PhD Program in Translational and Molecular Medicine DIMET

University of Milano-Bicocca School of Medicine and School of Science

Proteomics characterization of the role of FUS/TLS in human cells: from pre-mRNA splicing to DNA damage response

Coordinator: Prof. Andrea Biondi

Tutor: Prof. Silvia Maria Luisa Barabino

Giuseppe Filosa - No.760753

XXVII Cycle, Academic year 2013-2014

Table of contents

Chapter 2

Proteomic characterization of FUS/TLS interactome reveals a new role in pre-mRNA splicing

Chapter 1

General Introduction

1.1 FUS/TLS

FUS/TLS or FUS (fused in sarcoma/translocated in liposarcoma) is a protein that was firstly identified in the context of a chimeric oncoprotein in myxoid liposarcomas (MLS). In MLS and other cancers, chromosomal translocation events result in aberrant transcription factors, formed by a fusion between the N-terminus of FUS and the DNA-binding domain of an endogenous transcription factor such as CHOP (C/EBP homology protein), ERG (ETS-related gene), ATF1 (activation transcription factor 1), BBF2H7 (BBF2 human homolog on chromosome 7) (Ichikawa et al., 1994; Rabbitts et al., 1993; Storlazzi et al., 2003; Waters et al., 2000).

FUS has also been found mutated in Amyotrophic Lateral Sclerosis (ALS) (Vance et al., 2009) and Fronto Temporal Lobar Degeneration (FTLD) (Neumann et al., 2009). ALS is a progressive motor neuron disease that culminates in paralysis and death within 3 to 5 years of symptom onset. A majority (about 90%) of ALS cases are sporadic in nature with an unknown aetiology, while the remaining 10% of cases are attributed to inheritable genetic defects; mutations in the gene encoding FUS account for 3% to 5% of inherited, or familial, ALS (fALS) (Vance et al., 2009). To date, it is not clear whether ALS-linked mutations cause a loss of normal FUS functions (i.e. inhibition of FUS' nuclear functions) or induce a gain of toxic function (i.e. toxicity of the cytoplasmic FUS aggregates), or a combination of both. FTLD is characterized by progressive decline in behaviour, personality, or language, symptoms that are attributed to the

degeneration of the frontal and temporal lobes. The 25% to 50% of cases have a family history, and disease pathology is often characterized by neuronal inclusions of disease-specific proteins (Rademakers and Mackenzie, 2013). Although mutated FUS is detected in both fALS-FUS and FTLD-FUS, the majority of disease-causing mutations within FUS are associated with fALS-FUS cases (Mackenzie et al., 2010).

1.2 FUS structure and biological functions

FUS is a 53 kDa protein belonging to the FET family of proteins, which in mammals also includes EWS (Ewing sarcoma) and TAF15 (TATA box-binding protein-associated factor 68kDa). This family represents a class of proteins that function at all stages of gene expression from transcription to protein translation thanks to their ability to interact with DNA, RNA, and proteins (Tan and Manley,

2009). FET proteins comprise several conserved domains: an N-terminal glutamine-glycine-serine-tyrosine (QGSY)-rich domain, a glycine-rich region, a RNA-recognition motif (RRM), a zinc-finger domain, and a C-terminal arginine-glycine-glycine (RGG)-rich domains [Figure 1]. FET proteins are ubiquitously expressed in most tissues and are predominantly localized in the nucleus although they engage in nucleocytoplasmic shuttling and thus play important roles also in the cytoplasm (Andersson et al., 2008; Zinszner et al., 1997). Specifically for FUS, the functional domains allow to perform a large array of biological functions, which span from RNA metabolism, to transcription, DNA damage repair (DDR), stress control, and others [Figure 2].

1.2.1 The binding of FUS to RNA molecules

In a proteomic study aimed to identify protein complexes assembled on mRNA precursors, the

hnRNP P2 was identified by mass spectrometry as the product of the FUS gene, thus suggesting the role of FUS as a RNA-binding protein (Calvio et al., 1995).

Early SELEX (Systematic Evolution of Ligands by EXponential enrichment) and EMSA (Electrophoretic Mobility Shift Assays) analyses demonstrated that recombinant FUS selectively binds RNAs containing a GGUG motif with nanomolar affinity *in vitro* (Lerga et al., 2001). However, while one report confirms an enrichment of GUGGU-rich sequences bound by FUS in naive mouse and in human brains (Lagier-Tourenne et al., 2012), others report limited sequence specificity (Rogelj et al., 2012). One study also demonstrated that FUS binds AU-rich RNA stem loops structures, with 15-fold higher affinity than a GGU repeat RNA (Hoell et al., 2011). Therefore, FUS is able to bind GU-rich sequences both *in vitro* and *in vivo*, but it

appears that such sequences are neither sufficient nor required for interactions between FUS and RNA. Moreover, a consistent finding across most of the studies is the binding of FUS to long introns (Hoell et al., 2011; Ishigaki et al., 2012; Lagier-Tourenne et al., 2012; Rogelj et al., 2012). FUS exhibits preferential binding toward the 5' end of long introns, indicative of FUS deposition on nascent transcripts during transcription elongation (Lagier-Tourenne et al., 2012; Rogelj et al., 2012). FUS-binding sites were also identified within the 3' UTR of target genes (Lagier-Tourenne et al., 2012).

As described earlier, the FUS protein has a single RRM motif that is generally known for binding to the RNA, even though it was also described as mediator of DNA and protein interactions (Blatter et al., 2011). This domain is important for the binding of FUS to RNA molecules and it was described that mutagenesis of four phenylalanine

residues (F305L, F341L, F359L, and F368L) within the RRM of human FUS effectively abolishes interactions between FUS and RNA (Daigle et al., 2013). There are additional motifs within FUS that also bind RNA. Experiments with isolated recombinant FUS domains expressed in mammalian cells (Bentmann et al., 2012) do not support a strong interaction between RNA and the FUS RRM, but instead point to the RGG1/2 and the zinc finger domains as mediating tight-binding interactions between FUS and RNA. These observations suggest that FUS interactions with RNA are complex. Indeed, it may be that multiple domains of FUS contribute simultaneously to the recognition of the mRNA (Schwartz et al., 2013).

Recent genome-wide approaches have aimed to identify all transcripts bound and potentially regulated by FUS. One PAR-CLIP analysis compared transcripts bound by WT FUS and two ALS-linked

FUS variants (R521G and R521H) that were predominately expressed in the cytoplasm of HEK293 cells (Hoell et al., 2011). Thousands of transcripts were cross-linked to WT FUS as well as to FUS variants. Interestingly, 80% of transcripts bound by mutants of FUS were also bound by wild type FUS. The authors propose that transcripts bound exclusively by FUS mutants might result from the cytoplasmic mis-localization of the mutants and not because the ALS-linked mutations themselves physically alter the binding between FUS and RNA, supporting a gain of toxic function for mutant FUS with respect to RNA binding and processing. Gene categories related to proteostasis, including the unfolded protein response (UPR) and endoplasmic reticulum (ER), as well as protein binding and mitochondrion were overrepresented amongst transcripts uniquely bound by cytoplasmic FUS mutants in this study (Hoell et al., 2011). However,

UPR-associated transcripts were also bound to wild type FUS in an RNA immunoprecipitation and chip (RIP–CHIP) analysis in mouse NSC-34 cells, likely because this protocol enriched for FUS in the cytoplasmic fraction; additional functional categories and pathways for FUS mRNA targets in NSC-34 cells included regulation of transcription, cell cycle, ribosome genesis, spliceosome assembly, RNA processing, and DNA repair (Colombrita et al., 2012). Despite the different methodologies and cell types employed, Colombrita and colleagues and Hoell and colleagues reported a 63% overlap in the FUS mRNA targets between these two studies (Hoell et al., 2011). The effect of FUS on the expression of genes important for neuronal function, including synaptic genes, was revealed through similar analyses in mouse and human brain tissue and may bear more relevance to neurodegenerative disorders caused by FUS.

Additional mRNA targets of FUS that may be relevant to ALS and FTLD include those encoding SOD1, medium and heavy chains of neurofilament (NEFL, NEFM, NEFH), glutamate transporter (EAAT2), ubiquilin 1 and 2, and the FUS protein itself. Importantly, a comparison of FUS mRNA targets in mouse versus human brain revealed a relatively high degree (69%) of overlap, indicating that the FUS–RNA interactomes are conserved between these species (Lagier-Tourenne et al., 2012).

It has also been described that FUS can bind long non-coding RNA that are involved in subnuclear structures formation, in particular NEAT-1 and NEAT-2 (Hoell et al., 2011). These RNAs, together with several protein factors, such as SFPQ, contributes to form RNA scaffold on which the paraspeckles nucleate (Naganuma and Hirose, 2013; Naganuma et al., 2012). Paraspeckles are

particular nuclear structures that have been proposed to control several biological processes, such as stress responses and cellular differentiation, even though their specific functions still remain unclear (Nakagawa and Hirose, 2012). Interestingly, ALS-linked FUS mutations alter the paraspeckles assembly and FUS deficiency leads to loss of paraspeckles, thus confirming the fundamental role of FUS in the formation of these structures (Shelkovnikova et al., 2014).

1.2.2 pre-mRNA splicing and FUS

Pre-mRNA splicing is a process in which introns are removed from an mRNA precursor. Splicing consists of two trans-esterification steps, each involving a nucleophilic attack on terminal phosphodiester bonds of the intron. In the first step this is carried out by the 2' hydroxyl of the branch point (usually adenosine) and in the second step by the 3'

hydroxyl of the upstream (5') exon. This process is carried out in the spliceosome, a dynamic molecular machine the assembly of which involves sequential binding and release of small nuclear ribonucleoprotein particles (snRNPs) and numerous protein factors as well as the formation and disruption of RNA–RNA, protein– RNA and protein–protein interactions. Briefly, the process begins with the base pairing of U1 snRNA to the 5' splice site (ss) and the binding of splicing factor 1 (SF1) to the branch point in an ATP-independent manner to form the E' complex. The E' complex can be converted into the E complex by the recruitment of U2 auxiliary factor (U2AF) heterodimer (comprising U2AF65 and U2AF35) to the polypyrimidine tract and 3' terminal AG. The ATP-independent E complex is converted into the ATP-dependent pre-spliceosome A complex by the replacement of SF1 by U2 snRNP at the branch point. Further

recruitment of the U4/U6–U5 tri-snRNP leads to the formation of the B complex, which contains all spliceosomal subunits that carry out pre-mRNA splicing. This is followed by extensive conformational changes and remodelling, including the loss of U1 and U4 snRNPs, ultimately resulting in the formation of the C complex, which is the catalytically active spliceosome [Figure 3] (Matera and Wang, 2014). Alternative splicing (as opposed to constitutive splicing) refers to variations in splice site selection resulting in an mRNA species that contains or lacks a certain exon. It results in protein isoforms that differ in their peptide sequence and therefore can have different chemical and biological characteristic (Roy et al., 2013).

The spliceosomes are classified into the “major” or the “minor” spliceosome, according to the consensus sequence of acceptor and donor sites of pre-mRNA splicing. The so-called “major”

spliceosome regulates the splicing of the U2-type introns, the vast majority of the introns, that are characterized by typical GT–AG termini and have relatively variable sequences at the 5'ss. In contrast, the "minor" one regulates the splicing of the U12-type introns, which represent only a small percentage (less than 1%) of all the introns and are characterized by a unusual non-consensus AT–AC termini and a high degree of conservation at the 5' ss [Figure 4]. The assembly and the reactions of the two spliceosomes are similar in most of the steps, while there are several differences in the composition of the snRNPs that form the two complexes. The snRNPs of the "major" spliceosome are referred as U1 and U2, while the ones of the "minor" are the U11/U12 (summarized in Table 1) (Turunen et al., 2013).

Several lines of evidence suggest a role of FUS in pre-mRNA splicing. For example, in different studies

was observed that FUS associates with components of the spliceosome (Kameoka et al., 2004; Meissner et al., 2003) and in a recent large-scale yeast-two-hybrid screening it was reported that FUS associates with component of the A complex of the spliceosome (Hegele et al., 2012). FUS was also described to regulate the 5'-splice site selection in E1A pre-mRNA (Lerga et al., 2001). The global effect of FUS on alternative splicing has been recently revealed through several genome-wide exon array analyses. For example, an Affymetrix Mouse Exon array on primary cortical neurons with knocked-down FUS expression identified more than 3,202 exons that were altered, many associated with genes having neuronal functions or linked to neurodegeneration (Ishigaki et al., 2012). Significant changes in the splicing patterns of ribosome- and spliceosome-related genes were also reported in non-neuronal cells demonstrating that FUS likely

plays a general role in splicing in various cell types (van Blitterswijk et al., 2013). Moreover, different studies report the binding of FUS to its own transcript, which suggests an autoregulatory mechanism for FUS expression, that may be also relevant to ALS pathogenesis (Lagier-Tourenne et al., 2012; Nakaya et al., 2013). In support of this evidence, a recent study demonstrated that FUS regulates splicing of exon 7, but that this splicing activity is impaired for FUS variants that mislocalize to the cytoplasm. A misregulation of FUS expression may in turn contribute to the pathogenic accumulation of FUS in disease (Zhou et al., 2013). Indeed, a recent study reports that in ALS patient fibroblasts harbouring mutations in the FUS nuclear localization signal (NLS) the U1 snRNP, on of the most abundant FUS interactor, mislocalizes in the cytoplasm together with mutated FUS (Yu et al., 2015).

1.2.3 The role of FUS in transcriptional regulation

FUS may promote genomic integrity by functioning as a transcriptional regulator. For example, it has been demonstrated by yeast-two-hybrid assay that FUS directly interacts with PGC-1α (proliferator-activated receptor g-coactivator 1α), a transcriptional coactivator of oxidative stress protection genes and in FUS -/- MEFs there's a reduction of PGC-1α which results in a increase of reactive oxidative species (ROS) (Sánchez-Ramos et al., 2011).

FUS may also influence the DNA damage response through transcriptional regulation of cell cycle arrest genes. In response to DNA damage, the cell cycle arrests allowing cells to repair the lesions. It has been shown that, upon ionizing radiation (IR), FUS is recruited to the promoter of the cyclin D1 gene, a key regulator of cell cycle progression, and

reduces its transcription by blocking the histone acetyltransferase activity of the transcriptional coactivators CREB-binding protein and p300 (Wang et al., 2008). FUS also interacts with other transcription factors and can regulated indirectly the expression of their target genes; these factors include PRMT1 (Du et al., 2011), Spi-1/PU.1 (Hallier et al., 1998), β-catenin (Sato et al., 2005), p65 subunit of NF-kB (Uranishi et al., 2001).

In addition to interacting with transcription factors, FUS also affects gene expression through regulation of RNA-polimerase II. FUS binds the C-terminal domain (CTD) of the RNA-polimerase II and regulates its phosphorylation at Ser2; this results in either activation or repression of specific target genes (Schwartz et al., 2012). The ability of FUS to interact with the RNA polymerase II suggests a general role for FUS in cellular transcriptional regulation. Indeed, chromatin immunoprecipitation

of HeLa lysates with antibodies against FUS, followed by promoter microarray analyses, revealed that FUS might function as a general regulator of transcription by directly binding DNA in promoter regions (Tan et al., 2012). The study found that FUS bound to 1,161 promoter regions for genes involved in various cellular processes, including gene expression, cell cycle, and neuronal functions. Moreover, it was demonstrated that FUS self-assembles in the nucleus and binds to active chromatin region, through its N-terminal QGSY (glutamine-glycine-serine-tyrosine)-rich domain and ALS-related mutations in this domain abolish this interaction (Yang et al., 2014).

1.2.4 FUS in the DNA damage response and repair

All living organisms are constantly exposed to genotoxic stress, and DNA thus needs to be repaired to preserve the information that it

encodes. DNA Damage Response (DDR) signalling is specific of the kind of DNA damage that occurs, and the pathways that are activated are determined by the activation of the PI3K-like kinases (PIKKs) ataxia-telangiectasia-mutated (ATM), ataxia-telangiectasia and Rad3-related (ATR) and DNA-dependent protein kinase catalytic subunit (DNA-PKcs), which consequently phosphorylate and thus activate various proteins that coordinate the arrest of cell cycle progression and DNA repair pathways to preserve genome integrity. Specifically, the DDR pathway is composed of two main DNA damage sensors: the MRE11–RAD50–NBS1 (MRN) complex that detects DNA double-strand breaks (DSBs); and replication protein A (RPA) and the RAD9–RAD1–HUS1 (9-1-1) complex that detects exposed regions of single-stranded DNA. These sensors recruit the apical kinases ataxia-telangiectasia mutated (ATM) (through the MRN complex) and ataxia

telangiectasia and Rad3-related (ATR) (through RPA and the 9-1-1 complex), which is bound by ATR-interacting protein (ATRIP). These in turn phosphorylate the histone variant H2AX on Ser139 (known as γH2AX) in the region proximal to the DNA lesion. Thus, although ATM is predominantly activated by DSBs, ATR responds to the type of genotoxic stress that is caused by DNA replication stress, which is also caused by oncogenes. γH2AX is required to recruit mediator of DNA damage checkpoint 1 (MDC1) that further sustains and amplifies DDR signalling by enforcing further accumulation of the MRN complex and activation of ATM. BRCA1 is recruited at sites of DNA damage upon phosphorylation by ATM and ATR. p53-binding protein 1 (53BP1) is also involved in sustaining DDR signalling by enhancing ATM activation. DDR signalling relies on additional mechanisms that are based on ubiquitylation.

Eventually, DDR signalling spreads away from the damaged locus owing to the engagement of diffusible kinases CHK2 (which is mainly phosphorylated by ATM) and CHK1 (which is mainly phosphorylated by ATR) with signalling converging on downstream effectors such as p53 and the cell division cycle 25 phosphatase (CDC25). DDR-mediated cellular outcomes may be cell death by apoptosis; transient cell cycle arrest followed by repair of DNA damage and resumption of proliferation; or cellular senescence caused by the persistence of unrepaired DNA damage [Figure 5] (Sulli et al., 2012).

