UNIVERSITÀ
DI PAVIA

PhD IN BIOMEDICAL SCIENCES

DEPARTMENT OF BRAIN AND BEHAVIORAL SCIENCES

UNIT OF NEUROPHYSIOLOGY

Regulatory role of TDP-43 on R-loop formation in Amyotrophic Lateral Sclerosis (ALS)

PhD Tutor: Prof.ssa Cristina Cereda

PhD Tutor: Dott.ssa Daisy Sproviero

PhD Supervisor: Dr. Emanuele Buratti

PhD Supervisor: Prof.ssa Serena Carra

PhD dissertation of

Marta Giannini

a.a. 2019/2020

Abstract

R-loops are three-strand nucleic acid structures constituted by an RNA-DNA hybrid plus a displaced DNA strand (ssDNA) that are naturally involved in cellular processes such as DNA transcription, mitochondrial and nuclear DNA replication or immunoglobulin (Ig) class switching. Their development is favored by specific areas of genome enriched in GC clusters, called GC skew, such as promoter regions, thanks to the thermodynamic stability of RNA-DNA hybrid structure. A certain number of reports suggest that R-loops are formed at a higher frequency in transcribed sites of genomic loci and can act as cotranscriptional intermediates. In some pathological conditions, physiological factors that prevent formation of R-loops are impaired and hence R-loops' persistent formation can become a risky outcome with deleterious effects on genome integrity. The consequence of their accumulation is strongly related to genomic instability enhancing the development of ectopic homologous DNA sequences rearrangements, DNA hyper recombination, chromosomal rearrangements or, when abortive, to chromosome loss. It was shown that R-loops are associated with neuroinflammatory autoimmune disease such as Aicardi-Goutière Syndrome (AGS) or with neurodegenerative disorders such as fragile X-associated tremor/ataxia syndrome, while their relation with Amyotrophic Lateral Sclerosis (ALS) is still under study. ALS is the most common adult-onset motor neuron disease, and is characterized by the progressive loss of upper and lower motor neurons from the spinal cord, brain stem, and motor cortex, leading to muscle weakness and eventual respiratory failure. Despite the multifactorial nature of this pathology, TDP-43 was recently identified as the major pathological protein in sporadic ALS and in the most common pathological subtype of FTD. In TDP-43-positive cytoplasmic inclusions, the protein is abnormally phosphorylated, ubiquitinated and truncated resulting in the accumulation of 25-kDa and 35-kDa C-terminal fragment (CTF25 and CTF35), that avoid full length TDP-43 from exerting its nuclear physiological function related to RNA metabolism, including splicing, translation, and transport. Alternatively, TDP-43 mutations (1-2% cases), which mainly cluster in the region encoding the C-terminus glycine-rich low-complexity region (LC), may be neurotoxic through a novel gain of-function mechanism, developing features that can be directly connected with the pathology, such as increased aggregation propensity and half-life, altered sub-cellular localization and protein–protein interactions. Recently, emerging evidence in literature suggests the involvement of TDP-43 even in DNA damage response (DDR) as a critical component of the nonhomologous end joining (NHEJ)-mediated DNA double-strand break (DSB) repair pathway. Colocalization of TDP-43 with active RNA polymerase II at sites of DNA damage along with the DNA damage repair protein, BRCA1, demonstrates its participation in the prevention or repair of R loop associated DNA damage. As a consequence, TDP-43 depletion leads to increased sensitivity to various forms of transcription-associated DNA damage such as R-loops-

associated damage and the mutation in C-terminus glycine-rich low-complexity region, LC domains associated with the loss of its important nuclear function, is known to be associated with excessively stable R-loops, stalled transcription machinery and stable open DNA. So due to impairments or mutations of this ribonucleoprotein, cellular processes involved in RNA modulation and regulation cooperating with resolving or suppressing RNA-DNA hybrid factors are affected, leading to accumulation of R-loops. In turn, this mechanism increases the amount of aberrant transcripts, which further promotes cellular neurodegeneration. The aim of this project was based on the investigation of the regulatory role of TDP-43 in R-loops' formation using different "in vitro" cellular model such as neuroblastoma cell line SH-SY5Y, HeLa cervical cancer cells and lymphoblastoid cell line LCLs derived from ALS patients. We decided to study specifically p.A382T TDP-43 as this is one of the most commonly occurring and very well studied missense mutation, associated with an increased nuclear TDP-43 translocation into cellular cytoplasm according with the literature and with our data. We observed R-loops accumulation in SH-SY5Y expressing p.A382T missense mutation of TDP-43 and in silenced TDP-43 HeLa cells by immunofluorescence and DRIP-qPCR, with a significative increase in RPL13 gene, encoding for large ribosomal subunit EL13 protein as member of cytoplasmic/translation TDP-43 interactors cluster. We were able to confirm the obtained results in LCLs derived from ALS patients carrying p.A382T TDP-43 mutation by immunofluorescence and flow cytometry analysis. In LCLs we proved the interaction between TDP-43 and R-loops at chromatin level in control LCLs and in sALS derived LCLs, but less signal was detected in p.A382T mutated TDP-43 LCLs, indicating that TDP-43 cytoplasmic myslocalization occurred also in these cells avoiding the execution of its nuclear physiological function linked to R-loops' regulation. In p.A382T SH-SY5Y and siTDP-43 HeLa we also detected an increase in one of the members of the Fanconi anemia complementation group FANCD2, which is involved in homology-directed DNA repair controlled by R-loops' resolution and in DNA damage response colocalizing in nuclear foci together with BRCA1 and BRCA2. In the same cells, the activation of DDR was confirmed by the accumulation of γH2AX foci that was proved as dependent from R-loops' accumulation through the use of RNaseH enzyme, which treatment suppressed DNA DSBs signal thanks to the specific degradation of these hybrids. In siTDP-43 HeLa we also detected a significative increase in foci number of another R-loops related factor called H3S10P, which accumulation is associated with replication fork stalling, transcription– replication collisions. Finally, the presence of stress granules in the cytoplasm of p.A382T TDP-43 LCLs colocalizing with RNA-DNA hybrids confirmed the previously mentioned mechanism of relation between aberrant transcripts accumulation due to R-loops' increase, in turn connected with impairment of factors involved in hybrids' suppression or resolution pathway. The results obtained from these experiments confirmed from one side that in

particular p.A382T mutation is responsible for the loss of TDP-43 physiological nuclear function involved in the resolution of genomic R-loops due to its cytoplasmic translocation. At the same time p.A382T mutation affecting the aggregation capacity of this protein caused the sequestration of wild type TDP-43 in cytoplasmic aggregates such as stress granules gaining a new neurotoxical function related to the persistence of prolonged or repeated neuronal damage conditions. Anyway, a lot of questions remained open for example about the relation of distinct mutations in disease-associated genes with aberrant R loops' formation or about the consequences in the arising of different pathologies or about the particular sensitivity of the nervous system and motor neurons to defects in R-loops metabolism and genome stability. The knowledge of this molecular mechanism will shed some light on R-loops pathways and could represent a step forward in the understanding of human diseases and our ability to treat them. This could provide a possible therapeutic strategy not only for rare diseases but also other neurodegenerative disorders that present deficit in RNA metabolism, developing the possibility to treat R-loops as therapeutic targets, providing additional benefits to ALS patients by individualization of a new biomarker.

Table of content

1 Introduction

1.1 ALS

Amyotrophic lateral sclerosis (ALS) is an adult onset, rapidly progressive neurodegenerative disorder, caused by the selective loss of upper and lower motor neurons in the cerebral cortex, brainstem and spinal cord. Neuronal degeneration leads to weakness, muscular atrophy, and spasticity that evolve to paralysis; the disease is fatal within 3-5 years of onset, generally due to respiratory failure[1]. The features of ALS were first described as a clinic-pathological entity by the French neurologist Jean Martin Charcot in 1869[2]. "Amyotrophic" refers to the atrophy of muscle fibers, which are denervated as their corresponding anterior horn cells degenerate, leading to weakness of affected muscles and visible fasciculation. "Lateral sclerosis" refers to hardening of the anterior and lateral cortico-spinal tracts as motor neurons degenerate in these areas and are replaced by gliosis. ALS has recently been recognized as a multi-system disorder rather than a disease limited to motor neurons. Some ALS patients may show extrapyramidal features such as tremor, rigidity, propulsion, and impaired postural reflexes. In about one quarter of ALS patients, the disease is associated with subtle cognitive deficits. In addition, 3–5% of ALS patients are diagnosed with frontotemporal dementia (FTD), a dementia of non-Alzheimer's type with symptoms of behavioral changes, frontal executive deficit, and impaired handling of language[3]. The pathological hallmark of ALS is denervation and atrophy of muscle due to loss of spinal motor neurons. Swelling of the perikaryal and proximal axons is also observed, as the accumulation of phosphorylated neurofilaments, Bunina bodies and Lewy body-like inclusions, and the deposition of inclusions of ubiquitinated material in these axons[4] (Figure 1).

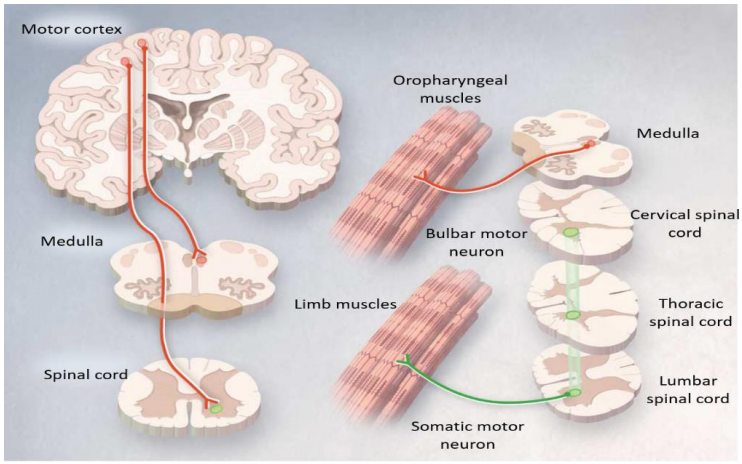

Figure 1. Degeneration of motor neurons in the motor cortex and degeneration of motor neurons in the brain stem and spinal cord selectively affected in ALS patients. (Modified by Rowlad and Shneider, 2001[25])

Unfortunately, there is no primary therapy for this disorder and the Riluzole, a presumed glutamate antagonist, is the only drug approved by the US Food and Drug Administration for the treatment of ALS, but the exact mechanism of action of Riluzole is still unclear. It appears to prolong ALS survival by a few months on average, although when given at an early stage or to younger patients, it might prove more effective[3]. Symptomatic measures (for example, feeding tube and respiratory support) are the mainstay of management of this disorder[1]. In the last years, over 60 molecules have been investigated as a possible treatment for ALS[4] and clinical trials (CTs) are ongoing, using compounds divided according to their principal mechanism of action in anti-glutamatergic, anti-inflammatory, anti-oxidative, neuroprotective, neurotrophic factors and others. Among all the tested substances two potentially efficacious treatments are claiming to demonstrate clinical benefit enriching statistical significance on secondary endpoints of human CTs: masitinib, a highly selective tyrosine kinase inhibitor which was shown to prevent CNS neuroinflammation, and edaravone, an intravenous free radical scavenger which eliminates lipid peroxide and hydroxyl radicals[5]. Otherwise on the EU Clinical Trials Register are reported 23 ongoing clinical trial, that followed EudraCT (European Union Drug Regulating Authorities Clinical Trials) protocol, such as a phase II trial on colchicine, which as anti-mitotic drug is able to block mitotic cells in metaphase by formation of tubulin-colchicine complexes and to induce the expression of the HSPB8 molecular chaperone, showing protective effects in cellular and animal models of ALS[6]. Another drug under investigation is guanabenz, which is an antihypertensive agent acting as central alpha-2 adrenergic agonist able to protect cells by prolonging translation attenuation, thus enhancing the capacity of the protein quality control (PQC) system to restore protein homeostasis upon conditions that promote the accumulation of unfolded and misfolded proteins[7]. The primary objective of Mirocals phase II study is based on the evaluation of the clinical efficacy and safety of the experimental drug (ld IL-2) over an 18 months period in order to establish the proof of concept (PoC) that immune responses modification through the enhancement of regulatory T cells modifies the rate of ALS disease progression. The validation of this new phase-II study could improve the efficiency of drug development in ALS with early determination of drug response using established biomarkers (BMs).

