Adam Mickiewicz University in Poznan
Faculty of Biology
Institute of Molecular Biology and Biotechnology
Department of Gene Expression
Laboratory of RNA Processing

Doctoral thesis

The effect of ALS-associated FUS mutations on U7 snRNP activity and the expression of core canonical histone genes in neuronal cells

Ankur Gadgil

Promotor: prof. UAM, dr. hab. Katarzyna Dorota Raczyńska

Poznan, Poland 2022

ACKNOWLEDGEMENTS

With all of my heart, I would like to thank:

Professor UAM, dr. hab. Katarzyna Dorota Raczyńska, for giving me the opportunity to be her student and to experience a truly exceptional atmosphere of both hard work and great joy in making scientific discoveries; for her outstanding support, understanding and inspiration to excel.

Professor Zofia Szweykowska-Kulińska and Professor Artur Jarmołowski for providing the laboratory work space in the initial years of my studies, for their constructive discussions, criticism and overall support.

I would also like to thank Dr. Izabela Sierocka being one of the interview panelists and providing positive assessment of my experience and capabilities. Importantly, I want to thank all my friends, colleagues, co-workers, professors, and administration support staff from the Adam Mickiewicz University, especially to the people from the Laboratory of RNA processing and the Department of Gene Expression for their invaluable help, scientific as well as social support, encouragement, and inspiration.

All members of our collaborators group of Dr. Marc-David Ruepp at the King's College London and group of Dr. Leena Latonen at the University of Eastern Finland.

Basia Marshałek, her family and all my friends in Poznań who all played a pivotal role for my integration into the Polish culture.

My family and my wife for being with me always.

Dedicated to my wife

CONTENT

STRESZCZENIE

Stwardnienie zanikowe boczne (ang. ALS, amyotrophic lateral sclerosis) to choroba neurodegeneracyjna, która polega na postępującej utracie górnych i dolnych neuronów ruchowych układu nerwowego. W 90 do 95% przypadków ALS jest klasyfikowany jako sporadyczny ALS (sALS), autosomalna dominująca postać rodzinnego ALS (fALS) stanowi pozostałe 5 do 10% przypadków. Za postać fALS odpowiadają mutacje w różnych genach, między innymi w genie *FUS*. FUS jest białkiem jądrowym podlegającym konstytutywnej ekspresji, zaangażowanym w naprawę DNA i regulację transkrypcji, splicing RNA i eksport RNA do cytoplazmy.

Mutacje FUS związane z ALS (ALS-FUS) występują głównie w zachowawczym regionie końca C białka i prowadzą do nieprawidłowej lokalizacji białka i tworzenia cytoplazmatycznych agregatów. Wcześniej wykazaliśmy, że białko FUS oddziałuje z cząstką U7 snRNP i bierze udział w regulacji transkrypcji i wydajności dojrzewaniu końca 3' genów histonów zależnych od replikacji (RDH, ang. replication dependant histones).

Bazując na dostępnych danych i naszych wcześniejszych obserwacjach, za cel pracy doktorskiej obrałem analizę wpływu mutacji ALS-FUS na aktywność U7 snRNP oraz efektywność transkrypcji i dojrzewanie pre-mRNA kanonicznych histonów rdzeniowych, jako mechanizmu molekularnego leżącego u podstaw ALS. Wyniki, jakie uzyskałem na modelu komórkowym i pierwotnych neuronów szczurzych wskazują, że mutacje ALS-FUS prowadzą do błędnej lokalizacji w agregatach cytoplazmatycznych wraz z białkiem FUS również cząstki U7 snRNP. Ta cytoplazmatyczna sekwestracja ALS-FUS wraz z U7 snRNP ma następnie znaczący wpływ na aktywność transkrypcyjną i nieprawidłowe dojrzewanie końca 3' pre-mRNA RDH. W proliferujących komórkach neuroblastomy transfekowanych mutantami ALS-FUS zaobserwowałem zahamowanie transkrypcji genów RDH i upośledzenie dojrzewania końca 3' ich pre-mRNA. Jednocześnie, w terminalnie zróżnicowanych komórkach nie zaobserwowałem zaburzenia dojrzewania końca 3' pre-mRNA, ale znaczne obniżenie poziomu transkryptów z uwagi na zahamowaną transkrypcję.

Uzyskane wyniki wskazują, że cząstka U7 snRNP jest jedną z cząstek snRNP, na których aktywność wpływają mutacje ALS-FUS, w efekcie prowadząc do szeregu zaburzeń, od nieefektywnego splicingu po zaburzoną obróbkę transkryptów RDH. Niewątpliwie, istnieje jeszcze więcej parametrów, które nie zostały wciąż zbadane. Niemniej, podsumowując dostępne dane, zwiększenie uszkodzeń DNA i zakłócone dojrzewanie pre-mRNA genów RDH spowodowane mutacjami ALS-FUS może prowadzić do niestabilności genomu i może być mechanizmem molekularnym leżącym u podstaw ALS.

Słowa kluczowe – ALS, FUS, U7 snRNP, histony zależne od replikacji

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder which involves the progressive loss of upper as well as the lower motor neurons of the nervous system. ALS is usually categorized as sporadic ALS (sALS) which constitutes for about 90 to 95% of the cases and autosomal dominant familial ALS (fALS) comprising of the remaining 5 to 10% cases. Different genes involving varied mutations have been discovered to be involved in fALS with mutations in *FUS* gene being one of them. FUS is a ubiquitously expressed predominantly nuclear protein involved in DNA repair and transcription regulation, RNA splicing and export to cytoplasm.

The ALS-linked FUS mutations (ALS-FUS) are mainly observed in the conserved C terminus region and are associated with mislocalization and cytoplasmic inclusion formation. We have previously shown that the FUS protein interacts with U7 snRNP and is involved in the regulation of transcription and 3' end processing efficiency of the replication-dependent histone (RDH) genes.

Based on the available data and our previous observation, the aim of the thesis was to analyze the effect of ALS-FUS mutations on U7 snRNP activity and the efficiency of transcription and processing of core canonical histone pre-mRNAs as the molecular mechanism underlying ALS. From the data obtained from our experimental results, it shows that ALS-FUS mutations along with itself mislocalize U7 snRNP in the cytoplasmic aggregates of cellular models and rat primary neurons. This cytoplasmic entrapment of ALS-FUS along with U7 snRNP has a significant impact on the transcriptional activity and aberrant 3' end processing of RDH pre-mRNAs. In proliferating neuroblastoma cells transfected with ALS-FUS mutants we observed inhibition of RDH gene transcription and impairment of the 3' end pre-mRNA maturation. At the same time, in terminally differentiated cells we observed no obvious impairment in the 3' end maturation but significant downregulation of transcript level due to inhibited transcription.

The obtained results indicate U7 snRNP is one of the snRNP whose activity is affected by ALS-FUS mutations resulting in a range of complications from inefficient splicing to disturbed RDH transcripts processing. Undoubtedly, there are even more parameters which are yet to be explored. But summarizing the available data, increased DNA damage and disrupted RDH pre-mRNA processing caused due to ALS-FUS mutations can cause genomic instability and may be the molecular mechanisms underlying in ALS.

Keywords – ALS, FUS, U7 snRNP, replication-dependent histones

LIST OF SCIENTIFIC WORKS INCLUDED IN THE DISSERTATION

The results of the experimental works are described in the following paper:

1. **Gadgil A**., Walczak A., Stępień A., Mechtersheimer J., Nishimura A.L., Shaw C.E., Ruepp M.D., Raczyńska K.D. (2021) ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells. *Scientific Reports*, 2021; 11(1):11868, doi:10.1038/s41598-021-91453-3
   PMID: 34088960, PMCID: PMC8178370
   Ministry points (MNiSW) - 100
   Impact factor (2021) – 4,379

In addition, the following review paper concerning U7 snRNA and its role in gene therapy is published.

2. **Gadgil A**., Raczyńska K.D. U7 snRNA: A tool for Gene therapy. *Journal of Gene Medicine* 2021;23(4): e3321, doi:10.1002/jgm.3321
   PMID: 33590603, PMCID: PMC8243935
   Ministry points (MNiSW) - 40
   Impact factor (2021) – 4,565

ABBREVIATIONS

RDH	–	Replication dependent histone
FUS	–	Fused in sarcoma
ALS	–	Amyotrophic lateral sclerosis
ALS-FUS	–	ALS linked FUS mutations
R495X	–	Truncation after arginine amino acid at position 495 in FUS amino acid chain
P525L	–	Proline to Leucine substitution at position 525 in FUS amino acid chain
WT	–	Wild type
FUS KO	–	FUS knockout
U snRNP	–	Uridine rich small nuclear ribonucleoprotein
snRNA	–	Small nuclear ribonucleic acid
fALS	–	Familial amyotrophic lateral sclerosis
sALS	–	Sporadic amyotrophic lateral sclerosis
hiPSC	–	Human induced pluripotent stem cells
SH-SY5Y	–	Neuroblastoma cell line
UTR	–	Untranslated region
CDS	–	Coding DNA sequence
SIC	–	Spreading initiation center
RNAP II	–	RNA polymerase II

Important achievements during PhD studies

1. Received the Włodziemierz Mozolowski award for oral presentation (young scientist under 30 years) at the international conference 'Congress BIO 2018' organized by the Polish Biochemical Society, Gdańsk, Sep 2018.
2. Recipient of the EMBO scientific exchange grant #8906. (Formerly known as EMBO short-term fellowship). During which I visited the University of Eastern Finland for a duration of 90 days. 04 January 2021 to 03 April 2021. The visit proved beneficial in establishing a new collaboration and extend the observations of my PhD studies to prostate cancer research.

Funding

This work was supported by the following sources

1. The Polish National Science Centre grant, NCN OPUS 2016/21/B/NZ1/00232 (to KDR).
2. The Polish National Science Centre grant, NCN SONATA BIS 2018/30/E/NZ2/00295 (to KDR)
3. The European Union: Passport to the future - Interdisciplinary doctoral studies at the Faculty of Biology, Adam Mickiewicz University. POWR.03.02.00-00-I022/16.
4. Faculty of Biology's 'Dean's grant', GDWB-01/2019.
5. EMBO scientific exchange grant, number 8906 (previously known as EMBO Short-term fellowship).
6. ID-UB grant number 003/13/UAM/0013.

Scientific collaboration

The experiments were performed at the Laboratory of RNA Processing, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland, in collaboration with:

1. Group of Dr. Marc-David Ruepp UK Dementia Research Institute Centre at King's College London, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK.

2. Group of Dr. Leena Latonen at the Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland.

Main thesis including results and summary

Aim of the study

In 2009, mutations in the *FUS* gene were identified in patients with an inherited form of amyotrophic lateral sclerosis. Mislocalization of this FUS protein is a hallmark of ALS wherein ALS-linked mutations in *FUS* genes disrupt nuclear import and trap the protein in cytoplasmic aggregates. Mislocalized FUS mutants sequester RNA-binding proteins and U snRNPs. Our laboratory previously showed that endogenous FUS can interact with U7 snRNP. FUS protein is also involved in the transcription of replication-dependent histone genes and the 3' end processing of their pre-mRNAs along with U7 snRNP.

Thus, the main goal of the project is to analyze the effect of ALS-linked FUS mutations on U7 snRNP activity and the efficiency of transcription and processing of core canonical histone pre-mRNAs as the molecular mechanism underlying ALS.

Experimental workplan included:

1. To check the colocalization of ALS-linked FUS mutants and U7 snRNA.
2. To check the colocalization of ALS-linked FUS mutants and Lsm11 protein.
3. To check the effect of ALS-linked FUS mutations on the RNA polymerase II (RNAP II) occupancy on histone genes.
4. To check the effect of ALS-linked FUS mutations on the transcription of replication-dependent histone genes.
5. To check the effect of ALS-linked FUS mutations on the 3' end processing of replication-dependent histone genes.

Background and introduction

ALS is a progressive neuromuscular disease characterized by the progressive degenerative of upper and lower motor neurons in the brain and spinal cord, respectively. Early symptoms include muscle weakness or stiffness, muscle twitching in arm, leg, muscle cramps, slurred or thick speech and difficulty in projecting the voice, loss of motor control in the arms and legs leading to tripping and falling. As the disease progresses, the affected individuals gradually lose their strength, ability to speak, swallow, and commonly the cause of death being inability to breathe. Diagnosis is usually performed using muscle and imaging tests, like blood/urine tests, muscle/nerve biopsy, nerve conduction study, magnetic resonance imaging (MRI), electromyography. Currently, there is no available treatment or cure to halt or reverse the progression of the disease. Although, aid can be provided to control symptoms, prevent unnecessary complications and make living with the disease easier.

ALS is neurodegenerative disease reported worldwide irrespective of any race or ethnic background. The cause of ALS is still unknown with about 90-95% cases occurring sporadically at random with no clearly associated risk factors and no family history of the disease. However, the remaining 5-10% of the cases are familial wherein a clear genetic history exists [1,2]. More than 25 genes are known to be associated with ALS of which often most mutated are *C9orf72*, *SOD1*, *TARDBP* and *FUS* [3].

Notably, in 2009, mutations in the *FUS* gene were identified in patients with an inherited form of ALS [4,5]. Most of these mutations are missense mutations localized around the conserved C terminus region which contains the nuclear localization signal and lead to almost abolished nuclear import of FUS protein. Moreover, the C terminus contains RNA recognition motif, Arginine-Glycine (RGG) repeats and zinc finger domain involved in RNA processing.

The ALS-FUS mutants thus tend to aggregate in the cytoplasm and a liquid to solid phase transition is accelerated by the disease mutation(s) [6]. When aggregating in the cytoplasm, the ALS-FUS mutants sequester other RNA-binding proteins, FUS binding proteins (hnRNP A1, hnRNP A2, SMN), poly(A) mRNAs and U snRNA/U snRNP ((U1 snRNPs, U2 snRNPs, U1 snRNA, U11 snRNA and U12 snRNA) [19-23].

Our lab previously reported an important observation which provided valuable information for forming the basis of my PhD studies. We observed that the endogenous FUS protein acts as a positive regulator of replication-dependent histone gene transcription and the 3'end processing of their pre-mRNAs during the S phase of the cell cycle. Furthermore, we suggested that FUS might be involved in U7 snRNP-dependent repression of histone gene expression outside of S phase, thereby preventing the synthesis of extra histones that would be harmful to the cells [7]. U7 snRNP is one of the key factors involved in the processing of RDH pre-mRNAs. It consists of the U7 snRNA and five Sm proteins that are shared with spliceosomal snRNPs (SmB/B', SmD3, SmE, SmF and SmG) and two unique Sm-like proteins: Lsm10 and Lsm11 [8,9]. These proteins along with other proteins form a histone cleavage complex, recognize the histone downstream element and perform a single endonucleolytic cleavage to produce mature RDH mRNA. By introducing controlled changes at the histone binding sequence and the Sm motif of the U7 snRNA, it can be utilized as a tool for gene therapy [25].

Our preliminary FISH experiment showed that the ALS-FUS mutants traps U7 snRNA/snRNP in the cytoplasm. Therefore, I hypothesized that a consequence of the loss of the nuclear function of ALS-linked FUS mutants and U7 snRNP mislocalization, is deregulation of RDH gene expression which might be relevant to altered glial cells and motor neurons homeostasis in ALS. In glial cells, which do proliferate, decreased histone gene transcription and decreased efficiency of the 3'end processing of pre-mRNAs might be observed. On the other hand, in terminally differentiated neurons, FUS mutations can lead to destabilized transcription control mechanism that in turn might results in aberrant activation of histone proteins synthesis. Nevertheless, disturbed histone gene expression resulting in genome

instability or toxic effect of excess of histones may be the molecular mechanisms underlying altered glial cells and motor neurons homeostasis in ALS.

The experiments to identify the effect of ALS-linked FUS mutations on U7 snRNP activity and on transcription and the 3'end processing of replication-dependent histone genes were performed during the PhD studies and the observations are published as a research article in the *Scientific reports* in 2021 [24].

After discussing the role of U7 snRNP in diseases we also were interested to elaborate the therapeutic importance of U7 snRNP. The different unique aspects of U7 snRNA, its advantages and limitations are discussed in our review U7 snRNA: A tool for gene therapy.

Moreover, we have also studied the role of FUS and U7 snRNP in prostate cancer and have presented the data in this thesis.

The main thesis and achievements of the work

Gadgil A., Walczak A., Stępień A., Mechtersheimer J., Nishimura A.L., Shaw C.E., Ruepp M.D., Raczyńska K.D. (2021) ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells. *Scientific Reports,* 2021; 11(1):11868, doi:10.1038/s41598-021-91453-3

The experiments were carried out using HeLa and SH-SY5Y FUS knockout cell lines rescued by transient transfection with wild type FUS and with ALS-associated FUS mutants. One plasmid carrying the P525L mutation and the other plasmid with FUS R495X mutation were used for the experiments. Importantly, the SH-SY5Y cell line was used as proliferating cells as well as terminally differentiated neuron-like cells. For differentiation, the SY-SY5Y proliferating cells were treated using retinoic acid for a period of 10 days, confirmed for terminal differentiation, and then transfected. The aim behind this strategy was to explore if the effect of ALS-FUS mutations depends on the proliferation status of the cell.

Further, we also used human induced pluripotent stem cells (hiPSC's) derived from ALS patients harboring FUS R514G mutation to perform all the experiments performed on cell lines. The experiments were focused to observe the effect of ALS-FUS mutations on localization of U7 snRNP and their effect on 3'end processing of replication-dependent histone mRNAs. However, we weren't able to observe any significant effect as in cell lines. This was due to both heterozygous and mild nature of R514G mutation. Therefore, these results were not included in the publication. Unfortunately, due to extreme difficulty of obtaining patient samples we could not proceed further with other planned experiments, including patient samples with R495X or P525L mutations.

Even then keeping up with high standards of quality of data produced, we then performed the localization studies in wild-type rat brain primary neurons expressing the ALS-FUS protein by transient transfection. Here we observed similar and significant observations as in the cell lines, further supporting our hypothesis.

In summary, in the localization studies performed using FUS knockout cells lines of HeLa, SH-SY5Y and samples of rat brain primary neurons, it was evident that ALS-FUS mutants form aggregates and sequester U7 snRNA along with it into the cytoplasm. We were then interested to check if the mislocalization is limited to the U7 snRNA only or it also affects the complete U7 snRNP. To address this question, we analyzed the localization of Lsm11 protein in correlation with ALS-FUS. Based on the results, it was evident that the Lsm11 protein is also mislocalized into the cytoplasm along with mutant ALS-FUS protein. Thus, these consistent observations in multiple cellular models as well as rat primary neurons proved that the ALS-FUS protein mislocalize the whole U7 snRNP complex into the cytoplasm.

With the first two questions being addressed we were then interested to check the effect of ALS-linked FUS mutations on the RNAP2 occupancy on RDH genes, H2AC, H2BJ and H4J. This question was necessary to address, because in our previous publication [7] we reported that decreased levels of total histone transcripts correlate with weaker binding of FUS to histone promoters and diminished levels of RNAP2 on histone genes. It was thus necessary to investigate the effect of ALS-FUS mutations on the activation of RDH gene transcription. We performed ChIP assay to address this question, however, we performed it only for proliferating cells. We also attempted to perform it on differentiated neuron-like cells, but we had difficulties with isolation of the required amount of DNA and thus we could not proceed further. RNAP2 occupancy was analyzed on the 5'UTR, open reading frame (CDS) and the 3'UTR for the three RDH genes by qPCR. The data revealed that RNAP2 occupancy is decreased in cells transfected with ALS-FUS mutants compared to FUS WT in all three regions of the three analyzed histone genes. We thus interpret it that in proliferating cells the ALS-FUS mutations cause inhibition of the transcription of RDH genes due to weak mounting and binding of RNAP2 on histone genes.

After addressing initial three questions, we then moved to our next aim, which was to analyze the effect of ALS-FUS mutations on the transcription efficiency and 3' end processing activity of the RDH pre-mRNAs. Based on the previous publication from our lab in 2015, it was proven that FUS interacts with U7 snRNP and participates in influencing the transcription activity and the 3'end processing of RDH pre-mRNAs [7]. Therefore, knowing that the ALS-FUS mutants mislocalize the whole U7 snRNP complex into the cytoplasm, we expected significant effect on the RDH pre-mRNA 3'end processing as well as pre-mRNAs. We selected the RDH genes based on our previous publication [7]. The replication-independent histone gene H2A.Z which undergoes cleavage and polyadenylation, was used as a reference gene. From these experiments, we observed that compared to the cells transfected with WT FUS, the proliferating SH-SY5Y FUS KO cells transfected with the ALS-FUS mutants showed significantly affected processing efficiency in the majority of the analyzed histones. We interpret that this effect resulted from an elevated level of extended transcripts, with the total level of transcripts significantly downregulated. The lower level of total histone transcripts confirmed our previous observation of decreased transcription efficiency of RDH genes, analyzed by ChIP. Along with FUS KO cell lines, we transiently transfected WT FUS and ALS-FUS in proliferating SH-SY5Y WT and HeLa WT cells, but the effect we observed was

more prominent in the FUS KO cells transfected with the ALS-linked FUS mutants. This difference could be due to high expression of endogenous FUS playing the role of negative autoregulation in wild-type cells[24] combined with low transfection efficiency achieved in neuroblastoma cell.

In real life scenario, ALS affects the motor neurons which are terminally differentiated cells. Thus, to study the real-life situation as precisely as possible we differentiated SH-SY5Y cells into neuron-like cells and confirmed it using qPCR for cellular differentiation markers, expression of MAP2 protein using western blotting, and actin phalloidin staining assessed with immunofluorescence to confirm the morphology. Interestingly, in these neuron-like cells, we observed no change in the 3'end maturation, although we observed significant downregulation of 'processing efficiency' when we compared the ratio of total to polyadenylated transcripts. However, we hypothesized this could be due to inhibited transcription rather than affected the 3'end processing. It is previously shown that terminally differentiated cells synthesize only polyadenylated histone transcripts in the process that is independent of U7 snRNP [10]. Therefore, we conclude that the results which we observed in neuron-like cells are due to the accumulation of mature polyadenylated mRNAs and inhibited transcription instead of the incorrectly processed transcripts.

The main thesis and achievements of the work

Gadgil A., Raczynska K.D. U7 snRNA: A tool for Gene therapy. *Journal of Gene Medicine* 2021;23(4):e3321, doi:10.1002/jgm.3321

With the thesis focused to study the effect of ALS-FUS mutations on U7 snRNP activity, RDH gene transcription and the 3'end processing of RDH pre-mRNAs, we found it interesting is necessary to shed light onto U7 snRNP, and its potential therapeutic importance. The U7 snRNP is not only a unique snRNP in terms of its cellular activity but its capacity to be used in therapy makes it commercially important too. U snRNPs are complexes involved in splicing of pre-mRNAs, but U7 snRNP is an exception since it is not involved in splicing but instead in the 3'end processing of RDH pre-mRNAs.

The U7 snRNP being such a unique molecule has unique uses too. By changing the unique Sm binding sites for Lsm10 and Lsm11, to consensus sequence for spliceosomal U snRNAs, it leads to the formation of modified U7 snRNP which efficiently accumulates in the nucleus and can be involved in pre-mRNA splicing. Modifying the unique binding sites of the U7 snRNA renders U7 snRNP particle nonfunctional in RDH pre-mRNA processing but can be targeted for correcting splicing and is of therapeutic importance in diseases that are an outcome of splicing defects.

In the review we have discussed elaborately about the studies conducted using modified U7 snRNA in treatment of myotonic dystrophy, Duchenne muscular dystrophy (DMD), ALS, β-thalassemia, HIV-1 infection, and spinal muscular atrophy (SMA). Some studies have shown

progress and have reached clinical trial stage; a few are already approved for treatment. We have in detail discussed about the benefits and limitations about using U7 snRNA as a tool for gene therapy. Wherein, we have tried to cover varied points such as reduced risk of immune response against U7 snRNA, ability to provide lifelong therapeutic effect, and the drawbacks in delivery, economic constraints, deciding the timepoint of therapy and the dose.

So overall, in this review, we have discussed about U7 snRNP, its role in RDH gene processing, the scientific background about designing it as a tool for gene therapy, detailed analysis about the diseases targeted and the advantages and limitations, thus trying to make it informative with a wider perspective.