FUS have also been involved in DNA damage repair due to its ability to interact with DDR-related proteins and to bind to DNA. For example, FUS directly binds both single- and double-stranded DNA (Liu et al., 2013), and this is important for two critical steps in homologous recombination (HR)

DNA repair pathway: the D-loop formation, and the homologous DNA pairing. When double-strand breaks occur in DNA, the 5’ end of the break is trimmed back to create a 3’ overhang of single-stranded DNA; this 3’ single-stranded DNA then binds a complementary sequence within duplex DNA of a homologous chromosome or sister chromatid (Li and Heyer, 2008). In addition, it was demonstrated that the RGG domain of FUS is able to bind G-quadruplex structures in both telomeric DNA and non-coding telomeric RNA (the so called Telomeric Repeat containing RNA, or TERRA; Wang et al., 2015). In particular, FUS binds the histone methyltransferase SUV4-20H2 and the overexpression of FUS leads to an increase in histone methylation and telomeres shortening (Takahama et al., 2013). Not only FUS promotes DNA pairing and the formation of the D-loops, but several studies directly correlates FUS to the DNA-

damage repair. For example, FUS can re-localize to the sites of the DNA damage upon laser-directed DSBs (Mastrocola et al., 2013) and this event occurs in the early stages of the DDR, and prior respect to key DNA-repair proteins, including NBS1, phospho-ATM, γH2AX, and Ku70. Both PARP1 and HDAC1 are proteins implicated in the mechanism associated with the localization of FUS at double-stranded DNA breaks, since their interaction with FUS is strengthen upon DNA damage (Mastrocola et al., 2013; Wang et al., 2013). This putative upstream role for FUS in DNA damage response was further demonstrated by the reduced localization of these proteins to DNA lesions when FUS expression was knocked down (Wang et al., 2013). On the other hand, while the inhibition of PARP activity prevents the recruitment of FUS to the damaged sites, inactivation of either ATM or DNA-PK, two other key regulators of the DNA damage response, had no

effect on the recruitment of FUS to laser-induced DNA damage sites (Mastrocola et al., 2013). Therefore, although ATM can phosphorylate FUS upon DSBs (Gardiner et al., 2008), this event does not appear to be required for its recruitment to the damaged sites. Moreover, the studies of Mastrocola and colleagues and Wang and colleagues, also demonstrated that FUS is required for an efficient DNA repair in both HR and NHEJ (non-homologous end joining) pathways, since the knockdown of FUS significantly reduces the ability of the cell to repair the lesions; moreover, the knockdown of FUS resulted in an increased DNA damage which also correlates with inefficient recruitment of γH2AX and 53BP1, two common markers of the damaged DNA (Wang et al., 2013). The studies discussed earlier have also tried to determine whether ALS-related mutants of FUS maintain the DNA repair activity and to what extent such defects might play a role in

the pathogenesis of ALS. In Wang et. al., it has been described that, upon knock down of the endogenous FUS in U2OS cells, the ALS mutants R244C, R514S, H517Q, and R521C are not able to restore the ability of the cells to repair the DNA lesion by HR (predominantly) and NHEJ; however, the H517Q mutant fully rescued the NHEJ repair activity upon the loss of expression of endogenous FUS wild-type. In addition, post-mortem brain sections from the motor cortex of ALS patients harbouring either FUS R521C or P525L mutations display increased levels of the γH2AX DNA damage marker relative to control brain sections (Wang et al., 2013). The role of FUS mutants in DNA damage repair was also studied taking advantage of several mouse models. For example, a transgenic mouse expressing FUS R521C, that exhibits severe motor defects and death 4 to 6 weeks after symptom onset, also exhibits elevated levels of DNA damage

markers (e.g., γH2AX, phosphorylated p53, and ATF3) in the CNS. Comet assays performed on isolated neurons supported this observation, with >50% of neurons from R521C mice showing comet tails compared with 20% from non-transgenic control (Qiu et al., 2014). Moreover, other two FUS knockdown (FUS -/-) mouse models reveal signs of genomic instability (Hicks et al., 2000; Kuroda et al., 2000). Furthermore, another study reports a novel function that FUS exerts in the context of the DDR, since it has been described that upon DNA damage, FUS can promote the SUMOylation of the onco-suppressor Ebp1 (or Proliferation-associated protein 2G4) protein by acting as SUMO E3 ligase; the post-translational modification of Ebp1, mediated by FUS, is important for its onco-suppressive activity (Oh et al., 2010). To date, this is the only study which describes the role of FUS as a SUMO E3 ligase, however it is possible to speculate

that FUS might promote the SUMOylation of other proteins during the DDR, since many proteins are modified by SUMO upon genotoxic stress (Sarangi and Zhao, 2015). Another evidence that correlates FUS to DDR is the fact that, in response to DNA damage, FUS binds to the long non-coding RNA ncRNA–CCND1, that is transcribed upstream of the CCND1 gene encoding the cyclin D1 protein, a cell cycle regulator repressed by DNA damage signals, and represses the transcription of CCND1 (Wang et al., 2008).

1.3 Figures and tables

Figure 1. The domains of the FET proteins. *Adapted from (Schwartz et al., 2014).*

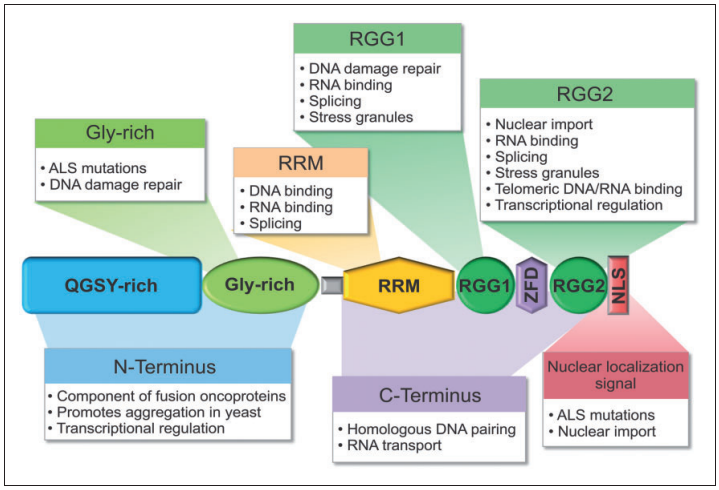


Figure 2. The functional domains of FUS. *Adapted from (Sama et al., 2014).*

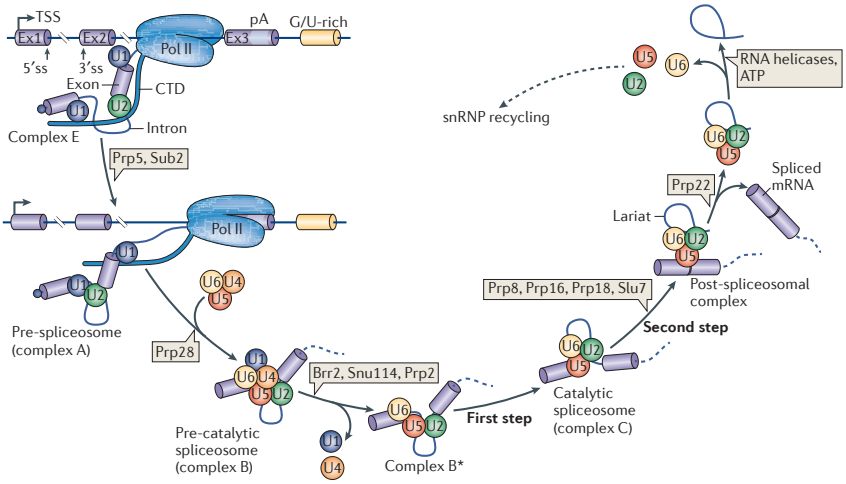


Figure 3. Stepwise assembly of the spliceosome.

Spliceosome assembly takes place at sites of transcription: the U1 and U2 snRNPs assemble in a co-transcriptional manner through recognition of the 5′ss and 3′ss, which is mediated by the CTD of Polymerase II. The U1 and U2 snRNPs interact with each other to form the pre-spliceosome (complex A). This process is dependent on DExD/H helicases pre-mRNA-processing 5 (Prp5) and Sub2. In a subsequent reaction catalysed by Prp28, the preassembled tri-snRNP U4–U6•U5 is recruited to form complex B. The resulting complex B undergoes a series of rearrangements to form a catalytically active complex B (complex B*), which requires multiple RNA helicases (Brr2, 114 kDa U5 small nuclear ribonucleoprotein component (Snu114) and Prp2) and results in the release of U4 and U1 snRNPs. Complex B* then carries out the first catalytic step of splicing, generating complex C, which contains free exon 1 (Ex1) and the intron–exon 2 lariat intermediate. Complex C undergoes additional rearrangements and then carries out the second catalytic step, resulting in a post-

spliceosomal complex that contains the lariat intron and spliced exons. Finally, the U2, U5 and U6 snRNPs are released from the mRNP particle and recycled for additional rounds of splicing. Release of the spliced product from the spliceosome is catalysed by the DExD/H helicase Prp22 (Matera and Wang, 2014).

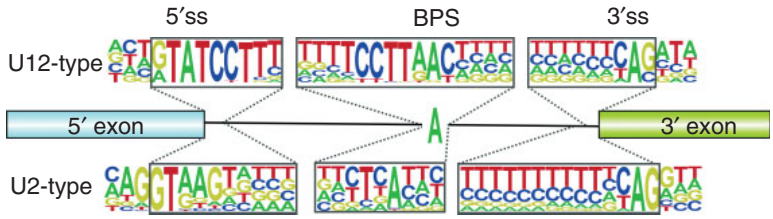


Figure 4. Consensus sequences of human U12- and U2-type introns. The height of the letters in each position indicates the relative frequency of individual nucleotides in that position (Turunen et al., 2013).

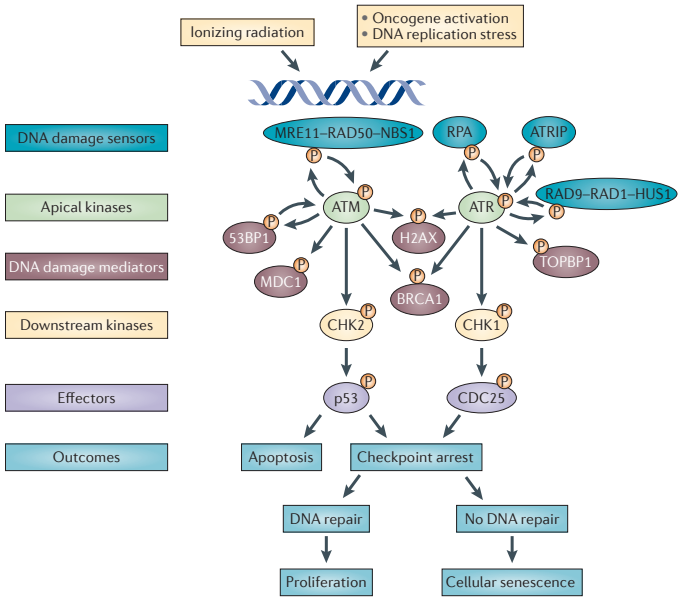


Figure 5. Snapshot of the DNA damage response. *(Sulli et al., 2012).*

Table 1. Proteins of the U1, U2, and U11/U12 snRNPs. *Adapted from (Turunen et al., 2013)*

12S U1	17S U2	18S U11/U12	Functions (with Selected References)
Sm proteins[1]	Sm proteins[1]	Sm proteins[1]	snRNP core components
U1-A (SNRPA)			Structural; RNA-binding
U1-C (SNRPC)			5′ss recognition
U1-70K (SNRNP70)			Structural; SR protein interactions
	U2A′ (SNRPA1)		Structural; RNA-binding
	U2B″ (SNRPB2)		Structural; RNA-binding
	SF3a complex[2]		BPS binding
	SF3b complex[2]	SF3b complex[2]	BPS binding
		20K (ZMAT5)	Unknown; homology to U1C
		25K (SNRNP25)	Unknown
		31K (ZCRB1)	Unknown; RNA-binding
		35K (SNRNP35)	SR protein interactions, homology to U1-70K
		48K (SNRNP48)	5′ss recognition
		59K (PDCD7)	Structural, binds 48K and 65K
		65K (RNPC3)	Structural, binds U12 snRNA
		Urp[3] (ZRSR2)	3′ss recognition
	hPrp43[4]	hPrp43 (DHX15)	
		Y Box-1[3] (YBX1)	

The Hugo names of proteins have been provided in parentheses.
[1]Sm proteins B/B′, D1, D2, D3, E, F, and G,
[2]Multi-subunit complexes.
[3]Also present in the major spliceosome, but not in U1 or U2 snRNPs.
[4]Almost stoichiometric presence in the 17S U2. Other proteins associated with U2 in substoichiometric amounts are omitted.

1.4 Scope of the thesis

The aim of my Ph.D. project was to obtain new insights into the diverse molecular pathways in which FUS is involved, with particular emphasis on elucidating its role in pre-mRNA splicing and DNA damage response.

Chapter 1

This chapter describes FUS/TLS proteins and underlines the most recent findings regarding the different biological functions in which it has been involved.

Chapter 2

In chapter 2, I report the results obtained for the identification of the interactors of FUS in human cells and the functional experiments aimed at characterizing its role in pre-mRNA splicing.

Chapter 3

In chapter 3, I report the preliminary results of the proteomic characterization of the role of FUS in DNA damage response.

Chapter 4

The last chapter summarizes the results obtained and underlines the possible future perspectives, focusing the attention on the putative translational applications of this research.

1.5 References

Andersson, M.K., Ståhlberg, A., Arvidsson, Y., Olofsson, A., Semb, H., Stenman, G., Nilsson, O., and Aman, P. (2008). 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. *9*, 37.

Bentmann, E., Neumann, M., Tahirovic, S., Rodde, R., Dormann, D., and Haass, C. (2012). Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). J. Biol. Chem. *287*, 23079–23094.

Blatter, M., Allain, H., and Cle, A. RNA recognition motifs : boring ? Not quite.

Van Blitterswijk, M., Wang, E.T., Friedman, B. a., Keagle, P.J., Lowe, P., Leclerc, A.L., van den Berg, L.H., Housman, D.E., Veldink, J.H., and Landers, J.E. (2013). Characterization of FUS Mutations in Amyotrophic Lateral Sclerosis Using RNA-Seq. PLoS One *8*, 1–8.

Calvio, C., Neubauer, G., and Mann, M. (1995). Identification of hnRNP P2 as TLS / FUS using electrospray mass spectrometry . 724–733.

Colombrita, C., Onesto, E., Megiorni, F., Pizzuti, A., Baralle, F.E., Buratti, E., Silani, V., and Ratti, A. (2012). TDP-43 and FUS RNA-binding proteins bind distinct sets of cytoplasmic messenger RNAs and differently regulate their post-transcriptional fate in motoneuron-like cells. J. Biol. Chem. *287*, 15635–15647.

Daigle, J.G., Lanson, N. a, Smith, R.B., Casci, I., Maltare, A., Monaghan, J., Nichols, C.D., Kryndushkin, D., Shewmaker, F., and Pandey, U.B. (2013). RNA-binding ability of FUS regulates neurodegeneration, cytoplasmic mislocalization and incorporation into stress granules associated with FUS carrying ALS-linked mutations. Hum. Mol. Genet. *22*, 1193–1205.

Du, K., Arai, S., Kawamura, T., Matsushita, A., and Kurokawa, R. (2011). TLS and PRMT1 synergistically coactivate transcription at the survivin promoter through TLS arginine methylation. Biochem. Biophys. Res. Commun. *404*, 991–996.

Gardiner, M., Toth, R., Vandermoere, F., Morrice, N. a, and Rouse, J. (2008). Identification and characterization of FUS/TLS as a new target of ATM. Biochem. J. *415*, 297–307.

Hallier, M., Lerga, A., and Tavitian, A. (1998). The Transcription Factor Spi-1 / PU . 1 Interacts with the Potential Splicing Factor TLS *. Biochemistry *273*, 4838–4842.

Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Lührmann, R., and Stelzl, U. (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell *45*, 567–580.

Hicks, G.G., Singh, N., Nashabi, a, Mai, S., Bozek, G., Klewes, L., Arapovic, D., White, E.K., Koury, M.J., Oltz, E.M., et al. (2000). Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. Nat. Genet. *24*, 175–179.

Hoell, J.I., Larsson, E., Runge, S., Nusbaum, J.D., Duggimpudi, S., Farazi, T. a, Hafner, M., Borkhardt, A., Sander, C., and Tuschl, T. (2011). RNA targets of wild-type and mutant FET family proteins. Nat. Struct. Mol. Biol. *18*, 1428–1431.

Ichikawa, H., Shimizu, K., and Hayashi, Y. (1994). An RNA-binding Protein Gene , TLS / FUS , Is Fused to ERG in Human Myeloid Leukemia with t (16 ; 21) Chromosomal Translocation. 2865–2868.

Ishigaki, S., Masuda, A., Fujioka, Y., Iguchi, Y., Katsuno, M., Shibata, A., Urano, F., Sobue, G., and Ohno, K. (2012). Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. Sci. Rep. *2*, 529.

Kameoka, S., Duque, P., and Konarska, M.M. (2004). P54(Nrb) Associates With the 5' Splice Site Within Large Transcription/Splicing Complexes. EMBO J. *23*, 1782–1791.

Kuroda, M., Sok, J., Webb, L., Baechtold, H., Urano, F., Yin, Y., Chung, P., de Rooij, D.G., Akhmedov, a, Ashley, T., et al. (2000). Male sterility and enhanced radiation sensitivity in TLS(-/-) mice. EMBO J. *19*, 453–462.

Lagier-Tourenne, C., Polymenidou, M., Hutt, K.R., Vu, A.Q., Baughn, M., Huelga, S.C., Clutario, K.M., Ling, S.-C., Liang, T.Y., Mazur, C., et al. (2012). Divergent roles of ALS-linked proteins FUS/TLS and TDP-43

intersect in processing long pre-mRNAs. Nat. Neurosci. *15*, 1488–1497.

Lerga, A., Hallier, M., Delva, L., Orvain, C., Gallais, I., Marie, J., and Moreau-Gachelin, F. (2001). Identification of an RNA Binding Specificity for the Potential Splicing Factor TLS. J. Biol. Chem. *276*, 6807–6816.

Li, X., and Heyer, W.-D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. *18*, 99–113.

Liu, X., Niu, C., Ren, J., Zhang, J., Xie, X., Zhu, H., Feng, W., and Gong, W. (2013). The RRM domain of human fused in sarcoma protein reveals a non-canonical nucleic acid binding site. Biochim. Biophys. Acta *1832*, 375–385.

Mackenzie, I.R. a, Rademakers, R., and Neumann, M. (2010). TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. Lancet Neurol. *9*, 995–1007.

Mastrocola, A.S., Kim, S.H., Trinh, A.T., Rodenkirch, L. a, and Tibbetts, R.S. (2013). The RNA-binding protein fused in sarcoma (FUS) functions downstream of poly(ADP-ribose) polymerase (PARP) in response to DNA damage. J. Biol. Chem. *288*, 24731–24741.

Matera, a G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat. Rev. Mol. Cell Biol. *15*, 108–121.

Meissner, M., Lopato, S., Gotzmann, J., Sauermann, G., and Barta, A. (2003). Proto-oncoprotein TLS/FUS is associated to the nuclear matrix and complexed with splicing factors PTB, SRm160, and SR proteins. Exp. Cell Res. *283*, 184–195.

Naganuma, T., and Hirose, T. (2013). Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol. *10*, 456–461.

Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N., and Hirose, T. (2012). Alternative 3'-end processing of long

noncoding RNA initiates construction of nuclear paraspeckles. EMBO J. *31*, 4020–4034.

Nakagawa, S., and Hirose, T. (2012). Paraspeckle nuclear bodies—useful uselessness? 3027–3036.

Nakaya, T., Alexiou, P., Maragkakis, M., Chang, A., and Mourelatos, Z. (2013). FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. RNA *19*, 498–509.

Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H. a., and MacKenzie, I.R. a (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain *132*, 2922–2931.