1.1.1 Epidemiology

The incidence of sporadic ALS shows little variation in the Western countries, ranging 2.2 new ALS cases per 100,000 person-years, with an estimated lifetime risk of 1 in 400 (considering 15,927 identified cases)[8]. ALS is rare before the age of 40 years and increases exponentially with age

thereafter, even if juvenile ALS forms are increasing. Mean age at onset is 55–65.1 years for sporadic ALS and 40–60 years for familial ALS, with a peak incidence in those aged 70–79 years. Men have a higher risk of ALS than women, leading to a male-to-female ratio of 3–2.4[9,10]. All the evidences converge to show that ALS risk is different across continents and ethnicities. During recent decades, population-based studies have measured a lowest ALS incidence in East Asia (0,89 per 100,000 person-year) and in South Asia (0,79 per 100,000 person-year), while a large part of Africa, Latin America and Asia does not have any population-based studies. Geographic foci of the Western Pacific form of ALS, mainly in Guam and the Kii Peninsula of Honshu Island, Japan, have been reported, with prevalence 50–100 times higher than in other parts of the world, but they rapidly disappeared in the last years underlining the importance of changes in lifestyle and environmental factors in disease development[8]. This kind of ALS was presented in three clinical forms, i.e., ALS, atypical Parkinsonism with dementia, and dementia alone, known collectively as the ALS-Parkinson's dementia complex (ALS-PDC). The cause of these aggregations remained elusive, and a decreasing prevalence of ALS-PDC has been described recently[3].

1.1.2 fALS and sALS

ALS can occur sporadically, without any family history (sALS; 90-95% of patients), while a small percentage of ALS cases are considered familial (fALS; 5-10%). It can be inherited either as an autosomal dominant or recessive trait. Adult onset autosomal dominant inheritance is more common than juvenile onset caused by recessive transmission. X-linked dominant inherited ALS has been reported in one family[11]. Thanks to increased massive parallel sequencing approaches such as whole-genome sequencing (WGS) and whole-exome sequencing (WES), more mutations have been discovered, adding more heterogeneity to the disease mechanisms. Among the genes reported in ALS pedigrees, Mutations in the major established causal ALS genes (SOD1, TARDBP, FUS, VCP, C9orf72, and PFN1) account for approximately 80% of fALS ALS and about 10% of sALS disease cases, with the GGGGCC hexanucleotide expansion mutation in the 50 noncoding region of C9orf72 being by far the biggest contributor (10-15% of all ALS patients)[12]. Gene mutations cause motor neuron death through different pathways: SOD1 mutations lead to oxidative stress; TARDBP, FUS and c9ORF72 induce disturbances in RNA machinery because their principal physiological functions include RNA processing, such as splicing, transport, and translation[13,14]. A single mutation can lead to different clinical presentations, suggesting that diverse mechanisms influence the outcome and that similar ALS phenotypes result from different mutations, implying that ALS is a syndrome of different causes that share similar pathophysiological pathways[15]. A genetic component is also thought to contribute to the pathogenesis of sporadic ALS, which accounts for the majority of ALS cases. A

number of observations suggest a role for genetic factors in sALS. A meta-analysis of three twin studies gives an estimate of sALS heritability of 0.61 (95% CI 0.38–0.78)[16]. Several groups have reported gene variants and association studies found in individuals with sporadic ALS. These studies linked to genetic variants sALS in particular account for a small number of the total cases reflecting a complex pattern of inheritance with very low penetrance, a high degree of heterogeneity and/or the existence of environmental factors predisposing to ALS[17]. ALS is relentlessly progressive, as motor neuron injury gradually spreads. Most patients with ALS die within 3 to 5 years after symptom onset, but the variability in clinical disease duration is large, with some patients dying within months after onset and others surviving for more than two decades. Large differences in survival and age at disease onset exist even between individuals from one family, in whom ALS is caused by exactly the same mutation, suggesting the existence of other factors that modify the phenotype[1]. Although ALS patients show some degree of heterogeneity as far as symptoms, age of onset, and disease duration are concerned, fALS cases are indistinguishable from sALS based on clinical and pathologic criteria. Thus, family history and genetics are the primary factors that discriminate between sALS and fALS. In the absence of family history, an early age of onset, atypical rapid or slow disease progression, pure lower motor neuron presentation or the presence of dementia may alert to a familial etiology[18]. All genes found mutated in fALS cases may also be mutated in sALS and first-degree relatives of patients with sALS have an increased risk of ALS and other neurodegenerative diseases[19]. Similar superoxide dismutase 1 (SOD1)-positive and TAR DNA-binding protein 43 (TDP43)-positive inclusions have been found both in sALS and both in fALS cases[20,21].

1.2 Pathogenesis of motor neurons degeneration in ALS

The exact molecular pathway causing motor neuron degeneration in ALS is still unknown, but as with other neurodegenerative diseases, is likely to be a complex interplay between multiple pathogenic cellular mechanisms which may not be mutually exclusive[22,23] including indeed different factors. Despite this, sporadic and familial ALS are clinically and pathologically similar, suggesting a common pathogenesis.

ALS, appearing as a multisystemic disease, is based on the involvement of different tissues and cell types, such as astrocytes and microglia, oligodendrocytes and muscles cells, among others, that are both targeted by disease processes and actively participate to the final illness outcome[1]. Motor neurons death stimulates a neuroinflammatory reaction by activating astrocytes and microglia that produce inflammatory cytokines. It has been reported that the expression of proinflammatory factors increases before the onset of the disease and during the active phase of the disease progression there is a considerable microglial activation[24]. A number of genes have been identified that are definitely

associated to ALS pathogenesis and that now cover a large fraction of the familial cases of the disease[25]. According to their known biological role, or to new acquired functions because of gene mutations, they are apparently connected to different physiological processes, implying that diverse pathogenic mechanisms, not necessarily related, might be involved in ALS onset and progression. One of these processes is the impairment of protein homeostasis which is characterized by the translation of misfolded proteins due to mutations in some genes and also by the change of native proteins into unfolding and misfolding molecules due to either defects in the PQC or to external/environmental stress, contributing to challenge proteostasis and playing an important role in disease pathogenesis. These abnormal or not normally localized proteins can directly or indirectly damage the ubiquitin–proteasome system or autophagic machinery of the cell, causing impaired protein turnover. Another dysfunction is the alteration of RNA metabolism that can be associated to neuronal injury. Mutations in TARDBP and FUS, as RNA binding proteins, can be implicated in this mechanism[26]. Other processes involved are the impaired DNA repair caused by mutations in FUS[27] or NEK1[28] and the neuroinflammation that can be observed post-mortem in patients' brains following the microglia activation[29]. Another potential factor that is associated to ALS is the oxidative stress, created by the failed balance between the generation and removal of ROS. ROS are generated by aerobic metabolism and they can damage cells by oxidizing many biomolecules, such as lipids, proteins and DNA. Several oxidative stress markers are elevated in ALS patients[30]. The presence of ROS may also alter protein conformation and structure, causing the aggregation of the abnormal protein inclusions. Endoplasmic Reticulum (ER) stress arises when proteins with an abnormal conformation accumulate in the ER lumen due to a dysregulation in protein quality control or an overload of newly synthesized polypeptides in the ER. The UPR endures ER stress by inducing adaptive responses in order to reduce the misfolding of proteins. However, failure in UPR or prolonged ER stress result in apoptotic cell death. The involvement of ER stress in the pathogenesis of ALS has been suggested in both SOD1-related ALS and sALS cases[31]. In addition, an analysis carried out on mutSOD1 transgenic mice reveals that the vulnerability of motor neurons in ALS is strictly related to chronic ER stress and these vulnerable motor neurons express ER stress markers[32]. A mitochondrial dysfunction and changes in mitochondrial morphology In ALS patients who have mutations in SOD1 can be observed[33]. Excessive intracellular calcium levels induce motor neuron death through several mechanisms including generation of reactive oxygen species (ROS), release of cytochrome c from the mitochondria and glutamate excitotoxicity. Motor neurons contain less mitochondrial density per volume compared to non-neuronal cells, thus these mechanisms make neurons more deficient in mitochondrial calcium buffering properties[34]. Mitochondrial impairment and calcium dysregulation handling together with protein aggregates formation coupled to

proteasome impairment and ER stress are major components of motor neuron injury. Impaired axonal transport may contribute to energy deficit in the distal axon and the dying-back axonopathy that is observed in ALS. Disrupted axonal transport and loss of metabolic support from myelin sheaths that are formed by oligodendrocytes seem to contribute to motor axon vulnerability. Non-cell autonomous processes, such as neuroinflammation and prion-like propagation of MutSOD1, might also damage motor neurons. Indeed, growing evidence suggests that a noxious property of SOD1 includes its prion-like propagation. This concept is supported by two notable characteristics of SOD1, namely cell-to-cell transmission and seeded aggregation. As an initiation event of intercellular transmission, both mutant and wild-type SOD1 aggregates can be released from injured cells and/or as part of exosomes, small vesicles involved in cell-cell trafficking[35]. MutSOD1 can be otherwise released via a main secretory pathway that is mediated by components of neurosecretory vesicles[34]. (Figure 2)