Unpublished data under the project funded by 'EMBO scientific exchange grant, #8906'

Although argued, the role of FUS protein in prostate cancer has been widely reported. It has been shown that FUS and some ribonucleoproteins are found in spreading initiation centers (SIC) - structures typical for early stages of cell spreading and responsible for cellular migration [11]. FUS and other FET family proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading [12]. These reported findings point out that apart from their housekeeping role, the FUS protein is involved in stress response, translational control, and adhesion, suggesting involvement in human cancer and neurodegeneration. Few publications mention the correlation of ALS and cancer mentioning that prostate cancer survivors have a reduced chance of ALS death [13,14,15]. Based on our observations related with ALS-FUS mutants and the reported publications, I further hypothesized that FUS may be the interlink in these observations.

Our collaborator, Leena Latonen group from the University of Eastern Finland, performed a preliminary study on prostate cancer samples and observed that FUS and histone proteins are significantly increased in castration resistant prostate cancer compared with the primary stage. Additionally, they also observed increased cytoplasmic accumulation of FUS in the cytoplasm of castration resistant prostate cancer cells as compared to primary stage prostate cancer cells.

Association of histone gene expression with cancer is reported, exhibiting upregulation of histone cluster 1 H1A and poor prognosis in individuals with higher expression of histone gene sets in breast cancer [16]. Another evidence for increased histone expression was reported in cervical cancer patients, and histone cluster 1 H2A, H2B, and H4 were found to be of prognostic importance [17]. Moreover, reduced FUS expression is reported to severely impair cellular proliferation and an increase in phosphorylated histone H3, a marker of mitotic arrest [18]. Thus, based on these observations, I hypothesized, that a surge in the number of RDH in prostate cancer may stimulate migration and invasion of prostate cancer cells.

My first question in the project was to check if U7 snRNP is colocalized in the cytoplasm along with FUS and is a part of the SICs. I hypothesized that if U7 snRNP is mislocalized into the cytoplasm, it would result in increase of wrongly processed

(polyadenylated) RDH mRNA transcripts. This would provide the cancer cells essential environment of genomic instability and providing continuous supply of RHD for highly proliferating cancer cells.

To check my hypothesis, I began with analyzing the cells in spreading stage, for which the procedure mentioned in de Hoog et al., 2004 was followed [11]. For staining, the combined immunofluorescence FISH protocol mentioned in our publication Gadgil et al., 2021 [24] was followed. For these experiments, I used PC3 and LNCaP cells which are androgen receptor negative and androgen receptor positive cell lines, respectively. First, I stained PC3 cells for FUS and vinculin (a marker for cell spreading) and observed that FUS and vinculin colocalize in the proposed SICs as shown previously (Fig. 1).

Fig. 1

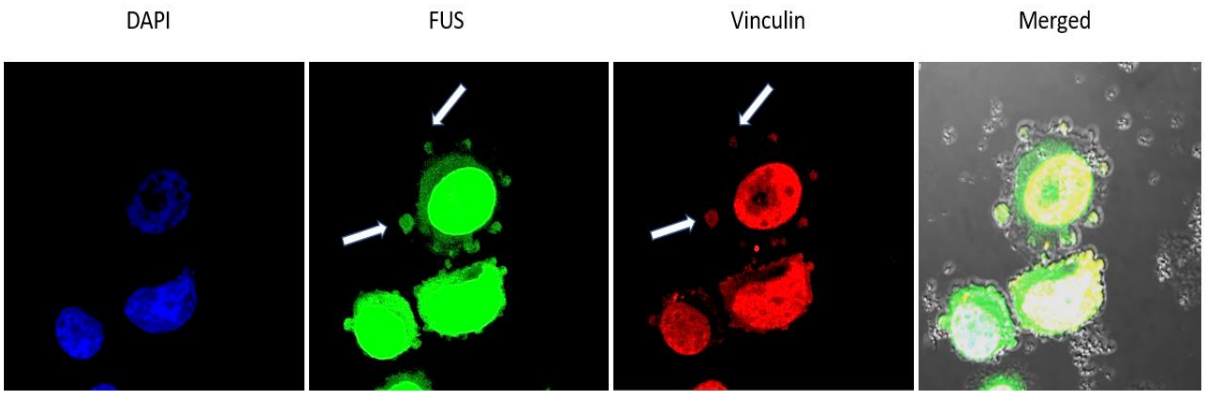


Figure 1 – PC3 cells stained with FUS and vinculin to observe colocalization in SICs.

Further, I stained PC3 cells for U7 snRNA (FISH) and vinculin (IF) and observed that U7 snRNA and vinculin colocalize in the proposed SICs as well. (Fig. 2)

Fig. 2

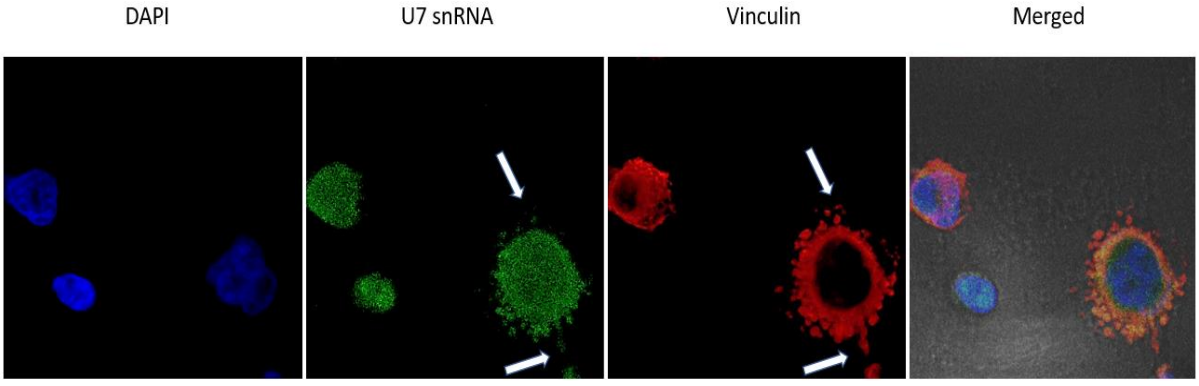


Figure 2 – PC3 cells stained with U7 snRNA and vinculin to observe colocalization in SICs.

Fig. 3

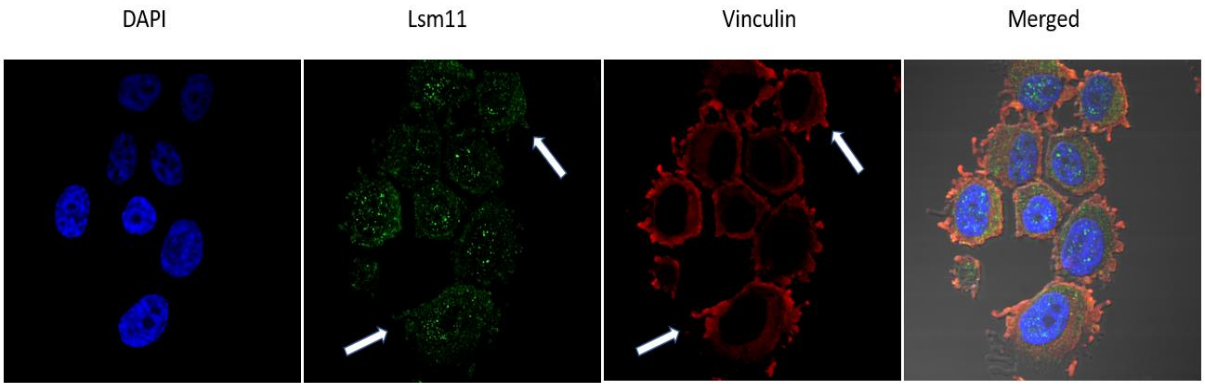


Figure 3 – PC3 cells stained with Lsm11 and vinculin to observe colocalization in SICs.

Importantly, I stained Lsm11 and vinculin to confirm if the whole U7 snRNP is colocalized in SICs. And indeed, I observed that the Lsm11 protein does colocalize with vinculin in the SICs (Fig. 3). It is an important observation showing that the whole U7 snRNP complex colocalizes with vinculin, similarly like FUS, and can be present in the SICs. Due to limitations in using and detecting multiple fluorescence markers, FUS, vinculin, and U7 snRNA together were not analyzed in SICs and needs to be done in future to make any concluding remark in this regard.

My next aim was to analyze the effect of FUS overexpression and knockdown on RDH gene expression. For this, I transfected LNCaP and PC-3 cells with plasmid for overexpression of *FUS* gene and another set of cells transfected with siFUS RNA for FUS knockdown. Cells stressed with sodium arsenite were used as a control for stressed condition. The extracted RNA was used for RT-qPCR experiments using the method as described in our publication Gadgil et al., 2021 [24]. We expected that since U7 snRNP is localized in the SIC it would elevate wrongly 3'end processed RDH pre-mRNA (polyadenylated) transcripts and increase cell proliferation. It would be a deciding factor in increased cell proliferation. Though we could observe some changes in the transcription levels and 3'end processing efficiency of RDH genes, there was no significant changes, and we are not able to have any final conclusion yet.

SUMMARY

The achievement of a doctoral dissertation prepared by me consists of one research publication which elucidates the molecular mechanism behind the pathogenesis of ALS. My PhD topic is focused to study the effect of ALS-FUS mutations on U7 snRNP activity and their effect on transcription of RDH genes and the 3'end processing of these gene transcripts. We found that ALS-FUS mutations do have an impact on U7 snRNP localization as well as their efficiency to process RDH pre-mRNAs. It was also necessary to write in detail about U7 snRNP which we discussed it in our review publication. Further, ALS is also one of the diseases that is under study for U7 snRNA-based therapy and has shown promising results. It was thus important to mention this in the review.

The obtained results published in the research paper encouraged me to write an EMBO scientific exchange grant to study the role of FUS protein and U7 snRNP in prostate cancer cells. Exploring the role of FUS and U7 snRNP in prostate cancer will add to our knowledge of the progression of the disease. Reports are published wherein the association of ALS and cancer is discussed involving the *FUS* gene. Some of these reports mention that prostate cancer survivors have a reduced chance of ALS death. Exploring the role of FUS and U7 snRNP may thus be a step towards understanding prostate cancer, and the association between prostate cancer survivors having a reduced chance of ALS death.

REFERENCES

1 - Maruyama H, Morino H, Ito H, et al. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature*. 2010;465(7295):223-226. doi:10.1038/nature08971

2 - Turner MR, Hardiman O, Benatar M, et al. Controversies and priorities in amyotrophic lateral sclerosis. *Lancet Neurol*. 2013;12(3):310-322. doi:10.1016/S1474-4422(13)70036-X

3 - Chen S, Sayana P, Zhang X, Le W. Genetics of amyotrophic lateral sclerosis: an update. *Mol Neurodegener*. 2013;8:28. Published 2013 Aug 13. doi:10.1186/1750-1326-8-28

4 - Vance C, Rogelj B, Hortobágyi T, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*. 2009;323(5918):1208-1211. doi:10.1126/science.1165942

5 - Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. 2009;323(5918):1205-1208. doi:10.1126/science.1166066

6 - Patel A, Lee HO, Jawerth L, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. 2015;162(5):1066-1077. doi:10.1016/j.cell.2015.07.047

7 - Raczynska KD, Ruepp MD, Brzek A, et al. FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors. *Nucleic Acids Res*. 2015;43(20):9711-9728. doi:10.1093/nar/gkv794

8 - Pillai RS, Grimmler M, Meister G, et al. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev*. 2003;17(18):2321-2333. doi:10.1101/gad.274403

9 - Pillai RS, Will CL, Lührmann R, Schümperli D, Müller B. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J*. 2001;20(19):5470-5479. doi:10.1093/emboj/20.19.5470

10 - Lyons SM, Cunningham CH, Welch JD, et al. A subset of replication-dependent histone mRNAs are expressed as polyadenylated RNAs in terminally differentiated tissues. *Nucleic Acids Res*. 2016;44(19):9190-9205. doi:10.1093/nar/gkw620

11 - de Hoog CL, Foster LJ, Mann M. RNA and RNA binding proteins participate in early stages of cell spreading through spreading initiation centers. *Cell*. 2004;117(5):649-662. doi:10.1016/s0092-8674(04)00456-8

12 - Andersson MK, Ståhlberg A, Arvidsson Y, et al. The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC Cell Biol*. 2008;9:37. Published 2008 Jul 11. doi:10.1186/1471-2121-9-37

13 - Freedman DM, Curtis RE, Daugherty SE, Goedert JJ, Kuncl RW, Tucker MA. The association between cancer and amyotrophic lateral sclerosis. *Cancer Causes Control*. 2013;24(1):55-60. doi:10.1007/s10552-012-0089-5

14 - Gibson SB, Abbott D, Farnham JM, et al. Population-based risks for cancer in patients with ALS. *Neurology*. 2016;87(3):289-294. doi:10.1212/WNL.0000000000002757

15 - Yamamoto I, Azuma Y, Yamaguchi M. Cancer-related genes and ALS. *Front Biosci (Landmark Ed)*. 2019;24(7):1241-1258. Published 2019 Jun 1. doi:10.2741/4777
16 - Xie W, Zhang J, Zhong P, et al. Expression and potential prognostic value of histone family gene signature in breast cancer. *Exp Ther Med*. 2019;18(6):4893-4903. doi:10.3892/etm.2019.8131
17 - Li X, Tian R, Gao H, et al. Identification of a histone family gene signature for predicting the prognosis of cervical cancer patients. *Sci Rep*. 2017;7(1):16495. Published 2017 Nov 28. doi:10.1038/s41598-017-16472-5
18 - Ward CL, Boggio KJ, Johnson BN, et al. A loss of FUS/TLS function leads to impaired cellular proliferation. *Cell Death Dis*. 2014;5(12):e1572. Published 2014 Dec 11. doi:10.1038/cddis.2014.508
19 - Sun, S., Ling, SC., Qiu, J. *et al.* ALS-causative mutations in FUS/TLS confer gain and loss of function by altered association with SMN and U1-snRNP. *Nat Commun* **6,** 6171 (2015). doi.org/10.1038/ncomms7171
20 - Yu, Y. et al. U1 snRNP is mislocalized in ALS patient fbroblasts bearing NLS mutations in FUS and is required for motor neuron outgrowth in zebrafsh. Nucleic Acids Res. 43(6), 3208–3218. doi.org/10.1093/nar/gkv157 (2015).
21 - Gerbino, V., Carri, M. T., Cozzolino, M. & Achsel, T. Mislocalised FUS mutants stall spliceosomal snRNPs in the cytoplasm. Neurobiol. Dis. 55, 120–128. doi.org/10.1016/j.nbd.2013.03.003 (2013).
22 - Takanashi, K. & Yamaguchi, A. Aggregation of ALS-linked FUS mutant sequesters RNA binding proteins and impairs RNA granules formation. Biochem. Biophys. Res. Commun. 452(3), 600–607. doi.org/10.1016/j.bbrc.2014.08.115 (2014).
23 - Takanashi, K. & Yamaguchi, A. Aggregation of ALS-linked FUS mutant sequesters RNA binding proteins and impairs RNA granules formation. Biochem. Biophys. Res. Commun. 452(3), 600–607. doi.org/10.1016/j.bbrc.2014.08.115 (2014).
24 - Gadgil, A., Walczak, A., Stępień, A. *et al.* ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells. *Sci Rep* 11, 11868 (2021). doi.org/10.1038/s41598-021-91453-3
25 - Gadgil A, Raczyńska KD. U7 snRNA: A tool for gene therapy. *J Gene Med*. 2021;23(4):e3321. doi:10.1002/jgm.3321

Publication

ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells

scientific reports

Check for updates

OPEN

ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells

Ankur Gadgil[1,2], Agnieszka Walczak[1], Agata Stępień[1], Jonas Mechtersheimer[3], Agnes Lumi Nishimura[3], Christopher E. Shaw[3,4], Marc-David Ruepp[3] & Katarzyna Dorota Raczyńska[1,2]✉

Genes encoding replication-dependent histones lack introns, and the mRNAs produced are a unique class of RNA polymerase II transcripts in eukaryotic cells that do not end in a polyadenylated tail. Mature mRNAs are thus formed by a single endonucleolytic cleavage that releases the pre-mRNA from the DNA and is the only processing event necessary. U7 snRNP is one of the key factors that determines the cleavage site within the 3′UTR of replication-dependent histone pre-mRNAs. We have previously showed that the FUS protein interacts with U7 snRNA/snRNP and regulates the expression of histone genes by stimulating transcription and 3′ end maturation. Mutations in the *FUS* gene first identified in patients with amyotrophic lateral sclerosis (ALS) lead to the accumulation of the FUS protein in cytoplasmic inclusions. Here, we report that mutations in FUS lead to disruption of the transcriptional activity of FUS and mislocalization of U7 snRNA/snRNP in cytoplasmic aggregates in cellular models and primary neurons. As a consequence, decreased transcriptional efficiency and aberrant 3′ end processing of histone pre-mRNAs were observed. This study highlights for the first time the deregulation of replication-dependent histone gene expression and its involvement in ALS.

During the evolution of eukaryotes, molecular mechanisms were established to ensure accurate assembly of newly replicated DNA into chromatin. Therefore, in the S phase of the cell cycle, DNA synthesis is tightly coupled with histone protein synthesis. These two processes are finely balanced, as any disturbance may result in dysregulation of gene expression, cell cycle arrest, and chromosome instability, which may lead to developmental failure[1]. This coordination is most sophisticated in higher eukaryotes.

The core histones H2A, H2B, H3, and H4 and the linker histone H1 are replication-dependent histones (RDH) and are responsible for DNA packaging. Their expression is strongly upregulated during the G1/S phase transition, with their mRNA levels increasing by ~ 35-fold in S phase due to activated transcription, efficient 3′ end processing and enhanced transcript stability[2,3]. At the end of the S phase, the availability of histones is repressed because an excess could be harmful to the cells[2]. In metazoan cells, replication-dependent histone mRNAs are not polyadenylated, and 3′ end processing relies only on a single cleavage event that is carried out by the endonuclease CPSF73 and mediated by a subset of specialized factors that recognize specific elements on the nascent transcripts[4–6]. Additionally, several of these factors are cell cycle-regulated, and their highest activity is in the S phase of the cell cycle[7–9].

One of the key factors required for replication-dependent histone pre-mRNAs is the U7 small nuclear ribonucleoprotein (U7 snRNP). The U7 snRNP consists of the U7 snRNA and five Sm proteins that are shared with spliceosomal snRNPs (SmB/B', SmD3, SmE, SmF and SmG) and two unique Sm-like proteins: Lsm10 and Lsm11[6,10,11]. Initially identified by Birnstiel and colleagues in 1983 in *Xenopus* oocytes[12], the U7 small nuclear RNA (U7 snRNA) encompasses 63 nucleotides in humans and is synthesized by RNA polymerase II (RNAP2)[12–14] with approximately 10^3–10^4 copies present in mammalian cells. The U7 snRNA 5′ end is complementary to the

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland. [2]Center for Advanced Technology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 10, 61-614 Poznan, Poland. [3]UK Dementia Research Institute Centre at King's College London, Institute of Psychiatry, Psychology and Neuroscience, King's College London, Strand, London WC2R 2LS, UK. [4]Centre for Brain Research, University of Auckland, 85 Park Road, Grafton, Auckland 1023, New Zealand. ✉email: doracz@amu.edu.pl

histone downstream element (HDE), which is located downstream of the cleavage site within the 3′UTR of histone pre-mRNAs. Binding of the U7 snRNP to the HDE is crucial for correct 3′ end processing[13,15,16].

Recently, we discovered that fused in sarcoma (FUS) is a linking factor that coordinates both RDH gene transcription and the 3′ end processing of their pre-mRNAs[17]. During the S phase of the cell cycle, FUS interacts with U7 snRNP and binds to histone promoters to enhance the binding of RNAP2 to histone genes[17]. FUS is predominantly nuclear, is ubiquitously expressed and is capable of binding to DNA and RNA. This protein plays a role in genomic maintenance and DNA recombination[18–20], in addition to regulating RNA metabolism and processing. These processes include transcription, splicing, alternative splicing, miRNA biogenesis and nucleocytoplasmic shuttling of mRNAs[20–24]. In human neurons, FUS is located at synaptic sites along dendrites and has an important function in transporting mRNA for local translation, as well as its involvement in synaptic plasticity[25,26].

In 2009, mutations in the *FUS* gene were identified in patients with amyotrophic lateral sclerosis (ALS)[27,28]. ALS is a chronic, progressive and uniformly fatal neurodegenerative disorder affecting motor neurons in the brain and spinal cord. Most FUS-linked ALS-causing mutations are missense mutations clustered in the highly conserved C-terminal nuclear localization signal (NLS) that disrupt the interaction with Transportin, which is required for nuclear import[29]. Mutations in the NLS can abolish or significantly reduce nuclear import of FUS, leading to its accumulation in the cytoplasm. Cytoplasmic inclusions containing FUS are found in neurons and glial cells in the brain and spinal cord of ALS patients[20,22,23,28–32]. In some cells, intranuclear inclusions containing FUS have also been described[33]. Splicing dysregulation and transcription impairment have been reported as a consequence of disruptions in FUS in mammalian cells as well as in ALS patients' fibroblasts[34–36]. In addition, studies suggest that FUS aggregates affect local mRNA translation at the synapse and contribute to synaptic loss in ALS motoneurons[33]. Interestingly, FUS cytoplasmic aggregates sequester U snRNPs (U1 snRNPs, U2 snRNPs, U1 snRNA, U11 snRNA and U12 snRNA), FUS-binding proteins (hnRNP A1, hnRNP A2, and SMN), and poly(A) mRNAs[37–44].

Here, we show that U7 snRNP is also trapped in cytoplasmic aggregates along with ALS-linked FUS mutants in cellular models and primary neurons. As a consequence of this mislocalization, we observed decreased transcriptional efficiency of RDH genes and aberrant 3′ end processing of their pre-mRNAs in SH-SY5Y cells. These changes can lead to histone instability and increased toxicity in neurons and motor neurons and thus can be involved in the pathogenesis of FUS-ALS.

Results

ALS-linked FUS mutations affect the cellular localization of U7 snRNA/snRNP. FUS knockout SH-SY5Y cells (FUS KO) were transiently transfected with vectors encoding ALS-linked FLAG-tagged wild-type FUS cDNA or cDNA from the ALS-linked FUS mutants FUS-P525L and R495X. Both mutations lead to a severe ALS clinical phenotype and a striking cytoplasmic accumulation of FUS[29,31]. Experiments were performed using both proliferating cells and cells differentiated to neuron-like cells by the addition of retinoic acid[45]. Two days after transfection, the cells were fixed and subjected to immunofluorescence (IF) and fluorescence in situ hybridization (FISH). As shown in Fig. 1, wild-type FUS was located in the nucleus in the proliferating and differentiated cells transfected with wild-type FUS (Fig. 1A,B, upper panels). In contrast, in the P525L- and R495X-transfected cells, the ALS-linked FUS proteins were localized in cytoplasmic aggregates. These cytoplasmic aggregates also contained U7 snRNA (Fig. 1A,B, middle and lower panel). We repeated this experiment in primary neurons isolated from rat brain, transfected with the FUS mutants P525L. In line with the previous results obtained from SH-SY5Y cells, we again observed that U7 snRNA was recruited into cytoplasmic aggregates along with the mutant FUS (Supplementary Fig. S1A). These experiments were also conducted in FUS-knockout HeLa cells, and similar results were obtained (Fig. 1C).

Based on these results, we investigated whether the ALS-linked FUS mutations are associated with the mislocalization of other U7 snRNP components. To answer this question, we analyzed the localization of Lsm11 in cells transfected with FUS mutants P525L or R495X and compared with that in cells transfected with wild-type FUS. Lsm11 is a unique protein for the U7 snRNP complex. When transfected with wild-type FUS, Lsm11 was detected as a focus inside the nucleus of differentiated SH-SY5Y FUS KO, primary neurons and HeLa FUS KO cells (Supplementary Figs.S1B, S2A,B, upper panels). However, in cells transfected with mutant FUS, Lsm11 was recruited into cytoplasmic aggregates along with FUS (Supplementary Figs. S1B, S2A,B, middle and lower panel).

Taken together, our data suggest that mutations in FUS lead to mislocalization of the U7 snRNP complex to cytoplasmic aggregates.