Oh, S.-M., Liu, Z., Okada, M., Jang, S.-W., Liu, X., Chan, C.-B., Luo, H., and Ye, K. (2010). Ebp1 sumoylation, regulated by TLS/FUS E3 ligase, is required for its anti-proliferative activity. Oncogene *29*, 1017–1030.

Orozco, D., Tahirovic, S., Rentzsch, K., Schwenk, B.M., Haass, C., and Edbauer, D. (2012). Loss of fused in sarcoma (FUS) promotes pathological Tau splicing. EMBO Rep. *13*, 759–764.

Qiu, H., Lee, S., Shang, Y., Wang, W.Y., Au, K.F., Kamiya, S., Barmada, S.J., Finkbeiner, S., Lui, H., Carlton, C.E., et al. (2014). ALS-associated mutation FUS-R521C causes DNA damage and RNA splicing defects. J. Clin. Invest. *124*, 981–999.

Rabbitts, T.H., Forster, a, Larson, R., and Nathan, P. (1993). Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. Nat. Genet. *4*, 175–180.

Rademakers, R., and Mackenzie, I.R. a (2013). Dementia. Nat. Rev. Neurol. *8*, 423–434.

Rogelj, B., Easton, L.E., Bogu, G.K., Stanton, L.W., Rot, G., Curk, T., Zupan, B., Sugimoto, Y., Modic, M., Haberman, N., et al. (2012). Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci. Rep. *2*, 1–10.

Roy, B., Haupt, L.M., and Griffiths, L.R. (2013). Review: Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity. Curr. Genomics *14*, 182–194.

Sama, R.R.K., Ward, C.L., and Bosco, D. a (2014). Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. ASN Neuro *6*.

Sánchez-Ramos, C., Tierrez, A., Fabregat-Andrés, O., Wild, B., Sánchez-Cabo, F., Arduini, A., Dopazo, A., and Monsalve, M. (2011). PGC-1α regulates translocated in liposarcoma activity: role in oxidative stress gene expression. Antioxid. Redox Signal. *15*, 325–337.

Sarangi, P., and Zhao, X. (2015). SUMO-mediated regulation of DNA damage repair and responses. Trends Biochem. Sci. *40*, 233–242.

Sato, S., Idogawa, M., Honda, K., Fujii, G., Kawashima, H., Takekuma, K., Hoshika, A., Hirohashi, S., and Yamada, T. (2005). ??-Catenin interacts with the FUS proto-oncogene product and regulates pre-mRNA splicing. Gastroenterology *129*, 1225–1236.

Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J., and Cech, T.R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. Genes Dev. *26*, 2690–2695.

Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA Seeds Higher-Order Assembly of FUS Protein. Cell Rep. *5*, 918–925.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2014). Biochemical Properties and Biological Functions of FET Proteins. Annu. Rev. Biochem. 1–25.

Shelkovnikova, T. a, Robinson, H.K., Troakes, C., Ninkina, N., and Buchman, V.L. (2014). Compromised paraspeckle formation as a pathogenic factor in FUSopathies. Hum. Mol. Genet. *23*, 2298–2312.

Storlazzi, C.T., Mertens, F., Nascimento, A., Isaksson, M., Wejde, J., Brosjö, O., Mandahl, N., and Panagopoulos, I. (2003). Fusion of the FUS and BBF2H7 genes in low grade fibromyxoid sarcoma. Hum. Mol. Genet. *12*, 2349–2358.

Sulli, G., Di Micco, R., and d'Adda di Fagagna, F. (2012). Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. Nat. Rev. Cancer *12*, 709–720.

Takahama, K., Takada, A., Tada, S., Shimizu, M., Sayama, K., Kurokawa, R., and Oyoshi, T. (2013). Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. Chem. Biol. *20*, 341–350.

Tan, A.Y., and Manley, J.L. (2009). The TET family of proteins: Functions and roles in disease. J. Mol. Cell Biol. *1*, 82–92.

Tan, A.Y., Riley, T.R., Coady, T., Bussemaker, H.J., and Manley, J.L. (2012). TLS / FUS (translocated in liposarcoma / fused in sarcoma) regulates target gene transcription via single-stranded DNA response elements. 1–6.

Turunen, J.J., Niemelä, E.H., Verma, B., and Frilander, M.J. (2013). The significant other: Splicing by the minor spliceosome. Wiley Interdiscip. Rev. RNA *4*, 61–76.

Uranishi, H., Tetsuka, T., Yamashita, M., Asamitsu, K., Shimizu, M., Itoh, M., and Okamoto, T. (2001). Involvement of the Pro-oncoprotein TLS (Translocated in Liposarcoma) in Nuclear Factor-??B p65-mediated Transcription as a Coactivator. J. Biol. Chem. *276*, 13395–13401.

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., et al. (2009).

Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science *323*, 1208–1211.

Wang, C., Zhao, L., and Lu, S. (2015). Role of TERRA in the Regulation of Telomere Length. Int. J. Biol. Sci. *11*, 316–323.

Wang, W.-Y., Pan, L., Su, S.C., Quinn, E.J., Sasaki, M., Jimenez, J.C., Mackenzie, I.R. a, Huang, E.J., and Tsai, L.-H. (2013). Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. Nat. Neurosci. *16*, 1383–1391.

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature *454*, 126–130.

Waters, B.L., Panagopoulos, I., and Allen, E.F. (2000). Genetic characterization of angiomatoid fibrous histiocytoma identifies fusion of the FUS and ATF-1 genes induced by a chromosomal translocation involving Bands 12q13 and 16p11. Cancer Genet. Cytogenet. *121*, 109–116.

Yang, L., Gal, J., Chen, J., and Zhu, H. (2014). Self-assembled FUS binds active chromatin and regulates gene transcription. Proc. Natl. Acad. Sci. *111*, 17809–17814.

Yu, Y., Chi, B., Xia, W., Gangopadhyay, J., Yamazaki, T., Winkelbauer-Hurt, M.E., Yin, S., Eliasse, Y., Adams, E., Shaw, C.E., et al. (2015). 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*, 3208–3218.

Zhou, Y., Liu, S., Liu, G., Oztürk, A., and Hicks, G.G. (2013). ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. PLoS Genet. *9*, e1003895.

Zinszner, H., Sok, J., Immanuel, D., Yin, Y., and Ron, D. (1997). TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. J. Cell Sci. *110 (Pt 1*, 1741–1750.

Chapter 2

Proteomic characterization of FUS/TLS interactome reveals a new role in pre-mRNA splicing

Manuscript in preparation

Giuseppe Filosa[1,2] , Stefan Reber[3], Jolanda Stettler[3], Angela Bachi[2], Silvia Barabino[1], Oliver Mühlemann[3], Marc-David Ruepp[3]

[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy; [2]IFOM, FIRC Institute of Molecular Oncology, via Adamello 16, 20139 Milan, Italy; [3]Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland

2.1 Abstract

FUS (*fused in* sarcoma, also known as TLS, *translocated in liposarcoma*) is a ubiquitously expressed RNA-binding protein that has been discovered as fused to transcription factors in several human sarcomas and found in protein aggregates in neurons of patients with an inherited form of Amyotrophic Lateral Sclerosis. FUS has been implicated in a variety of cellular processes such as transcriptional regulation, pre-mRNA splicing, and miRNA processing. The exact role that FUS exerts in these biological processes is still not well defined. Thus in order to shed light on the molecular functions of FUS, we investigated its interactome in human cells. FUS interacting proteins were isolated by immunoprecipitation from total extracts of HEK293T expressing wild type FUS-flagged protein in three different conditions: with or without RNase

treatment of the extract and with RNase and high salt washing. The interactors were identified by tandem mass spectrometry analysis, which resulted in the identification of 546 proteins in the untreated sample, 134 in the RNase treated sample and 53 in the RNase plus 750 mM NaCl wash sample. 40 interactors were common in all three samples most likely representing the strongest ones. The interactome analysis revealed an extensive cross talk with proteins that associate with RNA, consistent with previously described roles for FUS in RNA metabolism. Interestingly, among the most conserved FUS interactors there is a significant enrichment for U12-type splicing of the minor spliceosomal complex. The interaction between FUS and these particular splicing factors was validated and a tethered splicing approach demonstrated that FUS is able to bind to minor introns and to enhance their splicing.

2.2 Introduction

FUS/TLS is a ubiquitously expressed RNA-binding protein of the hnRNP family, which has been discovered as fused to transcription factors in several human malignancies and as part of protein aggregates in neurons of patients with an inherited form of Amyotrophic Lateral Sclerosis (Vance et al., 2013). FUS is a 53 kDa nuclear protein that contains structural domains, such as a RNA Recognition Motif (RRM) and a zinc finger motif, that give to FUS the ability to bind to both RNA and DNA sequences (as previously described in Chapter 1).

2.3 Results

2.3.1 Wild-type FUS interactome in human cells

In order to unravel the biological processes in which FUS is involved, a mass spectrometry-based

interactome analysis was performed. An affinity purification coupled to mass spectrometry (AP-MS) experiment was performed to identify wild-type FUS interacting proteins in HEK293T cells transiently expressing a flag-tagged recombinant version of the protein. Since FUS has the ability to bind to RNA sequences, the interactome analysis was performed in presence or absence of RNaseA treatment, therefore allowing the selection of only protein-to-protein interaction rather than interactions mediated by RNA intermediates. Moreover, in order to eliminate the contaminants of the anti-flag matrix, a control immunoprecipitation was performed on total protein extract from cells transfected with EBFP-flag (Enhanced Blue Fluorescent Protein). The immunoprecipitations resulted in a clean control and the RNase treated lane clearly showed a reduced amount of proteins, suggesting that many interactions were mediated

by the presence of RNA [Figure 1a]. The mass spectrometric analysis resulted in the identification of 583 potential interactors: 449 unique for untreated samples, 37 unique for RNase treated samples and 97 in common among the two conditions; all these proteins were not identified in the control [Figure 1b]. In addition, to distinguish the most conserved FUS interactors, an immunoprecipitation analysis was performed in presence of RNase treatment and with stronger washings conditions (750 mM of sodium chloride). A total number of 53 protein was identified, among these 40 are in common between all the conditions, thus representing the most conserved FUS interactors [Figure 1b; Table 1].

2.3.2 FUS preferentially interacts with proteins involved in RNA metabolism

Gene Ontology enrichment analysis was performed to highlight the biological processes in which FUS and its interactors participate. The analysis revealed that FUS and its interactors are implicated in different cellular processes, in particular to the ones related to RNA metabolism. Moreover, it highlighted a statistically significant enrichment of other functions, such as the DNA damage response and repair, for which the role of FUS still needs to be better clarified [Figure 2a]. The interaction between FUS and splicing factors was also demonstrated by a large-scale yeast-two-hybrid analysis aimed to map the protein-protein interaction of the different spliceosomal complexes (Hegele et al., 2012). In that study, FUS was identified to be part of the “A” spliceosomal complex. Our interactome analysis confirmed that FUS interacts with proteins of the “A spliceosomal complex”, but highlighted an interaction with

“core” proteins of other spliceosomal complexes, involved in almost all the catalytic steps, and with “non core” splicing regulating factors, such as HNRNPs and SR proteins [Figure 3]. These results suggest that FUS might be part of one of several spliceosome complexes and/or might be associated with the splicing factory throughout all the catalytic process.

The same analysis using as input the 40 most conserved FUS interactors, confirmed that the strongest FUS interactors are mostly involved in RNA metabolism and, interestingly, the annotation of the cellular component GO terms, revealed an enrichment for proteins that are part of the snRNPs of the U12-type spliceosomal complex, the so called “minor” spliceosomal complex [Figure 2b, 2c].

2.3.3 FUS interacts with the U11 snRNP

In order to better define the role of FUS in the splicing process, we first validated the interaction between FUS and the different snRNPs. We performed RNA-IPs from HeLa nuclear extract using antibodies directed against FUS or BSA (as negative control), followed by RT-qPCR [Figure 4a]. Consistent with the mass spectrometric analysis, the most enriched U snRNA was the U11 snRNA, a member of the minor spliceosome and constituent of the U11/U12 disnRNP. The second most enriched U snRNA was the U1 snRNA, whose interaction with FUS has been reported previously (Hackl and Lührmann, 1996; Yamazaki et al., 2012; Yu et al., 2015). These results are in accordance with previous findings of the Lührmann laboratory reporting the presence of FUS in the human spliceosomal complexes E, A, and B, in which the U1 and U11 snRNPs are present, but not in the $B^{activated}$ and C complexes, a stage at which U1 and U11

snRNPs have left the pre-mRNA (Behzadnia et al., 2007; Deckert et al., 2006). We then focused on the interaction of the minor spliceosome component U11 snRNP with FUS. This interaction was validated by performing pulldowns using a biotinylated antisense oligonucleotide against U11 snRNA. These U11 snRNP pulldowns were not only enriched for the U11/U12 di-snRNP specific factor U11-59K and the common spliceosomal Sm ring component SmD3 but also for FUS [Figure 4b].

2.3.4 FUS is involved in splicing of minor intron-containing reporter genes

The interaction of FUS with U11 snRNP suggested that FUS might be involved in regulating the splicing of minor introns, also called U12-dependent introns (Turunen et al., 2013). Therefore, we tested if FUS depletion affects the splicing of minor intron-containing reporter genes. We used two CTNND1-

derived minor intron minigenes and one SCN4A (generously provided by Dr. Mark T. McNally, Medical College of Wisconsin). The CTNND1 minigene consists of the shortened CTNND1 minor intron sequence flanked by exons 6 and 7 of CTNND1, whereas in the other case the minigene is a chimera termed NRS, consisting of a U11/U12 5' splice site from Rous Sarcoma Virus and a U11/U12 3' splice site from CTNND1 [Figure 5a]. HeLa cells were transfected with these minigenes along with knockdown-inducing pSUPuro plasmids expressing either a scrambled control shRNA (scrKD) or an shRNA targeting the FUS mRNA (FUS KD). Four days later, we analyzed the FUS protein levels by Western blotting [Figure 5b] and determined the ratio of spliced vs unspliced reporter RNA by RT-qPCR [Figure 5c]. We observed that FUS depletion results in a reduction of correctly spliced mRNA for all the three reporter genes. These results show

that FUS is involved in the splicing of minor introns, but it is not clear if the impairment of the splicing reaction is a direct or indirect consequence of FUS depletion.

2.3.5 FUS directly regulates the splicing of minor intron-containing reporter genes

To investigate if FUS has a direct effect on CTNND1 splicing, we used a CTNND1 minigene version in which the hnRNP H-binding G-rich tracts that are required to recruit the minor spliceosome were replaced by MS2 binding sites. This MS2 binding sites-containing mRNA is spliced very inefficiently unless a factor promoting the recruitment of the minor spliceosome (e.g. hnRNP H) is tethered to the MS2 binding sites by expressing it as a fusion protein with the MS2 coat protein (MS2CP). Tethering of proteins not associated with the minor spliceosome (e.g. hnRNP F or GFP) did not promote

splicing of this reporter RNA (McNally et al., 2006). Consistently, we observed that neither the tethering of MS2-GFP, nor the expression of flag-tagged FUS without the MS2CP moiety led to a significant increase in splicing, whereas tethering of MS2-hnRNPH or MS2-FUS produced more than 35 and 55 fold increase in spliced mRNA, respectively [Figure 6b], implicating that FUS plays a direct role in recruiting the U11/U12 di-snRNP to minor introns. Using the tethering system, we next decided to identify the parts of FUS required for promoting splicing. We created a series of truncation and deletion mutants of FUS, all comprising an N-terminal MS2CP fusion and the C-terminal FUS nuclear localization signal (NLS) to assure nuclear localization [Figure 6a]. Compared to the splicing activity obtained by tethering of full-length FUS, deletion of the first 284 amino acids (the QGSY-rich and RGG1 domains) almost

completely abrogated FUS' capacity to promote splicing of the p120 intron (Figure 6c). A series of C-terminal truncations identified that the first 165 amino acids (the QGSY-rich region) were sufficient to promote CTNND1 intron removal. Importantly, the FET binding motif (FETbm), which mediates homo- and heterodimer formation between the FET family members (Schwartz et al., 2014), is dispensable for the splicing activity of this 1-165 FUS fragment, indicating that the attraction of the minor spliceosome can be carried out by FUS itself and does not depend on an interaction with the other FET family members. Furthermore, the very C-terminus of FUS harbours its NLS and many ALS-associated mutations have been reported that affect the interaction of transportin with the NLS leads to a reduced concentration of FUS in the nucleus. We tested whether the reduction in nuclear FUS concentration caused by the NLS-

inactivating ALS-associated P525L mutation affects FUS' role in splicing. To this end, we expressed MS2-tagged FUS-P525L and performed a tethered splicing assay. MS2-FUS P525L expression indeed exhibited reduced splicing in the tethering assay. The splicing activity could be restored by appending the SV40 nuclear localization signal to MS2-FUS P525L, demonstrating that the reduced nuclear level of FUS P525L was responsible for its failure to efficiently promote CTNND1 intron splicing [Figure 6d]. Hence it is conceivable that the ALS-associated mutations in the NLS of FUS might contribute to the disease by altering the splicing pattern of mRNAs containing minor introns.

2.4 Discussion

The FUS protein was described to take part in numerous biological processes, however, for most of these pathways the exact role played by FUS is

not well defined. Therefore, our study was aimed at better characterizing its molecular functions, starting from an unbiased mass spectrometry-based proteomic approach to define its interactome in a human cell line. Since FUS is a RNA-binding protein, we performed the experiments either in the presence or in the absence of RNase treatment in order to distinguish between protein-to-protein and RNA-mediated interactions. In addition, we employed a high stringency purification of FUS and its interactors, by using high salt washes and RNase treatment. The interactome analysis resulted in the identification of more than 500 putative interactors, and in particular of a set of 40 proteins, identified in all the experimental conditions, that represents the most conserved FUS interactome. Gene ontology enrichment analysis highlighted a strong enrichment for most of the pathways in which FUS has already been implicated, especially the ones

related to RNA metabolism. In addition, the GO analysis performed on the 40 most conserved FUS interacting partners further confirmed the enrichment of splicing-related proteins in the dataset. Interestingly, the analysis revealed an interaction between FUS and some proteins involved in the splicing of the “minor” introns (U12-type introns), suggesting a role for FUS in the splicing of this particular set of introns, which represent only a small percentage (less than 1%) of all the introns (Turunen et al., 2013). This finding is particularly important since the splicing of the “minor” introns has been recently involved in the pathogenesis of neurodegenerative diseases, such as the ALS (as reviewed in Onodera et al., 2014), a pathology in which FUS is also implicated (as described earlier in Chapter 1). Therefore, we decided to investigate more in details the role of FUS in this specific pathway, in order to unravel

whether it might play a direct role on the splicing of the U12-type introns. We first demonstrated that FUS depletion leads to a reduced splicing of the minor introns, then, by a tethered splicing assay, we showed that FUS directly bind the minor introns and enhances their splicing. Moreover, the P525L ALS-causative FUS mutation, that abrogates the nuclear translocation of FUS, reduces the capacity of FUS to promote the splicing of the minor introns, suggesting that this specific role of FUS might be involved in the pathogenesis of the ALS.