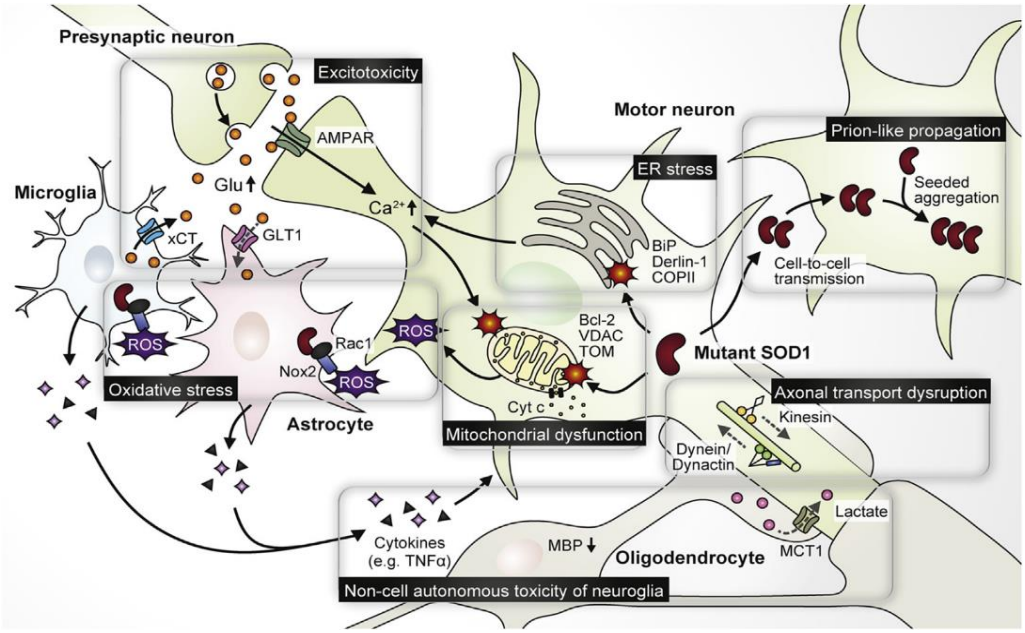

Figure 2. Overview of proposed neurotoxic mechanisms in SOD1-mediated ALS. Signalling and processes that are depicted by solid arrows are enhanced, while those that are depicted by dotted arrows are diminished in SOD1-mediated ALS. (AMPAR, AMPA receptor; Glu, glutamate; xCT, cystine/glutamate transporter. Modified by Hayashi et al. 2016[34])

1.2.1 Oxidative stress

Reactive oxygen species (ROS) are generated by aerobic metabolism and have the potential to damage cells by oxidizing various biomolecules, such as proteins, lipids, and DNA. To prevent this noxious signalling, antioxidants engage in the removal of ROS. The collapsed balance between the generation and removal of ROS is defined as oxidative stress and is implicated in various diseases,

including ALS. In fact, several oxidative stress markers are elevated in ALS patients[30]. Although SOD1 normally functions as an antioxidant enzyme by detoxifying superoxide ($O2^-$), oxidative damage by MutSOD1 is predominantly associated with a gain of toxicity via ROS upregulation rather than to a loss of enzymatic function. ROS can be generated through several mechanisms, including NADPH oxidase activity, mitochondrial respiratory chain impairment, MutSOD1 induced dysfunction of mitochondria[36] and glutathione (GSH) lower level in ALS patients motor cortex[37]. The relevance of ROS can also be related to the fact that it may also affect protein conformation and structure, leading to the accumulation of the abnormal protein inclusions that are extensively described in ALS mouse models and patient-derived tissue[38]. Interestingly, evidence suggests that also wild-type and mutant TDP-43 aggregation is caused by incorrect disulphide bonds involving Cys residues in one of its RNA recognition motifs, and that aggregation is promoted by ROS[39]. The formation of misfolded protein aggregates can interfere and exacerbate other molecular mechanisms involved in motor neuron degeneration as mitochondrial dysfunction, axonal transport, ER stress and RNA metabolism. Mitochondria, not only represent the main site of ROS production, but are a known target of ROS because of their high dependence on membrane integrity and because they possess their own DNA and RNA that may be damaged by oxidation as well[36]. Aggregation of TDP-43[40] and FUS[41] proceeds through the stress granules (SGs) pathway, as recently demonstrated using optogenetic[42]. However, recently, SG-independent mechanisms that lead to TDP-43 phase separation and aggregation were also described and likely contribute to disease[42]. SGs are highly dynamic structures that are formed upon OS and contain RNA-binding proteins, transcription factors, RNA helicases and nucleases that work as sorting granules for mRNAs undergoing degradation, storage or translation[43]. SGs are found as a consequence of FUS mutation and mutated FUS is more rapidly directed to SGs after oxidative stress than wild type FUS[43]. Furthermore, TDP-43 is recruited to SGs in conditions of oxidative stress[44]. These SGs may simply sequester a subset of mRNAs thus inducing cell dysfunction or serve as nucleation site for larger protein aggregates as the ones found in ALS patients[45].

1.2.2 Protein toxicity: protein aggregation, degradation and prion-like domains

Protein toxicity covers a primary role in ALS pathogenesis; indeed, several proteins (both wild-type and mutant) are dysfunctional in both fALS and sALS, as evidenced by the formation of aggregates, abnormal cleavage events, or distinctive post-translational modifications (e.g., ubiquitination or hyper-phosphorylation). These changes occur both as primary consequences of mutations in the affected proteins and as secondary phenomena induced by the underlying disease process[46].

Protein aggregation and inclusion bodies. Accumulation of dense aggregates of ubiquitinated proteins, in association with eosinophilic aggregates described as "Bunina bodies" are characteristic of later stages of motor neuron disease and ALS pathology. Whether these deposits are toxic or reflect a cellular response to a more primary pathology remains unclear. Indeed, the possibility that some aggregates may reflect beneficial, compensatory events has also been considered[46]. In most cases of sALS and fALS hyperphosphorylated, cleaved TDP-43 accumulates diffusely in the cytoplasm of neurons and glia, where it assembles into round and thread-like inclusions. The traslocation of TDP-43 from the nucleus to the cytoplasm in these disorders has suggested that TDP-43–mediated toxicity may reflect either loss of its function in the nucleus, an acquired adverse effect of its pathological presence in the cytoplasm (gain-of-function), or both[46]. Mutant FUS, like SOD1 and TDP-43, is detected in diverse types of intracellular inclusions. Notably, FUS is a major component of stress granules (SGs); mutations in FUS that lead to intracellular retention of this protein increase its propensity to form these cytoplasmic structures[46]. Mutant FUS is also a major component of ubiquitin and p62-positive cytoplasmic inclusions detected in neurons and glia of the brain and spinal cord of both fALS and FTLD patients. The C9orf72 expansion produces both nuclear RNA foci and cytoplasmic protein inclusions. At autopsy, C9orf72 mutant brains show widespread intra-nuclear RNA foci[47] generated from both sense and antisense transcripts across the expanded GGGCC segment. RNA Foci have been described in fibroblasts[48] and motor neurons derived from induced pluripotent cells generated from fibroblasts of C9orf72 ALS cases[49].

Proteins degradation. Misfolded proteins are generated in various cellular compartments, including the cytoplasm, nucleus and ER, are efficiently removed by quality control systems composed of the ubiquitin (Ub)-proteasome system (UPS), chaperone mediated autophagy (CMA) and macroautophagy [50]. Compared with proliferating cells, post-mitotic neurons are more sensitive to the accumulation of cytotoxic proteins because they cannot dilute toxic substances by cell division. Moreover, protein quality control is intrinsically challenging in neurons because of their cellular structure with dendrites and axons in which protein aggregates need to be packaged into autophagic vacuoles and send to the cell body, rich in lysosomes, for degradation[50]. Although young neurons can manage to clear cytotoxic proteins, this task becomes increasingly more difficult throughout the course of aging during which the components of the UPS, CMA and macroautophagy are downregulated in expression and activity[51,52,53]. In the affected neurons of many neurodegenerative diseases, pathogenic protein aggregates can further downregulate the activities of proteolytic

pathways[50]. The inability to metabolize abundant misfolded proteins will also impair the routinely turnover of other proteins and this may induce ER stress and the UPR[46].

Prion-like domains in ALS proteins. The possibility that the spread of the pathology and then motor neuron death in ALS might be prion-like mediated was first suggested by the observation that ALS usually begins focally and spreads in a pattern that implicates contiguous pools of motor neurons[46]. Cell-to-cell spread and propagated misfolding of both mutant[54] and wild type[55] SOD1 have been reported. Moreover, TDP-43 and FUS have low complexity, glycine-rich domains that enhance aggregate formation. Indeed, both proteins emerge in an *in silico* screen for proteins that harbor domains comparable to known yeast prion peptides[56]. Two heterogeneous nuclear ribonucleoproteins (hnRNPA1 and hnRNPA2B1) with prion-like domains have been genetically linked with fALS[57]. Interestingly, the ribonucleoprotein hnRNPA3 is reported to bind the C9orf72 fALS-associated C9orf72 repeat and accumulate in cytoplasmic inclusions unique to C9orf72 patients[58]. Many prion-like domains containing RNA-DNA binding proteins are components of ribonucleoprotein granules (RNP granules), which maintain RNA homeostasis during cellular stress[59] and stress granules (SGs)[46]. The interaction and self-assembly of the prion-like domains of cytosolic RNA-binding proteins facilitates the rapid assembly of stress granules, allowing the sequestration of RNA into these inclusions[60]. In response to stress, mutant form of TDP-43[45] and FUS[60] incorporate rapidly into persistent SGs or form small oligomeric aggregates that interact with the granules. An abnormally strong interaction of the mutant prion-like domains may prevent the disassembly of these granules, resulting in the persistent sequestration of mRNAs and inhibition of their translation. Alternatively, their incorporation into granules might facilitate the conversion of prion-like domains to amyloid states, seeding larger, fibrillary oligomers and inclusion bodies[46].

1.2.3 Exotoxicity and Glutamate response

Glutamate is the principal excitatory neurotransmitter of the central nervous system (CNS). Excessive release of glutamate from presynaptic neurons or delayed clearance from the synaptic cleft results in the sustained activation of postsynaptic receptors[61]. Activation of postsynaptic receptors provokes the Ca^{2+} overload in the postsynaptic cells and probably subsequent toxic events, such as mitochondrial dysfunction[34]. This glutamate-mediated neurotoxicity is called excitotoxicity and has been considered as a possible pathogenic mechanism of ALS even before the identification of SOD1 as a causative gene[62]. Glutamate transporter 1 (GLT1; also known as EAAT2) is expressed on astrocytes and is

implicated in glutamate reuptake. The selective loss of GLT1 has been reported in ALS and is considered to be a causative mechanism of excitotoxicity[63]. There are multiple hypotheses regarding the mechanism of GLT1 reduction, including alternative RNA editing, cleavage by caspase-3, altered miRNA expression, and defects in neuron-astrocyte communication[64]; however, the importance of GLT1 reduction in SOD1-related ALS pathogenesis remains controversial[65]. Another possible cause for excitotoxicity is the excessive glutamate efflux. It has been reported that glutamate release from spinal cord nerve terminals is promoted in SOD1 G93A mice[66]. Microglia also contribute to fatal excitotoxicity through its release of glutamate mostly by a cystine/glutamate antiporter that releases glutamate in exchange for the capturing of extracellular cysteine[67]. Excessive extracellular glutamate is supposed to increase Ca^{2+} level in postsynaptic neurons by the overactivation of Ca^{2+}- permeable glutamate receptors. Cytosolic Ca^{2+} level in motor neurons might be also raised by the release of stored Ca^{2+}. Inositol 1,4,5-triphosphate (IP3)-mediated Ca^{2+} efflux from the ER regulates various signalling pathways[68], and its defects have been reported to be involved in the neuronal toxicity. One of the inositol 1,4,5-triphosphate receptors (IP3R) subtypes, IP3R2 is upregulated in blood from ALS patients and the neuronal overexpression of IP3R2 shortened the lifespan of SOD1 G93A mice[69]. However, the precise mechanism of disturbances in neuronal calcium homeostasis in the ALS context is still unclear.