Mislocalization of FUS and U7 snRNA/snRNP leads to deregulated expression of RDH genes. Recently, we showed that FUS interacts with U7 snRNA/snRNP and participates in the 3′ end processing of replication-dependent histone pre-mRNAs. In cells with depleted FUS, we observed an elevated level of incorrectly processed, extended transcripts, which resulted from decreased processing efficiency[17]. To investigate the effect of mutant FUS on histone pre-mRNA processing, we isolated RNA from SH-SY5Y FUS KO cells transfected with FUS (WT) and FUS mutants (R495X and P525L) and analyzed the processing efficiency of different histones by RT-qPCR. The fraction of correctly cleaved histone mRNAs is displayed as the ratio of total to incorrectly processed (extended) transcripts, which represents the apparent processing efficiency, as described in[17]. The replication-independent histone gene H2A.Z, which undergoes cleavage and polyadenylation, was used as a reference. Indeed, compared to the cells transfected with WT FUS, the proliferating SH-SY5Y FUS KO cells transfected with the ALS-linked FUS mutants showed a significantly affected processing efficiency in the majority of histones analyzed (Fig. 2A). This effect resulted from an elevated level of extended transcripts; concomitantly, the level of total transcripts was significantly downregulated (Fig. 2B). Unfortunately, the low transfection efficiency prohibited the analogous experiment in primary neurons.

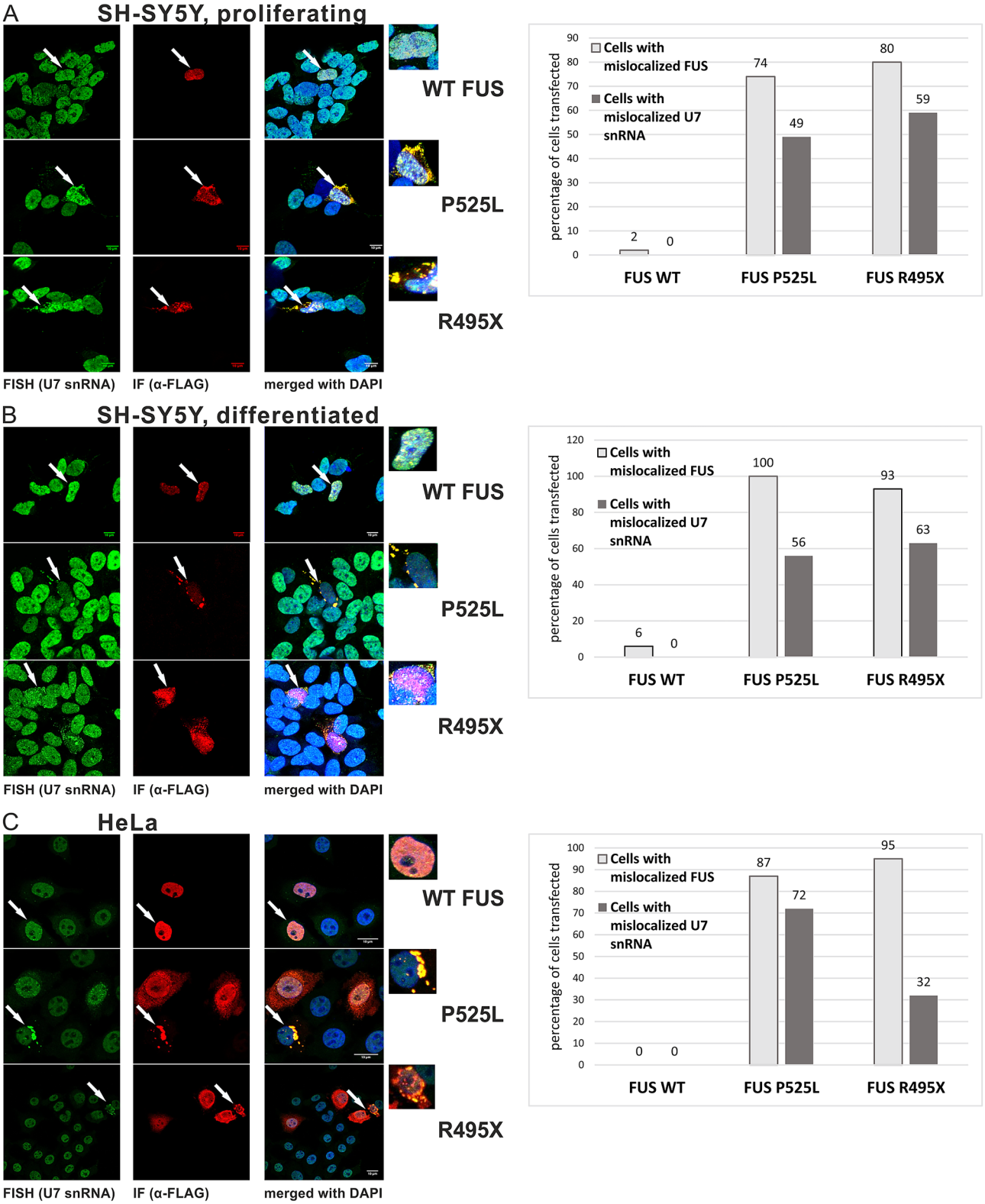


Figure 1. Localization of FUS and U7 snRNA in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations. Fluorescent in-situ hybridization (FISH) using a probe against U7 snRNA in combination with immunofluorescence (IF) using anti-FLAG antibodies was performed in SH-SY5Y FUS KO (**A**,**B**) and HeLa FUS KO (**C**) cells transfected with FLAG-tagged FUS. DAPI was used for nuclear staining. Graphs show the percentage of transfected cells with mislocalization of FUS and U7 snRNA. WT FUS—cells transfected with plasmids encoding the wild-type *FUS* gene. P525L and R495X—cells transfected with plasmids encoding the *FUS* gene with P525L and R495X mutations.

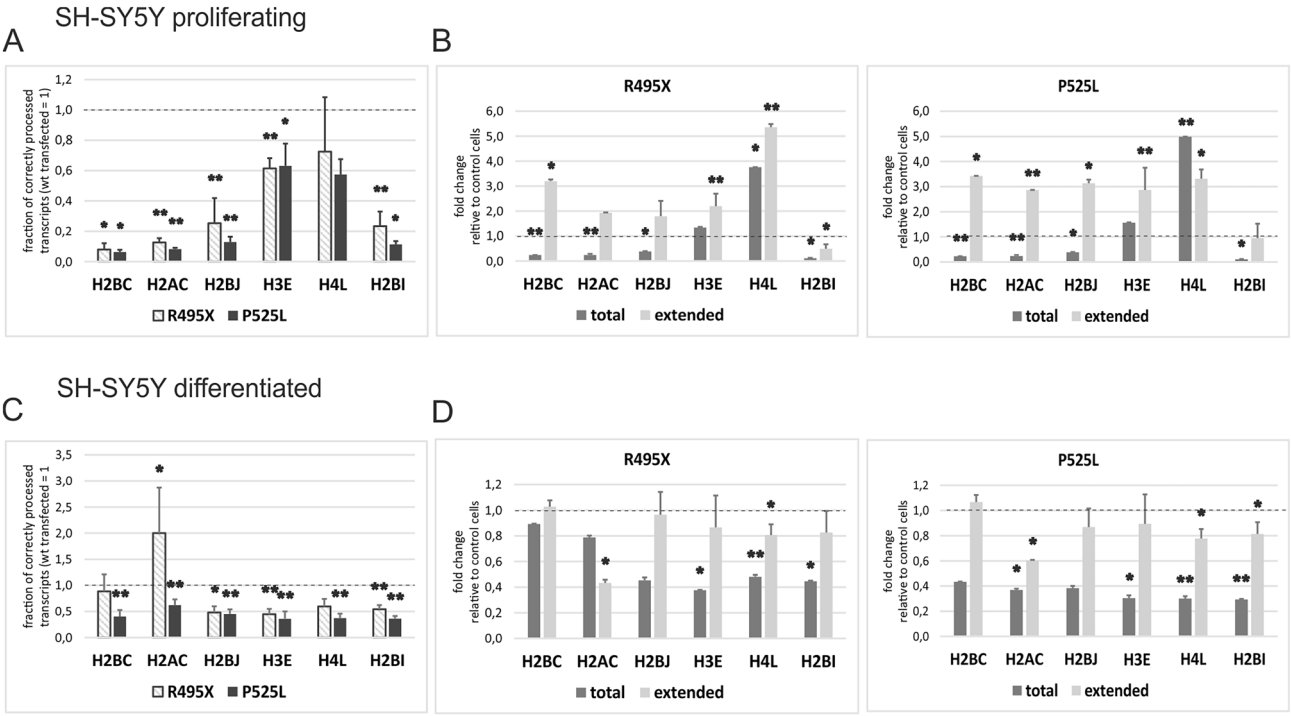


Figure 2. Replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations. (**A**,C) The fraction of correctly cleaved histone mRNAs in the proliferating (**A**) and differentiated (**C**) cells transfected with the FUS mutants compared to the cells transfected with wild-type FUS displayed as the ratio of total to unprocessed (extended transcripts). (**B**,**D**) Fold enrichment of total and extended transcripts in the cells transfected with the FUS mutants compared to the cells transfected with wild-type FUS. H2A.Z was used as a reference gene. Error bars represent the SD of three biological replicates. P-values were calculated using Student's t-test, and the statistical significance is represented as follows: *P ≤ 0.05; **P ≤ 0.01.

We have previously reported that decreased levels of total histone transcripts correlate with weaker binding of FUS to histone promoters and diminished levels of RNAP2 on histone genes[17]. To investigate whether the ALS-linked FUS mutations affect the activation of RDH gene transcription, we performed a ChIP assay in proliferating SH-SY5Y—FUS KO cells transfected with WT FUS or the P525L and R495X mutants. The RNAP2 occupancy on the promoter region/5′UTR, open reading frame (CDS) and 3′UTR of three RDH genes, H2AC, H2BJ and H4J, was analyzed by qPCR, as previously described[17]. The background signal was removed by subtracting the signals of control cells (non-transfected), and the fold enrichment of precipitated material from the cells transfected with the ALS-FUS mutants compared to the cells transfected with WT FUS was calculated. As shown in Supplementary Figure S3, mutations in FUS led to decreased RNAP2 occupancy in almost all three regions in the three histones analyzed. Collectively, our data suggest that mutations in FUS lead to inhibition of replication-dependent histone gene transcription in proliferating cells due to weak loading and binding of RNAP2 on histone genes.

Interestingly, no obvious impairment in 3′ end maturation was observed in the neuron-like SH-SY5Y FUS KO differentiated cells (Supplementary Fig. S3). Although "processing efficiency" was significantly downregulated in the mutant cells, this finding is probably due to inhibited transcription rather than affected 3′ end processing (Fig. 2C). Terminally differentiated cells were shown to synthesize only polyadenylated histone transcripts in a process that does not require U7 snRNP[46]. Therefore, the results in neuronal differentiated cells are due to accumulation of mature polyadenylated histone mRNAs instead of the incorrectly processed transcripts (Fig. 2D). We conclude that the "reduced processing efficiency" resulted from decreased levels of total pre-mRNAs and reduced transcription. This result may be due to the loss of nuclear function of the transcriptional activator FUS.

Discussion

Biological consequences of deregulation of RDH gene expression. We previously found that FUS is important in replication-dependent histone gene expression[17]. Here, we further investigated how ALS-linked mutations in the *FUS* gene can lead to disruption of the transcriptional activity of FUS and mislocalization of U7 snRNA/snRNP in cytoplasmic aggregates. As a consequence, we observed decreased transcriptional efficiency and aberrant 3′ end processing of histone pre-mRNAs, which is similar to what we observed in inducible HeLa FUS knockdown cells[17]. This study is the first to report deregulation of replication-dependent histone gene expression in the pathogenesis of ALS.

The involvement of FUS in DNA structure and chromosome maintenance has been previously shown[18–20]. FUS is involved in the formation of D-loops that are present during homologous recombination, DNA repair

and telomeres. This protein is important for the induction of the DNA damage response and recruitment of the repair complex at the sites of double-stranded breaks (DSBs)[18,20,24]. FUS knockout mice are characterized by high levels of genome instability, enhanced sensitivity to ionizing radiation and increased chromosomal disturbances in premeiotic spermatocytes[47,48]. Accumulated DNA damage caused by a reduction in nonhomologous end joining (NHEJ)-mediated DSB repair was observed in neurons of the postmortem motor cortex from patients with FUS-ALS[49]. The increased DNA damage was also suggested to cause neurodegeneration in motoneurons reprogrammed from induced pluripotent stem cells (iPSCs) derived from FUS-ALS patients[33]. Naumann and collaborators even proposed that DNA damage is an upstream event that enhances aggregate formation, cytoplasmic localization of FUS and neurodegeneration[50]. Taken together, increased DNA damage in human ALS patients harboring FUS mutations together with disrupted histone synthesis can cause genome instability and may be the molecular mechanisms underlying altered glial cell or motor neuron homeostasis in ALS.

To test the effect of the ALS-linked mutations of FUS on transcriptional efficiency and 3′ end processing of histone pre-mRNAs, we transiently expressed FUS in SH-SY5Y and HeLa cell lines, in both wild-type and FUS knockout cell lines. However, probably due to the high expression of endogenous FUS in wild-type cells, the effect that we observed was more prominent in the FUS KO cells transfected with the ALS-linked FUS mutants. Notably, due to the limited transfection efficiency, the result is representative of cells that contain exogenous FUS and of cells that do not contain FUS. Therefore, the effect of the ALS-linked FUS mutations on RDH gene expression might still be underestimated.

Higelin and colleagues have suggested that in iPSCs and motoneurons with ALS-linked FUS mutations, additional external factors such as DNA damage or hyperosmolar stress can induce changes in FUS mislocalization and stress granule formation. This change was accompanied by impaired DNA damage repair since FUS is a component of the DSB repair machinery. Interestingly, the researchers observed that the accumulation of DNA damage and the cellular response to DNA damage stressors were more pronounced in the postmitotic mutant FUS motoneurons than in the dividing pluripotent cells. Therefore, maturation, accumulation of DNA damage foci and later events in older motoneurons were supposed to increase the size and toxicity of FUS inclusions and finally induce neurodegeneration[33]. Under natural conditions, in motoneurons in the patient's brain, these factors might also enhance the negative effect of mutated FUS on U7 snRNA/snRNP function and RDH gene expression.

Biological consequences of U7 snRNA/snRNP mislocalization. As shown in Fig. 1 and Supplementary Figure S1, in HeLa FUS KO cells, in proliferating and differentiated SH-SY5Y FUS KO cells, and in primary neurons, the ALS-linked FUS mutations cause mislocalization of FUS and U7 snRNA in cytoplasmic aggregates. In the proliferating cells, this change led to deregulation of histone gene expression (transcription and processing of pre-mRNAs). As discussed above, this phenomenon can result in abrogated histone synthesis and, as a consequence, can cause genome instability. In contrast, in the differentiated neurons, replication-dependent histone gene expression was almost silenced [46] and necessary histones are eventually produced from extended, polyadenylated transcripts. For this reason, the effect of the ALS-linked FUS mutations on the unique process of 3′ end maturation of RDH pre-mRNAs was not detectable (Fig. 2D) However, the U7 snRNP in these cells was still present (Figs. 1B and S1A). The question arises as to which biological pathway is then affected by the mislocalized U7 snRNA/snRNP. The answer may shed new light on the molecular mechanisms underlying altered motor neuron homeostasis in ALS patients as well.

In summary, together with major and minor spliceosomal snRNPs, U7 snRNP is another snRNP in the cell, whose activity is affected by ALS-linked FUS mutations. As these snRNPs play different functions, the potential defects in ALS patients are expanded from splicing defects to replication-dependent histone processing defects. In this report, we have also expanded the known defects in snRNP/snRNA metabolism.

Materials and methods

Cell culture and differentiation. SH-SY5Y and HeLa cells were grown in Dulbecco's modified Eagle's medium containing L-glutamine and 4.5 g/L glucose (DMEM; Lonza) and supplemented with 10% fetal calf serum (Gibco) and antibiotics (100 U/ml penicillin, 100 µg/ml streptomycin (Sigma-Aldrich)) at 37 °C in a humidified atmosphere containing 5% CO_2. Rat primary cortical neurons were isolated from E18 pups and plated on 100 µg/ml PDL coated chamber slides (ibidi). Neurons were maintained in Neurobasal media supplemented with 1% Glutamax, 1% B27 supplement and 1% Penicillin/Streptomycin media (Life Technologies). SH-SY5Y and HeLa cells with FUS KO were prepared as described previously[39,51]. For cellular differentiation, SH-SY5Y FUS KO cells were subjected to retinoic acid (RA) treatment for a period of 10 days to transform them into neuron-like cells[52]. All-trans RA (USP, Tretinoin, 1674004) at a final concentration of 75 µM was added to DMEM with 10% FBS. The media along with RA were replaced every 3 days. After 10 days, fresh media without RA were replaced just before transfection. The differentiation efficiency was analyzed as follows: (i) level of differentiation markers: the transcription factor MYC and Growth Associated Protein 43 (GAP43) were tested by RT-qPCR (Supplementary Fig. S4A); (ii) actin staining was performed to show neuron-like cells (Supplementary Fig. S4B); and (iii) expression of Microtubule Associated Protein 2 (MAP2) in differentiated cells was confirmed by Western blots (Supplementary Fig. S4C).

Plasmid and cell transfection. The plasmids used in the experiments are based on the pCDNA vector encoding cDNA of the FUS WT plasmid, the FUS R495X plasmid where the amino acid sequence is truncated after arginine at position 495, and the FUS P525L plasmid where proline is mutated to leucine at position 525. The plasmids used for the transfection of SH-SY5Y and HeLa cells were kindly obtained from Don Cleveland's group (described in[37]), while the plasmids used for the transfection of rat cortical neurons were described in[39].

Transient transfections were performed with VIROMER RED (Lipocalyx) or Lipofectamine 2000 according to the manufacturer's instructions. Cells were harvested 48 h after transfection for further experiments.

Immunofluorescence and fluorescence in situ hybridization. SH-SY5Y FUS KO cells, HeLa FUS KO cells and primary neurons were grown on chambered coverslips. SH-SY5Y FUS KO cells and HeLa FUS KO cells were transfected with 250 ng of plasmid DNA (FUS WT, FUS P525L, FUS R495X), using VIROMER RED (Lipocalyx). Forty-eight hours after transfection, the cells were fixed with 4% PFA and permeabilized in 1× PBS pH 7.0 + 0.5% Triton X-100 (PBS-T). The rat primary cortical neurons were transfected on DIV7, with 175 ng of plasmid DNA pcDNA6F-FUS-wt and pcDNA6F-P525L using Lipofectamine 2000 following supplier's recommendations. Neurons were fixed on DIV9 with 4% PFA for 15 min at room temperature and pre-treated with 70% ethanol for 2 min at − 20 °C and finally stored in 100% ethanol at − 80 °C until staining. On the day of the staining, cells were washed twice with 70% ethanol for 2 min and three time with PBS for 5 min. Proceeding the protocol, the cells were incubated with blocking solution (1% BSA in PBS, 200 mM ribonucleoside vanadyl complex) for 30 min at RT. Primary antibody incubation for FUS was either with the mouse anti-FLAG antibody (1:200, Merck F3165) or anti-FUS antibody (Santa Cruz, SC 47711) for 1 h at RT. Cells were washed in blocking solution followed by secondary anti-mouse Alexa Fluor 555 antibody (Thermo Fisher Scientific A21422) or Alexa Fluor 546 (Thermo scientific A11030) in blocking solution for 45 min RT. Following immunofluorescence, the cells were washed 3 times with PBS and subsequently post-fixed with 4% PFA for 5 min at RT. The cells were then washed twice with 2× SSC (300 mM NaCl, 30 mM sodium citrate pH 7.0) and incubated in pre-hybridization buffer (15% formamide, 10 mM sodium phosphate, 2 mM RVC in 2× SSC) for 10 min at RT. U7 snRNA-specific probe was prepared with RNA Labeling Kit–488 (BaseClick BCK-RNA488-10). Briefly, in vitro transcription was carried out with the use of a plasmid template (full-length U7 snRNA complementary sequence cloned into the pBS plasmid) and the addition of 5-ethynyl-UTP. All the steps of in vitro transcription and labeling were performed according to the manufacturer's protocol. The probe was then diluted to 0.75 ng/µl in hybridization buffer (15% formamide, 10 mM sodium phosphate, 10% dextran sulfate, 0.2% BSA, 0.5 µg/µl salmon sperm DNA, 0.5 µg/µl *Escherichia coli* tRNA, 2 mM RVC in 2 × SSC) and hybridized to the cells overnight at 37 °C. The following day the cells were washed twice in pre-hybridization buffer for 30 min at 37 °C and three times for 10 in high stringency buffer (20% formamide, 2 mM RVC in 0.05× SSC). Subsequently, the cells were washed three times in 2× SSC for 2 min at RT. Washing steps were performed to remove unbound probes, and the slides/coverslips were mounted with ProLong Gold aqueous DAPI containing mounting medium or using Vectashield containing DAPI. For detection of LSM11, the cells were fixed, washed and blocked as described above. Rabbit anti-LSM11 antibody (1:200, Merck HPA039587) was applied followed by secondary anti-rabbit Alexa Fluor 555 antibody or anti-rabbit Alexa Fluor 546 antibody.

The images for SH-SY5Y FUS KO and HeLa FUS KO cells were acquired with a confocal scanning microscope (Nikon A1Rsi) using a 100×/1.4 or 63×/1.4 oil-immersion objective. Excitation was achieved with an argon laser at 488 nm (6-FAM: U7 snRNA) and with a diode laser at 561 nm (Alexa Fluor 555: FUS, Lsm11), 405 nm (DAPI). Images were analyzed using ImageJ open source software. Images for the rat cortical neurons were acquired using the VT-iSIM microscope (Nikon) using a 20×/dry or 100×/1.49 NA oil immersion lens. Deconvolution was performed with the NIS-Elements AR software (Ver 5.01) using the Richardson/Lucy algorithm with 15 iterations.

RNA isolation, cDNA preparation and qPCR. For RNA isolation, a Quick-RNA MiniPrep Kit (Zymo Research, R1055) was used according to the manufacturer's protocol. For cDNA preparation, 0.5 µg of random hexamer primer was used for 0.5–3 µg of RNA in a total volume of 25 µl. The samples were denatured at 65 °C for 5 min, and then, master mix consisting of 1× first strand buffer, 0.5 µM dNTPs, 1 µl of RNasin (40 U/µl stock concentration), 5 mM DTT and 200 U Super Script III Reverse transcriptase was added. The reaction was incubated for 1 h at 50 °C followed by a 10 min incubation at 75 °C. For qPCR, 1 µl of 3–5× diluted cDNA template, 0.2 µM primer mix (forward + reverse) and 5 µl of SYBR Green PCR master mix (Applied Biosystems) were added in a 10 µl reaction with the following conditions: denaturation for 10 min at 95 °C, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min (Applied Biosystems QuantStudio 7 Flex).

Western blotting analyses. For immunodetection, proteins were separated by SDS–polyacrylamide gel electrophoresis (PAGE), transferred to polyvinylidene difluoride (PVDF) membranes (Millipore), blocked with 5% skim milk diluted in PBS-T and then incubated for 1.5 h at room temperature (RT) with anti-MAP2 (Abcam, 5392, 1:8000 dilution in PBS-T) and anti-actin (MP Biomedicals, 691001, 1:50,000 dilution in PBS-T) primary antibodies. After 3 washes with PBS-T, the membrane was incubated for 1 h at RT with species-specific horseradish peroxidase (HRP)-coupled secondary antibody (Santa Cruz, goat anti-mouse, SC2005, 1:3000 dilution in PBS-T) followed by 3 washes as before. The signal was detected using the enhanced chemiluminescence method (ECL, GE Healthcare).