2.5 Figures and tables

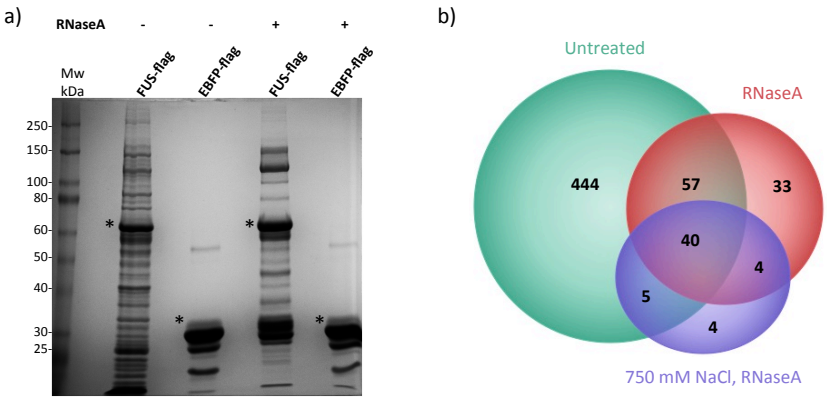


Figure 1. FUS-Flag Immunoprecipitation and MS analysis results. a) Representative coomassie stained SDS-PAGE of the elutions from the anti-flag immunoprecipitation in low stringent washing condition (150 mM NaCl), with/without RNaseA treatment and immunoprecipitation of the lysates from EBFP-flag transfected cells; asterisks indicate the baits (FUS-flag and EBFP-flag); b) Venn diagram representing the overlap of the proteins identified under low stringency (with and without RNaseA treatment) and high stringency plus RNaseA conditions; proteins identified in the control were not reported.

Table 1 – Most conserved FUS interactors

Protein IDs	Protein names	Gene names	Peptides Untreated	Peptides RNAse	Peptides 750 mM	Unique peptides Untreated	Unique peptides RNAse	Unique peptides 750 mM	Average Intensity
P35637-2	RNA-binding protein FUS	FUS	19	21	27	17	18	23	1393620000
O14744	Protein arginine N-methyltransferase 5	PRMT5	28	21	34	3	3	3	3209945367
P52272-2	Heterogeneous nuclear ribonucleoprotein M2	HNRNPM	26	45	29	19	26	20	2561431167
Q00839	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	46	50	22	37	39	18	1970919167
P51991	Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3	20	33	8	2	5	1	1566604917
J3KTA4	Probable ATP-dependent RNA helicase DDX5	DDX5	38	26	13	29	18	8	961125000
G8JLB6	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	20	23	18	14	16	13	699844333
Q9BQA1	Methylosome protein 50	WDR77	12	12	13	12	12	13	629266667
P98175	RNA-binding protein 10	RBM10	40	44	48	36	39	44	454733500
O00571	ATP-dependent RNA helicase DDX3X	DDX3X	36	17	8	33	14	6	397263072
O75688	Protein phosphatase 1B	PPM1B	21	19	25	2	2	3	378454167
O43143	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	DHX15	31	48	5	30	47	5	326095533
Q92499	ATP-dependent RNA helicase DDX1	DDX1	35	50	12	35	50	12	317085167
H0Y8G5	Heterogeneous nuclear ribonucleoprotein D0	HNRNPD	21	10	5	19	9	5	294154533
E7EX17	Eukaryotic translation initiation factor 4B	EIF4B	10	13	8	10	13	8	285325900
Q15393	Splicing factor 3B subunit 3	SF3B3	21	39	7	21	39	7	269009300
P61978-3	Heterogeneous nuclear ribonucleoprotein K	HNRNPK	24	18	4	24	18	4	259942833
O15042-2	U2 snRNP-associated SURP motif-containing protein	U2SURP	24	40	5	24	40	5	252276487
O75533	Splicing factor 3B subunit 1	SF3B1	26	61	2	26	61	2	249115772
H3BLZ8	Probable ATP-dependent RNA helicase DDX17	DDX17	33	22	10	24	14	6	240744172
Q9Y3I0	tRNA-splicing ligase RtcB homolog	RTCB	20	22	9	20	22	9	201878517
J3QK89	Calcium homeostasis endoplasmic reticulum protein	CHERP	17	19	6	17	19	6	164780433
Q15208	Serine/threonine-protein kinase 38	STK38	23	21	24	18	18	19	121617500
Q9BUJ2-2	Heterogeneous nuclear ribonucleoprotein U-like protein 1	HNRNPUL1	22	9	6	22	9	6	115160167
Q5W015	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3	PFKFB3	25	30	25	18	23	20	107720333
P52597	Heterogeneous nuclear ribonucleoprotein F	HNRNPF	15	11	4	13	10	2	104574510
Q8WVY3	U4/U6 small nuclear ribonucleoprotein Prp31	PRPF31	16	22	12	16	22	12	90981667
Q9BRS2	Serine/threonine-protein kinase RIO1	RIOK1	15	14	8	15	14	8	85641667
Q9Y2H1	Serine/threonine-protein kinase 38-like	STK38L	20	18	26	16	14	20	82633333
Q15424	Scaffold attachment factor B1	SAFB	4	25	7	2	16	3	64218133
A8MXP9	Matrin-3	MATR3	23	8	4	23	8	4	59883583
Q96I25	Splicing factor 45	RBM17	21	22	5	21	22	5	53688728
P55795	Heterogeneous nuclear ribonucleoprotein H2	HNRNPH2	13	14	6	8	8	1	37752932
P78332	RNA-binding protein 6	RBM6	14	24	9	13	22	7	32265500
P31153	S-adenosylmethionine synthase isoform type-2	MAT2A	9	10	3	9	10	3	24205933
P23458	Tyrosine-protein kinase JAK1	JAK1	9	19	17	9	19	17	22628683
O75688-2	Protein phosphatase 1B, isoform beta-2	PPM1B	21	18	24	2	1	2	10851770
Q9BUA3	Uncharacterized protein C11orf84	C11orf84	6	9	3	6	9	3	9113450
M0QZM1	Heterogeneous nuclear ribonucleoprotein M	HNRNPM	8	20	10	1	1	1	4223895
P42167	Thymopoietin	TMPO	4	3	3	4	3	3	3042405

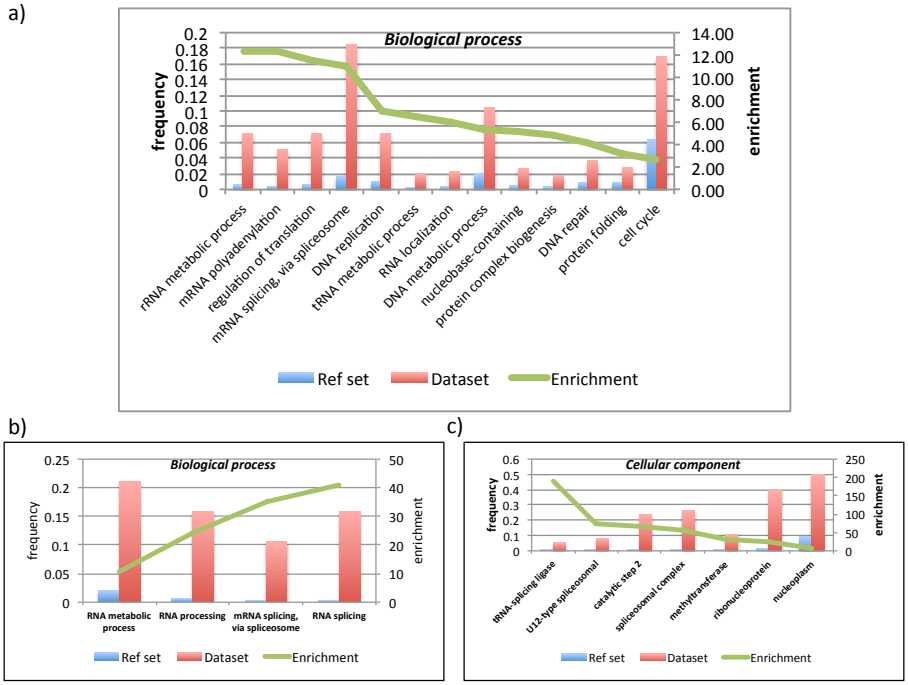


Figure 2. Gene Ontology enrichment analysis of FUS interactors. a) Gene Ontology enrichment analysis of the whole set of FUS interactors according to biological processes. b) and c) Gene Ontology enrichment analysis of the most conserved FUS interactors (listed in Table I) according to biological process (b) and cellular component (c). The frequency refers to the percentage of FUS interacting proteins annotated to a certain GO term in the dataset (red bar) and in the human reference set (blue bar); the enrichment value (green line) represents the ratio between the frequencies of the specific term in the FUS IPs and in the human genes reference dataset; all terms are statistically significantly enriched with a p-value <0.05 (after a Bonferroni correction for multiple testing).

Gene names	Category	Group
BUB3	core	A
CCAR1	core	A
DDX17	non-core	A
DDX5	core	A
FUS	core	A
LUC7L	non-core	A
RBM10	core	A
RBM39	non-core	A
SF1	core	A
THRAP3	core	B
SMU1	core	B
DDX41	core	C
TOE1	non-core	C2
DHX57	non-core	C2
GNB2L1	non-core	C2
MATR3	non-core	C2
RBM4	non-core	C2

Gene names	Category	Group
DDX3X	non-core	mRNA
ELAVL1	non-core	mRNA
ILF2	non-core	mRNA
NCBP1	core	mRNA
PABPC1	non-core	mRNA
SRRT	non-core	mRNA
YBX1	non-core	mRNA
SRRM1	non-core	SR
SRSF1	non-core	SR
SRSF10	non-core	SR
SRSF2	non-core	SR
SRSF4	non-core	SR
SRSF5	non-core	SR
SRSF6	non-core	SR
SRSF7	non-core	SR
SRSF9	non-core	SR
TRA2A	non-core	SR
TRA2B	non-core	SR

Gene names	Category	Group
SNRNP70	core	U1
SNRPA	core	U1
SRPK1	core	U1
SNRPB2	core	U2
SF3B4	core	U2
SF3B3	core	U2
SF3B2	core	U2
SF3B1	core	U2
SF3A3	core	U2
SF3A1	core	U2
CHERP	core	U2 rel
DHX15	core	U2 rel
RBM17	core	U2 rel
U2AF2	core	U2 rel
U2SURP	core	U2 rel

Gene names	Category	Group
PRPF3	core	U4/U6
PRPF31	core	U4/U6
PRPF4	core	U4/U6
EFTUD2	core	U5
PRPF6	core	U5
PRPF8	core	U5
SNRNP200	core	U5
SNRNP40	core	U5
SART1	core	U5*U4/U6
SRPK2	core	U5*U4/U6
BCAS2	core	Prp19
CDC5L	core	Prp19
PLRG1	core	Prp19
PRPF19	core	Prp19
XAB2	core	Prp19 rel
RBMX2	core	RES

Gene names	Category	Group
DHX38	core	2nd step
EIF4A3	core	EJC/TREX
DDX39B	non-core	EJC/TREX
THOC2	non-core	EJC/TREX
SYNCRIP	non-core	hnRNP
RBMX	non-core	hnRNP
RALY	non-core	hnRNP
PTBP1	non-core	hnRNP
PCBP2	non-core	hnRNP
PCBP1	non-core	hnRNP
HNRNPUL1	non-core	hnRNP
HNRNPU	non-core	hnRNP
HNRNPR	non-core	hnRNP
HNRNPM	non-core	hnRNP
HNRNPK	non-core	hnRNP
HNRNPH3	non-core	hnRNP
HNRNPH2	non-core	hnRNP
HNRNPH1	non-core	hnRNP
HNRNPF	non-core	hnRNP
HNRNPD	non-core	hnRNP
HNRNPC	non-core	hnRNP
HNRNPAB	non-core	hnRNP
HNRNPA3	non-core	hnRNP
HNRNPA2B1	non-core	hnRNP
HNRNPA1	non-core	hnRNP
HNRNPA0	non-core	hnRNP
CLNS1A	non-core	MISC
DHX9	non-core	MISC
KHDRBS1	non-core	MISC
KHDRBS3	non-core	MISC
PPP1CA	non-core	MISC
PRMT5	non-core	MISC
QKI	non-core	MISC
RBBP6	non-core	MISC
SMN1	non-core	MISC
WDR77	non-core	MISC

Figure 3. FUS interactors involved in splicing activities. The interactors of FUS that takes part in the splicing process are grouped according to snRNP association, function, presence in a stable heteromeric complex, or association with a particular spliceosomal complex, based on the commonly used nomenclature (Hegele et al., 2012; Wahl et al., 2009); they are also classified into “core” and “non-core” proteins based on their abundance in spliceosomal complexes (Agafonov et al., 2011).

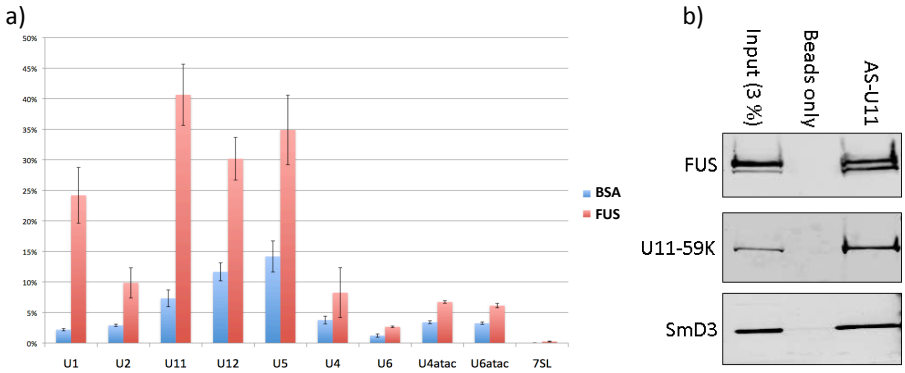


Figure 4. Quantification of U snRNAs interacting with FUS and validation of the interaction with U11 snRNP. a) RNA-IP reveals a strong enrichment of U1, U11 snRNA in the FUS immunoprecipitates. The bars indicate the relative percentages of snRNAs precipitated by FUS and BSA antibodies; the 7SL RNA served as negative control. b) Biotinylated antisense oligonucleotide pulldown. Biotinylated 2'O-methyl RNA complementary to the 5' end of the U11 snRNA coupled to magnetic streptavidin beads (AS-U11) or beads without any RNA oligonucleotide (Beads only) were incubated with HeLa nuclear extract. After blotting, proteins were detected by western blotting with antibodies against FUS, U11-59K, and SmD3.

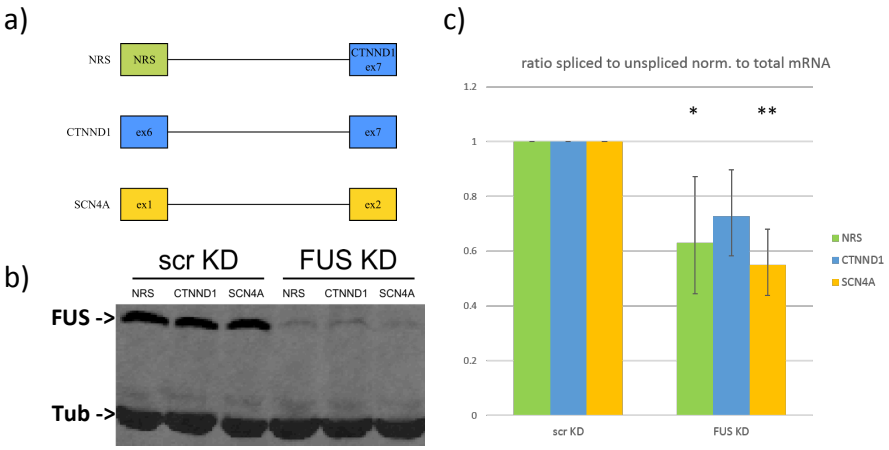


Figure 5. FUS knockdown influences the splicing of the minor introns. a) Schematic representation of the minor intron minigenes CTNND1, SCN4A and NRS. b) Western blot analysis with anti-FUS antibodies on total cell lysates from HeLa cells transfected with the minigenes and the pSUPuro plasmids targeting FUS or scrambled; the tubulin staining was used to normalized the protein amount. c) RT-qPCR results indicating the ratio of spliced to unspliced RNA resulting from the selected minigenes, normalized to the total amount of mRNA; single or double asterisk indicates the significance <0.05 and between 0.05 and 0.01 respectively.

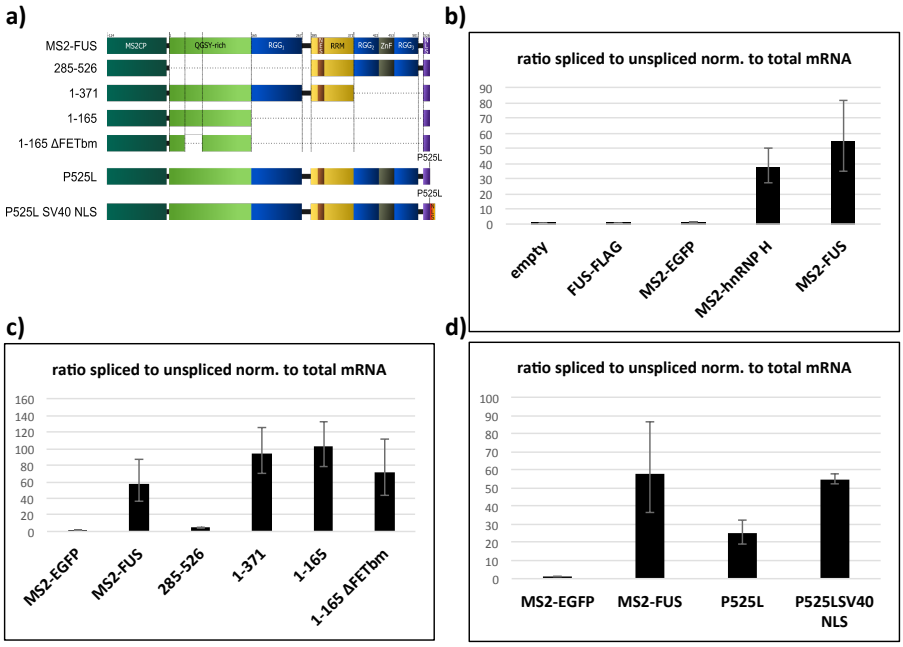


Figure 6. Tethering splicing assays with different FUS constructs on CTNND1 minigene. a) Schematic representation of the MS2-FUS constructs used for the tethered splicing assay. b) RT-qPCR results of the spliced or unspliced form of the CTNND1 reporter gene, comparing full-length wild-type MS2-FUS and MS2-HNRNPH proteins, or (c) full-length MS2-FUS respect to shortened FUS constructs, or (d) full-length wild-type MS2-FUS with FUS harbouring the P525L mutation or same mutant with the addition of the NLS of SV40 protein.

2.6 Materials and methods

Cell culture and transfections

HeLa and HEK293T cells were cultivated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FCS), penicillin (100 IU/mL), and streptomycin (100 µg/mL) and grown at 37°C and 5% CO2. Plasmid DNA transfections were performed with Dreamfect, Dogtor (OZ Biosciences), and Fugene HD (Promega) whereas siRNA transfections were performed using Lipofectamine2000 according to manufacturers instructions.