1.2.4 Mitochondrial abnormalities

Another target organelle for the mutSOD1 is the mitochondria[70]. In mutSOD1 transgenic mice, cardinal mitochondrial functions, such as respiratory chain activity and Ca^{2+} buffering capacity, are diminished by the time of disease onset[71]. Moreover, mutSOD1 selectively accumulates on the cytosolic face of spinal cord mitochondria, even at the presymptomatic stage, which implies that mitochondrial SOD1 aggregates contribute to disease initiation[72]. Mitochondrial homeostasis can be disrupted through the interaction of mutSOD1 with several target proteins. Primary targets are Bcl-2 and its binding partner, voltage-dependent anion channel 1 (VDAC1), both of which are integral membrane proteins that are embedded in the outer mitochondrial membrane. In contrast, another heory is that mutSOD1 directly binds and inhibits VDAC1[73]. Another possible target is the translocase of the outer mitochondrial membrane (TOM) complex. Although interactions between the mutSOD1 protein and the TOM complex were not detected, the mitochondrial protein import was partially impaired and was accompanied with an increase in several components, TOM20, TOM22, and TOM40, in SOD1 G93A transgenic rats[74]. Mitochondrial dysfunction may trigger apoptosis by its release of cytochrome c (Cytc) from the mitochondrial intermembrane space. Moreover, mitochondria play an important role in the intracellular calcium homeostasis as a calcium buffer,

accumulating or releasing calcium depending on the cytosolic levels. Abnormalities in mitochondrial calcium homeostasis were reported in ALS patients and in mutant SOD1 animals. Excessive intracellular calcium levels induce motor neuron death through several mechanisms including: generation of reactive oxygen species (ROS); release of cytochrome c from the mitochondria; glutamate excitotoxicity. As showed by Pansarasa et al., SOD1 mutated LCLs, FUS mutated LCLs and TDP-43 mutated LCLs derived from ALS patients presented mitochondrial morphological changes, mitochondrial dynamics alteration and differences in mitochondrial respiration at the expense of glycolytic flux[75]. All these mechanisms may have a special role in motor neurons because these cells contain less mitochondrial density per volume compared to non-neuronal cells, thus making neurons more deficient in mitochondrial calcium buffering properties. In addition, ALS patients show a deficiency in calcium binding proteins calbidin and paralbumin in cortical motor and spinal motor neurons. These two proteins regulate intracellular calcium levels and their deficiency may result in neuronal loss[76].

1.2.5 Endoplasmic Reticulum stress

Endoplasmic Reticulum (ER) stress occurs when proteins with an aberrant conformation accumulate in the ER lumen due to a disruption in protein quality control or to an overload of newly synthesized polypeptides in the ER. The unfolded protein response (UPR) copes with ER stress by inducing several adaptive responses to reduce the amount of misfolded proteins. However, failure in UPR or prolonged ER stress results in apoptotic cell death. The involvement of ER stress in the pathogenesis of ALS has been suggested in both SOD1-related ALS and SALS cases[31,76]. Moreover, a longitudinal analysis that uses mutSOD1 transgenic mice revealed that the vulnerability of motor neurons in ALS is closely related to chronic ER stress and that vulnerable (fast-fatigable) motor neurons express ER stress markers even at presymptomatic stages and long before resistant (fast fatigue-resistant or slow) motor neurons[32].

Several targets for mutSOD1 both inside and outside the ER have been proposed. In the ER lumen, mutant SOD1 aggregates bind to an ATP-dependent chaperone BiP[77]. ADP-bound BiP binds to substrates while its co-chaperone, SIL1, catalyses the conversion of an ADP- to an ATP-bound state of BiP, thereby facilitating its release of substrates[78]. A recent compelling study suggests that the expression level of SIL1 in motor neurons determines their vulnerability by modifying the intensity of UPR signalling[79]. The authors speculated that, in vulnerable motor neurons, a reduced SIL1 level impairs the ability of BiP to process misfolded proteins, which exacerbates ER stress. Further research should elucidate the precise mechanism of neurotoxicity that is caused by the reduction of SIL1. Cytosolic mutSOD1 interacts with Derlin-1, which serves as a critical component of ER-associated

degradation (ERAD), a machinery for eliminating misfolded proteins from the ER. This interaction inhibits ERAD and induces ER stress, which ultimately elicits apoptosis through the activation of an ER stress sensor, inositol-requiring enzyme 1 (IRE1) and its downstream effector, apoptosis signal-regulating kinase 1 (ASK1). Notably, in addition to many of the reported ALS related mutants, wtSOD1, under zinc-depleted conditions, interacts with Derlin-1[80]. This indicates the possibility that SOD1-Derlin-1 interaction may be involved in the pathogenesis of SOD1 mutation-negative ALS. Additionally, disrupted ER-Golgi trafficking can also produce ER stress possibly through the interaction of the MutSOD1 with coat protein II complex, that initiates budding process for protein transports from the rough endoplasmic reticulum to the Golgi apparatus[81].

1.2.6 Inflammatory dysfunction and contribution of non-neuronal cells

There is considerable evidence that inflammatory processes and non-neuronal cells may play a part in pathogenesis of ALS[82]. The motor neuron death process provokes a neuro-inflammatory reaction that recruits and activates astrocytes and microglia that produce inflammatory cytokines such as interleukins, cyclooxygenase- 2 (COX-2), tumor necrosis factor alpha (TNFα) and monocyte chemoattractant protein-1 (MCP-1), monocyte colony-stimulating factor (MCSF) and transforming growth factor 1 (TGF-1). Some cytokines can trigger microglia to produce copious reactive oxygen and nitrogen species through assembly of NADPH oxidase (NOX), induction of nitric oxide synthase (iNOS), and transcriptional upregulation of lipid-oxidizing enzymes such as COX2. Toxicity induced by ALS-derived astrocytes to motor neurons has been demonstrated *in vitro* using human astrocytes and motor neurons derived from induced pluripotent stem cells[42] with the involvement of signalling via prostaglandin receptors[83] Pro-inflammatory mediators including monocyte chemoattractant protein 1 and IL-10 are present in the CSF of patients with ALS and biochemical indices of immune-response activation are present in the blood[84]. Moreover, Cereda C. et al. reported higher level of TNF-alpha and its soluble receptors, TNF-R1 and TNF-R2, in plasma of patients affected by the sporadic form of ALS compared to normal subjects[85]. Compelling evidence of an association between inflammation and a progression of the disease arises also from studies in murine models of FALS both mutant SOD1 transgenic mice and rats. Several authors reported an increase in the expression of proinflammatory factors before the onset of the disease with sustained microglial activation throughout the active phase of the disease progression. This was confirmed by Beers et al. who studying double transgenic mice carrying mutant SOD1 and lacking CD4 showed development of a more aggressive ALS phenotype, still reversible by bone marrow transplantation[88]. However, the impact of microglia, which can be neuro-protective as well as toxic, is determined by many factors including the phenotype of incoming reactive T cells[46]. Reduced counts of $CD4^+$ $CD25^+$ regulatory

T (TREG) cells and monocytes ($CD14^+$ cells) are detected early in ALS, suggesting recruitment of these cells to the CNS early in the neurodegenerative process. TREG cells interact with CNS microglia, attenuating neuro-inflammation by stimulating secretion of anti-inflammatory cytokine[86]. Jenny S. Henkel et al. demonstrated that when the ALS patients were separated based on the rate of disease progression into rapidly versus slowly progressing ALS patients, the percent of $CD4^+$ $CD25^{high}$ TREGs were reduced in rapidly progressing patients compared with slowly progressing patients and reduced compared with control volunteers[87]. Finally, a study from Sproviero D. et al., showed a significative presence of leukocyte derived MVs (LMVs) compared to endothelial, platelet, erythrocyte derived MVs in plasma of ALS patients compared to AD patients and healthy donors, suggesting a possible role of these LMVs as biomarker of disease progression[89].

1.3 TDP-43

As previously reported, over 20 genes have been linked to ALS so far, with more expected in the future[90]. The most common ALS disease-causing gene mutations are found in: C9orf72 which is responsible for 10-15% of all ALS and contains a hexa nucleotide repeat expansion; SOD1 responsible for 2% of ALS; TARDBP responsible for 0.9% of ALS; and FUS responsible for 0.7% of ALS. Even if rare mutation in TARDBP gene is associated to ALS arise, ~95% of both fALS and sALS patients present inclusions positive to ubiquitin immunoreactive neuronal cytoplasmic inclusions in the degenerating motor neurons[91]. Phosphorylation and ubiquitination of TDP-43 have been identified and recognized to be the source of pathological protein aggregation, inclusion bodies formation and abnormal exosome secretion detected in ALS disease[92]. The pathogenic mechanism of ubiquitinated TDP-43 in ALS, including the origin and redistribution of pathological TDP-43, has been studied intensively in the past years, but more research is needed on the metabolic pathways of neurotoxic TDP-43 fragments for better understanding the relationship between TDP-43 inclusions formation and ALS pathogenesis.

1.3.1 Physiological role of TDP-43

TDP-43 is a 414 amino acid nuclear protein encoded by the TARDBP gene on chromosome 1 (reference sequence NM_007375.3). It was discovered in 1995 as repressor protein of HIV-1 transcription, becoming critical for the regulation of the viral gene expression by its binding to the trans-active response element DNA sequence of the viral genome[93]. Later on, TDP-43 was also reported to be involved in RNA splicing of cystic fibrosis transmembrane conductance regulator

(CFTR) exons and in the development of monosymptomatic and full forms of cystic fibrosis exerting an inhibitory effect on the recognition of CFTR exon 9[94]. TDP-43 is a highly conserved and ubiquitously expressed RNA-DNA-binding protein which belongs to the large heterogeneous nuclear ribonucleoprotein (hnRNP) family[95], that includes many of the most common and powerful splicing modulators known so far, such as hRNP A1/A2, PTB (hnRNP I), and hnRNP H[96]. As the others members, TDP-43 is mainly involved in regulation of RNA splicing, stability, transport, thanks to its ability to bind to RNA with considerable sequence-specificity achieved through the presence of one or more, highly conserved, RNA recognition motifs (RRMs)[97,98]. Through these domains, formed by five β-strands and two α-helices arranged in the β1-α1-β2-β3-α2-β4-β5 pattern, TDP-43 recognizes single-stranded DNA (ssDNA) or single-stranded RNA (ssRNA) with a preferential bind to (UG)n-enriched sequences[99]. TDP-43 is able to actively recognize the 3′ untranslated regions (UTRs) of several thousand mRNA transcripts, and moreover even to its own mRNA as an autoregulation mechanism to control its own cellular concentration and possibly also its solubility[100]. Even if this protein is predominantly localized in the nucleus, it is able to move to the cytoplasm as mRNA shuttle agent[101]. Its mitochondrial localization is based on internal motifs M1 (aa 35–41), M3 (aa 146–150), and M5 (aa 294–300), which are formed by repeated hydrophobic amino acids[102].