Chromatin immunoprecipitation. Neuroblastoma SH-SY5Y FUS KO cells, control cells (non-transfected) and cells transfected for 48 h with FUS WT, FUS R495X and FUS P525L plasmids were trypsinized, washed with PBS and crosslinked using 10 ml of 1% formaldehyde for 10 min at RT. The crosslinking reaction was stopped by adding 800 µl of 1 M glycine and incubating for 3 min at RT with gentle swirling. The cells were centrifuged at 1000*g* for 1 min at 4 °C, and the pellet was washed with 1X PBS and then lysed in lysis buffer (72.3 mM NaCl, 5 mM EDTA pH 8.0, 0.5% NP-40, 50 mM Tris pH 8.0, 1X EDTA-free protease inhibitor (Roche)), pipetting to make a homogenous suspension. After lysis, the cells were centrifuged at 12,000*g* for 1 min at 4 °C, and the pellet was resuspended in 1 ml of sonication buffer (1% SDS, 10 mM EDTA pH 8.0, 50 mM Tris pH 8.0 and 1 protease inhibitor cocktail tablet/10 ml buffer). The cells were sonicated for 30 cycles at high

intensity with 30 s ON/30 s OFF at 4 °C using a Bioruptor Plus sonicator (Diagenode) to generate fragments of DNA between 200 and 700 bp. The sizes of the fragmented DNA were verified by agarose gel electrophoresis. Before agarose gel analysis, the samples were treated with 1 U of RNase A (Thermo Scientific, EN0601) at 37 °C for 30 min. Next, 20 ng of proteinase K was added, and samples were incubated for 2 h at 45 °C at 1000 rpm. After sonication, the cell debris was removed by centrifuging the cells at 11,200*g* for 10 min at 4 °C. The supernatant was transferred to a new tube; 10% of the supernatant was kept as input. One hundred microliters of solution was transferred to another tube and diluted in a 1:10 ratio using dilution buffer (1 mM EDTA, pH 8.0, 0.01% SDS, 1.1% Triton-X 100, 0.17 M NaCl) and further used as an IgG control. The remaining solution was also diluted in a 1:10 ratio using dilution buffer and further used for immunoprecipitation (IP). One microgram of IgG (Invitrogen 10500C) was added to the IgG control probe, while 5 µg of RNAP2 (Abcam, ab10332 and ab10338) was added to the sample tubes. For IP, 15 µl of Dynabeads Protein G (Life Technologies) was pre-blocked with 0.5% bovine serum albumin (BSA) by overnight incubation at 4 °C. Next, the samples were centrifuged at 3500 rpm for 20 min at 4 °C. After centrifugation, the top 90% of the solution was transferred to another tube for IP, and the bottom 10% was discarded. For IP, 190 µl of protein G beads was added into the solution and incubated for 6 h at 4 °C with head-to-tail rotation. Next, the samples were washed 6 times with lysis buffer and rotated for 3 min at 4 °C. After the samples were washed, the beads were resuspended in 1 ml of ice-cold Tris–EDTA (TE) buffer to make a slurry. TE buffer was removed, and the beads were retained by keeping the tubes on a magnetic rack. Two hundred microliters of elution buffer (1% SDS, 0.1 M $NaHCO_3$) was added to the beads as well as to the input and incubated for 15 min at 1000 rpm at room temperature. The eluted solution was transferred to new tubes, and 8 µl of 5 M NaCl and 4 µl of 0.5 M EDTA, pH 8.0 were added to each sample and incubated overnight at 65 °C at 1000 rpm to de-crosslink the immunoprecipitated chromatin. The probes were further digested by adding 10 µg of RNase A and incubating at 37 °C for 30 min. Next, 8 µl of 1 M Tris-HC, pH 8.0, 4 µl of 0.5 M EDTA and 20 ng of Proteinase K (Thermo Scientific) were added, and the samples were incubated for 2 h at 45 °C at 1000 rpm. The purified DNA was then used for qPCR.

Ethics statement. All experiments concerning laboratory animals were carried out following the ethical guidelines of the Animal research reporting of in vivo experiments (ARRIVE), UK Animals (Scientific Procedures) Act 1986, and were approved by the Kings College, London ethics review panel. We further confirm that all methods were carried out in accordance with relevant guidelines and regulations.

Received: 25 September 2020; Accepted: 18 May 2021
Published online: 04 June 2021

References

1. Ghule, P. N. *et al.* Fidelity of histone gene regulation is obligatory for genome replication and stability. *Mol. Cell. Biol.* **34**(14), 2650–2659. https://doi.org/10.1128/MCB.01567-13 (2014).
2. Harris, M. E. *et al.* Regulation of histone mRNA in the unperturbed cell cycle: Evidence suggesting control at two posttranscriptional steps. *Mol. Cell. Biol.* **11**(5), 2416–2424. https://doi.org/10.1128/mcb.11.5.2416 (1991).
3. Luscher, B. & Schumperli, D. RNA 3′ processing regulates histone mRNA levels in a mammalian cell cycle mutant. A processing factor becomes limiting in G1-arrested cells. *EMBO J.* **6**(6), 1721–1726 (1987).
4. Dominski, Z. & Marzluff, W. F. Formation of the 3′ end of histone mRNA: Getting closer to the end. *Gene* **396**(2), 373–390. https://doi.org/10.1016/j.gene.2007.04.021 (2007).
5. Yang, X. C. *et al.* A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3′-end processing. *Mol. Cell Biol.* **33**(1), 28–37. https://doi.org/10.1128/MCB.00653-12 (2013).
6. Schumperli, D. & Pillai, R. S. The special Sm core structure of the U7 snRNP: Far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol. Life Sci.* **61**(19–20), 2560–2570. https://doi.org/10.1007/s00018-004-4190-0 (2004).
7. Ma, T. *et al.* Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev.* **14**(18), 2298–2313. https://doi.org/10.1101/gad.829500 (2000).
8. Romeo, V., Griesbach, E. & Schumperli, D. CstF64: Cell cycle regulation and functional role in 3′ end processing of replication-dependent histone mRNAs. *Mol. Cell Biol.* **34**(23), 4272–4284. https://doi.org/10.1128/MCB.00791-14 (2014).
9. Whitfield, M. L. *et al.* Stem-loop binding protein, the protein that binds the 3′ end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol. Cell Biol.* **20**(12), 4188–4198 (2000).
10. Pillai, R. S. *et al.* Unique Sm core structure of U7 snRNPs: Assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* **17**(18), 2321–2333. https://doi.org/10.1101/gad.274403 (2003).
11. Pillai, R. S., Will, C. L., Luhrmann, R., Schumperli, D. & Muller, B. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J.* **20**(19), 5470–5479. https://doi.org/10.1093/emboj/20.19.5470 (2001).
12. Galli, G., Hofstetter, H., Stunnenberg, H. G. & Birnstiel, M. L. Biochemical complementation with RNA in the Xenopus oocyte: A small RNA is required for the generation of 3′ histone mRNA termini. *Cell* **34**(3), 823–828. https://doi.org/10.1016/0092-8674(83)90539-1 (1983).
13. Mowry, K. L. & Steitz, J. A. Identification of the human U7 snRNP as one of several factors involved in the 3′ end maturation of histone premessenger RNA's. *Science* **238**(4834), 1682–1687 (1987).
14. Soldati, D. & Schumperli, D. Structural and functional characterization of mouse U7 small nuclear RNA active in 3′ processing of histone pre-mRNA. *Mol. Cell Biol.* **8**(4), 1518–1524. https://doi.org/10.1128/mcb.8.4.1518 (1988).
15. Bond, U. M., Yario, T. A. & Steitz, J. A. Multiple processing-defective mutations in a mammalian histone pre-mRNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. *Genes Dev.* **5**(9), 1709–1722. https://doi.org/10.1101/gad.5.9.1709 (1991).
16. Schaufele, F., Gilmartin, G. M., Bannwarth, W. & Birnstiel, M. L. Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3′ end of H3 messenger RNA. *Nature* **323**(6091), 777–781. https://doi.org/10.1038/323777a0 (1986).
17. Raczynska, K. D. *et al.* FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors. *Nucleic Acids Res.* **43**(20), 9711–9728. https://doi.org/10.1093/nar/gkv794 (2015).

18. Baechtold, H. *et al.* Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation. *J. Biol. Chem.* **274**(48), 34337–34342. https://doi.org/10.1074/jbc.274.48.34337 (1999).
19. Bertrand, P., Akhmedov, A. T., Delacote, F., Durrbach, A. & Lopez, B. S. Human POMp75 is identified as the pro-oncoprotein TLS/FUS: Both POMp75 and POMp100 DNA homologous pairing activities are associated to cell proliferation. *Oncogene* **18**(31), 4515–4521. https://doi.org/10.1038/sj.onc.1203048 (1999).
20. Sama, R. R., Ward, C. L. & Bosco, D. A. Functions of FUS/TLS from DNA repair to stress response: Implications for ALS. *ASN Neuro* https://doi.org/10.1177/1759091414544472 (2014).
21. Zinszner, H., Sok, J., Immanuel, D., Yin, Y. & Ron, D. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J. Cell Sci.* **110**(Pt 15), 1741–1750 (1997).
22. Dormann, D. & Haass, C. Fused in sarcoma (FUS): An oncogene goes awry in neurodegeneration. *Mol. Cell Neurosci.* **56**, 475–486. https://doi.org/10.1016/j.mcn.2013.03.006 (2013).
23. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D. W. TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **19**(R1), R46-64. https://doi.org/10.1093/hmg/ddq137 (2010).
24. Yang, S., Warraich, S. T., Nicholson, G. A. & Blair, I. P. Fused in sarcoma/translocated in liposarcoma: A multifunctional DNA/RNA binding protein. *Int. J. Biochem. Cell Biol.* **42**(9), 1408–1411. https://doi.org/10.1016/j.biocel.2010.06.003 (2010).
25. Fujii, R. *et al.* The RNA binding protein TLS is translocated to dendritic spines by mGluR5 activation and regulates spine morphology. *Curr. Biol.* **15**(6), 587–593. https://doi.org/10.1016/j.cub.2005.01.058 (2005).
26. Fujii, R. & Takumi, T. TLS facilitates transport of mRNA encoding an actin-stabilizing protein to dendritic spines. *J. Cell Sci.* **118**(Pt 24), 5755–5765. https://doi.org/10.1242/jcs.02692 (2005).
27. Kwiatkowski, T. J. Jr. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**(5918), 1205–1208. https://doi.org/10.1126/science.1166066 (2009).
28. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**(5918), 1208–1211. https://doi.org/10.1126/science.1165942 (2009).
29. Dormann, D. *et al.* ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *EMBO J.* **29**(16), 2841–2857. https://doi.org/10.1038/emboj.2010.143 (2010).
30. Bentmann, E., Haass, C. & Dormann, D. Stress granules in neurodegeneration–lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. *FEBS J.* **280**(18), 4348–4370. https://doi.org/10.1111/febs.12287 (2013).
31. Bosco, D. A. *et al.* Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum. Mol. Genet.* **19**(21), 4160–4175. https://doi.org/10.1093/hmg/ddq335 (2010).
32. Vance, C. *et al.* ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. *Hum. Mol. Genet.* **22**(13), 2676–2688. https://doi.org/10.1093/hmg/ddt117 (2013).
33. Higelin, J. *et al.* FUS Mislocalization and vulnerability to DNA damage in ALS patients derived hiPSCs and aging motoneurons. *Front. Cell Neurosci.* **10**, 290. https://doi.org/10.3389/fncel.2016.00290 (2016).
34. Yang, L., Gal, J., Chen, J. & Zhu, H. Self-assembled FUS binds active chromatin and regulates gene transcription. *Proc. Natl. Acad. Sci. USA* **111**(50), 17809–17814. https://doi.org/10.1073/pnas.1414004111 (2014).
35. Schwartz, J. C. *et al.* FUS is sequestered in nuclear aggregates in ALS patient fibroblasts. *Mol. Biol. Cell* **25**(17), 2571–2578. https://doi.org/10.1091/mbc.E14-05-1007 (2014).
36. Kino, Y. *et al.* Intracellular localization and splicing regulation of FUS/TLS are variably affected by amyotrophic lateral sclerosis-linked mutations. *Nucleic Acids Res.* **39**(7), 2781–2798. https://doi.org/10.1093/nar/gkq1162 (2011).
37. Sun, S. *et al.* ALS-causative mutations in FUS/TLS confer gain and loss of function by altered association with SMN and U1-snRNP. *Nat. Commun.* **6**, 6171. https://doi.org/10.1038/ncomms7171 (2015).
38. Yu, Y. *et al.* U1 snRNP is mislocalized in ALS patient fibroblasts bearing NLS mutations in FUS and is required for motor neuron outgrowth in zebrafish. *Nucleic Acids Res.* **43**(6), 3208–3218. https://doi.org/10.1093/nar/gkv157 (2015).
39. Reber, S. *et al.* Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. *EMBO J.* **35**(14), 1504–1521. https://doi.org/10.15252/embj.201593791 (2016).
40. Gerbino, V., Carri, M. T., Cozzolino, M. & Achsel, T. Mislocalised FUS mutants stall spliceosomal snRNPs in the cytoplasm. *Neurobiol. Dis.* **55**, 120–128. https://doi.org/10.1016/j.nbd.2013.03.003 (2013).
41. Groen, E. J. *et al.* ALS-associated mutations in FUS disrupt the axonal distribution and function of SMN. *Hum. Mol. Genet.* **22**(18), 3690–3704. https://doi.org/10.1093/hmg/ddt222 (2013).
42. Takanashi, K. & Yamaguchi, A. Aggregation of ALS-linked FUS mutant sequesters RNA binding proteins and impairs RNA granules formation. *Biochem. Biophys. Res. Commun.* **452**(3), 600–607. https://doi.org/10.1016/j.bbrc.2014.08.115 (2014).
43. Yamazaki, T. *et al.* FUS-SMN protein interactions link the motor neuron diseases ALS and SMA. *Cell Rep.* **2**(4), 799–806. https://doi.org/10.1016/j.celrep.2012.08.025 (2012).
44. Jutzi, D. *et al.* Aberrant interaction of FUS with the U1 snRNA provides a molecular mechanism of FUS induced amyotrophic lateral sclerosis. *Nat. Commun.* **11**(1), 6341. https://doi.org/10.1038/s41467-020-20191-3 (2020).
45. Constantinescu, R., Constantinescu, A. T., Reichmann, H. & Janetzky, B. Neuronal differentiation and long-term culture of the human neuroblastoma line SH-SY5Y. *J. Neural Transm. Suppl.* **72**, 17–28. https://doi.org/10.1007/978-3-211-73574-9_3 (2007).
46. Lyons, S. M. *et al.* A subset of replication-dependent histone mRNAs are expressed as polyadenylated RNAs in terminally differentiated tissues. *Nucleic Acids Res.* **44**(19), 9190–9205. https://doi.org/10.1093/nar/gkw620 (2016).
47. Hicks, G. G. *et al.* Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat. Genet.* **24**(2), 175–179. https://doi.org/10.1038/72842 (2000).
48. Kuroda, M. *et al.* Male sterility and enhanced radiation sensitivity in TLS(−/−) mice. *EMBO J.* **19**(3), 453–462. https://doi.org/10.1093/emboj/19.3.453 (2000).
49. Wang, W. Y. *et al.* Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. *Nat. Neurosci.* **16**(10), 1383–1391. https://doi.org/10.1038/nn.3514 (2013).
50. Naumann, M. *et al.* Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nat. Commun.* **9**(1), 335. https://doi.org/10.1038/s41467-017-02299-1 (2018).
51. Reber, S. *et al.* CRISPR-Trap: A clean approach for the generation of gene knockouts and gene replacements in human cells. *Mol. Biol. Cell* **29**(2), 75–83. https://doi.org/10.1091/mbc.E17-05-0288 (2018).
52. Simpson, P. B. *et al.* Retinoic acid evoked-differentiation of neuroblastoma cells predominates over growth factor stimulation: An automated image capture and quantitation approach to neuritogenesis. *Anal. Biochem.* **298**(2), 163–169. https://doi.org/10.1006/abio.2001.5346 (2001).

Acknowledgements

We would like to thank Don Cleveland's group for the plasmids encoding wild-type FUS and the ALS-linked FUS mutants R495X and P525L. We would also like to acknowledge Katarzyna Błaszczyk (Adam Mickiewicz University in Poland) for her invaluable help in conducting the ChIP experiment. We thank Aafra Aabdien for the rat primary cortical neurons. We thank George Chennell and Chen Liang of the Wohl Cellular Imaging Centre at King's College London for technical assistance.

Author contributions

A.G. and K.D.R. designed research; A.G., A.W., A.S., J.M., and M.D.R. performed research; A.G., A.S., A.L.N., J.M., and K.D.R. analyzed data; A.G., A.S., A.L.N., K.D.R. wrote the manuscript with J.M., C.E.S., and M.D.R. revising the manuscript.

Funding

The work of Ankur Gadgil, Agnieszka Walczak, Agata Stępień, and Katarzyna Dorota Raczyńska was funded by Polish National Science Centre (NCN) under the grant UMO-2016/21/B/NZ1/00232 to Katarzyna Dorota Raczyńska. The work of Jonas Mechtersheimer, Agnes Lumi Nishimura, Christopher E. Shaw, and Marc-David Ruepp is supported by the UK Dementia Research Institute, which receives its funding from UK DRI, Ltd., funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-91453-3.

Correspondence and requests for materials should be addressed to K.D.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information for

Manuscript title
ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells

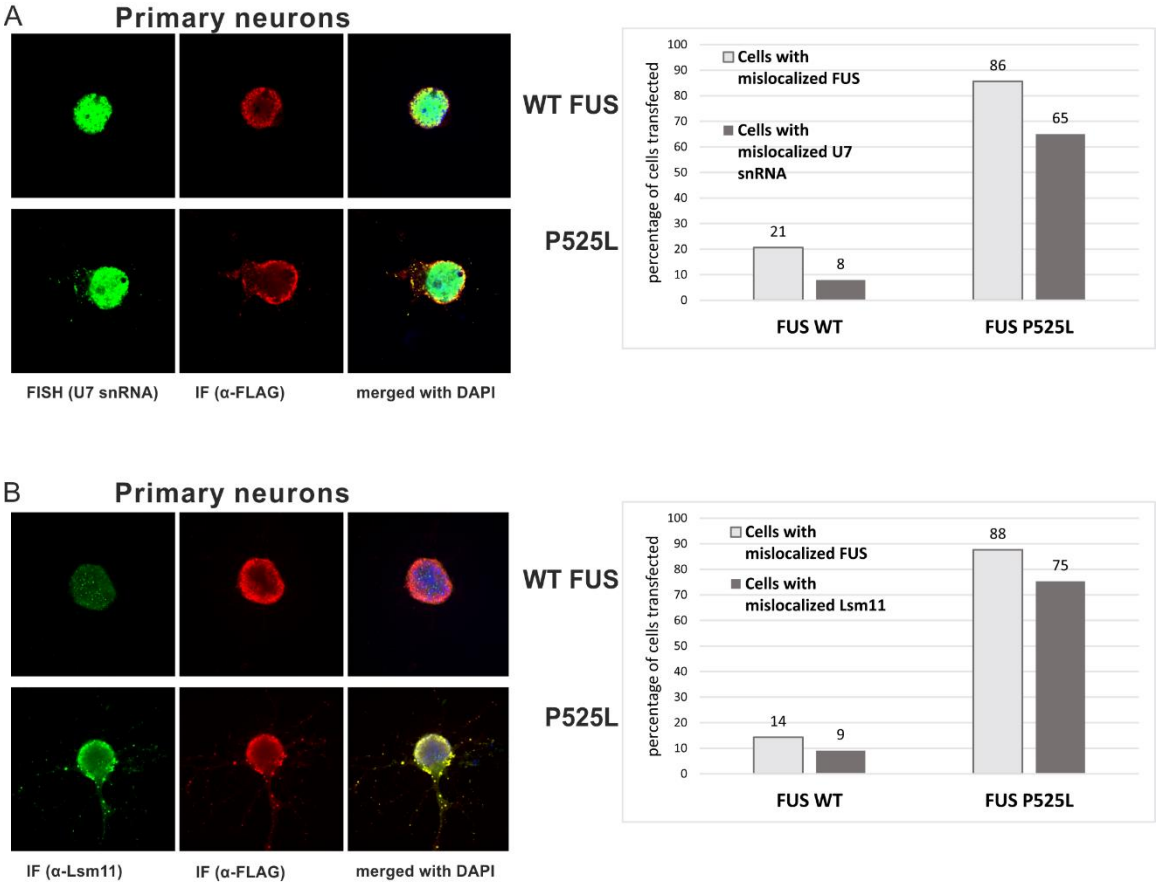
Authors
Ankur Gadgil[1,2], Agnieszka Walczak[1], Agata Stępień[1], Jonas Mechtersheimer[3], Agnes Lumi Nishimura[3], Christopher E. Shaw[3,4], Marc-David Ruepp[3], Katarzyna Dorota Raczyńska[1,2]*

*To whom correspondence should be addressed:
Katarzyna Dorota Raczynska, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland, phone: +48618291901,
e-mail: doracz@amu.edu.pl

This PDF file includes:

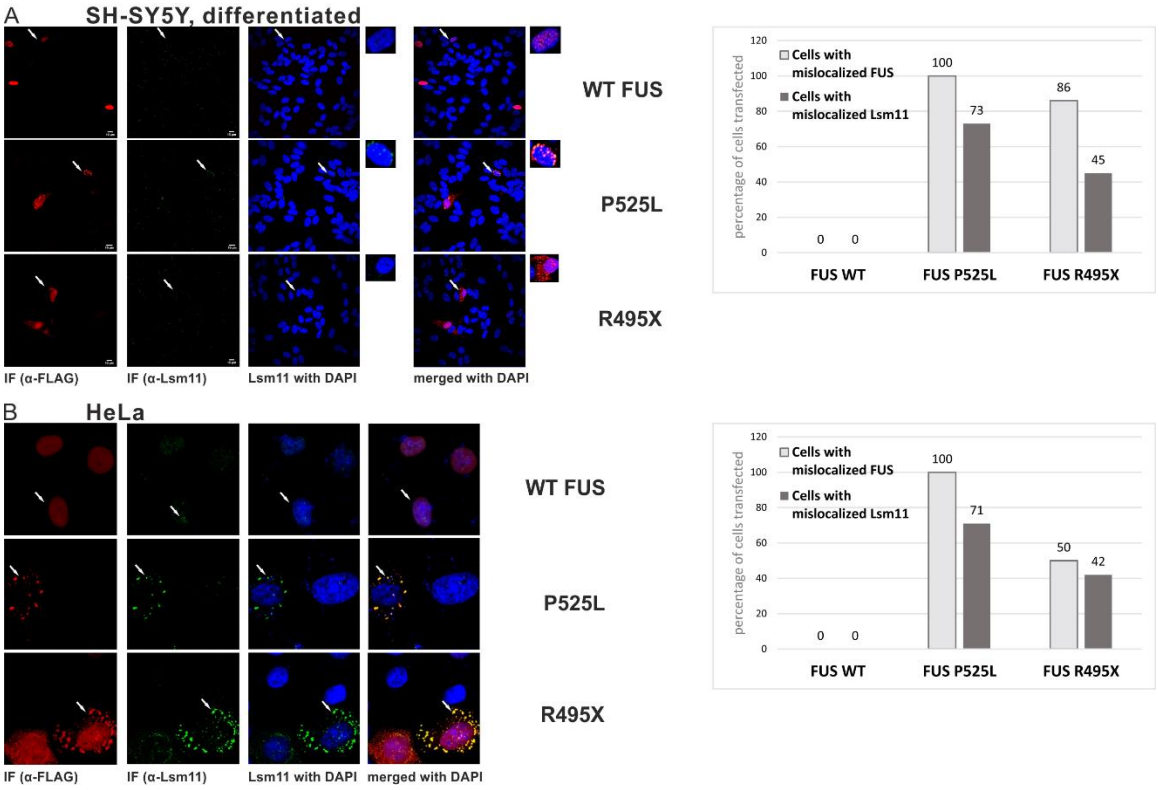
Supplementary
Figures S1 to S4

Supplementary Figure 1



Supplementary Figure S1
Localization of FUS, U7 snRNA and Lsm11 proteins. (A) Fluorescent in-situ hybridization (FISH) using a probe against U7 snRNA in combination with immunofluorescence (IF) using anti-FLAG antibodies, and (B) IF using anti-FLAG and anti-Lsm11 antibodies was performed in primary neurons transfected with FLAG-tagged FUS. DAPI was used for nuclear staining. Graphs show the percentage of transfected cells with mislocalization of FUS and U7 snRNA or Lsm11. WT FUS – cells transfected with plasmids encoding the wild-type *FUS* gene, P525L – cells transfected with plasmid encoding the *FUS* gene with P525L mutation.