Plasmids

pSUPuro-FUS was cloned by inserting double-stranded oligos into pSUPERpuro between the BglII and HindIII sites as described (Brummelkamp et al., 2002; Paillusson et al., 2005). The shRNA expressed from pSUPuro FUS is targeting nucleotides 535-553

of FUS mRNA, numbering according to NM_004960, GGACAGCAGCAAAGCTATA). As control a scrambled shRNA sequence was used (as described in Bühler et al., 2006). To create the C-terminally flag-tagged expression constructs, an entry construct was generated by inserting a double stranded oligo coding for Glycin-Serine $(Glycine)_{15}$-flag into the HindIII, XbaI sites of pcDNA3. The double stranded oligo contained XhoI and BamHI sites, separated by seven nucleotides, upstream of the linker and the flag tag for subsequent insertion of the cDNAs. PCR amplified FUS and EBFP cDNAs were then cloned into the XhoI and BamHI sites of this vector. The FUS PCR product was amplified from the full length ImaGenes Clone IRAUp969F059D. To create the N-terminally MS2 tagged expression constructs, a fragment coding for the MS2 coat protein was PCR-amplified from pGMV-PABPN1-MS2-HA (as described in Eberle et al., 2008) and subsequently

cloned into the HindIII, XhoI sites of pcDNA3 and the NheI, HindIII sites of pcDNA3.1(+) respectively. Full length and truncated FUS coding sequences where PCR amplified from the pcDNA3-FUS-GSG15-flag construct described above and cloned subsequently into the XhoI, ApaI sites of pcDNA3-MS2 or the HindIII, XbaI sites of pcDNA3.1-MS2 respectively. pcDNA3-MS2-FUS P525L was created by PCR-based QuickChange mutagenesis with the QuickChange Lightning Multi Site-Directed Mutagenesis Kit [Agilent] using pcDNA3-MS2-FUS as template according to the manufacturers manual. To create pcDNA3.1-MS2-FUS P525L SV40 NLS, FUS P525L SV40 NLS was PCR-amplified from pcDNA3-FUS-GSG15-flag using the forward primer 5'-AAAAAAGCTTGCCACCATGGCCTCAAAC-3' and the mutagenic reverse primer 5'-GATTGGGCCCTTCACTTGTCCTCCACTTTGCGTTTCTTT TTGGGATACAGCCTCTCCCTGCGATCC-3' which adds

the coding sequence of the SV40 nuclear localization signal, and subsequently cloned into the HindIII, ApaI site of pcDNA3.1-MS2. The pcDNA6F-meG, pcDNA6F-FUS wt and pcDNA6F-FUS-P525L plasmids were created as follows: A synthethic double stranded oligo (GeneArt) containing the chimeric intron from pCI-neo followed by a flag tagged monomeric EGFP residing between XbaI and BamHI cloning sites was cloned into the SacI, PmeI sites from pcDNA6/TR thereby exchaning the Tet-Repressor with flag-mEGFP. The mEGFP was later on excised with XbaI and BamHI and PCR amplified FUS or FUS-P525L was cloned into the linearized vector.

Western Blot and antibodies

After gel electrophoresis, proteins were electrotransferred on a nitrocellulose membrane (HybondTM-ECLTM Amersham Biosciences) for 1

hour at 390 mA in Tris-Glycine Transfer Buffer (Trizma base 25 mM, Glycine 192 mM, Methanol 20%, SDS 0.1%). Then nitrocellulose membrane was stained for 5 minutes with Red Ponceau Solution in order to control the transfer. The membrane was washed once with water and then was saturated with 5% non-fat milk (Regilait) in TBS (Tris-HCl 20 mM, NaCl 0.5 M) 1 hour at room temperature or 16 hours at 4°C. After a single washing of 5 minutes with TBS-T (Tris-HCl 20 mM, NaCl 0.5 M, Tween 20 0.1%), the membrane was incubated with the primary antibody dissolved in a 5% milk-TBS solution at different dilution based on the experiment, for 1 hour at room temperature or 16 hours at 4°C. After three washings of 15 minutes with TBS-T, the membrane was incubated with a secondary antibody HRP-conjugated dissolved in a 5% milk-TBS solution. After exhaustive washings with TBS-T, bands were visualized after enhanced

chemiluminescence reaction, using ECL-PlusTM Western Blotting Reagents (Amersham Biosciences), followed by autoradiography or CCD detection (ChemiDoc™ MP System, Biorad). The polyclonal rabbit anti-FUS antibody was prepared as follows: a cDNA fragment, amplified by PCR, encoding the first 286 amino acids of FUS was cloned between the EcoRI and XhoI sites of pET28a. The recombinant protein was expressed in BL21(DE3) Codon Plus RIPL and purified under denaturing conditions over Ni-NTA beads according to the manufacturer's instructions. The purified protein was dialyzed against PBS, and rabbits were immunized with the purified protein in combination with GERBU Adjuvant LQ. The monoclonal FUS antibody (19B2) was produced by Paratopes Ltd against recombinant hexahistine-tagged FUS, produced by Ramesh S. Pillai (EMBL, Grenoble).

Immunoprecipitations for mass spectrometric analysis

HEK293T cells transfected with pcDNA3-FUS-GSG15-flag or pcDNA3-EBFP-GSG15-flag were harvested by trypsinization, followed by centrifugation for 5 minutes at 200 x g at 4°C. The cell pellets were washed once with PBS and cells were suspended with ice cold hypotonic gentle lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 2 mM EDTA, 0.1% Triton X-100, 1 x Halt Protease Inhibitor (Pierce)) in the absence (RNase free) or in the presence of 0.2 mg/ml RNaseA (RNase treated) to a final concentration of 1 x 10^7 cells/mL. Cells were lysed for 10 minutes on ice, followed by supplementation of NaCl to 150 mM final concentration and further incubation on ice for 5 minutes. The lysate was then cleared from insoluble particles by centrifugation (15 minutes at 16100 x g at 4 °C). Supernatant was recovered and incubated

with anti-flag™ M2 Affinity Gel (20 μl/1x107 cells; Sigma) for 1.5 h head over tail at 4 °C. The solution was centrifuged (5 minutes at 1000 x g at 4 °C). The affinity gel was suspended in 1 mL NET-2 (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Triton-X-100) and washed five times by subsequent suspension and centrifugation steps. For the high salt interactors (750 mM), the precipitate from RNaseA treated extracts were washed three times with NET-2 supplemented with NaCl to a final concentration of 750 mM. After the last wash, the buffer was completely removed with a syringe and to elute the precipitated proteins from the affinity gel, the resin was incubated with 1 bed volume of elution buffer (NET-2, 0.5 x protease inhibitor, 1 mg/ml flag peptide) and incubated 30 minutes head over tail at 4°C. Eluates were mixed with SDS-Gel loading buffer, boiled for 5 minutes at 95 °C and loaded on an 8% SDS-polyacrylamide gel. After

electrophoresis, gels were stained with Colloidal Coomassie (0.08% Coomassie Brilliant Blue G-250, 20% Ethanol, 8% Ammonium Sulfate, 1.598% Phosphoric Acid) and destained with distilled water.

Sample preparation and mass spectrometric analysis

For in-gel tryptic digestion, the lanes of interest were excised from the Coomassie-stained gels and fractionated in 10 gel slices. Disulphide bonds were reduced with 10 mM dithiothreitol, 30 minutes at 56 °C, alkylated with 55 mM iodoacetamide, 20 minutes at room temperature, and digested overnight at 37 °C with bovine trypsin (Sigma-Aldrich), as previously described (Shevchenko et al., 2006). The resulting peptide mixture was purified from the excess of salt and concentrated using C18 StageTips procedure, as previously described (Rappsilber et al., 2007). Mass spectrometry

analysis was performed by nano liquid chromatography–tandem MS (nLC–ESI-MS/MS) using a LTQ-Orbitrap mass spectrometer (Thermo Scientific), as following: 5 µl of purified peptide mixture were injected in a chromatographic system (EasyLC, Proxeon Biosystems) and peptides were separated on a 25 cm fused silica capillary (75 µm inner diameter and 360 µm outer diameter, Proxeon Biosystems) filled with Reprosil-Pur C18 3 µm resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) using a pressurised packing bomb. Peptides were eluted with a 95 min gradient from 7% to 70% of buffer B (80% ACN, 0.5% acetic acid) at a flow rate of 200 nl/min. The LC system was connected to the LTQ-Orbitrap equipped with a nano electrospray ion source. Full-scan mass spectra were acquired in the LTQ-Orbitrap mass spectrometer in the mass range m/z 350–1750 Da and with resolution set to 60000. The lock-mass

option was used to internally calibrate mass spectra for most accurate mass measurements. The 10 most intense doubly and multiply charged ions were automatically selected and fragmented in the ion trap with a CID set to 35%.

Peptides and proteins identification by database searching

Raw data files were analyzed using the peptide search engine Andromeda, which is integrated into the MaxQuant software environment (version 1.5.2.8) (Cox et al., 2011), with the following parameters: uniprot_cp_hum_2014_01 as protein database, Oxidation (M), Acetyl (Protein N-term), as variable modifications, Carbamidomethyl (C) as fixed modifications, peptide false discovery rate (FDR) 0.01, maximum peptide posterior error probability (PEP) 1, protein FDR 0.01, minimum peptides 2, at least 1 unique, minimum length

peptide 6 amino acids. Two biological replicates were performed. Only proteins identified in both replicates were selected for further analysis.

Gene Ontology enrichment analysis

Gene Ontology (GO) terms enrichment analysis was performed with the PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system (Mi et al., 2013), using the gene names of the identified proteins as queries for the statistical overrepresentation test and the most updated (at the time of analysis) Homo sapiens genes annotations as reference set. The electronically inferred annotations were excluded and only the over-represented categories with a p-value <0.05 (after Bonferroni correction for multiple testing) have been chosen to be reported in the graphs.

RNA-Immunoprecipitation

180 μl of Protein G Dynabeads were coupled to 40 μg of FUS antibody in TBS-0.05% NP-40 for 2 hours head over tail at 4°C followed by two washes with TBS-NP40 to remove uncoupled antibodies. To each sample fresh TBS-NP40 and 60 μl of HeLa nuclear extract (IPRACELL, Mons, Belgium) were added and incubated head over tail for 1.5 hours at 4°C. After five washes with TBS-NP40 the beads were resuspended in 1 ml TRIZOL and RNA was isolated according to the manufacturer's instructions. As Input, 60 μl of nuclear extract were directly added to 1 ml of TRIZOL. The snRNA analysis was essentially performed as in (Nizzardo et al., 2014). 2 μg of RNA were reverse transcribed at 37°C in 50 μl in 1 x small RNA RT buffer (10 mM Tris-HCl pH 8, 75 mM potassium chloride, 10 mM dithiothreitol, 70 mM magnesium chloride, 0.8 mM anchored universal reverse transcription primer-GCTGTCAACGATACGCTACGTAACGGCATGACAGTGT

TTTTTTTTTTTTTTTTVN-, 2 U/µL of RiboLock [Fermentas], 10 mM dNTPs, and 2.5 mM rATP) supplemented with 5U of Escherichia coli Poly(A) Polymerase (New England Biolabs) and 1 µL of AffinityScript reverse transcriptase (Agilent). Reactions were heat inactivated for 10 minutes at 85°C. Reverse transcribed material corresponding to 18 ng of RNA was amplified with MESA GREEN qPCR Master Mix Plus for SYBR (Eurogentec) and the appropriate primers (600 nM each) in a total volume of 20 µL, using the Rotor Gene 6000 rotary analyser (Corbett).

Biotinylated antisense oligo pulldown

200 µl of magnetic streptavidin dynabeads slurry (MyOne T1, Life technologies) were washed three times with PBS and blocked for 30 minutes at 4 °C in blocking buffer (250 µg/ml tRNA, 250 µg/ml glycogen, 250 µg/ml BSA in PBS). 400 µl of buffer

D+/+ (20 mM Hepes pH 7.9, 100 mM KCl, 20% Glycerol, 0.25 mM EDTA, 0.5 mM DTT, 1x Protease Inhibitor, 1x Phosphatase Inhibitor, 0.01% NP-40) were supplemented with 100 µl of HeLa nuclear extract (HNE, Ipracell) and 200 pmol biotinylated U11 antisense oligo (ACGACAGAAGCCCUUUU-Bio-Bio-Bio, Microsynth). As negative control the antisense oligonucleotide was omitted. The oligonucleotides were hybridized for 45 minutes head over tail at 30 °C. The reactions were cleared by centrifugation (5 minutes at 16100 x g), and supernatants were transferred to new eppendorf tubes. 100 µl of blocked streptavidin beads were added to the supernatants each and incubated for 45 minutes at 4 °C head over tail. Subsequently, the beads were washed five times with 1 ml of buffer D+/+ supplemented with 0.05 % NP-40. With the last wash the beads were transferred to new eppendorf tubes, wash buffer was removed and

beads were resuspended in 50 µl of 2x LDS-loading buffer, boiled for 10 minutes at 70 °C and loaded on a 4-12% NuPage Gel (Life-technologies).

U12-dependent splicing under FUS knockdown conditions

2.5 x 105 HeLa cells where seeded into a well of a 6well plate in DMEM (day 0). On day 1, the cells were co-transfected with 100 ng reporter construct (NRS, CTNND1, SCN4A) and 500 ng pSUPuro-FUS or pSUPuro scrambled respectively using Dogtor. On day 2, the cells were split into a T25 flask in DMEM containing 2 µg/ml Puromycin [Santa Cruz]. On day 3, the medium was replaced by DMEM+/+ containing no Puromycin. 6-8 h later, the cells were harvested and relative mRNA levels were quantified by RT-qPCR.

Tethered splicing assay

2.5 x 10^5 HeLa cells where seeded into a well of a 6well plate in DMEM (day 0). On day 1, the cells were co-transfected with 300-800 ng plasmid coding for the MS2-fusion protein and 100-400 ng CTNND1 reporter construct using Dogtor. On day 2, the cells were split into a T25 flask. On day 3 the cells were harvested and relative mRNA levels were quantified by RT-qPCR.

RT-qPCR

HeLa cells from splicing assays were harvested by trypsinization. 2 x 10^5 cells where removed to verify protein expression and/or knockdown efficiency respectively by western blotting. The RNA of the remaining cells was isolated using guanidium thiocyanate:phenol:chlorophorm extraction (Metze et al., 2013). The RNA samples were DNase treated according to the TURBO DNA-free™ Kit [Life Technologies]. 1 µg total RNA was reverse

transcribed using 450 ng random hexamers [Microsynth], 1 x AffinityScript RT buffer, 1 µl AffinityScript Multiple Temperature Reverse Transcriptase, 0.4 mM dNTPs, and 10 mM DTT [Agilent] according to the manufacturers manual. To confirm successful DNase digestion, controls lacking reverse transcriptase were made. The cDNA was diluted to a RNA concentration of 8 ng/µl. Quantitative real-time PCR was performed using 3 µl cDNA, 1 x MESA GREEN qPCR Mastermix Plus for SYBR® Assay No ROX [Eurogentec] and each 8 µM forward and reversed primer respectively in a total volume of 15 µl. Samples were measured in duplicates using Rotorgene6000 [Corbett] using the following cycling conditions: 95 °C, 5 minutes; 95 °C, 15 seconds; 60 °C 1 minutes; 40 cycles. A melting curve was recorded from a temperature gradient from 65 °C to 95 °C, 5 sec/°C. Analysis was performed as described in (Metze et al., 2013). The

statistical significance of qPCR results was determined by a Welch's t-test.

2.7 References

Agafonov, D.E., Deckert, J., Wolf, E., Odenwälder, P., Bessonov, S., Will, C.L., Urlaub, H., and Lührmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. Mol. Cell. Biol. *31*, 2667–2682.

Behzadnia, N., Golas, M.M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C.L., Urlaub, H., Stark, H., et al. (2007). Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. EMBO J. *26*, 1737–1748.

Brummelkamp, T., Bernards, R., and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. Science (80-.). *296*, 550–553.

Bühler, M., Steiner, S., Mohn, F., Paillusson, A., and Mühlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nat. Struct. Mol. Biol. *13*, 462–464.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. a, Olsen, J. V, and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2006). Protein composition and electron microscopy structure of affinity-purified

human spliceosomal B complexes isolated under physiological conditions. Mol. Cell. Biol. *26*, 5528–5543.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol. *6*, 849–859.

Hackl, W., and Lührmann, R. (1996). Molecular cloning and subcellular localisation of the snRNP-associated protein 69KD, a structural homologue of the proto-oncoproteins TLS and EWS with RNA and DNA-binding properties. J. Mol. Biol. *264*, 843–851.

Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Lührmann, R., and Stelzl, U. (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell *45*, 567–580.

McNally, L.M., Yee, L., and McNally, M.T. (2006). Heterogeneous nuclear ribonucleoprotein H is required for optimal U11 small nuclear ribonucleoprotein binding to a retroviral RNA-processing control element: Implications for U12-dependent RNA splicing. J. Biol. Chem. *281*, 2478–2488.

Metze, S., Herzog, V. a, Ruepp, M.-D., and Mühlemann, O. (2013). Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA *19*, 1432–1448.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. *8*, 1551–1566.

Nizzardo, M., Simone, C., Salani, S., Ruepp, M.D., Rizzo, F., Ruggieri, M., Zanetta, C., Brajkovic, S., Moulton, H.M., Müehlemann, O., et al. (2014). Effect of combined systemic and local morpholino treatment on the spinal muscular atrophy δ7 mouse model phenotype. Clin. Ther. *36*, 340–356.

Onodera, O., Ishihara, T., Shiga, A., Ariizumi, Y., Yokoseki, A., and Nishizawa, M. (2014). Minor splicing pathway is not minor any more: Implications for the pathogenesis of motor neuron diseases. Neuropathology *34*, 99–107.

Paillusson, A., Hirschi, N., Vallan, C., Azzalin, C.M., and Mühlemann, O. (2005). A GFP-based reporter system to monitor nonsense-mediated mRNA decay. Nucleic Acids Res. *33*, 1–12.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2014). Biochemical Properties and Biological Functions of FET Proteins. Annu. Rev. Biochem. 1–25.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. *1*, 2856–2860.

Turunen, J.J., Niemelä, E.H., Verma, B., and Frilander, M.J. (2013). The significant other: Splicing by the minor spliceosome. Wiley Interdiscip. Rev. RNA *4*, 61–76.

Vance, C., Scotter, E.L., Nishimura, A.L., Troakes, C., Mitchell, J.C., Kathe, C., Urwin, H., Manser, C., Miller, C.C., Hortobágyi, T., et al. (2013). ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. Hum. Mol. Genet. *22*, 2676–2688.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. Cell *136*, 701–718.

Yamazaki, T., Chen, S., Yu, Y., Yan, B., Haertlein, T.C., Carrasco, M. a., Tapia, J.C., Zhai, B., Das, R., Lalancette-Hebert, M., et al. (2012). FUS-SMN Protein Interactions Link the Motor Neuron Diseases ALS and SMA. Cell Rep. *2*, 799–806.