The physiological structure of TDP43 comprises of an N-terminal region (aa 1–102) with a nuclear localization signal (NLS, aa 82–98), two RNA recognition motifs: RRM1 (aa 104–176) and RRM2 (aa 192–262), a nuclear export signal (NES, aa 239–250), a C-terminal region (aa 274–414) which is comprehensive of a prion-like glutamine/asparagine-rich (Q/N) domain (aa 345–366) and a glycine-rich region (aa 366–414)[103] (Figure 3).

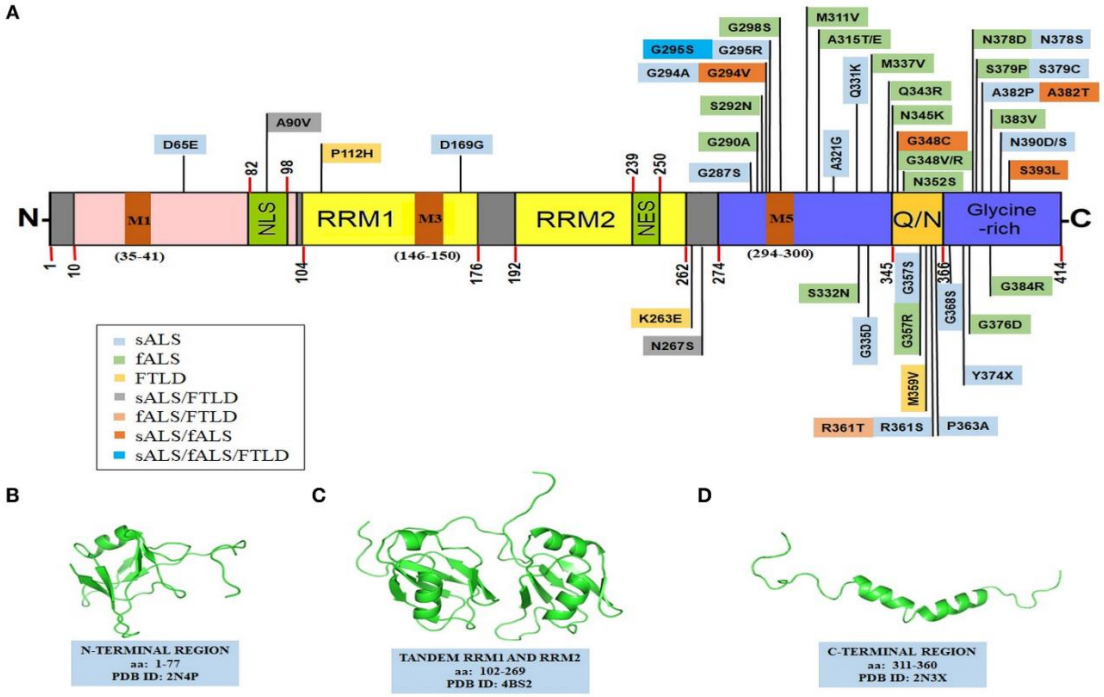

Figure 3. TDP-43 molecular structure. **3.A)** TDP-43's domain organization with ALS and FTLD-linked mutations. Protein Data Bank structures of: **3.B)** N-terminal region (PDB id-2N4P); **3.C)** a tandem RRM1 and RRM2 segment (PDB id-4BS2); **3.D)** a C-terminal region (aa: 311–360) (PDB id-2N3X)[104].

Accumulating evidence suggests that TDP-43 natively dimerization occurs through interactions of the N-terminal residues, that presents a ubiquitin-like fold, (α-helix and six β-strands in the β1- β2-α1-β3-β4-β5-β6 arrangement) [105]. It has been reported that the first ten residues of the NTD are crucial for the formation of the functional homodimers and also in the aggregation of the full-length TDP-43[106], forming TDP-43 dimers by head-to-head interaction of the two NTDs with the RRM2 domains that are extended outwards[107].

The C-terminal region of TDP-43 assumes a particular role for the pathological behaviour of TDP-43, because it is intrinsically aggregation prone and a highly disordered glycine-rich region, enriched in uncharged polar amino acids, glutamine and asparagine (Q/N) as prion-like domains[108]. Moreover, it comprehends most of the ALS-associated TARDBP mutations and phosphorylation sites and it is subject of aberrant activity of caspases for formation of C-terminal fragments (~25–35 kDa), which are highly cytotoxic and are the prominent species found in the inclusion bodies identified from the ALS-affected brains[109].

Like other hnRNP proteins, TDP-43 binds to nascent pre-mRNA molecules when they are released from the RNA Polymerase II (RNApol II) and regulates their maturation steps either through sequential interactions with binding partners or in collaboration/antagonism with specific RNA binding factors[110]. Besides its role in splicing, in mRNA life cycle such as transcription, translation, mRNA transport and stabilization, it is now under investigation its involvement into regulation of non-coding RNAs such as miRNAs, thanks to its interactions with the Drosha and Dicer complexes and lncRNAs by genome-wide studies[111,112]. From genome-wide RNA immunoprecipitation techniques (CLIP-seq), it was proved TDP-43 regulation of more than 6,000 mRNA targets, which would be nearly 30% of the entire transcriptome, and among them they were identified even mRNAs involved in the development of neurons and embryos[113,114]. Another important specific target regulated by TDP-43 at mRNA level that play a role in ALS disease is rasGAP SH3 domain binding protein 1 (G3BP) and T cell-restricted intracellular antigen-1 (TIA-1)[115], which are key component of stress granules (SG) and represent a very crucial target because alterations in SG dynamics have been suggested to play a key role in the TDP-43 protein aggregation process[116] (Figure 4).

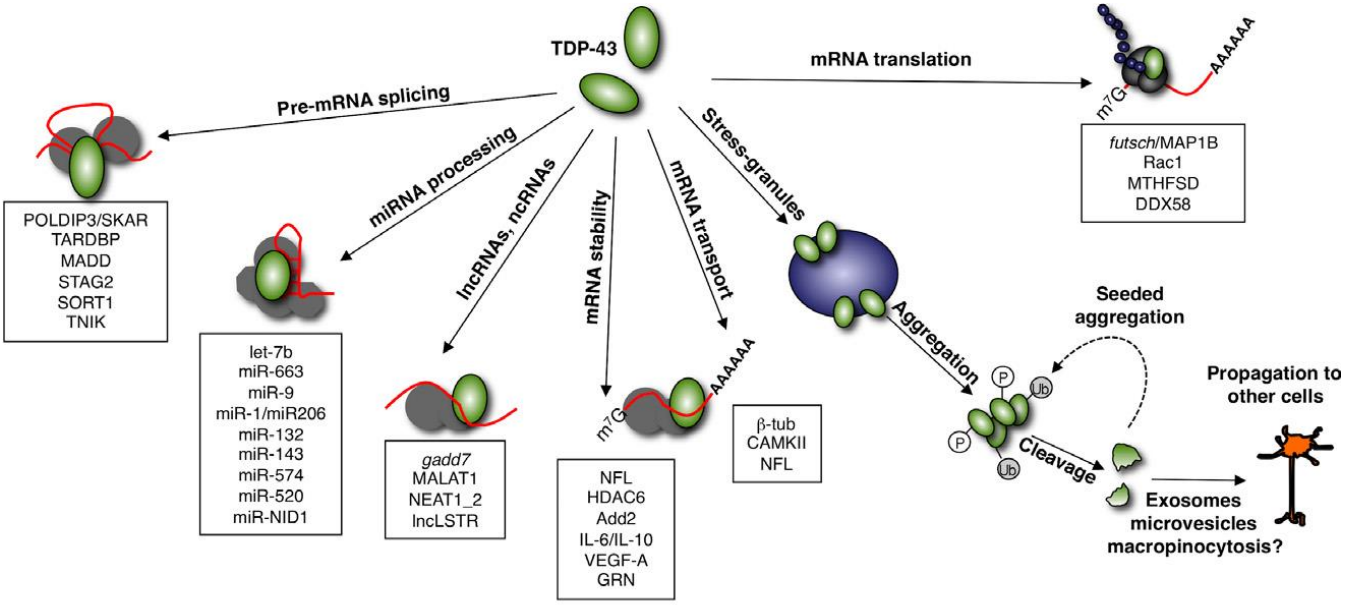

Figure 4. TAR DNA binding protein 43 kDa (TDP-43)-regulated cellular functions in the nuclear and cytoplasmic compartments[117].

1.3.2 TDP-43 mediated pathogenesis in ALS

The pathological hallmarks of TDP-43 proteinopathies include different modifications that have been described to affect TDP-43 in ALS patients: i) nucleus to cytoplasmic mislocalization[118], ii) post-translational modifications leading to deposition of ubiquitinated, acetylated, hyper-phosphorylated, sumoylated protein into inclusion bodies[119], iii) TDP-43 cleavage leading to formation of toxic C-terminal TDP-43 fragments (CTFs) of about 25kDa or 35kDa that tend to induce proteic aggregation[107,120]. The contributions of genetic and environmental factors to the etiology of ALS are complex and until today they are not totally understood. Currently, genetic causes are known for approximately 15 % of all ALS cases; accounting for 11 % of sporadic ALS and 68 % of familial ALS[121], so mutations in TARDBP are a rare cause of ALS accounting for approximately 1-2 % of total cases[122]. Despite to this low incidence, sporadic or familial mutations can increase the previously reported damaging effects and can cause early disease-onset, proving that both loss of physiological function and gain of proteinopathic transformation of mutant TDP-43, associated to an altered proteic structure (i.e., misfolding), is not simply a cellular response to disease but is considered as pathogenic. It has been suggested that the leading events are based on the lost capability of RNA processing and axonal transport function by mutant TDP-43[123, 124], and then on its failure to rescue motor neuron defects caused by knockout of endogenous TDP-43[125]. At the same time, it is also promoted C-terminal fragmentation[126,127], cytoplasmic mislocalization[128,129], aggregation[130,131], and altered proteostasis[132,133]. TDP-43 proteostasis is normally maintained by the coordinated action of the UPS

and autophagy, which is particularly important for clearing TDP-43 oligomers and aggregates[134,135], however, due to ALS-linked mutations in SQSTM1[136], VCP[137] and UBQLN2[138], these protein degradation systems fail to prevent the accumulation of TDP-43, instead inducing the accumulation of large protein complexes called aggresomes. The most common ALS-linked mutation is an intronic GGGGCC repeat expansion in C9ORF72[139,140]. C9ORF72 mutated ALS patients present TDP-43 inclusions in the motor cortex and spinal cord due to the sequestration of RNA-binding proteins, presenting an affection of RNA processing and an accumulation of aggregated TDP-43, associated with ubiquitin[140], p62[141], and/ or ubiquilin 2[142]. These inclusions contain dipeptide-repeat proteins (DPRs) translated in all 6 frames from repeat-containing RNA[143,144], that due to TDP-43 deposition became toxic and cause neurodegeneration. Mutations in the RNA-binding protein genes MATR3 and hnRNPA1 are also associated with ALS and TDP-43 proteinopathy, potentially through direct binding of the affected RBP to TDP-43[145,146], instead FUS mutations are thought to cause ALS, independent of TDP-43 proteinopathy, via impaired processing of transcripts that may be common to those targeted by TDP-43[147,148] (Figure 5).