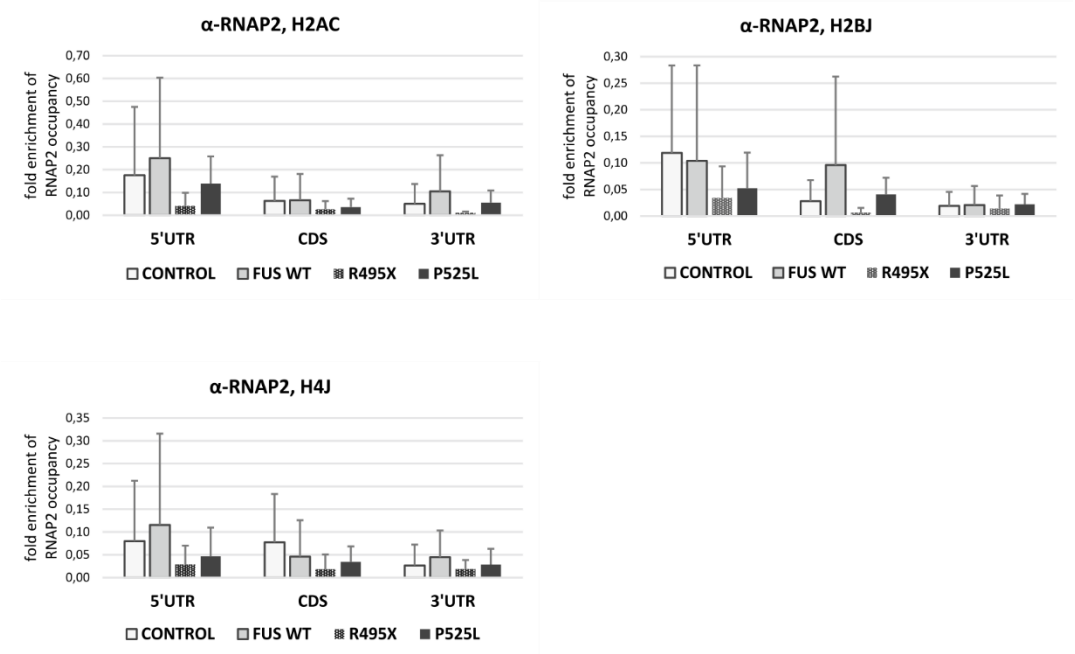
Supplementary Figure 2



Supplementary Figure S2

Localization of FUS and Lsm11 proteins. IF using anti-FUS and anti-Lsm11 antibodies was performed on SH-SY5Y FUS KO (A) and HeLa FUS KO (B) cells. DAPI was used for nuclear staining. Graphs show the percentage of transfected cells with mislocalization of FUS and U7 snRNA. WT FUS – cells transfected with plasmids encoding the wild-type *FUS* gene, P525L and R495X – cells transfected with plasmid encoding the *FUS* gene with P525L and R495X mutations.

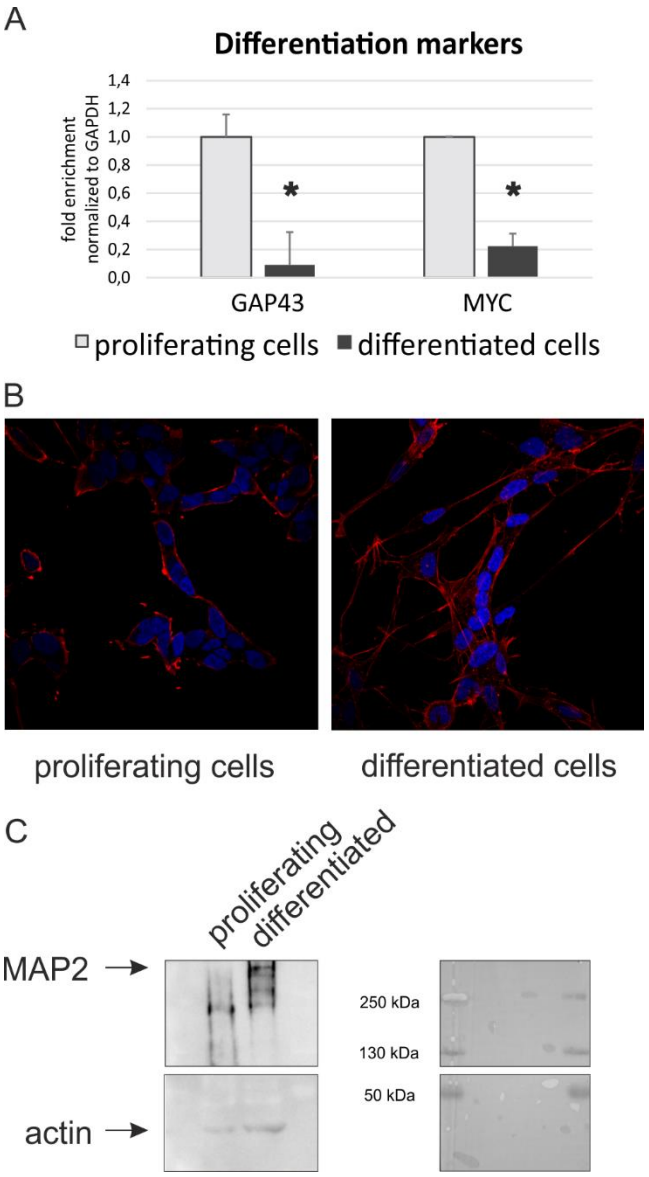
Supplementary figure 3



Supplementary Figure S3

RNA Polymerase II (RNAPII) occupancy on histone genes. ChIP-seq analyses followed by real-time quantitative PCR was performed on SH-SY5Y FUS KO cells (control, non-transfected) and the cells transfected with wild-type FUS (FUS WT), R495X and P525L FUS mutants. Charts represent the mean fold change value (n = 3). Graphs show RNAPII occupancy on the 5'UTR, CDS and 3'UTR for the replication-dependent histone genes H2AC, H2BJ, and H4J. The CT values from the non-transfected control cells were used to nullify the background. Input CT values were used as normalizers. Error bars indicate standard deviations (SD) of three biological replicates. P-values were calculated on percent of input values using Student's t-test.

Supplementary figure 4



Supplementary Figure S4

Confirmation of differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells.
A) The levels of the cellular proliferation markers growth-associated protein 43 (GAP43) and MYC were tested by real-time quantitative PCR. The marker level in proliferating cells was arbitrarily set at 1. Error bars indicate standard deviations (SD) of three biological replicates. P-values were calculated using Student's t-test, and the statistical significance was represented as follows: *P ≤ 0.05.
B) Actin phalloidin staining showing dendrite and axon growth (observed in red) in SH-SY5Y cells differentiated into neuron-like cells; proliferating SH-SY5Y cells show reduced dendrite and axon growth. The nucleus was stained with DAPI.
C) Microtubule-associated protein 2 (MAP2), which is activated in cells differentiated into neuron-

like cells, was detected by Western blots and immunodetection. Actin was used as a loading control

Publication

U7 snRNA: A tool for gene therapy

Received: 9 June 2020 | Revised: 22 January 2021 | Accepted: 9 February 2021

DOI: 10.1002/jgm.3321

REVIEW ARTICLE

WILEY

U7 snRNA: A tool for gene therapy

Ankur Gadgil[1,2] | Katarzyna Dorota Raczyńska[1,2]

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland

[2]Center for Advanced Technology, Adam Mickiewicz University, Poznan, Poland

Correspondence
Katarzyna Dorota Raczynska, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznanskiego 10, 61-614, Poznan, Poland.
Email: doracz@amu.edu.pl

Funding information
Narodowe Centrum Nauki, Grant/Award Numbers: UMO-2016/21/B/NZ1/00232, UMO-2018/30/E/NZ2/00295

Abstract

Most U-rich small nuclear ribonucleoproteins (snRNPs) are complexes that mediate the splicing of pre-mRNAs. U7 snRNP is an exception in that it is not involved in splicing but is a key factor in the unique 3′ end processing of replication-dependent histone mRNAs. However, by introducing controlled changes in the U7 snRNA histone binding sequence and in the Sm motif, it can be used as an effective tool for gene therapy. The modified U7 snRNP (U7 Sm OPT) is thus not involved in the processing of replication-dependent histone pre-mRNA but targets splicing by inducing efficient skipping or inclusion of selected exons. U7 Sm OPT is of therapeutic importance in diseases that are an outcome of splicing defects, such as myotonic dystrophy, Duchenne muscular dystrophy, amyotrophic lateral sclerosis, β-thalassemia, HIV-1 infection and spinal muscular atrophy. The benefits of using U7 Sm OPT for gene therapy are its compact size, ability to accumulate in the nucleus without causing any toxic effects in the cells, and no immunoreactivity. The risk of transgene misregulation by using U7 Sm OPT is also low because it is involved in correcting the expression of an endogenous gene controlled by its own regulatory elements. Altogether, using U7 Sm OPT as a tool in gene therapy can ensure lifelong treatment, whereas an oligonucleotide or other drug/compound would require repeated administration. It would thus be strategic to harness these unique properties of U7 snRNP and deploy it as a tool in gene therapy.

KEYWORDS
adenoassociated virus, gene-editing, gene-therapy, HIV, muscular dystrophy, neurodegenerative disease, RNA-technologies, stem/progenitor cell research

1 | INTRODUCTION

Small nuclear ribonucleoproteins (snRNPs) are complexes composed of small nuclear RNA (snRNA) and proteins in specific structures. Initially, the term snRNA was introduced by Weinberg and Penman in 1968.[1] It was further observed that some of these snRNAs were uridine rich compared to ribosomal or messenger RNAs and thus were described as U snRNAs. These U snRNAs were further numbered (e.g. U1 snRNA) by their order of discovery, and not by size, location or abundance. Valuable research was published by the group led by Joan Steitz in 1980, where they initially reported the role of snRNPs in splicing.[2] The spliceosomes, comprising large complexes that catalyse splicing, are divided into major and minor spliceosomes: U1, U2, U4, U5, U6 and U11, U12, U4atac, U5atac and U6atac snRNP, respectively. Subsequently, our knowledge about U snRNPs that interact with pre-messenger RNA (pre-mRNA) and other proteins forming the spliceosome has expanded.

In simple terms, splicing is the removal of introns and joining of exons from pre-mRNA. Splicing is divided into constitutive and alternative splicing (Figure 1) with alternative splicing considered responsible for protein diversity.[3–6] The interaction between *cis*-acting elements and *trans*-acting factors during alternative splicing define the exons that would be in the mature mRNA.[3] Exonic and intronic splicing enhancers are *cis*-acting elements bound by positive *trans*-active factors such as the nuclear phosphoprotein family (serine/arginine-rich). By contrast, exonic and intronic splicing silencers are bound by negatively acting factors such as heterogeneous ribonucleoproteins (hnRNP).[3] In constitutive splicing, the enhancing elements are dominant, whereas the silencers play a prominent role in controlling alternative splicing. Collaboration between these elements results in the promotion or inhibition of spliceosome assembly at the weak splice sites.[3,7,8]

Exceptionally, one of the U snRNPs, U7 snRNP, is not involved in splicing but is a key factor in the unique 3′ end processing of replication-dependent histone (RDH) pre-mRNAs. Moreover, recent studies showed U7 snRNA as an important tool in therapeutic studies, with reports based on modified U7 snRNP (U7 Sm OPT) targeting splicing to induce efficient skipping or inclusion of selected exons. U7 Sm OPT is designed by changing the histone binding sequence at the 5′ region of U7 snRNA to the complementary sequence of the gene to be modified. Further modifications include changing of U7 snRNP specific proteins, Lsm10 and Lsm11, to the consensus protein ring of spliceosomal snRNPs.

In U7 Sm OPT-based therapy, an antisense oligonucleotide is incorporated into the U7 snRNA. Antisense oligonucleotide and its importance was published for the first time by Stephenson Zamecnik in 1978.[9] Antisense oligonucleotides are short single-stranded nucleotide sequences binding the mRNA to alter gene expression by degradation of the transcript or by inhibition of translation.[10] They do so by various mechanisms, which include ribonuclease H mediated decay of pre-mRNA or steric hindrance. An alternative way of using antisense oligonucleotides is splicing modulation in which these short nucleotides can bind to pre-mRNA splicing elements and disrupt the recognition of splicing regulators.[11] For this purpose, as a tool of manipulation of pre-mRNA splicing, antisense oligonucleotides are used in the U7 Sm OPT therapeutic strategy. Although directly using antisense oligonucleotide has certain benefits, it has some limitations because it is sensitive to degradation, may cause immunoreactivity and usually needs repeated dosage. Thus, incorporating antisense

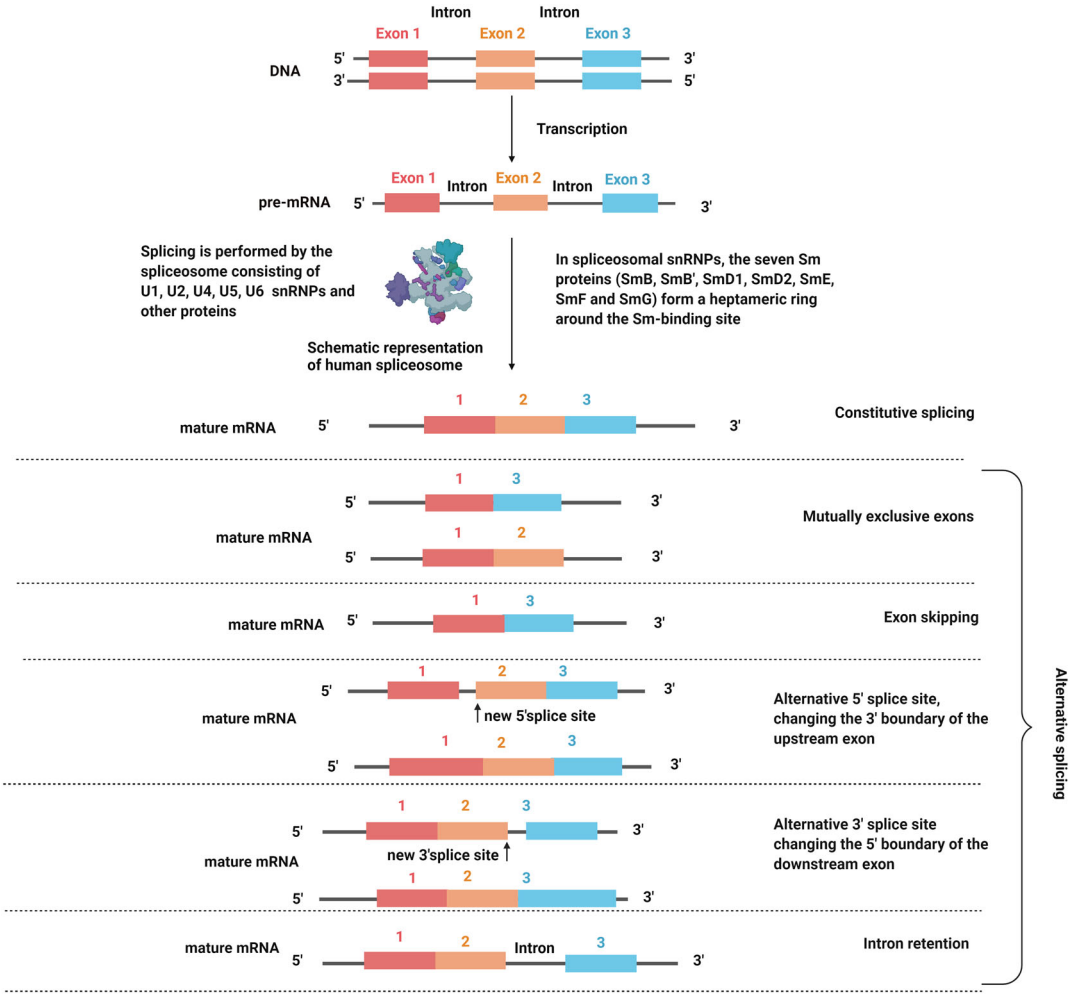


FIGURE 1 Constitutive and alternative splicing

oligonucleotides into U7 snRNP and delivering it using viral vectors overcomes many limitations. U7 Sm OPT snRNP is compact in size, accumulates in the nucleus, is non-toxic to the cells and overcomes the issue of repeated administration of oligonucleotides. Therefore, using U7 snRNA as a tool in gene therapy enables the desired splicing correction to be achieved along with stable antisense oligonucleotide levels. All of these factors are the basis of gene therapy and U7 snRNP clearly appears to have the upper hand compared to other candidates.

The use of U7 snRNP as a tool for gene therapy has shown promise in pre-clinical studies and in human clinical trials. The representative diseases include myotonic dystrophy, Duchenne muscular dystrophy (DMD), amyotrophic lateral sclerosis (ALS), β-thalassemia, HIV-1 infection, and spinal muscular atrophy (SMA). The aim of this review is to discuss the role of U7 snRNA as a tool for gene therapy.

2 | HISTONE GENES AND THEIR PROCESSING

Histone proteins comprise two major classes: canonical RDHs and proteins called histone variants. The genes of RDH are in clusters and are intronless. Interestingly, RDH transcripts are the only mRNAs in metazoan cells that are not polyadenylated.[12,13] By contrast, the genes encoding histone variants contain introns, are not arranged in clusters and produce polyadenylated mRNAs.[13,14] The expression of RDH genes is activated at the S phase of the cell cycle.[15] During the 3′ end processing, RDH pre-mRNAs undergo a single endonucleolytic cleavage after a specific stem-loop structure, located at the 3′ end and recognized by the stem-loop binding protein (SLBP).[16,17] Downstream of the stem-loop structure is a purine-rich conserved sequence, known as the histone downstream element (HDE). HDE is recognized by the 5′ end of U7 snRNA.[18–20] The stem-loop structure, HDE, SLBP and U7 snRNP are respective *cis* and *trans* elements in the post-transcription regulation of RDH mRNAs. These elements determine the cell-cycle regulated expression of canonical histone genes.[15,21,22] The binding of U7 snRNP to HDE aids in the recruitment of other factors involved in processing, known as the histone cleavage complex. Cleavage occurs between the 3′ stem-loop and the HDE, and is catalysed by endonuclease CPSF73[22–27] (Figure 2). It remains unclear which of the mechanisms, polyadenylation or RDH 3′ end processing, developed first in evolution, although it is agreed that U7 snRNP based RDH 3′ end processing was lost in protozoa, plants and fungi, thus making it unique to metazoan cells.[13]

3 | U7 SNRNP

In humans, U7 snRNP comprises 63 nucleotides of U7 snRNA and a protein core consisting of five common Sm proteins, SmB/B′, SmD3, SmE, SmF and SmG, as well as two U7 snRNP-specific Sm-like proteins, Lsm10 and Lsm11.[28] Lsm10 and Lsm11 replace SmD1 and SmD2, respectively, which are found in other U snRNPs.[28–30]

The 5′ part of the U7 snRNA is complementary to the HDE motif of the 3′ UTR of RDH pre-mRNA, and the 3′ region is occupied by a non-canonical Sm binding site (recognized by the Sm/Lsm protein core), followed by a conserved stem-loop secondary structure required for its stability[31] (Figure 2). It has been reported in a

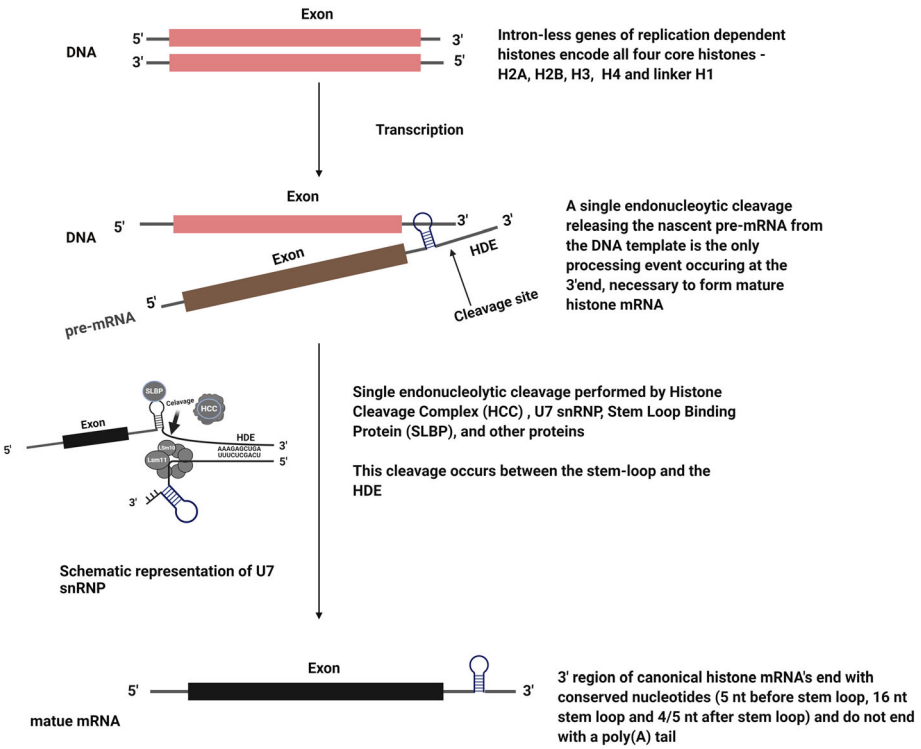


FIGURE 2 Formation of mature RDH transcripts

Drosophila model that blocking the 5′ end of the U7 snRNA with a complementary oligonucleotide specifically blocks the processing of RDH pre-mRNA.[32] This is one such primary finding that is further implemented with respect to using U7 snRNA as a tool for gene therapy.

U7 snRNP follows the same maturation pathway as spliceosomal U snRNPs. During biogenesis, newly transcribed U7 snRNA is exported to the cytoplasm, where it is assembled with core proteins already present in the cytoplasm (SmB/B′, SmD3, SmE, SmF, SmG and Lsm10 and Lsm11) in a process mediated by the survival motor neuron (SMN) complex and protein arginine methyltransferase-5 (PRMT5) complex.[33,34] After assembly, the U7 snRNP is imported into the nucleus and localizes in histone locus bodies.[28,33,35]

In comparison with major spliceosomal snRNPs, the level of U7 snRNP in the cell is very low, reaching approximately 500 molecules.[36] The non-canonical Sm binding site and Lsm10 and Lsm11 are assumed to be critical factors of U7 snRNP, and replacement of the U7 Sm binding site with the consensus sequence leads to a higher production of U7 Sm OPT, reaching a level comparable to that of other U snRNPs.[28,36]

4 | U7 SNRNA IN GENE THERAPY

Factors that make U7 snRNA such an interesting choice as a tool for treating genetic disorders are its small size, good stability and ability to accumulate in the nucleus. However, to further improve its activity, the wild-type U7 snRNA needs to be modified.

Schuemperli and colleagues have shown that the non-canonical Sm binding site of U7 snRNA (AAUUUGUCUAG; U7 Sm WT) can be converted into the consensus sequence derived from major spliceosomal U snRNPs (AAUUUUUGGAG; U7 Sm OPT), leading to the formation of a spliceosomal-type heptameric protein core wrapped around U7 Sm OPT. This change results in the augmented expression of U7 Sm OPT particles, which efficiently accumulate in the nucleus and can be redirected to the sites of pre-mRNA splicing. Moreover, because U7 Sm OPT is unable to bind U7-specific proteins, Lsm10 and Lsm11, U7 Sm OPT particles are non-functional in RDH mRNA processing and thus do not affect this process.[28,36,37] Next, modification of the sequence motif of U7 snRNA, which is complementary to HDE within RDH pre-mRNA, can make the snRNP particle hybridize to almost any RNA sequence within the nucleoplasm. U7 Sm OPT, even when expressed permanently, does not elicit immunological reactions or toxic effects in the cells. This modification enables U7 snRNA, with its original 3′ elements, to be used in gene therapy (Figure 3).

The diseases targeted for U7 snRNA gene therapy are predominantly an outcome of splicing defects. Mutations disrupt splice sites, leading to the creation of cryptic acceptor or donor sites, and aberrant splicing leads to the synthesis of premature terminated or truncated proteins. One disease can be caused by different mutations, thus requiring different therapeutic approaches. Therefore, using antisense oligonucleotides incorporated into U7 Sm OPT proves to be advantageous because it has varied therapeutic applications, such as exon skipping to restore expression, exon skipping to introduce a stop codon, and exon inclusion or displacement of mRNA binding factors (Figure 4).