Yu, Y., Chi, B., Xia, W., Gangopadhyay, J., Yamazaki, T., Winkelbauer-Hurt, M.E., Yin, S., Eliasse, Y., Adams, E., Shaw, C.E., et al. (2015). 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*, 3208–3218.

Chapter 3

Proteomic characterization of the role of FUS/TLS in DNA damage response

Manuscript in preparation

Giuseppe Filosa[1,2], Angela Bachi[2], Silvia Barabino[1]

[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy; [2]IFOM, FIRC Institute of Molecular Oncology, via Adamello 16, 20139 Milan, Italy

3.1 Abstract

FUS is a ubiquitously expressed RNA-binding protein that has been discovered as fused to transcription factors in several human sarcomas

and found in protein aggregates in neurons of patients with an inherited form of Amyotrophic Lateral Sclerosis. FUS has been implicated in a variety of cellular processes such as gene expression control, transcriptional regulation, pre-mRNA splicing and miRNA processing. In addition, several evidences link FUS to genome stability control and DNA damage response. Previous interactome analysis of wild type FUS protein in a human cell line highlighted the interaction with proteins involved in DNA damage response, therefore with the present study we aim to investigate the role of FUS in DNA damage response by dissecting its interactome in presence of genotoxic stress. HEK293T cells stably expressing a flag-tagged version of wild-type FUS protein were treated with Etoposide to generate DNA double-strand breaks and FUS interacting proteins have been identified and quantified by anti-flag

immunoprecipitation followed by label-free quantitative mass spectrometry. The results highlighted an enhanced interaction between FUS and nucleic acids binding proteins, in particular to RNA-binding proteins involved in splicing. This evidence suggests that FUS might play a role in the regulation of the DDR by recruiting specific RNA-binding proteins that have been recently characterized as key regulators of the DDR.

3.2 Introduction

FUS/TLS (fused in sarcoma/translocated in liposarcoma, hereafter indicated as FUS) is a ubiquitously expressed RNA-binding protein of the hnRNP family, which has been discovered as fused to transcription, factors in several human malignancies and found in protein aggregates in neurons of patients with an inherited form of Amyotrophic Lateral Sclerosis (Vance et al., 2009).

FUS is a 53 kDa nuclear protein of the FET family that contains structural domains, such as a RNA Recognition Motif (RRM) and a zinc finger motif, that give to FUS the ability to bind to both RNA and DNA sequences (Schwartz et al., 2014). FUS has been implicated in a variety of cellular processes, such as pre-mRNA splicing, miRNA processing, gene expression control and transcriptional regulation (Sama et al., 2014). Moreover, some evidences link FUS to genome stability control and DNA damage response (DDR). For example, mice lacking FUS are hypersensitive to ionizing radiation (IR) and show high levels of chromosome instability (Kuroda et al., 2000) and, in response to double-strand breaks (DSBs), FUS is phosphorylated by the protein kinase ATM (Gardiner et al., 2008). In addition, the presence of a RanBP2-type zinc finger confers to FUS the ability to act as E3 SUMO ligase (SUMO, Small Ubiquitin-like MOdifier), and it has been

described that upon DNA damage stress, FUS mediates Ebp1 (ErbB3 receptor-binding protein) SUMOylation, and this post-translational modification is required for its oncosuppressive activity (Oh et al., 2010). Despite these evidences, the exact role of FUS in these cited pathways is still unclear. Therefore, in order to shed light on the biological role that FUS exerts in the context of the DDR, we applied a quantitative mass spectrometry approach to identify whether its interacting partner might change upon DNA damage.

3.3 Results

3.3.1 FUS interacts with DDR-related proteins

As described previously (Chapter 2), we performed a mass spectrometric-based interactome analysis on FUS wild type proteins in order to dissect the biological pathways in which it is involved. Briefly,

the analysis resulted in the identification of about six hundreds potential interactors, arising from different conditions, with/without RNaseA treatment and with low or high stringent washings with NaCl. A Gene Ontology enrichment analysis, performed on the whole set of interactors, revealed an enrichment of several pathways in which FUS was described to take part, especially the ones related to RNA metabolism. Among these, the DNA damage repair biological process also resulted to be enriched in the dataset with respect to the human reference set [Figure 1]. This is consistent with what has been reported in literature (as described in Chapter 1) and with our experimental evidence that several DDR-related proteins were identified in the FUS interactome [listed in Table 1]. The interaction between FUS and these proteins is favoured by the presence of RNA molecules, since most of them were identified in the samples without RNaseA

treatment (proteins whose interaction with FUS is dependent from the presence of the RNA are marked with “+” in Table I).

In order to confirm this evidence, we validated the interaction of FUS with some selected DDR-related interactors by reverse immunoprecipitation and western blot analysis. HEK293 cells were transfected with recombinant flag-tagged version of wild type XRCC5, XRCC6, NONO, SFPQ, and HDAC1. Cell lysis was performed with and without RNaseA treatment. Then we performed an anti-flag immunoprecipitation followed by Western blot analysis to probe for the presence of FUS in the elutions from the different immunoprecipitations. The results confirmed the interaction of these proteins with FUS, in particular in presence of the RNA, since in the RNaseA treated conditions the signal of FUS in the elution is very low (consistently with the mass spectrometric data) [Table I; Figure

2a, 2b]. Moreover, we also confirmed the interaction of FUS with HDAC1, as reported by Wang and colleagues (Wang et al., 2013), although we were not able to detect this protein in the previous interactome analysis.

3.3.2 Proteomic characterization of FUS interactome upon DNA damage

In order to investigate the biological significance of these interactions and to clarify the role that FUS exerts in the DDR, we sought to dissect the interactome of FUS in presence of genotoxic stress. To generate genotoxic stress we decided to use Etoposide, as reported in Wang and colleagues, 2013. Etoposide is a cytotoxic anticancer drug that belongs to the topoisomerase inhibitor drug class. It blocks topoisomerase II by forming a ternary complex with DNA and the topoisomerase II enzyme (which aids in DNA unwinding) (Pommier et

al., 2010); this prevents the re-ligation of the DNA strands causing DNA strands breaks (Montecucco and Biamonti, 2007). In order to choose the best conditions to cause genotoxic stress in HEK293T FUS-flag expressing cells (that will be used for the interactome analysis) we tested a standard concentration of drug (10 μM), which is believed not to interfere with cell viability, at different time points, and we follow the activation of the DDR. Cell lysis was performed with the addition of the Benzonase DNase enzyme, which digest DNA thus helping to recover the fraction of FUS which is bound to the chromatin (Yang et al., 2014). Lysates were subjected to SDS-PAGE and western blot analysis to follow the activation of the DDR by probing for the phosphorylated form of histone H2AX (γH2AX). The results showed that at 15 minutes there is DNA damage similar to the basal level, while at 30 and 60 minutes there is a

significant activation of the DDR. As expected, the amount of FUS remains stable and the genotoxic stress does not affect the total amount of the protein [Figure 3].

In order to test whether the Etoposide treatment (10 μM for 1 hour) might affect the subcellular localization of FUS, we performed a Western blot analysis, on fractionated cell lysates, by keeping separated the cytosol, the nucleus and the chromatin fractions. The analysis clearly showed that FUS is mostly enriched at chromatin level, due to its capacity to bind DNA and active chromatin regions (as previously described in Chapter 1), and that the treatment does not affect its abundance on chromatin [Figure 4].

In order to characterize whether the interactome of FUS changes upon DNA damage we decided to apply a label-free quantitative mass spectrometry approach to identify and relatively quantify FUS

interacting proteins whose binding affinity changes upon genotoxic stress. To induce DDR, HEK293T cells stably expressing wild type FUS-flag were treated with Etoposide 10 µM for 1 h (DMSO was added to the control cells). Cells were lysed and total extract was used as input for anti-flag immunoprecipitation. Proteins eluted from the immunoprecipitations were subjected to proteolytic digestion, by the FASP protocol (Filter Aided Sample Preparation, Wiśniewski et al., 2009). Peptides were injected in triplicate in the mass spectrometer and proteins were identified and quantified with the MaxQuant and Perseus software (Cox and Mann, 2008) [Figure 5a]; two biological replicates of the experiment were performed and only proteins quantified in both experiments were processed for further analysis. Western blot analysis of the inputs and elutions of the immunoprecipitations showed that the immunoprecipitation successfully enriched

for FUS, and that the treatment with the Etoposide induced genotoxic stress as indicated by the expression of the DDR marker γH2AHX [Figure 5b]. The mass spectrometric analysis resulted in the identification of most of the FUS interactors previously identified in our interactome analysis. To relatively quantify changes in FUS interactome with and without genotoxic stress we applied t-test statistical analysis on the normalized intensity values from the label-free analysis (LFQ-intensities) (Cox et al., 2014). The results of the analysis are reported in the volcano plot, in which the $\log_2$ ratio between the averaged LFQ-intensities of the proteins identified in the damaged condition (ETOP) versus the normal condition (DMSO) are plotted against the negative $\log_{10}$ p-value (after a Benjamini-Hochberg false discovery rate correction) of the test. The proteins with a fold change of at least 1.3 and a corrected p-value <0.05 were

considered significantly changing in the two conditions [Figure 5c].

3.3.3 DDR enhances the interaction between FUS and RNA-binding proteins

For the purpose of the experiment we then focused our attention on the proteins involved in the DDR. Surprisingly, for most of the DDR-related FUS interactors (listed in Table 1) we did not observe significant changes in the affinity for FUS upon genotoxic stress. In addition, from the quantitative analysis we identified only few proteins (consistently in both biological replicates) with a lower affinity for FUS upon the Etoposide treatment; these proteins are CDK9 (Cyclin-dependent kinase 9), Histone H2A, Lamin B1, and SPTAN1 (Spectrin alpha chain, non-erythrocytic 1). In contrast, the quantitative analysis resulted in the identification of a set of proteins with an enhanced

affinity for FUS upon DNA damage (listed in Table 2). Almost all these proteins were identified as FUS interactors from our previous analysis, except for ILF3 (Interleukin enhancer-binding factor 3), DNAJC8 (DnaJ homolog subfamily C member 8) and PCNP (PEST proteolytic signal-containing nuclear protein). Interestingly, we noticed that most of the proteins with higher affinity for FUS are nucleic acids binding proteins, especially RNA-binding proteins [Figure 6]. In order to confirm this evidence, we performed on this dataset of proteins a Gene Ontology enrichment analysis, by searching for the molecular functions and the cellular component to which they are annotated. The analysis confirmed the over-representation of nucleic acids binding proteins and in particular of RNA-binding proteins in the dataset [Figure 7a]; in addition there is a significant enrichment of proteins that are part of ribonucleoprotein

complexes and therefore are directly involved in the splicing process [Figure 7b].

3.4 Discussion

Several lines of evidence link FUS to the DNA damage response and repair (as discussed earlier in Chapter 1). However, the molecular mechanism through which FUS exerts its functions is still not clear. Therefore we applied mass spectrometric-based proteomic approaches first to identify its interacting partners in basal condition, then to dissect its interactome upon DNA damage in order to highlight the possible pathway in which FUS is involved during the DDR. We demonstrated an interaction between FUS and several well-known DDR-related proteins, which are involved in most of the steps of the DDR [Table 1]. However, from the quantitative analysis we did not notice significant changes in the interaction with these proteins upon

genotoxic stress. In contrast, we characterized that Etoposide-induced DNA damage enhances the interaction between FUS and a subset of RNA-binding proteins, most of them involved directly in the splicing pathway. This evidence is biologically relevant in the context of the DDR since from recent literature is now evident that RNA-binding proteins represent the major players in the prevention of genome instability, because of several reasons: 1) RBPs prevent harmful RNA/DNA hybrids and are involved in the DNA damage response, 2) RBPs allow the selective regulation of DDR genes at multiple post-transcriptional levels (from pre-mRNA splicing/polyadenylation to mRNA stability translation) and are directly involved in DNA repair; 3) RBPs bind to mRNAs, nascent transcripts, noncoding RNAs, and damaged DNA. Moreover, several RBPs, included FUS, upon genotoxic stress,

change their cellular localization and move towards the DDR sites (as reviewed in Dutertre et al., 2014). Recently, the splicing and the alternative splicing pathways have been involved in the DDR. In fact alternative splicing alterations can activate the DDR or inactivate the function of genes directly involved in such response, and this can also lead to the accumulation of DNA damage (reviewed in Lenzken et al., 2013). In addition, it has been demonstrated that cellular stress induces the formation of nuclear bodies, such as paraspeckles, that nucleate on scaffolds consisting of non-coding RNAs, such as NEAT1_2; these structures consist of RBPs and RNAs that together regulate several function within the cell (Naganuma and Hirose, 2013). RBPs identified in our experiments as interactors of FUS (NONO, SFPQ, RBM14, HNRNPK and HNRNPH3) have also a role in the assembly and stabilization of the paraspeckles. However, the genotoxic stress

didn’t result in an increased interaction with these proteins, with the only exception of HNRNPH3. Moreover, it has been demonstrated that low-complexity regions in various RNA-binding proteins, including FUS, are thought to be critical for the formation of membrane-less subcellular structures and mediate the subsequent recruitment of additional proteins and RNAs to these structures (Han et al., 2012). In addition, PAR-CLIP analysis showed that FUS binds to both NEAT1_1 and NEAT1_2 (Hoell et al., 2011; Naganuma et al., 2012). Importantly, some of the proteins identified as higher affinity interactors (Table 1, Chapter 2) have been recently involved in DDR. Among these, the RBMX protein, an hnRNP involved in alternative splicing (Heinrich et al., 2009), was found in a genome-wide siRNA-based screen to detect regulators of homologous recombination (HR). RBMX regulates HR in a positive manner,

accumulates at sites of DNA damage in a PARP1-dependent manner and promotes resistance to several DNA damaging agents (Adamson et al., 2012). Moreover, the ILF2/ILF3 heterodimer (also known as NF45/NF90), interacts with DNA-PK, can module its functions, and cells depleted from ILF2/ILF3 showed a reduced non-homologous end-joining (NHEJ) activity *in vivo* (Shamanna et al., 2011). All these evidences, deriving from both our experiments and prior knowledge, lead to the hypothesis that upon DNA damage, the cell starts a DDR response, which involves the recruitment of RBPs and formation of nuclear bodies to the damaged site; such NBs nucleate on scaffold ncRNAs (e.g. NEAT1_2 for paraspeckles) and are constituted by RBPs. In this context, FUS might play both a direct role in regulating the expression of DDR genes or function as scaffold protein by

recruiting partner RBPs and ncRNAs to the sites of the DNA damage.

3.5 Figures and tables

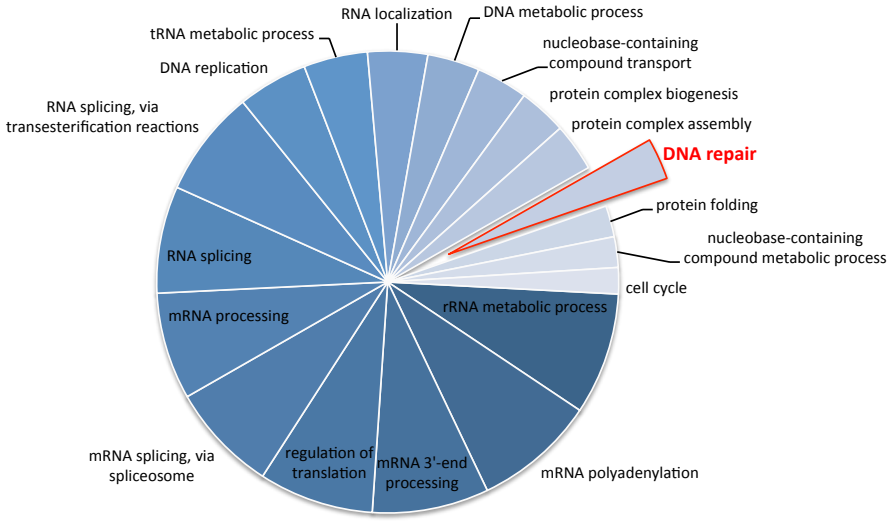
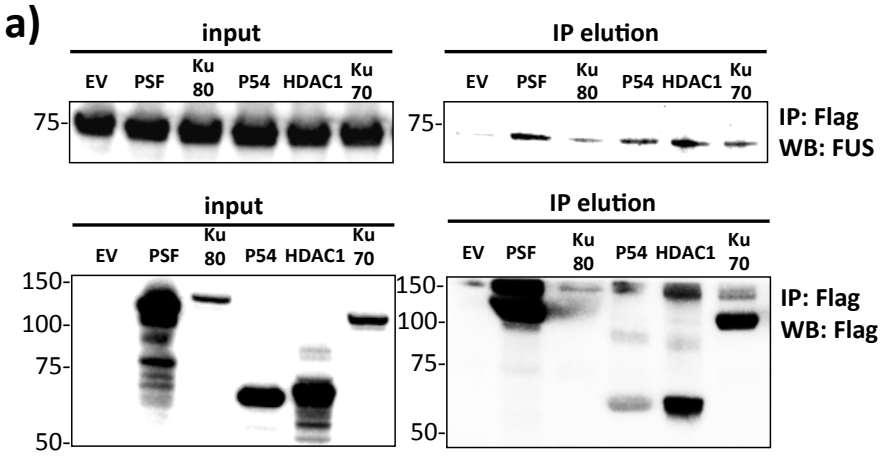


Figure 1. Gene Ontology analysis of FUS interactors. Gene Ontology Biological process analysis was performed on the total FUS interactors; the colour gradient indicates the fold change enrichment of the GO term in the dataset respect to the human reference set; in red the DNA repair process is highlighted. All terms are statistically significantly enriched with a p-value <0.05 (after a Bonferroni correction for multiple testing).