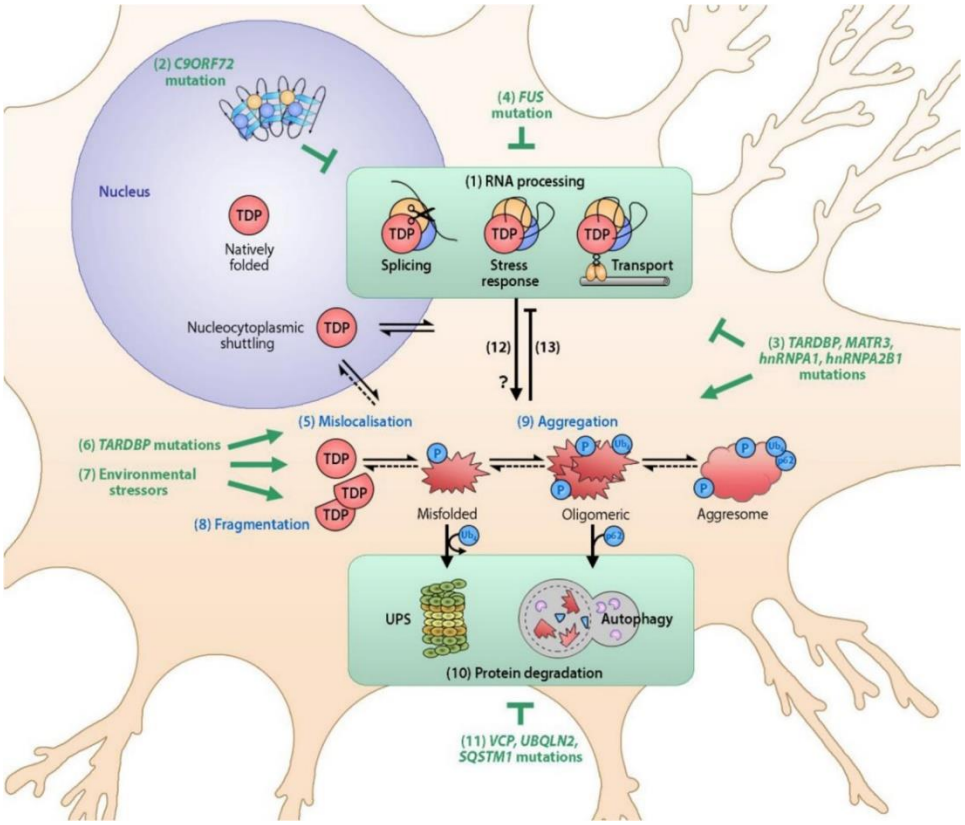

Figure 5. TDP-43 proteinopathy and all the mechanisms involved into amyotrophic lateral sclerosis (ALS) pathogenesis[149].

1.3.3 TARDBP genetic mutations

Most of the ALS associated mutations appear in the exon 6 of the TARDBP gene, which encodes for the C-terminal glycine-rich region of TDP-43 patient case of fALS, enhancing its intrinsic aggregation propensity[150]. The most commonly occurring missense mutations are A382T and M337V and some of the most well-studied mutations are A315T, Q331K, M337V, D169G, G294A/V, and Q343R etc., for which several ALS-disease models have also been established. Mutations, such as G294V, G348C, A328T, and S393L are found in both the sporadic as well as familial cases of ALS[151]. Certain mutations in TDP-43 like G294V, A315T, M337V, A382T, and G376D, are also found to enhance the cytoplasmic mislocalization of TDP-43[152]. In this project we focused our attention on the study of p.A382T mutation which is the most common mutation of the TARDBP gene in ALS disease. This mutation was initially described in two familial patients with ALS of French origin[125] and subsequently identified in large studies carried out on patients of Italian and French origin[153,154]. However, it is in Sardinia that a founder effect accounts for the highest frequency in the world, involving approximately 30% of patients with ALS[155,156], and 20% of fALS and of FTD[157]. p.A382T mutation presents an incomplete penetrance (60% at 70 years) with an increased risk for males[156] and it has been established in several studies as pathogenic[158]. Recently, this mutation was reported to cause reduction of endoplasmic reticulum Ca^{2+} signalling since it was found associated with an increased translocation of TDP-43 in cytoplasm after endoplasmic reticulum (ER) stress induced by the calcium modifying drug thapsigargin[159]. The p.A382T mutation has also been analyzed for its potential effect on protein turnover but the results reported in the literature are contradictory, because from one side it is reported that A382T together with seven other mutations extended the half-life of the protein and caused cytotoxicity as a result of accumulation, cleavage and insolubility in the cytoplasm[160]. On the other hand, another more recent study found faster protein turnover for the p.A382T mutation whereas subcellular analysis of TDP-43 distribution did not reveal significant differences in comparison to the distribution of the WT protein[161]. Another study on TDP-43 distribution in primary fibroblasts of ALS patients with A382T mutation found a significantly accumulation of the protein in both the nucleus and cytoplasm compared to controls and ALS patients with SOD1 mutation[162].

1.3.4 TDP-43 effect on stress granules formation

Eukaryotic cells have developed several mechanisms that protect them against cellular insults such as oxidative stress, heat shock, viral infection and chemical exposure. One example is represented by the formation of stress granules (SG), that are membrane-less cytoplasmic foci of sizes ≤5μm, that

act as safe “storage and sorting stations” for RNA binding proteins, translationally stalled mRNAs and arrested pre-initiation complexes[165]. SG formation is a highly regulated and reversible process that is initiated by the self-oligomerization of the core SG proteins G3BP (RasGAP SH3 domain binding protein and TIA-1 (T-cell intracellular antigen) within 15-20 min of stress exposure[166,167]. By 2 hours post-stress exposure they dissolve allowing cellular processes to return to normality[168,169] and giving translational priority to specific transcripts that are necessary for the cell to overcome the encountered stress[168]. ALS-linked mutations can also influence stress granule dynamics in different ways. Interestingly, TDP-43 C-terminal region can undergo liquid-liquid phase separation (LLPS) to form dynamic protein droplets, where C-terminal residues show mild transient interactions, that appear crucial for stress granule formation[108]. Mutations, persistent stress conditions, or aging, are proposed to cause these droplets to undergo a liquid-to-solid phase separation (LSPS), thereby forming irreversible pathological aggregates[109]. Under sorbitol-induced osmotic stress, the G348C mutant TDP-43 was found to be localized into progressively larger stress granules[163]. On the contrary, the R361S mutant of TDP-43 was shown to disrupt the stress granule assembly[115]. Orrù et al. showed that p.A382T mutation caused a reduction in the ability of cells to respond to stress through loss of TDP-43 function in SG nucleation and reduction in the number of SGs per cell with a regulatory effect on the G3BP1 core protein[164]. TDP-43 is capable of assembling into stress granules, indicating its protective role against cellular insults[165,170] and it is involved in both assembly and maintenance of SGs by the regulation in the expression of G3BP and TIA-1[115]. It still remains unclear whether TDP-43 exerts a pathogenetic action at the cytoplasmic level by disrupting the correct assembly/disassembly of SGs through direct participation to SG composition, or at nuclear level through control of proteins essential to SG formation. A research published by Aulas A. et al. investigated TDP-43 involvement in SG dynamics, showing that especially SG assembly in si-TDP43 cells was impaired by decrease of TIA aggregation tendency. Lastly, they observed that SG composition and/or kinetics are not identical in different cell types as previously reported[172], because while it is possible that HeLa cells have evolved other means to avoid cell death in the absence of normal SG kinetics, in neuronal-like SK-N-SH cells the disruption of this mechanism could be an underlying element of motor neuron vulnerability in ALS, evidencing how the proper control of SG dynamics is essential to neuronal survival[171]. Indeed, considering the sensibility and vulnerability of neuronal cells to conditions of prolonged or repeated stress, in presence of a defective stress the conversion of SGs may become irreversible inducing pathological inclusion bodies as seen in the ALS and FTLD-affected brains[45,170,173].

1.3.5 Emerging role of TDP- 43 in DNA damage response

Many studies have associated TDP-43 toxicity to different cellular pathways, including autophagy, loss of synaptic transmission, inflammation, and microglia infiltration as cause of motor neuron death and degeneration in both familial and sporadic forms of ALS[174,175]. However, these processes are not able to induce motor neurons death by themselves and neither their intervention is sufficient to rescue degenerating neurons. Hence, further investigation is required to identify other physiological functions of TDP-43 involved in the survival of motor neurons. In addition to its RNA-binding activity, TDP-43 also binds to DNA[176,177] and, thanks to this capacity, that is still not totally well understood, new studies are arising for analyzing and investigating its role in DNA transactions.

The DNA damage response (DDR) in mammalian cells is a complex and highly orchestrated signalling process. Even while there are little differences in the early events, following DNA DSB damage, ATM together with the Mre11/Rad50/NBS1 complex is the first that move to the site of DNA damage. Then, activated ATM (autophosphorylation at serine 1981) phosphorylates the H2AX bound to the DSB site, which then facilitates recruitment of other DDR proteins, including 53BP1, in proximity of the DNA break[178]. Repair of double strand brakes (DSBs) can occur via one of three sub pathways, two of which are homologous recombination (HR) in S/G2 cells, and NHEJ, which represents the major pathway for repair of DSBs in the postmitotic neurons mediated by DNAPK holoenzyme (DNA-PKcs/Ku70/Ku80), a DNA polymerase (μor λ), and XRCC4/DNA Lig4[179,180]. A recent study from Hedge et al. documented TDP-43's involvement in DDR as a key component of NHEJ, showing a significant DSB accumulation and reduced NHEJ levels in TDP-43–depleted human neural stem cell-derived motor neurons, as well as in sporadic ALS patients' spinal cord specimens with TDP-43 pathology[181]. Researchers were able to identify TDP-43 as an integral component of early DDR by a strong association with key DDR marker proteins such as γH2AX, pATM, and p53BP1 and as main regulator in the recruitment/activity of the DSB break-sealing ligation complex XRCC4-DNA ligase 4 in healthy neurons[181,182] (Figure 6).

