATGCGGTTTTGGC
AGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG
TTTGTTTTGCACCAAAATCAACGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGG
CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACCTG

TTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGATTCGAATCCCGGCCGGGAACGGTGCATTGG
AACGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTTATAGGCCCACAAAAAATGCTTTCTTCTT
TTAATATACTTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTTCAGGGCAATAATGATACAATGTATC
ATGCCTCTTTGCACCATTTAAAGAATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATTTCTGCATATAAATA
TTTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACATTCTGCTTTT
ATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATC
TTCCTCCCACAGCTCCTGGGCAACGTGCTGGTCTGTGGCTGGCCCATCACTTTGGCAAAGAATTGGGATTCGAACAT
CGATTGAATTCCCCGGGGATCCctagtaacggccgccagtgtgctggaattcatgccgtctgagggtcgctgctgggagaccttgaaggccct
acgcagttccgacaaaggtcgcctttgctactaccgcgactggctgctgcggcgcgaggatgttttagaagaatgtatgtctcttcccaagctatcttctta
ttctggatgggtggtagagcacgtcctaccccatatgcaggagaaccaacctctgtctgagacttcgccatcctctacgtcagcttcagccctagatcaac
cctcatttgttcccaaatctcctgacgcaagctctgccttttccccagcctcccctgcaacaccaaatggaaccaagggcaaagatgagtcccagcacac
agaatctatggtacttcagtcctcacgggggatcaaagtggaaggctgcgtccgaatgtacgaactggtacacagaatgaaaggaacagagggcctg
aggctatggcaggaggagcaggagaggaaggtgcaagccctctcggagatggcatctgaacaactgaagcggtttgatgaatggaaggaactgaag
cagcataaagaattccaggacttgcgggaagtaatggagaagagctccagagaagccttgggacaccaagagaagctaaaagctgagcaccgtcac
agagcaaagattctcaacctgaagctgcgggaagcagagcagcagcgcgtgaagcaagcagaacaggagcggcttcggaaggaagaaggccagat
ccgcctgcgggccctctatgctctgcaggaggagatgctgcagctcagccagcagctggatgcctctgagcagcacaaagccctgcttaaggtcgacct
ggctgccttccagacccgaggcaaccagctgtgcagcctcatctcagggatcatccgggcctcttcagagagcagctatcccacagcagagagtcaag
ctgaggctgagcgagctctgcgggaaatgcgggacctcctgatgaacttggggcaggagatcaccagagcctgcgaagacaagaggaggcaggatg
aagaagaggcccaggtaaagctgcaagaggcacagatgcagcagggaccagaggcccacaaagagcccccagctcccagccagggcccaggagg
gaaacagaatgaagacctccaggtgaaggtacaagacattacaatgcagtggtaccagcagctgcaggatgcttccatgcagtgtgtgttgacctttga
gggcctgaccaacagcaaggacagtcaggccaaaaagataaagatggacctccagaaggctgctaccatcccagtgagccaaatctctaccattgca
ggctcaaaactgaaggagatctttgacaagatccacagcctgctctctggaaaacctgttcaatctggtgggcgctctgtgtctgtcacacttaacccac
aggggctggactttgttcaatacaaactggcagagaaatttgtgaaacaaggcgaggaggaagtggcctctcaccatgaagcagcattccccattgca
gttgtggcatccgggatctgggagctccaccccagagtgggggacctcattcttgctcatctacataagaagtgtccttactctgttcctttctatcccactt
tcaaggagggaatggctttggaagactatcagaggatgcttggttaccaagtaaaggattccaaagtggagcagcaagacaactttctaaaacgcatg
tcagggatgatccgtctctacgctgctatcatccagctccggtggccatatggaaaccgacaggagattcaccctcatggcttaaatcatggatggcgct
ggttggcacagatcttaaacatggagcccttgtcagatgtgacagccaccctcctctttgacttcctggaggtgtgtgggaatgccctcatgaagcaatac
caggttcagttctggaagatgctaattctcatcaaagaggactactttcccagaattgaagctatcacaagctcaggacagatgggctccttcatacgcc
tcaagcagttcttggagaaatgtttgcaacacaaggacattcctgtccccaagggctttctgacttcctccttctggcgctcctgatgtcactccatcaccc
accatcaccgctgctgcaaagaggcaataataaaggaactgaagacagctgtattgggagaagtcatgtcatccatcacactggcggccgCTCGAG
AGATCTACGGGTGGCATCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCA
GCCTTGTCCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGAGGGGGG
TGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGT
GCAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGT
TGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCACCATATTGGCC
AGGCTGGTCTCCAACTCTAATCTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACC
ACTGCTCCCTTCCCTGTCCTTCTGATTTTGTAGGTAACCACGTGCGGACCGAGCGGCCGCAGGAACCCTAGTGATGG
AGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTT
TGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCGCAGGGGCGCCTGATGCGGTATTTTCTCCTTACG
CATCTGTGCGGTATTTCACACCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGC
GGGTGTGGTGGTTACGCCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCC
TTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTAGTGCTTTAC
GGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGC
CCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGC
TATTCTTTTGATTTATAAGGGATTTTGCCGATTTGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACG
CGAATTTTAACAAAATATTAACGTTTACAATTTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCAGCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAA
GCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAACGCGCGAGACGAAAGGGCCT
CGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAA
ATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATA
AATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGC
ATTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGT
GGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCGAAGAACGTTTTCCAATGATGA
GCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA

CACTATTCTCAGAATGACTTGGTTGGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGA
ATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAG
AGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGC
CATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACTTGCGCAAACTATTAACTGGCGAAC
TACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG
GCCCTTCCGGCTGGCTGGTTTATTGTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACT
GGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
TGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAGACAGATCGCTGAGATAGGTGCCTCACTGAT
TAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGG
ATCTAGGTGAAGATCCTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACC
CCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAAAAAAAACCAC
CGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCG
CAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGAAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCT
ACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAGGCGGACAGGTA
TCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATA
GTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAA
AACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT

Outcomes of PhD Programme

Conferences

- **Oral Presentations**
    - Northeast Postgraduate Conference (NEPG), Newcastle UK 2018: "Investigating the role of the RNA exporter protein Gle1 in motor neuron diseases" Emily Graves, Eva Karyka, Laura Ferraiuolo, Guillaume Hautbergue, Mimoun Azzouz.
- **Poster Presentations**
    - NEPG, Newcastle 2019: "Investigating the role of the RNA exporter protein Gle1 in motor neuron diseases" Emily Graves, Eva Karyka, Laura Ferraiuolo, Guillaume Hautbergue, Mimoun Azzouz.
    - Medical School Research Day 2019, Sheffield: "Investigating the role of the RNA exporter protein Gle1 in motor neuron diseases" Emily Graves, Eva Karyka, Laura Ferraiuolo, Guillaume Hautbergue, Mimoun Azzouz.
    - DiMeN Knowledge Exchange 2019, Sheffield: "Investigating the role of the RNA exporter protein Gle1 in motor neuron diseases" Emily Graves, Eva Karyka, Laura Ferraiuolo, Guillaume Hautbergue, Mimoun Azzouz.
- **Conferences Attended**
    - American Society of Gene and Cell Therapy (ASGCT) conference online 2020
    - SMA Research Symposium online 2020
- **Conferences Organised/Chaired**
    - Member of NEPG organising committee 2020,
        - Chaired Science Communication workshop and Neuroscience Oral Presentation session

Outreach

- I'm a Scientist Online 2020
- Tutor for the Brilliant Club 2020/2021
- MRC Festival 2018/2019
- Writer for PreLights 2019

Outside Training/Opportunities

- BioTech YES competition 2018

- 4 month internship at CamBioScience, 2020.

Publications

1. Evangelia Karyka, Nelly Berrueta Ramirez, Christopher P. Webster, Paolo M. Marchi, **Emily J. Graves,** Vinay K. Godena, Lara Marrone, Anushka Bhargava, Swagat Ray, Ke Ning, Guillaume M. Hautbergue, Sherif F. El-Khamisy and Mimoun Azzouz. SMN-deficient cells exhibit increased ribosomal DNA damage *Life Science Alliance* April 2022
2. Tobias Moll, **Emily Graves**, John Franklin, Adrian Higginbottom, Mimoun Azzouz, Johnathan Cooper-Knock & Pamela J Shaw ALS-associated GLT8D1 mutations cause loss of motor neurons via disrupted neurotrophic signalling (In Preparation)
3. **Emily J Graves**, Matthew A Roach, Mimoun Azzouz. What's new in gene therapy for ALS? *Expert Opinion in Biological Research* (In Preparation)