Table 1 – FUS interactors involved in DNA damage response

Protein names	Gene names	DDR pathway	RNA dependent
Activating signal cointegrator 1 complex subunit 3	ASCC3	DNA dealkylation involved in DNA repair	+
Bromodomain adjacent to zinc finger domain protein 1A	BAZ1A	DNA repair	+
Bcl-2-associated transcription factor 1	BCLAF1	positive regulation of response to DNA damage stimulus	+
Bloom syndrome protein	BLM	DNA repair	+
Cell division cycle 5-like protein	CDC5L	signal transduction involved in DNA damage checkpoint	+
Cyclin-dependent kinase 9	CDK9	regulation of DNA repair	+
Casein kinase I isoform epsilon	CSNK1E	DNA repair	+
ATP-dependent RNA helicase DDX1	DDX1	DNA duplex unwinding; double-strand break repair	
Probable ATP-dependent RNA helicase DDX17	DDX17	double-strand break repair	
ATP-dependent RNA helicase DDX3X	DDX3X	DNA duplex unwinding	
Probable ATP-dependent RNA helicase DDX5	DDX5	positive regulation of DNA damage response, signal transduction by p53 class mediator	
DNA excision repair protein ERCC-6	ERCC6	base-excision repair	+
Ras GTPase-activating protein-binding protein 1	G3BP1	DNA duplex unwinding	+
Interleukin enhancer-binding factor 2	ILF2	double-strand break repair via nonhomologous end joining	
Matrin-3	MATR3	DNA repair	
DNA-3-methyladenine glycosylase	MPG	base-excision repair	+
28S ribosomal protein S35, mitochondrial	MRPS35	DNA damage response, detection of DNA damage	+
28S ribosomal protein S9, mitochondrial	MRPS9	DNA damage response, detection of DNA damage	+
Non-POU domain-containing octamer-binding protein	NONO	DNA repair	+
Nucleophosmin	NPM1	DNA repair	+
Proliferation-associated protein 2G4	PA2G4	cell cycle arrest; cell prolferation	+
Poly [ADP-ribose] polymerase 1	PARP1	DNA repair	+
Poly [ADP-ribose] polymerase 2	PARP2	DNA repair	+
DNA-dependent protein kinase catalytic subunit	PRKDC	signal transduction involved in mitotic G1 DNA damage checkpoint	+
Pre-mRNA-processing factor 19	PRPF19	double-strand break repair via nonhomologous end joining	
Proteasome subunit alpha type-4	PSMA4	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	+
26S protease regulatory subunit 4	PSMC1	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	
26S protease regulatory subunit 8	PSMC5	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	
26S proteasome non-ATPase regulatory subunit 3	PSMD3	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	
RNA-binding protein 14	RBM14	DNA repair	+
Replication factor C subunit 1	RFC1	nucleotide-excision repair	+
Replication factor C subunit 2	RFC2	nucleotide-excision repair	+
Replication factor C subunit 3	RFC3	DNA synthesis involved in DNA repair	+
Replication factor C subunit 4	RFC4	DNA repair	+
Replication factor C subunit 5	RFC5	DNA repair	+
40S ribosomal protein S3	RPS3	cellular response to DNA damage stimulus	+
RuvB-like 2	RUVBL2	cellular response to UV	+
Splicing factor, proline- and glutamine-rich	SFPQ	double-strand break repair via homologous recombination	
DNA topoisomerase 2-alpha	TOP2A	cellular response to DNA damage stimulus	+
Transcription intermediary factor 1-beta	TRIM28	DNA repair	+
E3 ubiquitin-protein ligase TRIP12	TRIP12	cellular response to DNA damage stimulus	+
Ubiquitin carboxyl-terminal hydrolase 10	USP10	DNA damage response, signal transduction by p53 class mediator	+
pre-mRNA 3 end processing protein WDR33	WDR33	postreplication repair	+
Pre-mRNA-splicing factor SYF1	XAB2	transcription-coupled nucleotide-excision repair	
X-ray repair cross-complementing protein 5	XRCC5	double-strand break repair via nonhomologous end joining	+
X-ray repair cross-complementing protein 6	XRCC6	double-strand break repair via nonhomologous end joining	+



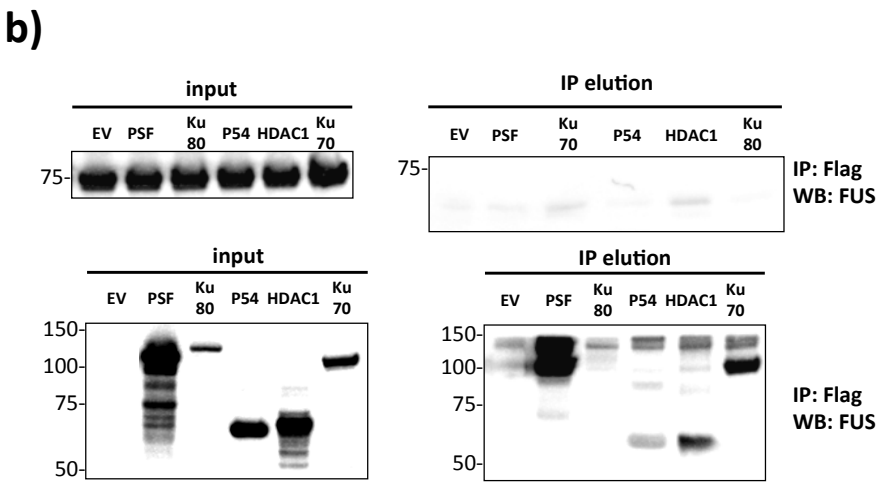


Figure 2. FUS interacts with proteins involved in DNA repair process. Validation of the interaction by reverse immunoprecipitation with selected DNA repair-related proteins. HEK293T cells were transiently transfected with flag-tagged PSF, p54, Ku80, Ku70, HDAC1, and empty vector. Cell lysis was performed with (a) or without (b) RNaseA treatment and upon anti-flag immunoprecipitation the inputs and the resulting elutions were probed for the presence of the endogenous FUS by anti-FUS immunoblotting and for the effective immunoprecipitation of the flag-tagged proteins by anti-Flag immunoblotting.

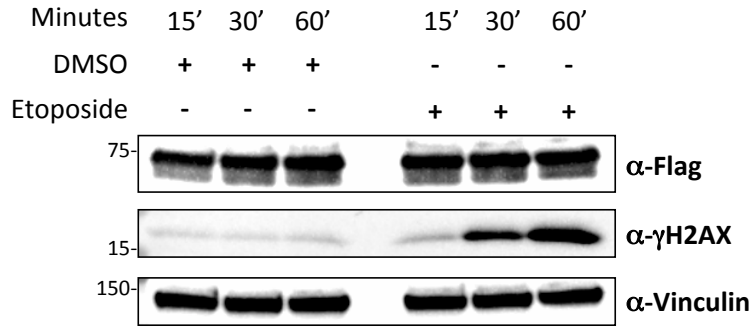


Figure 3. Etoposide treatment of HEK293T FUS-flag expressing cells. Western blot analysis of total cell lysate of HEK293T FUS-flag expressing cells treated with Etoposide 10 mM (or DMSO as control), for 15, 30 and 60 minutes. Anti-Vinculin staining was used to normalize the signal of FUS and γH2AX.

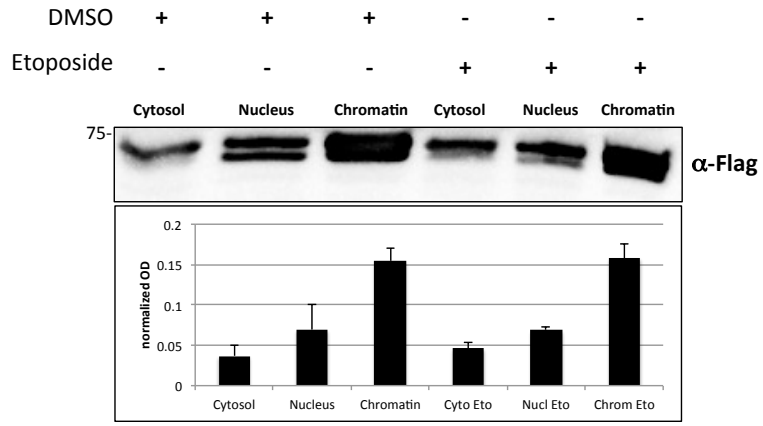


Figure 4. Subcellular localization of FUS upon Etoposide treatment. Western blot analysis of a fractionated cell lysate of HEK293T FUS-flag expressing cells treated with Etoposide 10 μM (or DMSO as control) for 1 hour. The optical density of the signal of FUS-flag in the

different fractions of the lysates were normalized on the total protein amount of each fraction and represents the average of three biological replicates (standard deviation as error bars).

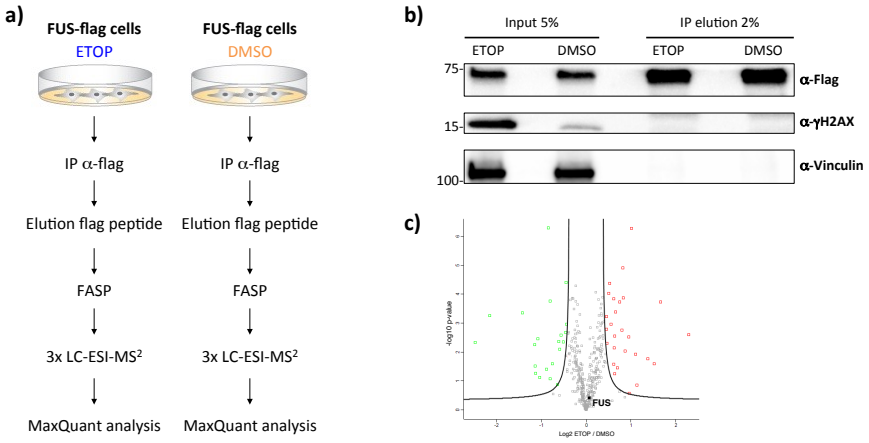


Figure 5. FUS interactome analysis upon DNA-damage. a) Schematic workflow of the label-free interactome analysis with/without DNA damage. HEK293T cells stably expressing FUS-flag protein were treated with 10 μM Etoposide for 1 hour (ETOP) or DMSO as control. FUS interactors were purified from the total cell extract by anti-flag IP and the elution was processed with FASP protocol prior to LC-ESI-MS2 in triplicate and bioinformatics analysis. b) Western blot analysis of the anti-flag IP; the anti-Vinculin immunoblotting was used to normalize the input; c) Representative Volcano-plot of the result of the t-test comparing the intensities of the proteins identified in Etoposide treated and DMSO IPs; log2 ETOP/DMSO indicates the

ratio between the mean protein intensities; values are plotted against the negative logarithmized p-value (after FDR Benjamini-Hochberg correction); proteins with a log2 ratio ETOP/DMSO >0.5 were defined as “higher affinity interactors” (red square) or <0.5 were defined as “lower affinity interactors” (green square).

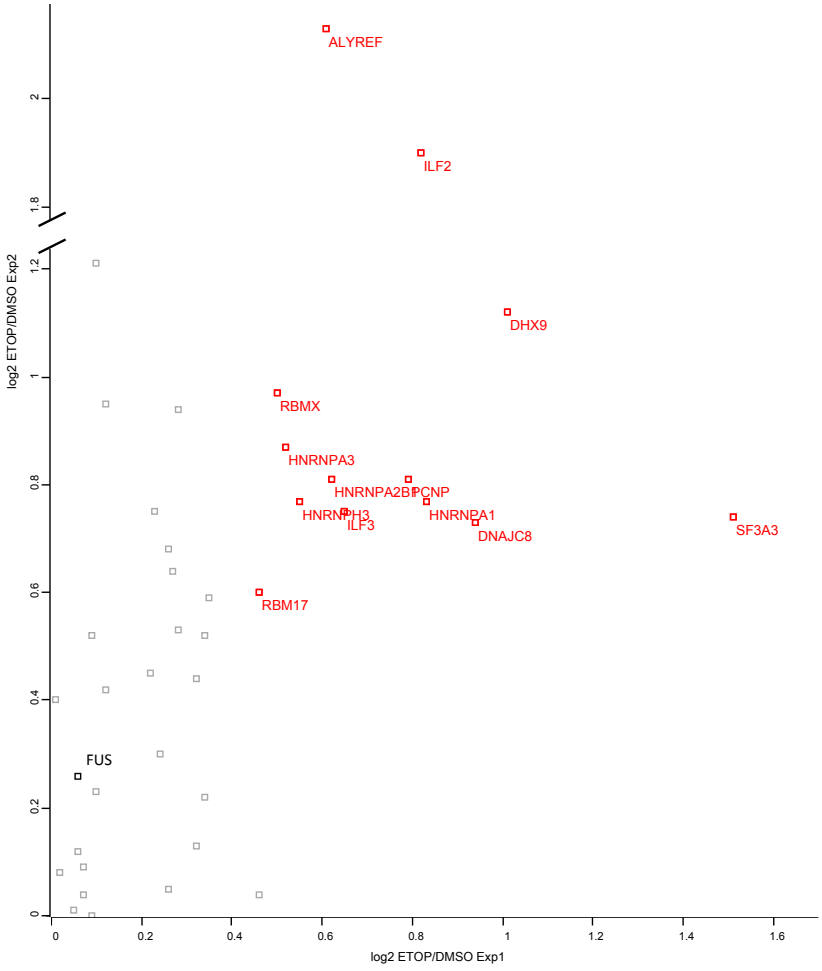


Figure 6. FUS interactors with higher affinity for FUS upon DNA damage. The scatter plot highlights in red the proteins (labelled with

the corresponding gene name) that have a significantly higher affinity for FUS upon DNA damage in both biological replica.

Table 2 – Proteins with higher affinity for FUS upon DNA damage

Proteins	Gene names	Ratio ETOP/DMSO	Unique peptides DMSO	Unique peptides ETOP	FUS Interactome
THO complex subunit 4	ALYREF	2.95	3	4	+
Interleukin enhancer-binding factor 2	ILF2	2.75	5	7	+
Splicing factor 3A subunit 3	SF3A3	2.26	2	2	+
ATP-dependent RNA helicase A	DHX9	2.09	32	41	+
DnaJ homolog subfamily C member 8	DNAJC8	1.79	6	6	
Heterogeneous nuclear ribonucleoprotein A1	HNRNPA1	1.74	11	12	+
PEST proteolytic signal-containing nuclear protein	PCNP	1.74	3	4	
RNA-binding motif protein, X chromosome	RBMX	1.69	14	17	+
Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	1.64	17	20	+
Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3	1.63	10	13	+
Interleukin enhancer-binding factor 3	ILF3	1.62	17	20	
Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	1.58	4	6	+
Splicing factor 45	RBM17	1.44	6	6	+

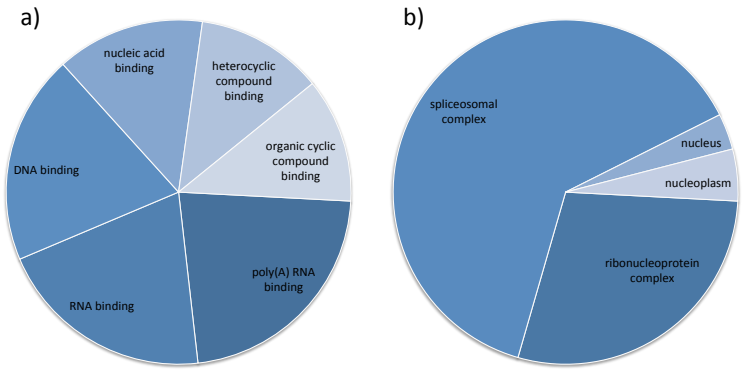


Figure 7. Gene Ontology analysis of FUS interactors with higher affinity upon DNA damage. Gene Ontology Biological process analysis was performed on the FUS interactors with higher affinity for FUS upon Etoposide treatment, according to molecular function (a)

and cellular component (b) terms; the colour gradient indicates the fold change enrichment of the GO term in the dataset with respect to the human reference set. All terms are statistically significantly enriched with a p-value <0.05 (after Bonferroni correction for multiple testing).

3.6 Materials and methods

Cell culture, transfection, and lentiviral infection

Parental HEK293T and lentivirally infected HEK293T cells stably expressing C-terminally flag tagged wild-type FUS (FUS-flag) were cultivated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FCS), penicillin (100 IU/mL), and streptomycin (100 μg/mL) and grown at 37°C and 5% CO2. Plasmid DNA transfections were performed with the calcium phosphate method (Baldi et al., 2005), with a total amount of 10 μg of plasmidic DNA for a 50% confluent 10 cm^2 cell culture dish, and cells were harvested by trypsinization 24 hours post-transfection. To create

the HEK293T cell line stably expressing wild-type FUS-Flag cell line, a lentiviral transduction system was employed. Briefly, to generate the viruses carrying the FUS-Flag vector, HEK293T packaging cells were transfected with plasmids encoding the lentiviral proteins VSV-G, REV, GAG, and POL together with the FUS-Flag gene cloned into a lentiviral vector pCDH-CuO-MCS-EF1-RFP (System Biosciences), by calcium phosphate (day one). On day two, the cell culture media was replaced with a half volume of culture medium to concentrate the viruses. On day three, HEK293T target cells were infected with the cell culture medium from the packaging cells. The efficacy of the infection was evaluated by *in vivo* fluorescent microscopy (EVOS® Cell Imaging Systems, Life-technologies), according to the constitutive expression of the RFP (Red Fluorescent Protein) present in the lentiviral vector encoding FUS-flag gene. HEK293T FUS-flag cells

were cultured in the conditions previously reported and the stable expression of the FUS-flag recombinant protein was evaluated by western blotting.

Plasmids

The DNA plasmids used for transient transfections of HEK293T cells are the following: pCMVmyc-flag-p54 (NONO), pcDNA3.1(+)-HDAC1-flag, pCMVmyc-flag-PSF (SFPQ), pEGFP-C1-flag-Ku70 (XRCC6), pEGFP-C1-flag-Ku80 (XRCC5). All plasmids were purchased from the Addgene plasmid repository.

Western Blot and antibodies

After gel electrophoresis, proteins were electrotransferred on a nitrocellulose membrane (HybondTM-ECLTM Amersham Biosciences) for 1 hour at 390 mA in Tris-Glycine Transfer Buffer (Trizma base 25 mM, Glycine 192 mM, Methanol

20%, SDS 0.1%). Then nitrocellulose membrane was stained for 5 minutes with Red Ponceau Solution in order to control the transfer. The membrane was washed one time with water and then was saturated with 5% non-fat milk (Regilait) in TBS (Tris-HCl 20 mM, NaCl 0.5 M) 1 hour at room temperature or 16 hours at 4°C. After a single washing of 5 minutes with TBS-T (Tris-HCl 20 mM, NaCl 0.5 M, Tween 20 0.1%), the membrane was incubated with the primary antibody dissolved in a 5% milk-TBS solution at different dilution based on the experiment, for 1 hour at room temperature or 16 hours at 4°C. After three washings of 15 minutes with TBS-T, the membrane was incubated with a secondary antibody HRP-conjugated dissolved in a 5% milk-TBS solution. After exhaustive washings with TBS-T, bands were visualized after enhanced chemiluminescence reaction, using ECL-PlusTM Western Blotting Reagents (Amersham

Biosciences), followed by autoradiography or CCD detection (ChemiDoc™ MP System, Biorad). The following primary antibodies were used: custom rabbit polyclonal anti FUS antibody, 1:3000 (as described earlier), rabbit polyclonal anti Flag tag (Antibodies-online) 1:2000, mouse monoclonal anti phospho-Histone H2AX (Millipore) 1:1000, mouse monoclonal anti Vinculin (Sigma) 1:2000. Secondary HRP-conjugated anti mouse and anti rabbit antibodies were purchased from Biorad and used at the dilution of 1:4000.