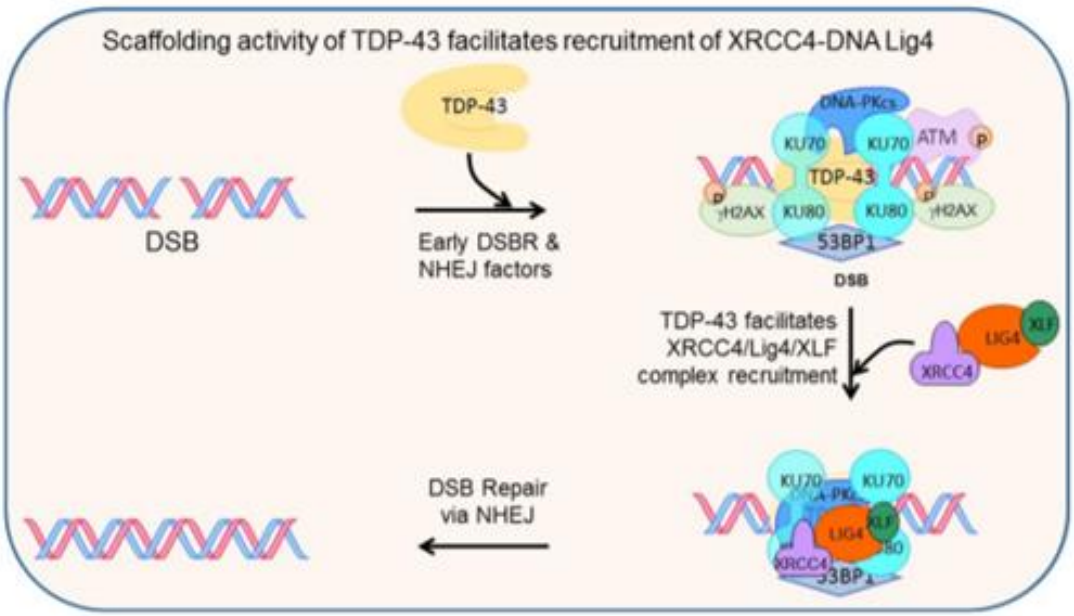

Figure 6. Schematic model showing how TDP-43 facilitates recruitment and activity of XLF/XRCC4/Lig4 complex for NHEJ repair in neurons[181].

In another study from Hedge et al., researchers showed how TDP-43 mutations affect its DNA repair function, in particular by the investigation of the role of the Q331K mutation in neuronal genome instability[185]. The Q331K mutation, first identified in sporadic ALS patients[186], was demonstrated as responsible of an abnormal nucleic acid binding, resulting in an increased aggregation rate of the protein[187], which affect the DNA repair functions of this protein.

A significant accumulation of DNA strand breaks and an increase in levels of DDR markers γH2AX, p-53BP1 and p-ATM were recovered both in genomic DNA isolated from ALS patient spinal cord tissue harboring the Q331K mutation and in neuronal cells ectopically expressing Q331K. This suggested an innovative role of mutant Q331K TDP-43 in promoting DNA damage or affecting DNA repair process in neurons, by promoting ROS stress with gain of toxicity. Moreover, it was seen that the mutant TDP-43 mislocalization to the cytosol is able to inhibit the nuclear translocation of XRCC4, providing an important molecular insight into how mutant TDP-43 affects DSB repair. The role of TDP-43 in regulating the function of DSB is specifically linked to XRCC4-DNA ligase 4 complex, because the same association with XRCC4-DNA ligase 4 was not achieved for another RNA binding protein FUS, which instead is able to bind to XRCC1-DNA ligase 3[188].

1.4 R-loops

1.4.1 R-loops as regulators of genome dynamics

While RNA-DNA hybrids develop naturally during replication and transcription, R-loops are rare events where the nascent RNA transcript invades the DNA duplex as soon as it exits the RNA polymerase II (Pol II) forming a three-strand nucleic acid structure constituted by an RNA-DNA hybrid plus a displaced DNA strand (ssDNA), identical to the RNA molecule. The described "thread-back'' model is the most accepted mechanism for R-loops formation[194], supported by the crystallographic structure of Pol II that demonstrates the exit of DNA and RNA molecules through different channels. This model is in contrast with the ''extended RNA-DNA hybrid'' model, which suggests that RNA-DNA hybrid duplex could arise by threading back the RNA before the two strands of the DNA duplex reanneal within the transcription bubble[195]. According to extensive in vitro studies from the Lieber laboratory[196], R-loops formation are favoured by G clusters and DNA nicks downstream from the promoter on the non-template DNA strand, whereas subsequent RNA-DNA hybrid extension and stabilization are enhanced by high G density and negative supercoiling[197]. Once formed, R-loops are thermodynamically more stable than DNA/DNA interactions[193,198] adopting a conformation that is an intermediate between the A form of a dsRNA and the B form of a DNA duplex[199]. R-loops form naturally as key intermediates in specific cellular-processes, such as E. coli plasmid replication [200,201,202], mitochondrial-DNA replication[203], or immunoglobulin (Ig)-class switching[189,204]. In the latter case, R-loops formation is involved in facilitating class switch recombination (CSR) that generates diverse antibody isotypes thank to the stabilization effect from G quadruplex (G4), formed on the single-stranded exposed strand[205].

Due to the fact that formation of R-loops is an evolutionarily conserved mechanism[190], organisms have evolved several mechanisms to regulate their formation[194]. (Figure 7)

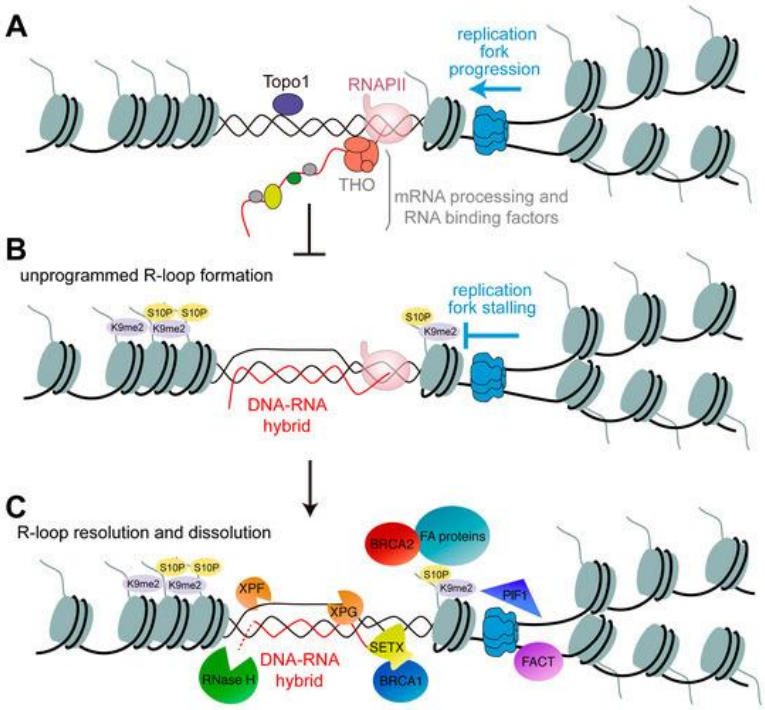

Figure 7. Factors involved in the prevention or in the resolution/dissolution of the of harmful R-loops accumulation **7.A)** Topoisomerase I (TopoI) resolutive action on negative supercoiling together with the cotranscriptional processes of mRNP biogenesis and several RNA binding factors prevent RNA-DNA hybridization. **7.B)** In the absence of either of these factors, the nascent RNA can reinvade the DNA double helix to form a structure termed the R-loops. Chromatin alterations constitute a roadblock for incoming replication forks. **7.C)** RNaseH and helicases, such as Senataxin control R-loops levels in cells by degrading the RNA part of the R-loops. BRCA1 recruits SETX to R-loops sites. FACT, PIF1, BRCA2 and FA factors are involved in to enrich R-loops' resolution[206].

One actively involved factor in RNA-DNA hybrids removal is RNase H1 enzyme, which endonucleolytically cleaves the RNA within the hybrid in a sequence-independent manner (Figure 7.C). Eukaryotic RNase H1 consists of a single polypeptide, with the N-terminal domain being responsible for binding to the RNA-DNA hybrid (hybrid-binding domain or HBD), and the CTD containing the RNase H active site[207]. This enzyme is present both in mitochondria, where it is essential for mitochondrial replication, and even in the nucleus, where its overexpression has been widely used to experimentally remove R-loops, representing the only well studied and efficient way to decrease the cellular levels of these structures[208]. RNA-DNA helicases such as the yeast Sen1 or homologous human Senataxin[209,210] and the human DHX9 helicase, which even acts on G4 structures[211], are involved in R-loops elimination and in prevention of genomic instability by R-loops-mediated DNA damage[209] (Figure 7.C). Besides its role in the resolution of R-loops at G-rich termination pause sites[210], senataxin involvement in DNA repair is under investigation[212,213] thanks

to increasing evidence about formation of nuclear foci in S/G2 phase in response to DNA damage and impaired DNA replication which decreased significantly after R-loops resolution or transcriptional inhibition[213]. Defects in DNA repair enzymes are strongly linked to neurodegenerative disorders[214] and this occur for example in ataxia oculomotor apraxia 2 (AOA2) and in amyotrophic lateral sclerosis type 4 (ALS4), that are caused by mutations mapped in SETX gene[215,216]. Several factors are instead involved in prevention of R-loops formation, such as proteins involved in regulation of mRNA biogenesis and processing proteins. Topoisomerase DNA TOP1 is an evolutionarily conserved factor that, through its ability to relax negative DNA supercoiling, suppresses R-loops formation[217,218](Figure 7.A). THO/TREX complex is responsible for packaging of pre-mRNA with RNA-binding proteins and depletion of the human THO complex results in accumulated DNA breaks and consequent R-loops dependent genomic instability[219] (Figure 7.A). Moreover, genome-wide data suggested that RNA processing factors protect from genomic instability[220], pointing towards a powerful interplay between pre-mRNA processing and R-loops-dependent genomic instability. Eukaryotic Rad51 protein plays a major role in HR during DNA repair of DSBs by stimulation of ssDNA protein RPA in the site of DNA break and promotes strand exchange (invasion of ssDNA into duplex DNA) by forming nucleoprotein filaments[221]. Rad51 is able to promote cis and trans R-loops-mediated DNA damage and genomic instability[222] (Figure 7.A). Rad51 is regulated by Rad52, which is required for Rad51 binding to ssDNA, and by Srs2, which acts as an antagonist and consequently prevents R-loops formation. Moreover, recent studies indicate that during replication stress (RS), dormant origins, represented by inactive origins at the transcription termination sites, start to be duplicated in the opposite direction, and the replication fork makes a head-on collision with transcription in the gene body[223]. In mammalian cells, replication–transcription collisions in long human genes (>800 kb) have been shown to be associated with R-loops accumulation, which in turn causes instability in specific areas called common fragile sites (CFSs)[224], characterised by slower or incomplete replication due to compaction of chromatin[225]. A marker of chromatin compaction is histone 3 Ser10 phosphorylation (H3S10P)[226], that has been linked to R-loops accumulation in S. cerevisiae, C. elegans, and human cells (Figure 7.C). R-loops trigger formation of the H3S10P mark, which in turn could cause replication fork stalling, transcription– replication collisions, and, ultimately, DSBs in the newly synthesized DNA[226]. Even FANCD2, a member of Fanconi anemia complementation group (FA proteins), is required for the stability of common fragile sites (CFSs) under mild replication stress (RS) and it is also reported that is required for efficient DNA replication across CFSs[227](Figure 7.C). It is unclear whether R-loops-FANCD2 contributes/facilitates these pathways or not, but it is expected that central CFS regions at G2 replication stage are bound by FANCD2 and contain R-loops in RS. FANCD2 functions are not

only linked to R-loops suppression but also to protection and rescue of stalled replication forks, thereby preventing overall genome instability. After entering M phase, if this region is still under-replication stage, it may undergo DNA synthesis (termed MiDAS) depending on factors such as SLX4, RAD52, MUS81 and POLD3[228]. Alternatively, the under-replicated DNA may generate the ultrafine DNA bridges colocalizing with FANCD2 foci that have been observed during anaphase leading to genome instability such as micronuclei or 53BP1 bodies at the next G1 phase[229].