In modifications made to U7 snRNP for blocking splicing, the binding sites to U1 and U2 snRNP are known to comprise the most effective targets.[38] A stable stem-loop structure as short as 7 bp in an RNA transcript has been shown to abolish enhancer activity, and an exonic splicing enhancer was found to act as intronic splicing enhancer when located in the intron instead of the exon.[39–41] It is further shown that a looping out pre-mRNA leads to exonic sequestration from the rest of pre-mRNA transcript.[3,42]

An initial report based on data obtained from experiments conducted on cell lines mentioned the use of double-target antisense U7 Sm OPT in the treatment of β-thalassemia. This treatment was

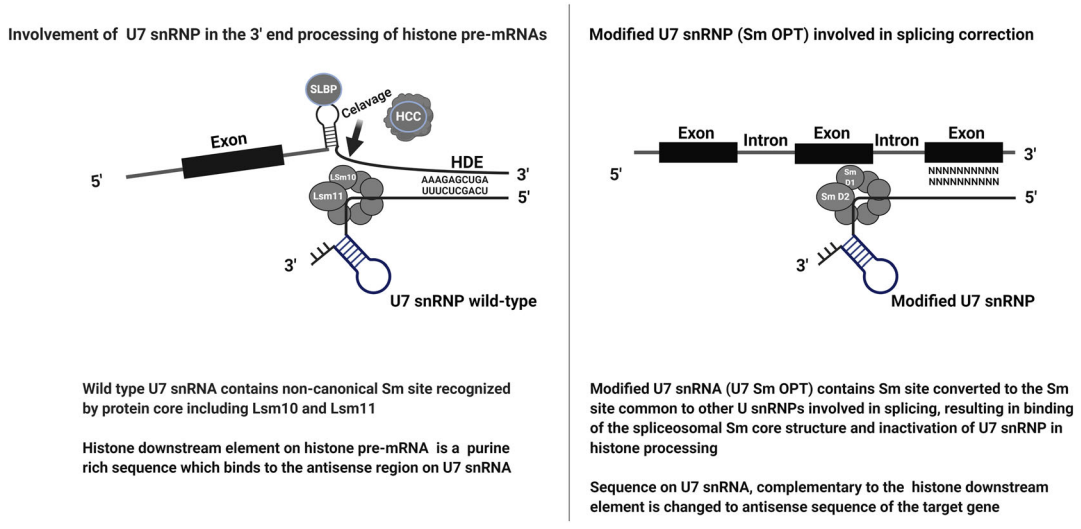


FIGURE 3 Modifications made to U7 snRNP to be used as a tool in gene therapy

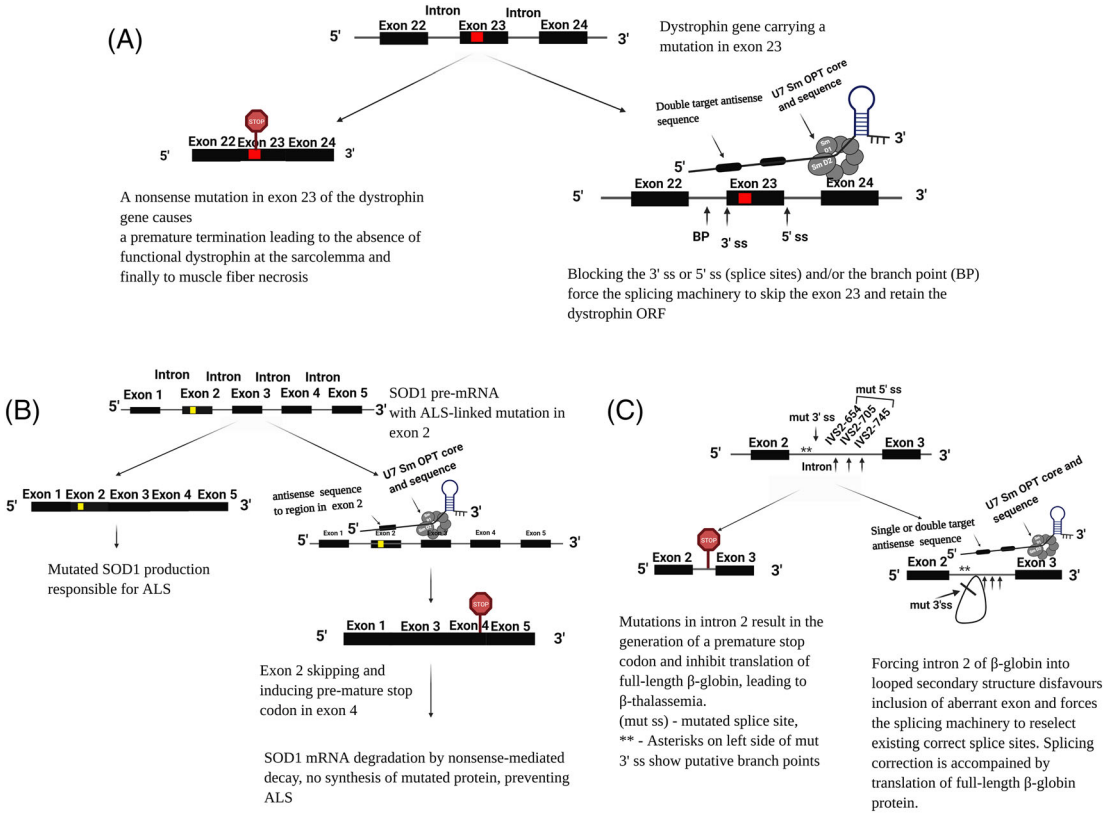


FIGURE 4 Examples of U7 snRNA-based splicing modulation for therapy of (a) DMD, (B) ALS and (C) β-thalassemia

achieved by forcing the aberrant exon into a looped secondary structure and strongly promoting its exclusion from the mRNA.[43] This approach is the fundamental observation reported by the Schuemperli group. Research was conducted on myoblasts obtained from DMD patients,[44] as well as fibroblasts from an SMA patient,[45] and extended to a SMA mouse model.[46] The efficiency of U7 Sm OPT was increased by adding sequences acting as splicing silencers or enhancers.[44–46]

The delivery of engineered constructs is an important aspect in gene therapy and is rapidly evolving subsequent to the study of viral vectors. Adeno-associated viruses (AAV) are non-enveloped, single-stranded DNA, replication defective viruses from the *Parvoviridae* family.[47–49] AAV are currently not known to cause any disease and thus adeno-associated viral vectors can be used to deliver U7 Sm OPT to the nervous system, as well as other targets organs, as a result of the availability of numerous serotypes with different tissue tropism.[50,51] The delivery of U7 Sm OPT by AAV is now broadly used because it ensures high efficiency gene transfer and relatively stable expression.

5 | DISEASES TARGETED BY U7 SNRNA GENE THERAPY

U7 Sm OPT is of therapeutic importance in diseases that are an outcome of splicing defects, such as myotonic dystrophy, DMD, ALS, β-thalassemia, HIV-1 infection and SMA, as described in detail below. Apart from this, using U7 Sm OPT for correction of aberrant splicing of three genes, *PTCH1*, *BRCA1* and *CYP11A*, where mutations cause nevoid basal cell carcinoma syndrome, breast cancer and congenital adrenal insufficiency, respectively, is also reported.[52] A brief summary of the diseases targeted by U7 snRNA gene therapy is provided in Table 1.

5.1 | U7 snRNA therapy for treating myotonic dystrophy

Myotonic dystrophy type 1 caused by expanded CTG repeats in the 3′ UTR of *DMPK* (myotonic dystrophy protein kinase) gene is the most prevalent form of adult muscular dystrophy.[56] Research was conducted on skeletal muscle cells isolated from individuals with myotonic dystrophy 1 using U7 Sm OPT containing 15 CAG repeats. The construct targeted the expanded CUG repeats of mutant *DMPK* transcripts in myotonic dystrophy.[60] This resulted in the subsequent release of mRNA binding factors and MBNL1 from the foci, improving splicing abnormalities and differentiation defects. Thus, permanent targeted degradation of pathogenic *DMPK* mRNA was ensured without affecting the products of wild-type *DMPK* allele over subsequent cell divisions.

5.2 | U7 snRNA therapy for treating DMD

DMD is an X-linked recessive muscle-wasting severe muscular disorder caused by mutations in the dystrophin gene, leading to

TABLE 1 Summary of different research studying the use of U7 snRNP in gene therapy

Condition	Description
Nevoid basal cell carcinoma syndrome[52]	**Cause** – mutation c.584G>A in the *PCTH1* gene at the 3′ end of exon 3 leading to insertion of a 37 bp intronic sequence between exon 3 and exon 4 and premature termination of PCTH1 protein **Study**[52] – used HeLa cell line. U7 Sm OPT contains a sequence complementary to the cryptic donor splice site of *PCTH1* intron 3. U7 Sm OPT was transfected along with a minigene containing the sequence for exon 3, intron 3 and exon 4. Authentic splicing restored
Breast cancer[52]	**Cause** – mutation IVS16+6 T>C, located in intron 16 of *BRCA1* gene leading to a 65 bp insertion of the 5′ end of intron 16 and premature termination of BRCA1 protein **Study**[52] – used HeLa cell line. U7 Sm OPT contains a complementary sequence encompassing activated cryptic splice site in intron 16. Transfected along with a minigene containing the sequence of exon 16, intron 16 and exon 17. Observed partial correction of splicing and premature termination.
Congenital adrenal insufficiency[52]	**Cause** – mutation c.566C>T in the exon 3 of *CYP11A* gene leading to a 61-bp deletion of exon 3 **Study**[52] – used HeLa cell line. U7 Sm OPT contains anti-CYP11A sequence targeting the cryptic donor splice site in exon 3. Transfected along with minigenes containing exon 3. Targeted exon deletion to various degrees rather than restoring splicing was observed
Myotonic dystrophy 1 (DM1)[57]	**Cause** – expanded CTG repeats in the 3′ UTR of the *DMPK* gene **Study**[57] – used skeletal muscle cells isolated from individuals with DM1, containing various CTG expansions, and myoblasts. U7 Sm OPT contains antisense oligonucleotide with 15 CAG repeats. Long-lasting selective destruction of deleterious CUG^{exp} RNAs in DM1 cells and in transduced wild-type myoblasts in a dose-dependent manner was observed
DMD[44,59,60,64,65,70,71]	**Cause(s) –** mutations in the *dystrophin* gene resulting in a premature stop codon leading to the absence of functional dystrophin at the sarcolemma and finally to muscle fibre necrosis. Mutations include, for example: (i) deletion or duplication of exon 2[53]; (ii) nonsense mutation in exon 23 in *mdx* mice[53] (iii) C7360A mutation in exon 51[53]; and (iv) single base change (A to G) in the 3′ splice site of intron 6 of the dystrophin gene, which cause skipping of exon 7 and termination of the ORF within N-terminus of exon 8[62] **Study**[64] – used HEK 293 cells, HSMM cells (wild-type or with exon 2 duplication), fibroblast cell lines, myoblasts and DMD mouse model (Dup2). U7 Sm OPT targeted exon 2 of the *dystrophin* gene contains antisense oligonucleotides targeting splice donor and splice acceptor sites. Exon 2 skipping resulted in generation of a truncated reading frame upstream of the IRES, which led to the synthesis of a functional N-truncated isoform in both human subject–derived cell lines (HEK293, HSMM) and in Dup2 model. Expression of the truncated isoform protected the muscle from contraction-induced injury and corrected muscle force to the same level as that observed in control mice **Study**[59] – used dystrophin deficient mouse models, *mdx* mice. Double target U7 Sm OPT (AAV-U7-SD23/BP22)[59] contains: (i) a 24-nucleotide sequence located across the splicing branching point in intron 22 (BP22) and (ii) a 20-nucleotide sequence in intron 23 that corresponds to the U1 binding region at the donor site (SD23). Reported significant stability of exon 23 skipping in comparison to other U7 snRNA-based constructs in myoblasts or oligonucleotides injected *in vivo* **Study**[65] – used myoblasts from H-2Kb-tsA58 *mdx* mice and C2C12 cells. Double target U7 Sm OPT targeted to block the 5′ splice site and the branch point upstream of exon 23. Using a double target molecule yielded a maximum effect but not an additive effect as in the case of β-thalassemia[65] **Study**[60] – used *mdx* mice and dystrophin/utrophin double-knockout mice. Used the construct as described by Le *et al*.[60] for treatment and showed that arrest of dystrophin process is crucial for maintaining viral genomes. Further reported that nontherapeutic U7 Sm OPT resulted in the loss of AAV genome within three weeks which correlates with dystrophin loss but except in the heart **Study**[44] – bifunctional U7 Sm OPT containing antisense sequences to exon 51, targeting either the acceptor or donor splice site, was transduced into immortalized myoblasts from healthy individuals and from DMD patient carrying a deletion of exons 49 and 50 to restore ORF and dystrophin expression. The second construct carried a 20-nucleotide sequence complementary to exon 51 and a free tail harboring high-affinity binding site for hnRNPA1.[59]The study suggests that either the exonic splicing enhancer, proper internal secondary RNA structure, or combination of two play an important role in correct splicing of exon 51. Remarkably, the hnRNPA1-tailed U7 Sm OPT induced complete exon 51 skipping in patient cells with restoration of dystrophin expression to almost wild-type level **Study**[70] – used double knockout utrophin/dystrophin mice, single dose injected intravenously. The construct described in reference[59] was used. Restored near-normal levels of dystrophin in all muscles, including heart; the treated muscles showed dystrophin detection even 1 year after injection **Study**[71] – Golden retriever muscular dystrophy dog (GRMD), injected intramuscularly with proximal muscles as a test and with contralateral muscles as control. Three-week olf GRMD puppies were injected intramuscularly in their whole left cranial leg muscle compartment. Two GRMD dogs were given high pressure intravenous injection in one of the forelimbs by locoregional delivery. Used constructs AAV1-U7E6 and AAV1-U7E8ref.[71] The constructs contained 2′ O-methylated 22 bp and 24 bp antisense

TABLE 1 (Continued)

Condition	Description
	oligonucleotides against exon-splicing enhancers of exon 6 and exon 8, respectively. Partially functional truncated dystrophin was restored
ALS[80]	**Cause** – single base substitutions such as H48Q,[77] insertion mutation such as 132insTT,[77] CA dinucleotide repeat (D21S223) in exon 2,[78] C to A substitution at codon 41 in exon 2[79] in human *SOD1* gene **Study**[80] – used ALS mouse model B6SJLTg (SOD1*G93A)1Gur/J. U7 Sm OPT contains two steric blocking RNA-based antisense oligonucleotides masking the splicing acceptor site in intron 1 or the exon splicing enhancer in exon 2 to promote efficient skipping of exon 2. U7 Sm OPT was administered using AAV vector by intravenous and intracerebroventricular routes combined. Therapy at birth or at 50 days of age delayed disease onset, prevented weight loss, prevented the decline of neuromuscular junction, and increased life expectancy by 92% and 58%, respectively
β-Thalassemia/HbE disorder HbE; $\alpha_2\beta_2^{26Glu\ to\ Lys83}$	**Cause** – G to A mutation (glutamic acid is changed to lysine) at codon 26 in exon 1 of the human *β-globin* gene, leading to the activation of the cryptic 5′ splice site at codon 25 and generation of aberrantly spliced β^E-*globin* mRNA with a premature termination codon at position 55, leading to the reduction of β^E-globin chains **Study**[83] – used HeLa β^E-cell model and erythroid progenitor cells from β-thalassemia/HbE patients. The U7 Sm OPT (U7 βE4+1) (UCCACUUGCACCUACUUCAACCACC) targeted 102–127 nucleotides of exon 1. Observed near complete splicing correction for 5 months in HeLa cells. Furthermore, observed improved erythroid cell pathology
β-Thalassemia[43,85,86,88,91]	**Cause** – mutation at position 654 (C to T), or 705 (T to G), or 745 (C to G) in intron 2 of *β-globin* gene creates an aberrant 5′ splice site at different positions but a common cryptic 3′ splice site at the nucleotide 579 in the *β-globin* intron 2 **Study**[43] – used HeLa cell line stably expressing plasmids carrying *β-globin* gene with each of the three mutations. Efficient and permanent correction of aberrant splicing and production of β-globin levels similar to cells expressing wild-type gene could be obtained by stable expression of a double target construct U7-BP+5′654. U7–3′[c] and U7–3′/24[c] are other potent candidates **Study**[88] – used HeLa cell line carrying the thalassemic IVS2–705 human *β-globin* gene and cell lines stably expressing U7 Sm OPT, containing a sequence antisense to either the 5′ splice site created by the 705 mutation (U7.5) or to the cryptic 3′ splice site activated in the aberrant splicing pathway (U7.3 and its derivatives). The approach was effective at restoring correct splicing. 65% and 55% correction in cell lines **Study**[86] – used iPSCs derived from mesenchymal stromal cells from a patient with IVS2–654 β-thalassemia mutation. U7 Sm OPT (U7.623) (reference[91]) carrying antisense oligonucleotide (UGUUAUUCUUUAGAAUGGUGCAAAG) targeted the 623 position of intron 2 of the IVS2–654 β-globin pre-mRNA. Erythroblasts generated from these iPSCs expressed ∼80% restoration in splicing compared to healthy cells. Initial report of combined use of U7 Sm OPT with patient-specific iPSCs together to treat patients **Study**[91] – used HeLa cells expressing IVS2–654, IVS2–705 and IVS2–745 human thalassemic *β -globin* genes. U7 Sm OPTU7.623 targeted the 623 position of intron 2 of *β-globin* mRNA and U7.324 targeted cryptic 3′ splice site activated by IVS2–654 mutation in the *β* -globin gene. It increased the levels of correctly spliced *β -globin* mRNA and protein by for at least 6 months. It showed therapeutic potential in haematopoietic stem cells and erythroid progenitor cells from a patient with IVS2–745/IVS2–1 thalassemia. 25-fold correction in patient cells after 12 days of transduction was observed **Study**[85] – used HeLa IVS2–654 cells and erythroid progenitor cells from patients carrying $\beta^{IVS2-654}$ thalassemia mutation. U7. BP+623 targeted the cryptic branch point site and the exonic splicing enhancer in intron 2 of *β-globin* pre-mRNA. Therapeutic potential was shown in both cell models
HIV-1 infection/AIDS[93–95]	**Cause** – HIV type 1 **Study**[93] – used HEK 293 T, HeLa and human T-cell lines CEM-SS or CEM. U7 Sm OPT constructs contain antisense sequences targeting flanking internal HIV-1 exons to reduce tat and rev expression. A merged construct with an additional exonic splicing enhancer and upstream splice donor, named as ESE/SD4, proved 40–50% effectiveness in the context of 'real' HIV-1 replication in human T cells of CEM-SS line **Study**[94] – used HEK 293 T, HeLa, CEM-SS, CEM, P4.2 and $CD4^+$ T cells. Triple combination therapy was used: (i) shRNA targeting nucleotides 330–348 of human *cyclophilin A* mRNA; (ii) shRNA targeting nucleotides 423–443 and 479–498 of the *vif* ORF of HIV; and (iii) U7 Sm OPT ESE/SD4 as in reference.[93] Complete inhibition of viral multiplication in semi-permissive CEM T cells was observed **Study**[95] – used cell lines HEK 293 T, HeLa and human T-cell lines Jurkat and CEM-SS. U7 Sm OPT double target constructs targeted 5′ and 3′ splice site of exon 3 or exon 4 of *cyclophilin A* gene to eliminate either exon or both and inhibit the interaction of cyclophilin A with the capsid protein of the virus. In addition, siRNAs targeting the region between nucleotides 265 to 283 and 330 to 348 of *cyclophilin A* mRNA were used. Treated CEM-SS cell line showed delayed and reduced HIV-1 multiplication
SMA[45,46,104,107,108]	**Cause** – homozygous exon 7 deletion or inactivation of *SMN1* gene. Furthermore, C to T transition in *SMN2* gene located at position 6 of exon 7 causes exon 7 deletion in mRNA, leading to the synthesis of truncated protein

(Continues)

TABLE 1 (Continued)

Condition	Description
	Study[104] – used HeLa cells. U7 Sm OPT contains antisense nucleotide complementary to 36 nucleotides upstream and 34 nucleotides downstream of the intron7/exon 8 junction to target the 3′ splice site of exon 8. Anti-SMN U7 Sm OPT G as the most potent along with anti-SMN U7 Sm OPTs targeting intronic silencer or exon 7 was found. These potent molecules contained ∼20 nucleotides hybridization regions with high G/C content. The length of the most potent anti-SMN RNAs was between 18 to 22 nucleotides. In addition to length and G/C content, competing secondary structures of anti-SMN U7 snRNAs are important. Inclusion of exon 7 is present as long as the expression cassette is retained in the cell **Study** [45] – used HEK 293 T cells, HeLa cells and immortalized human fibroblasts from SMA type 1 patient and a healthy individual. bifunctional U7 Sm OPT (U7-ESE-B) (sequence B – GUGCUCACAUUCCUUAAAU)[54,55] carries an antisense sequence to exon 7 of SMN2 gene together with an exon splicing enhancer or serine arginine repeat. The approach prolonged SMN protein restoration ensuring its localization in gems **Study**[46] – used SMA mice strain FVB.Cg-Tg (SMN2)89Ahmb smn1tm1Msd/J, stock number: 005024 from Jackson laboratories. Used U7-ESE-B construct from reference.[45] Introducing therapeutic U7 snRNA by germline transgenesis resulted in efficient rescue of exon 7 in the most severe SMA mouse model **Study**[105] – used severe SMA mouse strain (Burghes severe model) with stillborn or death by postnatal 4–6 days, stock number: 005204 from Jackson laboratories. Neuromuscular junctions of the diaphragm and soleus muscles having the discrete function in breathing and locomotion were selected for study. Used U7-ESE-B construct from reference.[45] Neuromuscular junction in treated mice showed correct SMN2 splicing, with delayed or no SMA symptoms **Study**[107] – used HeLa cells and SMA1-patient-derived fibroblast. Five U7 Sm OPT constructs were designed to target exon 8 of *SMN* gene. Upregulation of SMN levels similar to control cells was observed **Study**[108] – used HeLa S2 and HeLa cells stably transformed with human SMN2 minigene and SMN Δ7 mice, stock number: 00525 from Jackson laboratories. Used U7-ESE-B construct from reference,[45] delivered by intracerebroventricular injection. Introduction into motoneurons significantly increased life span and improved muscle function. Therapeutic U7 snRNA was also observed to be expressed in the heart and liver

aberrant or reduced levels of the dystrophin protein. The majority of mutations causing DMD are deletion mutations, which disrupt the open reading frame (ORF) and lead to the synthesis of prematurely terminated proteins.[58] Approximately 70% of the mutations lead to the absence of dystrophin protein and thus to severe DMD phenotype.[59] The presence of qualitatively and quantitively altered dystrophin protein results in Becker muscular dystrophy, a mild form of DMD. In Becker muscular dystrophy, by preserving the reading frame, a truncated, partially functional dystrophin protein is produced that contains deletions in the rod domain with an intact N- and C-terminus. For treatment, research is being carried out that aims to use antisense oligonucleotides to induce specific exon skipping at the dystrophin pre-mRNA level. The target for therapy is the central rod-domain because it is known to tolerate large internal deletions. The aim is to convert the "out-of-frame" mutation into an "in frame" mutation, giving rise to internally deleted but still functional dystrophin.[60,61] Moreover, deletion or duplication of exon 2, nonsense mutation in exon 23, C7360A mutation in exon 51, and A to G mutation in the 3′ splice site of intron 6 of dystrophin gene leading to exon 7 skipping are some of the examples of the cause of DMD.[62]

Patients with exon 2 deletions are either asymptomatic or show mild symptoms as a result of the expression of N-truncated isoform of dystrophin.[63] For treatment, two copies each of antisense oligonucleotides targeting splice donor and splice acceptor were introduced into U7 Sm OPT and delivered using AAV. Research conducted on patient fibroblasts, myoblasts and mice using this construct showed efficient skipping of exon 2 resulting in alternative translation initiation in exon 6 (via an internal ribosome entry sequence), leading to the expression of a functional N-truncated protein.[64] Based on these studies, a recent clinical trial "AAV9 U7 snRNA gene therapy to treat boys with DMD exon 2 duplications" is now underway, starting from January 2020, registered at NIH with the clinical trials.gov identifier NCT04240314. This clinical trial focuses on skipping of exon 2 leading to either mRNA containing only one copy of exon 2 mRNA (error-free) or no copy of exon 2, thus giving rise to a highly functional isoform.