Etoposide treatment and subcellular fractionation

HEK293T FUS-Flag cells were treated with Etoposide (Enzo lifesciences), dissolved in DMSO, which was added to culture media at final concentration of 10 μM (same volume of DMSO was used for untreated cells). Cells were washed with PBS and harvested by trypsinization after 15, 30 and 60 minutes. Cell

pellets were lysed in RIPA buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.5% Sodium Deoxicholate, 1% NP40, 0.1% SDS, 1x Halt Protease Inhibitor (Pierce), 1x Phosphates Inhibitor (Phostop tablets, Roche)), with the addition of 10 mM $MgCl_2$ together with Benzonase DNase enzyme 0.25 U/μl (Millipore) and lysates were incubated at 4 °C for 30 minutes on rotation head-over-tails. The lysates were centrifuged 15 min at 16'100 x g at 4 °C and the clarified supernatants were quantified by Bradford assay (Biorad). A total protein amount of 40 μg was loaded on a 4-15% pre-cast gel (Biorad) and subjected to SDS-PAGE and western blot analysis. To perform subcellular fractionation of HEK293T FUS-flag cells, treated with Etoposide (10 μM for 1 hour) or untreated (DMSO), cell pellet was resuspended in 5 volumes of Buffer-I (Hepes 10 mM pH 7.5, KCl 10 mM, $MgCl_2$ 1.5 mM, EDTA 1 mM, Triton-X100 0.5%, Protease and Phosphates

inhibitors) and incubated for 20 minutes on ice. Samples were centrifuged at 6000 g for 10 minutes at 4° C, and the supernatants were kept as “cytosolic fraction”. The nuclear pellet was washed twice with Buffer-I (without Triton-X100) and, after centrifugation, was resuspended 2 volumes of Solution-II (Hepes 20 mM pH 7.5, NaCl 100 mM, EDTA 5 mM, Triton-X100 1%, Protease and Phosphates inhibitors) and incubated on ice for 10 minutes. Samples were centrifuged at 16100 g for 10 minutes at 4° C, and the supernatants were kept as “nuclear fraction”. The remaining chromatin pellet was finally lysed in 1 volume of RIPA buffer with the addition of the Benzonase and, after centrifugation at 16100 g for 10 minutes at 4 °C, the supernatants were kept as “chromatin fraction”. The extracted fractions were quantified by Bradford assay and same amount of total proteins (40 μg) for fraction was loaded on a 4-15% pre-cast gel and

subjected to SDS-PAGE and western blot analysis. The optical densities of the western blot bands were obtained using the freely available ImageJ software.

Immunoprecipitations for interactions validation and label-free interactome analysis

HEK293T cells transfected with plasmids encoding the flag-tagged DDR related proteins or HEK293T FUS-flag Etoposide treated (ETOP) or untreated (DMSO) cells were harvested by trypsinization, followed by centrifugation for 5 min at 200 x g at 4°C. The cell pellets were washed once with PBS and cells were suspended with ice-cold hypotonic gentle lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 2 mM EDTA, 0.1% Triton X-100, 1x Halt Protease Inhibitor, 1x Phosphates Inhibitors) in the absence (RNase free) or in the presence of 0.2 mg/ml RNaseA (RNase treated). Cells were lysed for

10 min on ice, followed by supplementation of NaCl to 150 mM final concentration and further incubation on ice for 5 minutes. Then, $MgCl_2$ 10 mM was added together with Benzonase DNase enzyme 0.25 U/μl and lysate was incubated at 4 °C for 30 minutes on rotation head-over-tails. The lysate was then cleared from insoluble particles by centrifugation (15 minutes at 16100 x g at 4 °C). Supernatant was recovered and incubated with anti-flag™ M2 Affinity Gel (20 μl/1x10^7 cells; Sigma) for 1.5 h head over tail at 4 °C. The solution was centrifuged (5 minutes at 1000 x g at 4 °C). The affinity gel was suspended in 1 mL NET-2 (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Triton-X-100) and washed five times by subsequent suspension and centrifugation steps. After the last wash, the buffer was completely removed with a syringe and to elute the precipitated proteins from the affinity gel, the resin was incubated with 1 bed volume of

elution buffer (NET-2, 0.5 x protease inhibitor, 1 mg/ml flag peptide) and incubated 30 min head over tail at 4°C. Eluates were mixed with SDS-Gel loading buffer, boiled for 5 minutes at 95 °C and loaded on an 4-15% pre-casted SDS-polyacrylamide gel (Biorad). After electrophoresis, gels were processed for western blot analysis (as described earlier). For the label-free interactome analysis upon genotoxic stress, the 2% (in volume) of the eluates from the immunoprecipitations was combined with loading buffer and processed for SDS-PAGE and western blotting, while the remaining part was directly processed for mass spectrometric analysis (see below).

Sample preparation and mass spectrometric analysis

Proteins eluted from flag-matrix were digested in-solution with Lys-C and Trypsin using the FASP

protocol (Filter Aided Sample Preparation, Wiśniewski et al., 2009), with spin ultrafiltration units of nominal molecular weight cut off of 10 kDa. Briefly, proteins from each different sample were transferred to YM-10 Microcon filters (Cat No. MRCF0R010, Millipore) and centrifuged at 14000 × g for 20 minutes. 60 μl of UA buffer (8 M urea, 0.1 M Tris-HCl, pH 8.5) were added and samples were reduced by adding 10 µL of 100 mM DTT for 30 min at room temperature. Samples were centrifuged at 14,000 × g for 20 minutes, then 60 µL of UA and 10 µL of 55 mM iodoacetamide were added to the filters and incubated in the dark for 20 minutes. Filters were washed three times with 100 µL of UA, and 1 µg of Lys-C (Wako) was added and incubated for 3 hours at room temperature. Subsequently, 1 µg of trypsin was added, after four-fold dilution of the urea with Tris-HCl 0.1 M pH 8.5, and incubated overnight at room temperature. Then samples were

centrifuged at 14000 × g for 20 minutes and finally 50 μl of 0.5 M NaCl were added to the filter and the released peptides were collected by centrifugation. The resulting peptide mixture was purified from the excess of salt and concentrated using home made C18 StageTips, as previously described (Rappsilber et al., 2007). Mass spectrometry analysis was performed by nano liquid chromatography–tandem MS (nLC–ESI-MS/MS) on a quadrupole Orbitrap Q-Exactive mass spectrometer (Thermo Scientific). Peptides separation was achieved on a linear gradient from 95% solvent A (2 % ACN, 0.1% formic acid) to 40% solvent B (80% acetonitrile, 0.1% formic acid) over 90 minutes and from 40% to 100% solvent B in 2 min at a constant flow rate of 250 nL/min on UHPLC Easy-nLC 1000 (Thermo Scientific) where the LC system was connected to a 25 cm fused-silica emitter of 75 μm inner diameter (New Objective, Inc. Woburn, MA USA), packed in-house

with ReproSil-Pur C18-AQ 1.9 µm beads (Dr. Maisch Gmbh, Ammerbuch, Germany) using a high-pressure bomb loader (Proxeon, Odense, Denmark). MS data were acquired using a data-dependent top10 method for HCD fragmentation. Survey full scan MS spectra (300–1750 Th) were acquired in the Orbitrap with 70000 resolution, AGC target 1e6, IT 120ms. For HCD spectra resolution was set to 35000, AGC target 1e5, IT 120ms; normalized Collision energy 25% and isolation width 3.0 m/z. Three technical replicates of each sample were injected.

Peptides and proteins identification by database searching and statistical analysis

Raw data files were analyzed using the peptide search engine Andromeda, which is integrated into the MaxQuant software environment (version 1.5.2.8) (Cox et al., 2011), with the following

parameters: uniprot_cp_hum_2015_03 as protein database, Oxidation (M), Acetyl (Protein N-term), as variable modifications, Carbamidomethyl (C) as fixed modifications, peptide false discovery rate (FDR) 0.01, maximum peptide posterior error probability (PEP) 1, protein FDR 0.01, minimum peptides 2, at least 1 unique, minimum length peptide 7 amino acids, and trypsin specificity was used with up to two missed cleavages allowed. The lists of identified proteins were filtered to eliminate reverse hits and known contaminants. Label-free analysis was performed including the “match between run” option (time window of 1 min). A minimum LFQ ratio count of 2 was considered and the “LFQ intensities” which are the intensity values normalized across the entire dataset were used. Only proteins identified in both replicates were selected for further analysis. Statistical t-test analysis was performed using Perseus program

(version 1.5.1.6). Missing LFQ-intensities values were replaced by random numbers drawn from a normal distribution by the function “imputation” (width 0.3, Down shift 1.8, separately for each column). For the statistical significance a Benjamini-Hochberg FDR of 0.05 was applied with a permutation test (500 randomizations).

Gene Ontology enrichment analysis

Gene Ontology (GO) terms enrichment analysis was performed with the PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system (Mi et al., 2013), using the gene names of the identified proteins as queries for the statistical overrepresentation test and the most updated (at the time of analysis) Homo sapiens genes annotations as reference set. The electronically inferred annotations were excluded and only the over-represented categories with a p-value <0.05

(after Bonferroni correction for multiple testing) have been chosen to be reported in the graphs.

3.5 References

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., and Elledge, S.J. (2012). Response. *14*, 318–328.

Baldi, L., Muller, N., Picasso, S., Jacquet, R., Girard, P., Thanh, H.P., Derow, E., and Wurm, F.M. (2005). Transient gene expression in suspension HEK-293 cells: application to large-scale protein production. Biotechnol. Prog. *21*, 148–153.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. a, Olsen, J. V, and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction.

Dutertre, M., Lambert, S., Carreira, A., Amor-Guéret, M., and Vagner, S. (2014). DNA damage: RNA-binding proteins protect from near and far. Trends Biochem. Sci. *39*, 141–149.

Gardiner, M., Toth, R., Vandermoere, F., Morrice, N. a, and Rouse, J. (2008). Identification and characterization of FUS/TLS as a new target of ATM. Biochem. J. *415*, 297–307.

Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., et al. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell *149*, 768–779.

Heinrich, B., Zhang, Z., Raitskin, O., Hiller, M., Benderska, N., Hartmann, A.M., Bracco, L., Elliott, D., Ben-Ari, S., Soreq, H., et al. (2009). Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. J. Biol. Chem. *284*, 14303–14315.

Hoell, J.I., Larsson, E., Runge, S., Nusbaum, J.D., Duggimpudi, S., Farazi, T. a, Hafner, M., Borkhardt, A., Sander, C., and Tuschl, T. (2011). RNA targets of wild-type and mutant FET family proteins. Nat. Struct. Mol. Biol. *18*, 1428–1431.

Kuroda, M., Sok, J., Webb, L., Baechtold, H., Urano, F., Yin, Y., Chung, P., de Rooij, D.G., Akhmedov, a, Ashley, T., et al. (2000). Male sterility and enhanced radiation sensitivity in TLS(-/-) mice. EMBO J. *19*, 453–462.

Lenzken, S.C., Loffreda, A., and Barabino, S.M.L. (2013). RNA splicing: a new player in the DNA damage response. Int. J. Cell Biol. *2013*, 153634.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. *8*, 1551–1566.

Montecucco, A., and Biamonti, G. (2007). Cellular response to etoposide treatment. Cancer Lett. *252*, 9–18.

Naganuma, T., and Hirose, T. (2013). Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol. *10*, 456–461.

Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N., and Hirose, T. (2012). Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. EMBO J. *31*, 4020–4034.

Oh, S.-M., Liu, Z., Okada, M., Jang, S.-W., Liu, X., Chan, C.-B., Luo, H., and Ye, K. (2010). Ebp1 sumoylation, regulated by TLS/FUS E3 ligase, is required for its anti-proliferative activity. Oncogene *29*, 1017–1030.

Pommier, Y., Leo, E., Zhang, H., and Marchand, C. (2010). DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem. Biol. *17*, 421–433.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Sama, R.R.K., Ward, C.L., and Bosco, D. a (2014). Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. ASN Neuro *6*.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2014). Biochemical Properties and Biological Functions of FET Proteins. Annu. Rev. Biochem. 1–25.

Shamanna, R. a, Hoque, M., Lewis-Antes, A., Azzam, E.I., Lagunoff, D., Pe'ery, T., and Mathews, M.B. (2011). The NF90/NF45 complex participates in DNA break repair via nonhomologous end joining. Mol. Cell. Biol. *31*, 4832–4843.

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., et al. (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science *323*, 1208–1211.

Wang, W.-Y., Pan, L., Su, S.C., Quinn, E.J., Sasaki, M., Jimenez, J.C., Mackenzie, I.R. a, Huang, E.J., and Tsai, L.-H. (2013). Interaction of

FUS and HDAC1 regulates DNA damage response and repair in neurons. Nat. Neurosci. *16*, 1383–1391.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat. Methods *6*, 359–362.

Yang, L., Gal, J., Chen, J., and Zhu, H. (2014). Self-assembled FUS binds active chromatin and regulates gene transcription. Proc. Natl. Acad. Sci. *111*, 17809–17814.

Chapter 4

Conclusions

4.1 Summary

In the present thesis, I report the results obtained from my research activity carried out during the Ph.D program. The principal aim of the project was to obtain new insights into the diverse molecular pathways in which FUS/TLS protein is involved in human cells. FUS is a member of the TET (TLS/EWS/TAF15) protein family of DNA/RNA-binding proteins (Schwartz et al., 2014). It is characterized by a N-terminal domain enriched in serine, tyrosine, glutamine and glycine residues (SYQG region), a glycine-rich region, an RNA recognition motif (RRM), multiple RGG repeats implicated in RNA binding, a C-terminal zinc finger

motif and a highly conserved extreme C-terminal region. FUS was described to take part in several different processes in the cell, such as cell cycle regulation, gene expression, DNA damage response, pre-mRNA splicing, and others (Sama et al., 2014). Since the exact role of FUS in these pathways is still not well elucidated, we employed a high-throughput mass spectrometry-based interactome analysis to identify the protein to which FUS is interacting in order to exert its functions. The interactome analysis resulted in the identification of more than 500 interacting proteins, the majority involved in the biological process in which FUS was involved, confirming the experimental evidences reported in literature. In particular, there is enrichment for the biological processes related to the RNA metabolism, especially within the set of proteins that resulted from our experiments to be the most conserved FUS interactors. Among the

processes, we noticed a statistically significant enrichment for the U12-type dependent splicing, which is a less frequent splicing event that regulates the splicing of a sub-class of introns (the so called “minor introns”). We therefore decided to investigate about the role of FUS in this particular splicing event and we found that it binds to the snRNPs of the U12-type spliceosome and directly to the minor introns enhancing their splicing. Moreover, we demonstrated that an ALS causative mutation of FUS (P525L), that abrogates the nuclear translocation of FUS, leading to a cytosolic accumulation of the protein, causes a reduced activity of FUS on the minor introns.

Another important aspect of the FUS biology is its involvement in the DNA damage response and repair (DDR) pathways (as earlier described in this thesis), therefore we investigated about this role from a proteomic point-of-view, since the

interactome analysis led to identification of a list of interactors that have been related to DDR, and this pathway resulted significantly enriched in our dataset. We first validated the interaction with selected proteins of the DDR by immunoprecipitation and western blotting analysis, and then we applied a high-throughput label-free quantitative mass spectrometry approach to characterize changes in the interactome of FUS upon genotoxic stress. The analysis resulted in the identification of most of the interactors of FUS, which have been characterized by our previous interactome analysis, and although the vast majority didn't show changes in the interaction affinity for FUS, we identified a set of RNA-binding proteins whose interaction with FUS is augmented upon the DNA damage. This finding is coherent with novel evidences that describe the RNA-binding proteins as key regulator of the DDR, primarily due

to their ability to modulate the gene expression of the DDR-related factors (Dutertre et al., 2014).

4.2 Future perspectives

From the present doctoral thesis, two principal evidences arose: the first is that FUS interacts with numerous proteins mostly involved in RNA metabolism, particularly in the splicing pathway, and can enhance the splicing of selected U12-type introns; secondly the interaction of FUS with a set of RNA-binding proteins is augmented upon genotoxic stress. However, still the exact molecular mechanisms through which FUS exerts its functions in the splicing machinery and in the DDR are not clear and further experiments are needed in order to better elucidate these aspects. One opening question is whether FUS can regulate the splicing of a numerous set of “minor” introns, since we only characterized this function with selected minigenes.

The answer to this question would be particularly important for the translational aspect of this study, since the U12-dependent splicing was recently implicated in the pathogenesis of motor neuron disease, such as the ALS (Onodera et al., 2014). FUS has been extensively related to this neurodegenerative disease, therefore a high-throughput approach, aimed to identify the “minor” splicing events that FUS regulate, might help to elucidate novel mechanisms through which FUS is involved in the pathogenesis of the ALS.

Furthermore, the evidence that FUS show a higher affinity for a set of RNA-binding proteins, upon DNA damage, does still not explain how FUS is involved in the DDR, whether it has a regulating role by recruiting other factors to the site of the damaged DNA, or being directly involved in the repair of the lesions through its ability to bind to DNA molecules. One hypothesis, that might be worth of further

investigations, is the possibility that FUS is recruited to the site of DNA damage in order to promote the SUMOylation proteins that are involved in the DDR, as for example the RNA-binding proteins that interact with FUS. The ability of FUS to act as a SUMO E3 ligase has been described already by Oh and colleagues, in the context of the DDR, but directed to a specific onco-suppressor protein (Ebp1) (Oh et al., 2010), while a global screen of eventual proteins whose SUMOylation might be regulated by FUS during the DDR is still missing. However, from our proteomics results, we noticed the interaction of FUS with several proteins that have been described as SUMOylated and involved in DDR, such as SFPQ, NONO, PARP1, XRCC5, and others; in addition, SUMOylation itself have been extensively involved in the DDR (as reviewed in Dou et al., 2011). The elucidation of this aspect might be relevant for the translational value of this thesis,

since defects in the DDR in neuronal cells have been reported to be involved in the neurodegeneration and the SUMOylation of the RNA-binding proteins was described to be involved in neurodegenerative disease (as reviewed in Filosa et al., 2013)

4.3 References

Dou, H., Huang, C., Van Nguyen, T., Lu, L.-S., and Yeh, E.T.H. (2011). SUMOylation and de-SUMOylation in response to DNA damage. FEBS Lett. *585*, 2891–2896.

Dutertre, M., Lambert, S., Carreira, A., Amor-Guéret, M., and Vagner, S. (2014). DNA damage: RNA-binding proteins protect from near and far. Trends Biochem. Sci. *39*, 141–149.

Filosa, G., Barabino, S.M.L., and Bachi, A. (2013). Proteomics strategies to identify sumo targets and acceptor sites: A survey of RNA-Binding proteins SUMO ylation. NeuroMolecular Med. *15*, 661–676.

Oh, S.-M., Liu, Z., Okada, M., Jang, S.-W., Liu, X., Chan, C.-B., Luo, H., and Ye, K. (2010). Ebp1 sumoylation, regulated by TLS/FUS E3 ligase, is required for its anti-proliferative activity. Oncogene *29*, 1017–1030.

Onodera, O., Ishihara, T., Shiga, A., Ariizumi, Y., Yokoseki, A., and Nishizawa, M. (2014). Minor splicing pathway is not minor any more: Implications for the pathogenesis of motor neuron diseases. Neuropathology *34*, 99–107.

Sama, R.R.K., Ward, C.L., and Bosco, D. a (2014). Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. ASN Neuro *6*.

Schwartz, J.C., Cech, T.R., and Parker, R.R. (2014). Biochemical Properties and Biological Functions of FET Proteins. Annu. Rev. Biochem. 1–25.