One of the earliest research studies on the rescue of DMD via U7 snRNA-mediated exon skipping reported using a 'double-target' U7 Sm OPT to skip exon 23 of *mdx* dystrophin mRNA[59] (Figure 4A). In this study, Goyenvalle and colleagues used U7 Sm OPT equipped with two sequences: (i) a 24-nucleotide sequence complementary to the splicing branch point within intron 22 and (ii) a 20-nucleotide sequence complementary to U1 snRNA binding region within intron 23.[65,66] The engineered U7 snRNA was introduced into skeletal muscles of *mdx* mouse using AAV vector.[67] The observed level and stability of exon skipping were significantly higher than those obtained using other U7 snRNA-based constructs in myoblast cultures or using oligonucleotides injected *in vivo*. Histology performed after treatment showed healthy morphology in corrected muscles.[59] In another study performed on *mdx* mice, AAV vector was encoded with U7 Sm OPT as reported previously by Goyenvalle *et al*.[59] This construct allowed efficient exon 23 skipping, therefore rescuing dystrophin. However, it was observed that, over the course of time, loss of AAV vector genome from muscle fibres led to a decrease in exon-skipping therapeutic effect.[60]

Next, Goyenvalle and colleagues reported bifunctional U7 Sm OPT as a promising tool for DMD therapy.[44,68] The bifunctional

U7 Sm OPT was constructed as follows: first, the U7 Sm OPT was equipped with a splicing silencer sequence complementary to the exonic splicing enhancer located within exon 51 of the human *dystrophin* gene, aiming to promote efficient skipping of exon 51. Furthermore, the U7 Sm OPT carrying splicing silencer was extended by a tail harbouring canonical binding sites for the heterogeneous ribonucleoprotein A1 (hnRNPA1), which naturally contributes to exon skipping. Enhanced exon 51 skipping, induced by this bifunctional U7 Sm OPT, was observed to restore a near wild-type level of dystrophin expression in myoblasts derived from DMD patients. The efficacy of the U7 Sm OPT construct was further confirmed by the observation of positive results after subsequent injection into the tibialis anterior muscle of a mouse model transgenic for the entire human dystrophin locus.[44,69] Based on this strategy, gene therapy employing bifunctional U7 Sm OPT showed positive results with respect to treating muscular dystrophy in both *mdx* mice and golden retriever muscular dystrophy dogs.[61,70]

A study performed on a golden retriever muscular dystrophy dog reported sustained correction of the dystrophic phenotype with U7 Sm OPT carrying antisense oligonucleotides designed for exon skipping therapy. In golden retriever muscular dystrophy dogs, a single base mutation changes the 3′ end of intron 6, leading to skipping of exon 7 and termination within the N-terminal domain of exon 8.[71,72] Using antisense sequence to skip exons 6 and 8, it is possible to restore partially functional truncated dystrophin. The researchers constructed two U7 Sm OPT cassettes for targeting exon 6 or exon 8 skipping. These two different constructs were delivered using AAV vector to treat the forelimb of a golden retriever muscular dystrophy dog by locoregional delivery. No immune rejection was observed after a 5-year follow-up.[71]

The research performed on DMD is very successful and medication such as eteplirsen (brand name Exondys 51) has been approved in the USA to treat DMD.[73] Ataluren is another drug that has been approved for use in certain cases.[74] Another study sheds light on using the *utrophin* gene as therapy for the treatment of DMD in animal models.[75] In this case, miniaturized utrophin, a highly functional and non-immunogenic substitute for dystrophin, is used, thus preventing the most deleterious histological and physiological aspects of DMD in small and large animal models.[75]

5.3 | U7 snRNA therapy for treating ALS

ALS is a neurodegenerative disease involving mostly mutations in *SOD1* (superoxide dismutase 1) gene. The human *SOD1* gene is located on chromosome 21 and encodes the superoxide dismutase 1 enzyme responsible for destroying potentially toxic free superoxide radicals in the body. Over 180 mutations in the *SOD1* gene are described as being involved in ALS.[76] Some of the mutations are single base substitutions such as H48Q,[77] insertion mutation such as 132insTT,[77] CA dinucleotide repeat (D21S223) in exon 2[78] and C to A substitution at codon 41 in exon 2.[79]

One study mentions using U7 Sm OPT to restore the function of SOD1 and prolong the survival of mice in an ALS mouse model.[80] For this, high copy $SOD1^{G93A}$ mice (Jackson SN 2726) were used.[80] U7 Sm OPT containing an antisense sequence targeting exon 2 of the human *SOD1* pre-mRNA was embedded in AAV vectors and administered by intravenous and intracerebroventricular routes combined.[80] It was designed to mask the splicing acceptor site in intron 1 or the exonic splicing enhancer in exon 2 to promote exon 2 skipping. This led to the production of mRNA with a premature stop codon, subjecting it to a RNA degradation pathway and thus reducing the level of toxic protein[80,81] (Figure 4B). This treatment, when initiated at birth or at 50 days of age, lead to a delay of disease onset and an increase in life expectancy of 92% and 58%, respectively, thus indicating the effectiveness of an exon-skipping approach in SOD1-ALS mice. Therefore, in ALS studies, exon 2 is targeted with the aim of producing mature mRNA containing a premature stop codon to inhibit protein synthesis and thus ultimately reduce the level of toxic proteins. This is in contrast to the discussed therapy for DMD (related to exon 2), in which exon 2 carrying the mutation was skipped to produce a wild-type protein or its isoform.

5.4 | U7 snRNA therapy for treating β-thalassemia

β-thalassemia is an autosomal recessive hereditary disease caused by defects in the β- chain, which affect hemoglobin (Hb) production. The hemoglobin E (HbE) $\alpha_2\beta_2 26$ variant, where glutamic acid is changed to lysine as a result of G to A mutation at codon 26 in exon 1 of the human *β-globin* (*HBB*) gene, is a widespread HbE variant found among the Southeast Asian population.[82,83] This substitution activates the cryptic 5′ splice site at codon 25, which generates an aberrantly-spliced β^E*-globin* mRNA containing a premature termination codon at position 55, leading to 16-nucleotide deletion of the 3′ end in exon 1 of *β-globin* pre-mRNA. In a study that focused on the treatment of β-thalassemia patients using engineered U7 snRNA, antisense oligonucleotides, spanning nucleotides from 102 to 130 of *β-globin* mRNA exon 1, were designed to target the 5′ cryptic splice site created by HbE mutation.[83] These oligonucleotides were incorporated into U7 Sm OPT and transiently introduced into cells to test the optimal target site for U7 Sm OPT to attain maximum splicing correction of *β-globin* mRNA. From the series of vectors tested, one of the most active engineered U7 Sm OPT vectors, U7 βE4+1 snRNA, was selected. It was then transduced into the HeLa-β^E model cell line and effectively restored the correctly spliced β^E*-globin* mRNA for at least 5 months. The U7 βE4+1 snRNA was further used to correctly splice β^E*-globin* mRNA in erythroid progenitor cells from β-thalassemia patients, resulting in phenotypic improvements of β-thalassemic erythroid cells.[83]

Aberrant splicing is also associated with mutations in intron 2 of the *β-globin* gene at nucleotide positions 654, 705 or 745. These mutations are not the same (in mutation IVS2–654, C is substituted to T; in mutation IVS2–705, T is substituted to G; in mutation IVS2–745, C is substituted to G) and generate aberrant 5′ splice sites at different positions while activating the same cryptic 3′ splice site at the

579 position, leading to the inclusion of intronic sequences in *β-globin* pre-mRNA.[84–86] Mutations in intron 2 result in the generation of a premature stop codon and inhibit translation into full-length β-globin, leading to β-thalassemia.[84] Antisense oligonucleotides targeting either 5′ or 3′ splice site were shown to repair β-globin mRNA in mammalian cells.[87] Schuemperli and colleagues first reported the use of the U7 Sm OPT-mediated approach in splicing correction of *β-globin* genes carrying mutations IVS2–654, IVS2–705[88] and IVS2–745, expressed from exogenous plasmids in HeLa cells.[43] Some key points from this foundational research were (i) an optimum sequence of approximately 24 nucleotides influenced the efficiency of antisense sequence and (ii) combining two antisense sequences directed against different target sites in intron 2 significantly enhanced the efficiency of splicing correction[43] (Figure 4C).

Mutation IVS2–654 creates an intron inclusion [nucleotides 580–652][89,90] and patients with this IVS2–654 mutation produce a minor amount of correctly spliced *β-globin* mRNAs with severe symptoms of β-thalassemia. An initial report mentioned combining U7 Sm OPT and induced pluripotent stem cells (iPSCs), leading to successful aberrant splicing reduction of the *β-globin* gene in IVS2–654 β-thalassemia disease.[86] In this study, the U7 Sm OPT, called U7.623 snRNA, was constructed as reported by Vacek *et al*.[91] It contained a 25-bp antisense oligonucleotide targeting the 623 position of intron 2 of *β-globin* mRNA[91] to restore correct splicing and protein levels. This U7 Sm OPT was transduced into iPSCs derived from mesenchymal stromal cells from a patient with the IVS2–654 mutation using a lentivirus system. The U7 Sm OPT stably integrated into the genome and maintained splicing correction over many passages. Erythroblasts were further generated from these transduced iPSCs, which expressed high levels of correctly spliced *β-globin* mRNA. These erythroblasts were further differentiated into haematopoietic stem cells and transplanted back into the same patient with a reduced chance of rejection.[86]

Another study reported the restoration of correct $\beta^{IVS2-654}$*-globin* mRNA splicing and hemoglobin production using U7 Sm OPT (U7.BP +623).[85] U7 BP+623 targeted the cryptic branch point and an exonic splicing enhancer. Therapy was performed using thalassemia patient erythroid progenitor cells, as well as HeLa cells carrying the $\beta^{IVS2-654}$ thalassemic mutation. Altogether, this approach showed a positive response by restoring correctly spliced *β-globin* mRNA and correcting thalassemic erythroid cell pathology.[85]

5.5 | U7 snRNA therapy for treating HIV infection

HIV is a single-stranded, positive sense, enveloped RNA retrovirus.[49] HIV cellular tropism is immune cells of the body, with $CD4^+$ T cells being the preferred target.[49,92] HIV-1 exploits alternative splicing, which makes it stand out from other retroviruses. In addition to the standard retroviral genes *gag*, *pol* and *env*, HIV-1 also codes for *tat*, *rev* and *nef*, expressed at an early stage from the integrated provirus genome along with some other proteins. Nef attunes the physiological status of the host cell, whereas Tat and Rev are RNA-binding proteins important for the synthesis of full-length genomic RNA, the regulation of expression of other viral genes and the production of progeny virions.[49] HIV-1 uses alternative splicing in which Tat strongly activates transcription and Rev channels unspliced and partially spliced RNAs into a nucleo-cytoplasmic export.[49]

In a study on HIV-1 inhibition, Schuemperli and colleagues employed U7 Sm OPT equipped with two antisense sequences directed against the *tat* and *rev* pre-mRNA internal exons.[93] Targeting the transcripts leads to exon skipping within *tat* and *rev* ORFs, inhibition of expression of both proteins and thus inhibition of transition into the late phase of the viral replication cycle.[43,93] This U7 snRNA-based approach resulted in the suppression of HIV-1 multiplication by up to 50% in human T cells. It was further suggested that the use of other antivirals or siRNA-based methods in combination with the U7 snRNA-based approach will lead to increased effects, ensuring a robust response and preventing the evolution of viral escape mutants.[93]

The inhibitory effect of the U7 Sm OPT was observed to be stronger for a viral infectivity factor (Vif)-deficient virus. Therefore, a new approach combining RNAi and U7 Sm OPT strategies was developed.[94] Three different cassettes containing antiviral RNAs were expressed from one triple lentiviral vector: shRNA against the host factor cyclophilin A (CyPA), U7 Sm OPT to modulate exon skipping of *tat* and *rev* viral pre-mRNA, and shRNA targeting the *Vif* ORF. This approach dramatically affects HIV-1 infection and allows for a strong inhibition of HIV-1 multiplication in human T cell lines. Moreover, all three therapeutic RNAs exhibit antiviral effects at early stages of the viral replication cycle. Using these strategies, no changes in cell proliferation or morphology have been observed, suggesting a low toxicity to cells.[94]

Inhibition of HIV-1 replication by targeting the cellular protein cyclophilin A is another strategy based on U7 Sm OPT.[95] Cyclophilin A is known to engage Gag polyprotein in HIV-1 replication and is essential for HIV-1 infectivity.[96] Importantly, this protein is neither crucial for early development, nor essential for cell survival.[97] In this research, U7 Sm OPT was exploited as an alternative splicer by inserting appropriate antisense sequences directed against the 3′ and 5′ splice sites of exons 3 and 4 of *cyclophilin A* pre-mRNA, respectively. Using this approach, efficient skipping of these exons was achieved, which in turn disturbed the protein ORF. As a consequence, a significantly reduced level of cyclophilin A protein and delayed multiplication of HIV-1 in T cells were observed.[95] The ability of the cells to sustain HIV-1 replication was impaired. In the same study, RNAi was used in addition to U7 Sm OPT to reduce cyclophilin A. Adding RNAi along with U7 Sm OPT further improved the survival of T cells. It was further hypothesized that using the aforesaid strategy to modify haematopoietic stem cells instead of T cells would be beneficial to achieve a long-lasting effect.[95]

5.6 | U7 snRNA therapy for treating SMA

SMA is an autosomal recessive neuromuscular disorder characterized by the degeneration of anterior horn cells of the spinal cord, leading

to symmetrical muscle weakness and atrophy.[98] SMA is linked to a genetic mutation in the *SMN1* gene,[99] which is unable to correctly code for the SMN protein, either as a result of a deletion at exon 7 or other point mutations.[100] Simultaneously, a *SMN1* paralogous gene, the *SMN2* gene, has identical coding potential but, because of single point mutation in exon 7, an exonic splicing enhancer is disrupted or an exon silencer element is created. This favours alternative splicing at the junction of intron 6 to exon 8, resulting in truncated protein (SMNΔ7).[101,102] It is thus hypothesized that correcting the splicing of exon 7 would help in the treatment of SMA. Because exclusion of exon 7 is a major reason for the pathogenesis of SMA and individuals with SMA usually lack the correct *SMN1* gene and have a functional *SMN2* gene, modulating the inclusion of exon 7 in the human *SMN2* gene represents an attractive therapeutic outlook.[103]

Exchange of the anti-histone region of U7 snRNA with a region complementary to the intron7/exon 8 junction of *SMN2* pre-mRNA increases the inclusion of exon 7.[104] Indeed, the administration of such anti-SMN U7 Sm OPT to a HeLa cell line expressing the *SMN2* minigene favoured exon 7 inclusion in *SMN2* mRNA and resulted in a higher concentration of full-length SMN protein.[104]

Next, a more permanent strategy using bifunctional U7 snRNA therapy to correct *SMN2* pre-mRNA splicing was developed.[45] In this approach, the U7 Sm OPT was modified both with an antisense sequence binding to exon 7 of *SMN2* pre-mRNA and with the splicing enhancer sequence, improving the recognition of the targeted exon. Addition of the splicing enhancer sequence to the U7 Sm OPT strongly stimulated *SMN2* exon 7 recognition and boosted exon inclusion, performing an almost complete re-inclusion of *SMN2* exon 7 in all systems tested. In fibroblasts from a type I SMA patient, this bifunctional approach induced prolonged expression of SMN protein, which was correctly localized within the cell.[45]

Schuemperli and colleagues confirmed the hypothesis that promoting *SMN2* exon 7 inclusion might help to prevent or benefit treatment of SMA.[46] They used a U7 snRNA-mediated splicing modulation approach to rescue a severe mouse model of SMA. The U7 Sm OPT construct was developed to target the 3′ end of exon 7 and carried an exon splicing enhancer to attract stimulatory splicing factors. When introduced into the mouse by germline transgenesis, this U7 Sm OPT construct proved efficient in rescuing a severe SMA mouse model. This process reversed the pathology to milder forms and, in some cases, restored neuromuscular functionality and life expectancy.[46]

U7 Sm OPT construct capable of stimulating the inclusion of *SMN2* exon 7 in severe SMA mouse models was also designed by Voigt *et al*.[105] They checked the effectiveness of the therapy using a severe mouse model where the mice contained two copies of the human *SMN2* transgene and one copy of the mouse *Smn* gene.[107] For this study, the diaphragm and soleus muscle, which have discrete functions in breathing and locomotion, respectively, were selected. It was observed that the diaphragm displayed prominent adverse morphological alterations in a severe mouse model. No significant changes were observed in either muscle of SMA mice undergoing treatment. Research was further extended by investigating neuromuscular junctions (NMJs). Unfortunately, the development of NMJs was incomplete in a severe SMA mouse model, as well as in treated individuals.[105]

To restore full-length SMN, Geib and Hertel[107] modified the histone binding region of the U7 snRNA to a sequence complementary to the 3′ splice site of *SMN* exon 8 to inhibit its recognition and induce exon 7 inclusion. The experiments were carried out in both HeLa cells and fibroblasts derived from SMA-1 patients. AAV vector delivered U7 Sm OPT therapy resulted in a higher rate of exon 7 inclusion, as well as an increase in the level of full-length SMN protein from 33% to 60%. Moreover, the percentage of gems significantly increased after AAV/U7 Sm OPT therapy, ensuring proper SMN nuclear localization.[107]

Recently, Schuemperli and colleagues performed somatic therapy on an SMNΔ7 mouse model (SMN1−/−, SMN2+/+, and SMNΔ7+/+).[108] The mice that would normally survive for only 14–17 days survived for a median of 150 days when treated with the U7 Sm OPT construct, prepared as described previously.[45,46] The SMNΔ7 mice, when injected with U7 Sm OPT, showed therapeutic U7 Sm OPT expressed in both the heart and liver. The mice grew smaller in size compared to wild-type mice and showed mild SMA symptoms, although they were able to feed and drink without any assistance. Additionally, the neuromuscular junctions were less affected compared to SMA models. In conclusion, somatic gene therapy using U7 Sm OPT is feasible for correcting the skipped exon and significantly improving the SMA phenotype.[108]

The research examples described above demonstrate that splicing correction of *SMN* mRNA by the U7 Sm OPT construct works for both transgenic and somatic gene therapy. Nusinersen (SPINRAZA™), an antisense oligonucleotide that acts as a modulator for alternative splicing of the *SMN2* gene and converts its functionality to the *SMN* gene, has been approved against SMA.[109–112] In May 2019, the US Food and Drug Administration approved Zolgensma, which represents the first gene therapy that has been approved to treat children aged less than 2 years who are diagnosed with SMA.[113,114]

5.7 | Limitations and benefits

U7 snRNP is not a spliceosomal U snRNP and thus requires additional modifications to function as a splicing modifier. For example, in U7 Sm OPT, the binding of unique Lsm proteins is changed to canonical Sm proteins. However, it can generate errors in the 3′ end processing of U7 snRNA and lead to the production of a truncated (20 nucleotides shorter) version of U7 snRNA that is unlikely to mediate correct pre-mRNA splicing. Because the higher level of truncated U7 snRNA was observed in mice compared to humans, it has been suggested that the processing of U7 Sm OPT differs in these two species and thus should be taken into consideration.[115,116] It must also be considered that the snRNA processing capacity of the cell is limited, and that going past this limit may lead to additional U7 snRNA by-products, which comprise both 3′ processed and unprocessed species. Understanding the formation of these potentially inactive by-products is necessary to improve the potency of U7 snRNP-based gene

therapy.[116] A limitation for using U7 snRNP in gene therapy is also the amount of vector that can be administered. Administering modified snRNA using the AAV vector requires very high doses that are potentially toxic.[116] Moreover, a study performed on *mdx* mice emphasizes how the immune response against AAV leads to the production of neutralizing antibodies and thus forbids reinjection, indicating it as a major drawback of the AAV vector approach.[60] Successful delivery is not only a hurdle for administering U7 Sm OPT but for all antisense therapies in general.[117,118] Even though all these difficulties exist, viral vectors including AAV vectors and bioreducible lipid nanoparticles comprise useful candidates for delivery, creating a fertile ground for antisense oligonucleotide incorporated U7 snRNA-based therapeutics.[50,119–121]

Some questions that remain to be answered are, for example, how to decide a time point to start the therapy? In the case of diseases that develop in a later stage of life, would individual screening for genetic diseases be practised? Nowadays, screening for breast cancer is performed for individuals with genetic background, and maybe this approach can be expanded to other diseases as well. The economic constraint is a hurdle that needs to be overcome. A further consideration is the number of U7 snRNA copies required to achieve the desired correction. High dosage would ensure improved correction but can pose a threat of potential side effects if the dosage increases beyond the cellular processing capacity. Low dosage treatment can have no or minimum side effects but, at the same time, it may not provide maximum correction. It can thus lead to repeated visits to the therapy centre. These decisions are very complicated because they involve physical, economical, psychological and, importantly, emotional factors concerning the affected individual and its family.

Neurodegenerative diseases still remain to be fully understood, with a number of different hypotheses describing the process and cause of neurodegeneration. SMN deficiency, found in SMA and ALS, leads to lower levels and defects in U snRNP assembly, including U7 snRNP. Therefore, the biogenesis and function of U7 snRNP can paradoxically be disrupted in diseases, and thus used as a potential tool in gene therapy.

RDH pre-mRNA 3′ end processing is undoubtedly the fundamental role of U7 snRNP but, by introducing controlled changes at the histone binding region and the binding site of Sm/Lsm proteins, it can be used as an effective tool for gene therapy. A modification such as converting U7 snRNA into U7 Sm OPT facilitates the accumulation of U7 snRNP in the nucleus and, at the same time, the particle is no longer functional for the processing of *RDH* pre-mRNAs, making it an advantage.[28,36] The benefits of using U7 snRNP in gene therapy include its compact size and the ability to accumulate in the nucleus without causing any toxic effects to the cells. U7 cassettes will fit into smaller vectors and do not carry the risk of transgene misregulation because the therapeutic RNA only corrects the expression of an endogenous gene that is controlled by its own regulatory elements. The benefit of U7 snRNA-based gene therapy is also that a single treatment can ensure a lifelong therapeutic effect, while an oligonucleotide or other drug/compound would require repeated administration. U7 snRNP is a unique resource available exclusively to metazoan cells and it would be strategic to use our resources to the fullest.

ACKNOWLEDGEMENTS

This work was supported by: Polish National Science Centre, Narodowe Centrum Nauki under Grant UMO-2016/21/B/NZ1/00232 and UMO-2018/30/E/NZ2/00295 (to KDR).

AUTHOR CONTRIBUTIONS

Ankur Gadgil prepared the main text of the manuscript, as well as Figures 1 to 4 and Table 1. Katarzyna Dorota Raczynska revised the manuscript.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article because no datasets were generated or analysed during the current study.

ORCID

Ankur Gadgil https://orcid.org/0000-0001-5155-4079
Katarzyna Dorota Raczyńska https://orcid.org/0000-0002-3712-3528

REFERENCES

1. Weinberg RA, Penman S. Small molecular weight monodisperse nuclear RNA. *J Mol Biol*. 1968;38:289-304.
2. Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA. Are snRNPs involved in splicing? *Nature*. 1980;283:220-224.
3. Wang Y, Liu J, Huang BO, et al. Mechanism of alternative splicing and its regulation. *Biomed Reports*. Mar 2015;3:152-158.
4. Gilbert W. Why genes in pieces? *Nature*. 1978;271:501-501.
5. Consortium CeS. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998;282:2012-2018.
6. International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. *Nature*. Oct 21 2004;431:931-945.
7. Wang Z, Burge CB. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*. 2008;14:802-813.
8. Wang Z, Xiao X, Van Nostrand E, Burge CB. General and specific functions of exonic splicing silencers in splicing control. *Mol Cell*. Jul 7 2006;23:61-70.
9. Stephenson ML, Zamecnik PC. Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proc Natl Acad Sci U S a*. 1978;75:285-288.
10. Shen X, Corey DR. Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res*. 2018;46:1584-1600.
11. Dominski Z, Kole R. Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc Natl Acad Sci U S A*. 1993;90:8673-8677.
12. Marzluff WF. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Curr Opin Cell Biol*. Jun 2005;17(3):274-280.
13. Davila Lopez M, Samuelsson T. Early evolution of histone mRNA 3′ end processing. *RNA*. 2008;14:1-10.
14. Ausio J. Histone variants--the structure behind the function. *Brief Funct Genomic Proteomic*. 2006;5:228-243.

15. Harris ME, Bohni R, Schneiderman MH, Ramamurthy L, Schumperli D, Marzluff WF. Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol Cell Biol*. 1991;11:2416-2424.
16. Muller B, Schumperli D. The U7 snRNP and the hairpin binding protein: key players in histone mRNA metabolism. *Semin Cell Dev Biol*. 1997;8:567-576.
17. Wang ZF, Whitfield ML, Ingledue TC 3rd, Dominski Z, Marzluff WF. The protein that binds the 3′ end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev*. 1996;10:3028-3040.
18. Mowry KL, Steitz JA. Identification of the human U7 snRNP as one of several factors involved in the 3′ end maturation of histone premessenger RNA's. *Science*. 1987;238:1682-1687.
19. Bond UM, Yario TA, Steitz JA. Multiple processing-defective mutations in a mammalian histone pre-mRNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. *Genes Dev*. 1991;5:1709-1722.
20. Schaufele F, Gilmartin GM, Bannwarth W, Birnstiel ML. Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3′ end of H3 messenger RNA. *Nature*. 1986;323:777-781.
21. Luscher B, Stauber C, Schindler R, Schumperli D. Faithful cell-cycle regulation of a recombinant mouse histone H4 gene is controlled by sequences in the 3′-terminal part of the gene. *Proc Natl Acad Sci U S a*. 1985;82:4389-4393.
22. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(a) tail. *Nat Rev Genet*. 2008;9:843-854.
23. Dominski Z, Yang XC, Marzluff WF. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell*. 2005;123:37-48.
24. Yang XC, Sabath I, Debski J, et al. A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3′-end processing. *Mol Cell Biol*. 2013;33:28-37.
25. Marzluff WF, Koreski KP. Birth and death of histone mRNAs. *Trends Gen*. 2017;33:745-759.
26. Romeo V, Schumperli D. Cycling in the nucleus: regulation of RNA 3′ processing and nuclear organization of replication-dependent histone genes. *Curr Opin Cell Biol*. 2016;40:23-31.
27. Ruepp MD, Schweingruber C, Kleinschmidt N, Schumperli D. Interactions of CstF-64, CstF-77, and symplekin: implications on localisation and function. *Mol Biol Cell*. 2011;22:91-104.
28. Schumperli D, Pillai RS. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol Life Sci*. 2004;61:2560-2570.
29. Pillai RS, Will CL, Luhrmann R, Schumperli D, Muller B. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO j*. 2001;20:5470-5479.
30. Pillai RS, Grimmler M, Meister G, et al. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev*. 2003;17:2321-2333.
31. Gilmartin GM, Schaufele F, Schaffner G, Birnstiel ML. Functional analysis of the sea urchin U7 small nuclear RNA. *Mol Cell Biol*. 1988;8:1076-1084.
32. Dominski Z, Yang XC, Purdy M, Marzluff WF. Cloning and characterization of the drosophila U7 small nuclear RNA. *Proc Natl Acad Sci U S a*. 2003;100:9422-9427.
33. Azzouz TN, Pillai RS, Dapp C, et al. Toward an assembly line for U7 snRNPs: interactions of U7-specific Lsm proteins with PRMT5 and SMN complexes. *J Biol Chem*. 2005;280:34435-34440.
34. Tisdale S, Lotti F, Saieva L, et al. SMN is essential for the biogenesis of U7 small nuclear ribonucleoprotein and 3′-end formation of histone mRNAs. *Cell Rep*. 2013;5:1187-1195.
35. Dominski Z, Marzluff WF. Formation of the 3′ end of histone mRNA: getting closer to the end. *Gene*. 2007;396:373-390.
36. Grimm C, Stefanovic B, Schumperli D. The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO j*. 1993;12:1229-1238.
37. Stefanovic B, Hackl W, Luhrmann R, Schumperli D. Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into Xenopus oocytes. *Nucleic Acids Res*. 1995;23:3141-3151.
38. Morcos PA. Achieving targeted and quantifiable alteration of mRNA splicing with Morpholino oligos. *Biochem Biophys Res Commun*. 2007;358:521-527.
39. Jin Y, Yang Y, Zhang P. New insights into RNA secondary structure in the alternative splicing of pre-mRNAs. *RNA Biol*. 2011;8:450-457.
40. Liu W, Zhou Y, Hu Z, et al. Regulation of splicing enhancer activities by RNA secondary structures. *FEBS Lett*. Nov 5 2010;584:4401-4407.
41. McManus CJ, Graveley BR. RNA structure and the mechanisms of alternative splicing. *Curr Opin Genet Dev*. 2011;21:373-379.
42. Nasim FU, Hutchison S, Cordeau M, Chabot B. High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5′ splice site selection in support of a common looping out and repression mechanism. *RNA*. 2002;8:1078-1089.
43. Suter D, Tomasini R, Reber U, Gorman L, Kole R, Schumperli D. Double-target antisense U7 snRNAs promote efficient skipping of an aberrant exon in three human beta-thalassemic mutations. *Hum Mol Genet*. 1999;8:2415-2423.
44. Goyenvalle A, Babbs A, van Ommen GJ, Garcia L, Davies KE. Enhanced exon-skipping induced by U7 snRNA carrying a splicing silencer sequence: promising tool for DMD therapy. *Mol Ther*. 2009;17:1234-1240.
45. Marquis J, Meyer K, Angehrn L, Kampfer SS, Rothen-Rutishauser B, Schumperli D. Spinal muscular atrophy: SMN2 pre-mRNA splicing corrected by a U7 snRNA derivative carrying a splicing enhancer sequence. *Mol Ther*. 2007;15:1479-1486.
46. Meyer K, Marquis J, Trub J, et al. Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. *Hum Mol Genet*. 2009;18:546-555.
47. Lai CM, Lai YK, Rakoczy PE. Adenovirus and adeno-associated virus vectors. *DNA Cell Biol*. 2002;21:895-913.
48. Akli S, Caillaud C, Vigne E, et al. Transfer of a foreign gene into the brain using adenovirus vectors. *Nat Genet*. 1993;3:224-228.
49. Fields BN, Knipe DM, Howley PM. Chapter 49 - Human Immunodeficiency Viruses: Replication. In: *Fields virology*. Vol. 2. 6th ed. 2013:1502-1560.
50. Balakrishnan B, Jayandharan GR. Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. *Curr Gene Ther*. 2014;14:86-100.
51. Nonnenmacher M, Weber T. Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther*. 2012;19:649-658.
52. Uchikawa H, Fujii K, Kohno Y, et al. U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites. *J Hum Genet*. 2007;52:891-897.
53. Vieitez I, Gallano P, Gonzalez-Quereda L, et al. Mutational spectrum of Duchenne muscular dystrophy in Spain: study of 284 cases. *Neurologia*. 2017;32:377-385.
54. Skordis LA, Dunckley MG, Yue B, Eperon IC, Muntoni F. Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proc Natl Acad Sci U S a*. 2003;100:4114-4119.

55. Cartegni L, Krainer AR. Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nat Struct Biol*. 2003;10: 120-125.
56. Brook JD, McCurrach ME, Harley HG, et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell*. 1992;68:799-808.
57. Francois V, Klein AF, Beley C, et al. Selective silencing of mutated mRNAs in DM1 by using modified hU7-snRNAs. *Nat Struct Mol Biol*. 2011;18:85-87.
58. Goyenvalle A, Davies KE. Engineering exon-skipping vectors expressing U7 snRNA constructs for Duchenne muscular dystrophy gene therapy. *Methods Mol Biol*. 2011;709:179-196.
59. Goyenvalle A, Vulin A, Fougerousse F, et al. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science*. 2004; 306:1796-1799.
60. Le Hir M, Goyenvalle A, Peccate C, et al. AAV genome loss from dystrophic mouse muscles during AAV-U7 snRNA-mediated exon-skipping therapy. *Mol Ther*. 2013;21:1551-1558.
61. Harper SQ, Hauser MA, DelloRusso C, et al. Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nat Med*. 2002;8:253-261.
62. White SJ, Aartsma-Rus A, Flanigan KM, et al. Duplications in the DMD gene. *Hum Mutat*. 2006;27:938-945.
63. Lattanzi A, Duguez S, Moiani A, et al. Correction of the exon 2 duplication in DMD myoblasts by a single CRISPR/Cas9 system. *Mol Ther Nucleic Acids*. 16 2017;7:11-19.
64. Wein N, Vulin A, Falzarano MS, et al. Translation from a DMD exon 5 IRES results in a functional dystrophin isoform that attenuates dystrophinopathy in humans and mice. *Nat Med*. 2014;20: 992-1000.
65. Brun C, Suter D, Pauli C, et al. U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. *Cell Mol Life Sci*. 2003;60:557-566.
66. Mann CJ, Honeyman K, McClorey G, Fletcher S, Wilton SD. Improved antisense oligonucleotide induced exon skipping in the mdx mouse model of muscular dystrophy. *J Gene Med*. 2002;4: 644-654.
67. Chao H, Liu Y, Rabinowitz J, Li C, Samulski RJ, Walsh CE. Several log increase in therapeutic transgene delivery by distinct adeno-associated viral serotype vectors. *Mol Ther*. 2000;2:619-623.
68. Goyenvalle A. Engineering U7snRNA gene to reframe transcripts. *Methods Mol Biol*. 2012;867:259-271.
69. Bremmer-Bout M, Aartsma-Rus A, de Meijer EJ, et al. Targeted exon skipping in transgenic hDMD mice: a model for direct preclinical screening of human-specific antisense oligonucleotides. *Mol Ther*. 2004;10:232-240.
70. Goyenvalle A, Babbs A, Wright J, et al. Rescue of severely affected dystrophin/utrophin-deficient mice through scAAV-U7snRNA-mediated exon skipping. *Hum Mol Genet*. 2012;21: 2559-2571.
71. Vulin A, Barthelemy I, Goyenvalle A, et al. Muscle function recovery in golden retriever muscular dystrophy after AAV1-U7 exon skipping. *Mol Ther*. 2012;20:2120-2133.
72. Sharp NJ, Kornegay JN, Van Camp SD, et al. An error in dystrophin mRNA processing in golden retriever muscular dystrophy, an animal homologue of Duchenne muscular dystrophy. *Genomics*. 1992;13: 115-121.
73. Syed YY. Eteplirsen: first global approval. *Drugs*. Nov 2016;76(17): 1699-1704.
74. D'Ambrosio P, Orsini C, Nigro V, Politano L. Therapeutic approach with Ataluren in Duchenne symptomatic carriers with nonsense mutations in dystrophin gene. Results of a 9-month follow-up in a case report. *Acta Myol*. 2018;37:272-274.
75. Song Y, Morales L, Malik AS, et al. Non-immunogenic utrophin gene therapy for the treatment of muscular dystrophy animal models. *Nat Med*. 2019;25:1505-1511.
76. Pansarasa O, Bordoni M, Diamanti L, Sproviero D, Gagliardi S, Cereda C. SOD1 in amyotrophic lateral sclerosis: 'ambivalent' behavior connected to the disease. *Int J Mol Sci*. 2018;19: 1345-1357.
77. Orrell RW, Habgood JJ, Malaspina A, et al. Clinical characteristics of SOD1 gene mutations in UK families with ALS. *J Neurol Sci*. 1999; 169:56-60.
78. Rosen DR, Siddique T, Patterson D, et al. Mutations in cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*. 1993;362:59-62.
79. Niu Q, Yi Y, Sun X, et al. The G41D mutation in the superoxide dismutase 1 gene is associated with slow motor neuron progression and mild cognitive impairment in a Chinese family with amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. 2016;87:788-789.
80. Biferi MG, Cohen-Tannoudji M, Cappelletto A, et al. A new AAV10-U7-mediated gene therapy prolongs survival and restores function in an ALS mouse model. *Mol Ther*. 2017;25:2038-2052.
81. Ward AJ, Norrbom M, Chun S, Bennett CF, Rigo F. Nonsense-mediated decay as a terminating mechanism for antisense oligonucleotides. *Nucleic Acids Res*. 2014;42:5871-5879.
82. Fucharoen S, Weatherall DJ. The hemoglobin E thalassemias. *Cold Spring Harb Perspect Med*. 2012;2:a011734.
83. Preedagasamzin S, Nualkaew T, Pongrujikorn T, et al. Engineered U7 snRNA mediates sustained splicing correction in erythroid cells from beta-thalassemia/HbE patients. *Biochem Biophys Res Commun*. 2018; 499:86-92.
84. Cheng TC, Orkin SH, Antonarakis SE, et al. Beta-thalassemia in Chinese: use of in vivo RNA analysis and oligonucleotide hybridization in systematic characterization of molecular defects. *Proc Natl Acad Sci U S a*. 1984;81:2821-2825.
85. Nualkaew T, Jearawiriyapaisarn N, Hongeng S, Fucharoen S, Kole R, Svasti S. Restoration of correct beta (IVS2–654)-globin mRNA splicing and HbA production by engineered U7 snRNA in beta-thalassaemia/HbE erythroid cells. *Sci Rep*. 2019;9:7672-7679.
86. Phanthong P, Borwornpinyo S, Kitiyanant N, et al. Enhancement of beta-globin gene expression in Thalassemic IVS2-654 induced pluripotent stem cell-derived erythroid cells by modified U7 snRNA. *Stem Cells Transl Med*. 2017;6:1059-1069.
87. Sierakowska H, Sambade MJ, Agrawal S, Kole R. Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci U S a*. Nov 12 1996;93: 12840-12844.
88. Gorman L, Suter D, Emerick V, Schumperli D, Kole R. Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proc Natl Acad Sci U S a*. 1998;95:4929-4934.
89. Zhang JZ, Cai SP, He X, et al. Molecular basis of beta thalassemia in South China. Strategy for DNA analysis. *Hum Genet*. 1988;78:37-40.
90. Huang SZ, Zeng FY, Ren ZR, et al. RNA transcripts of the beta-thalassaemia allele IVS-2-654 C-->T: a small amount of normally processed beta-globin mRNA is still produced from the mutant gene. *Br J Haematol*. 1994;88:541-546.
91. Vacek MM, Ma H, Gemignani F, Lacerra G, Kafri T, Kole R. High-level expression of hemoglobin a in human thalassemic erythroid progenitor cells following lentiviral vector delivery of an antisense snRNA. *Blood*. 2003;101:104-111.
92. Cunningham AL, Donaghy H, Harman AN, Kim M, Turville SG. Manipulation of dendritic cell function by viruses. *Curr Opin Microbiol*. 2010;13:524-529.
93. Asparuhova MB, Marti G, Liu S, Serhan F, Trono D, Schumperli D. Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing tat and rev exon skipping. *J Gene Med*. 2007;9:323-334.

94. Asparuhova MB, Barde I, Trono D, Schranz K, Schumperli D. Development and characterization of a triple combination gene therapy vector inhibiting HIV-1 multiplication. *J Gene Med*. 2008;10:1059-1070.
95. Liu S, Asparuhova M, Brondani V, Ziekau I, Klimkait T, Schumperli D. Inhibition of HIV-1 multiplication by antisense U7 snRNAs and siRNAs targeting cyclophilin a. *Nucleic Acids Res*. 2004;32:3752-3759.
96. Thali M, Bukovsky A, Kondo E, et al. Functional association of cyclophilin a with HIV-1 virions. *Nature*. 1994;372:363-365.
97. Colgan J, Asmal M, Luban J. Isolation, characterization and targeted disruption of mouse ppia: cyclophilin A is not essential for mammalian cell viability. *Genomics*. 2000;68:167-178.
98. Wirth B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum Mutat*. 2000;15:228-237.
99. Brzustowicz LM, Lehner T, Castilla LH, et al. Genetic mapping of chronic childhood-onset spinal muscular atrophy to chromosome 5q11.2–13.3. *Nature*. 1990;344:540-541.
100. Lefebvre S, Burglen L, Reboullet S, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;80:155-165.
101. Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S a*. 1999;96:6307-6311.
102. Lewin B. Genes for SMA: multum in parvo. *Cell*. 1995;80:1-5.
103. Nlend Nlend R, Meyer K, Schumperli D. Repair of pre-mRNA splicing: prospects for a therapy for spinal muscular atrophy. *RNA Biol*. 2010;7:430-440.
104. Madocsai C, Lim SR, Geib T, Lam BJ, Hertel KJ. Correction of SMN2 pre-mRNA splicing by antisense U7 small nuclear RNAs. *Mol Ther*. 2005;12:1013-1022.
105. Voigt T, Meyer K, Baum O, Schumperli D. Ultrastructural changes in diaphragm neuromuscular junctions in a severe mouse model for spinal muscular atrophy and their prevention by bifunctional U7 snRNA correcting SMN2 splicing. *Neuromuscul Disord*. 2010;20:744-752.
106. Monani UR, Sendtner M, Coovert DD, et al. 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*. 2000;9:333-339.
107. Geib T, Hertel KJ. Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. *PLoS One*. 2009;4:e8204.
108. Odermatt P, Trub J, Furrer L, Fricker R, Marti A, Schumperli D. Somatic therapy of a mouse SMA model with a U7 snRNA gene correcting SMN2 splicing. *Mol Ther*. 2016;24:1797-1805.
109. Hoy SM. Nusinersen: first global approval. *Drugs*. 2017;77:473-479.
110. Zanetta C, Nizzardo M, Simone C, et al. Molecular therapeutic strategies for spinal muscular atrophies: current and future clinical trials. *Clin Ther*. 2014;36:128-140.
111. Bennett CF, Krainer AR, Cleveland DW. Antisense oligonucleotide therapies for neurodegenerative diseases. *Annu Rev Neurosci*. 2019;42:385-406.
112. Mercuri E, Darras BT, Chiriboga CA, et al. Nusinersen versus sham control in later-onset spinal muscular atrophy. *N Engl J Med*. 2018;378:625-635.
113. Zolgensma - one-time gene therapy for spinal muscular atrophy. *Med Lett Drugs Ther*. 2019;61:113-114.
114. Mendell JR, Al-Zaidy S, Shell R, et al. Single-dose gene-replacement therapy for spinal muscular atrophy. *N Engl J Med*. 2017;377:1713-1722.
115. d'Arqom A, Nualkaew T, Jearawiriyapaisarn N, Kole R, Svasti S. Engineered U7 small nuclear RNA restores correct beta-globin pre-mRNA splicing in mouse beta (IVS2–654)-Thalassemic erythroid progenitor cells. *Hum Gene Ther*. 202010.1089/hum.2020.145. Online ahead of print
116. Eckenfelder A, Tordo J, Babbs A, Davies KE, Goyenvalle A, Danos O. The cellular processing capacity limits the amounts of chimeric U7 snRNA available for antisense delivery. *Mol Ther Nucleic Acids*. 2012;1:e31-e38.
117. Weyermann J, Lochmann D, Zimmer A. Comparison of antisense oligonucleotide drug delivery systems. *J Control Release*. 2004;100:411-423.
118. Juliano R, Alam MR, Dixit V, Kang H. Mechanisms and strategies for effective delivery of antisense and siRNA oligonucleotides. *Nucleic Acids Res*. 2008;36:4158-4171.
119. Yang L, Ma F, Liu F, Chen J, Zhao X, Xu Q. Efficient delivery of antisense oligonucleotides using bioreducible lipid nanoparticles in vitro and in vivo. *Mol Ther Nucleic Acids*. 2020;19:1357-1367.
120. Imbert M, Dias-Florencio G, Goyenvalle A. Viral vector-mediated antisense therapy for genetic diseases. *Gene*. 2017;8:51-69.
121. Roberts TC, Langer R, Wood MJA. Advances in oligonucleotide drug delivery. *Nat Rev Drug Discov*. 2020;19:673-694.

How to cite this article: Gadgil A, Raczyńska KD. U7 snRNA: A tool for gene therapy. *J Gene Med*. 2021;e3321. https://doi.org/10.1002/jgm.3321

Author statement

AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that the research article 'Gadgil, A. *et al.* **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3', is a part of my PhD dissertation. I, Ankur Gadgil, along with my supervisor Katarzyna Dorota Raczyńska designed the research. I, performed combined FISH and immunofluorescence staining to analyze localization of FUS, Lsm11 and U7 snRNA in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure, Fig. S2), I analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), I assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3), and I confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4). I wrote the manuscript, prepared figures, I performed data analyses and I did the editorial work.

Poznan, 15th April 2022

Name: Ankur Gadgil

AUTHOR STATEMENT

for the review article 'U7 snRNA: A tool for gene therapy'

I declare that the review article 'Gadgil A, Raczyńska KD. **U7 snRNA: A tool for gene therapy**. J Gene Med. 2021 Apr;23(4):e3321. doi: 10.1002/jgm.3321', is a part of my PhD dissertation. I, Ankur Gadgil, conceptualized the idea, wrote the manuscript, prepared all the figures Fig.1, Fig.2, Fig.3, Fig.4, and Table 1, and performed all the editorial work.

Poznan, 15th April 2022

Name: Ankur Gadgil

Co-author statement

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al.* **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022
Name: Agnieszka Walczak (AW)
Signature:

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al.* **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022
Name: Agata Stepien (AS)
Signature: Agata Stepien

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al*. **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022
Name: Jonas Mechtersheimer (JM)

Signature: 

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al*. **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022

Name: Agnes Lumi Nishimura (ALN)

Signature:

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al*. **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022

Name: Christopher E. Shaw (CES)

Signature:

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al.* **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), contributed in writing the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in HeLa FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

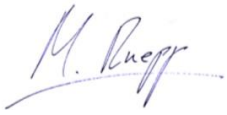
ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited, and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022
Name: Marc-David Ruepp (MDR)

Signature: 

CO-AUTHOR STATEMENT

for the research article 'ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells'

I declare that I am aware that the work in the research article 'Gadgil, A. *et al.* **ALS-linked FUS mutants affect the localization of U7 snRNP and replication-dependent histone gene expression in human cells**. *Sci Rep* 11, 11868 (2021). doi: /10.1038/s41598-021-91453-3' of which I am a co-author, is a part of the PhD dissertation by Ankur Gadgil who designed the study, performed combined FISH and immunofluorescence staining to analyze co-localization of FUS protein and U7 snRNA and FUS and Lsm11 proteins in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 1 A, and Fig. 1 B, and Supplementary Figure, Fig. S2), analyzed replication-dependent histone gene expression in cells transfected with plasmids carrying FUS WT and ALS-linked FUS mutations (Fig. 2), assessed the RNA Polymerase II (RNAPII) occupancy on histone genes by chromatin immunoprecipitation assay (Supplementary Figure. S3) and confirmed the differentiation of SH-SY5Y FUS KO proliferating cells to neuron-like cells using qPCR, immunofluorescence and western blot (Supplementary Figure Fig. S4), wrote the manuscript, prepared figures, performed data analyses and editorial work.

AW performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and Lsm11 proteins in SH-SY5Y FUS KO cells (Supplementary Figure Fig. S2)

AS performed combined FISH and immunofluorescence staining to analyze the co-localization of FUS and U7 snRNA in HeLa FUS KO cells (Fig. 1 C), and contributed in writing of the manuscript.

JM performed staining and analyzed data for the experiment of combined FISH and immunofluorescence staining to analyze the co-localization of FUS with U7 snRNA and Lsm11 in primary neurons (Supplementary Figure S1) and contributed to the revision of the manuscript.

ALN performed data analyses and contributed to the writing process.

CES provided material, lab space and contributed to the revision of the manuscript.

MDR designed and constructed the U7 snRNA probe for FISH, was involved in performing combined FISH and immunofluorescence staining (Supplementary Fig. S2) and contributed to the revision the manuscript.

KDR designed research, supervised all experiments, analyzed the data, edited and finalized all the images, shared funding and wrote the manuscript.

Date: 15th April 2022

Name: Katarzyna Dorota Raczyńska (KDR)

Signature

Raczyńska

CO-AUTHOR STATEMENT
for the review article 'U7 snRNA: A tool for gene therapy'

I declare that the review article 'Gadgil A, Raczyńska KD. **U7 snRNA: A tool for gene therapy**. J Gene Med. 2021 Apr;23(4):e3321. doi: 10.1002/jgm.3321', is a part of the PhD dissertation for my student Ankur Gadgil. In this article, he conceptualized the idea, wrote the manuscript, prepared all the figures (Fig.1-4), Table 1 and performed all the editorial work.

KDR supervised in conceptualizing the idea, finalizing all the figures and the table, writing the manuscript and editing it.

Date: 15th April 2022
Name: Katarzyna Dorota Raczyńska

Signature