1.4.2 R-loops: a cause of genomic instability in ALS

In recent years, an increasing number of studies[230,231] have highlighted an alteration in R-loops metabolism as a possible source of genome instability involved in the pathogenesis of a variety of diseases, including cancers and neurodegenerative disorders[232]. Defects in the mechanisms responsible for the resolution/prevention of R-loops formation or the presence of a G-rich sequence in the non-template DNA, that stabilizes R-loops by the formation of G-quadruplexes, can lead to aberrant and alteration of chromatin structure, transcriptional modification, DNA double-strand breaks, and consequent cellular damage[233,234]. A possible resolution for this detrimental situation has been proposed by Byrd et al., who proposed a new role for G-quadruplex. He suggested a signalling pathway whereby cells, through excision of damaged genomic DNA in response to oxidative stress, can regulate translation accumulating G-rich sequences in cytoplasmic stress granule[235].
Currently, the following main mechanisms responsible for non-physiological R-loops formation in neurodegenerative disorders can be related to 1) mutations in genes encoding factors playing a role in R-loops generation and resolution pathways or 2) can be associated with the presence of pathogenic repeat expansion. Considering these recent findings, an involvement of R-loops in the pathogenesis of the ALS was suggested and hypothesized[232] (Figure 8).

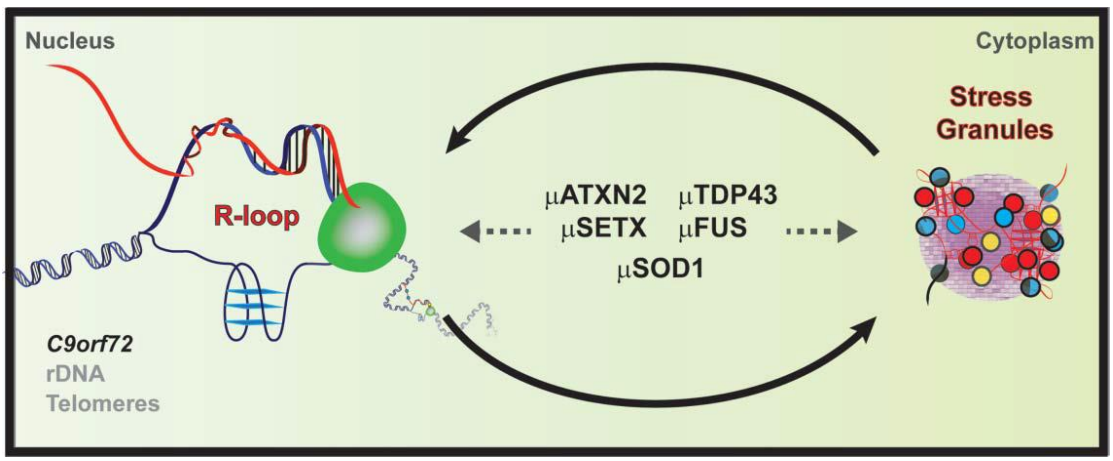

Figure 8. Connection between R-loops and stress granules in mutated linked-ALS genes at the molecular level. Impairment in ALS-linked factors can abrogate their ability to suppress R-loops formation leading to their aberrant sequestration in cytoplasmic stress granules. Increased R-loops accumulation in turn leads to genomic instability and aberrant transcript accumulations, which in turn further promotes stress granule formation that highly affect motor neurons[236].

SETX is an RNA-DNA helicase which is a cause of neurodegeneration in a dominant juvenile form of amyotrophic lateral sclerosis type 4 (ALS4) and in a recessive form of ataxia oculomotor apraxia type 2 (AOA2). SETX mutations can lead to a transcriptional termination defect, associated with accumulation of R-loops and genome instability[234,235,236,237]. SETX also protects genome integrity by managing meiotic recombination with transcription during spermatogenesis and gene silencing during meiotic sex chromosome inactivation. In fact, Setx knock-out mice accumulated DNA double strand breaks and R-loops[238,239], failing to disassemble Rad51 filaments and resulting in a failure to cross over due to collision between R-loops with Holliday junctions[240]. Studies in neuronal cells have demonstrated a role for SETX in neuronal differentiation through fibroblast growth factor 8 (FGF8) signalling, providing one explanation for the effects of loss-of function of AOA2 mutations[241]. In contrast, in a recent study from Lavin and colleagues, R-loops were enriched in proliferating cells (testes), but not in the brain tissues from Setx, Atm, Tdp1 or Aptx knock-out mice[242]. These results are in contrast with inducible R-loops detected in neuronal cells at the Snord116 locus, which is associated with the neurodevelopmental disorder Angelman syndrome[243]. So, it is possible that R-loops are regulated by different mechanisms in proliferating cells and postmitotic neurons, thereby leading to different R-loops kinetics and so preventing their detection in some model systems.

Although events leading to C9ORF72 hexanucleotide repeat expansion (HRE) remain unclear, it has been described that C9ORF72 HRE DNA sequences constitute G-quadruplex DNA-containing R-loops that interfere with RNA Pol II transcriptional elongation and lead to the generation of truncated G-quadruplex RNA-containing C9ORF72 transcripts[244,245]. Haeusler et al. observed that R-loops resolution after RNase H treatment reduces the accumulation of abortive transcripts, suggesting a direct involvement of G-quadruplex DNA-containing R-loops in the impairment of the transcriptional process[246]. These truncated transcripts bind to ribonucleoproteins including the nucleolar protein nucleolin, which in turn leads to the aggregation of aborted transcripts and associated factors within the nucleus and cytoplasm[244,245]. Consistent with this finding, C9ORF72 HRE ALS patient cortex tissues showed evidence of nucleolar stress and motor neurons derived from induced pluripotent stem cells of C9ORF72 HRE patients exhibit nucleolin mislocalization[244]. Although C9ORF72 HRE could explain the increase in R-loops in ALS patients' cells, there is another proposed mechanism, which

associates the reduction of C9ORF72 expression to R-loops formation[247]1) by transcriptional impairment through RNA polymerase pausing, antisense transcription, and histone methylation[248] and 2) by chromatin modification due to the presence of numerous R-loops observed in hypermethylated regions on the C9ORF72 gene[249,250].

Salvi et al. observed that Pbp1, the ortholog of ATXN2 in yeast, regulates R-loops generation by suppression of deleterious R-loops formation at G quadruplex-DNA sites in the intergenic spacer of repeated regions and at telomeres of the ribosomal DNA[251]. In humans, ATXN2 can also exert its inhibitor action on R-loops formation at some loci even if it is mainly a cytoplasmic protein[252].

Recently, intermediate ATXN2 polyQ expansions were identified as a major risk factor for ALS[253] and an intriguing possibility is that this may lead to R-loops accumulation within key coding or non-coding DNA sequences and loci with a central role in neural function. Besides this, ATXN2 can also physically interacts with TDP-43 in a RNA-dependent manner, influencing its aggregation tendency in a proportional way respect to ATXN2 level in the cell[254] and to the length of the polyQ expansions[255]. One possible model is the formation of stress granules or cytoplasmic protein aggregates causing the sequestration of proteins normally located in the nucleus, such as TDP-43, ATXN2, and FUS, altering their nuclear functions, in particular their roles in RNA splicing, processing and transport as well as R-loops metabolism and resolution. R loop accumulation can subsequently lead to genomic instability and accumulation of mutations that can promote stress granules formation, sustaining the cycle[236]. After stress induction, ATNX2 and TDP-43 both relocate to stress granules[39] together with FUS[252] altering their ability to bind DNA and RNA. Otherwise, mutated TDP-43 and FUS are potentially able to sequester ATXN2 into cytoplasmic stress granules, altering its possible function as a suppressor of R-loops formation[236].

More experiments are needed to clarify the molecular pathway that controls the relation between the genomic accumulation of detrimental R loops and the cytoplasmic sequestration of the previously cited mutated proteins in the stress granules, for better understanding the neurodegenerative mechanisms implicated in ALS pathogenesis.

1.4.3 R-loops proposed therapies

Considering the proposed molecular model based on the pathogenic pathway of DNA damage induction driven by R-loops as cause of neurodegeneration in ALS, R-loops manipulation strategies would have important implications, including a possible therapeutic target role in this multifactorial disorder. Individualising a conspicuous number of alterations in RNA metabolism in different neuromuscular diseases, the hypothesis of a common defect could induce the development of possible therapeutic strategies not only for rare diseases but also for other neurodegenerative disorders. R-

loops also warrant investigation as therapeutic targets for diseases, such as Spinal Muscular Atrophy (SMA), that already have an available approved therapy[296]. However, despite their apparent importance in different pathogenic pathways, R-loops have yet to be fully exploited in drug design[248]. There are different ligands that can target RNA-DNA hybrids, including ethidium bromide, the aminoglycosides neomycin and paramomycin, and the polyamides distamycin and netropsin[256] through intercalation and binding to the nucleic acid groove. Although exhibiting high binding affinities to RNA-DNA hybrids, many of these molecules also bind dsDNA and RNA and for this reason they become mutagenic, limiting their potential biological applications[256]. However, recent studies suggest that combining the properties of ligands such as those linking aminoglycosides to derivatives of ethidium bromide[257] subnanomolar affinity for RNA-DNA hybrids can be achieved, providing a possible approach for the development of potent and specific RNA-DNA hybrid ligands in future drug design efforts. Various compounds that modulate DNA supercoiling and inhibit DNA topoisomerases, such as topotecan and camptothecin, can also affect R-loops formation in vivo[258,259]. It has previously been proposed that R-loops in trinucleotide expansion diseases could be targeted to suppress repeat expansions or reactivate silenced genes[260]. R-loops in trinucleotide repeat expansions could be modulated by targeting the expansions themselves[260]. A recent study showed direct evidence that a small molecule is able to suppress R-loops formation at expanded CGG repeats in the FMR1 gene, thereby preventing FMR1 epigenetic silencing in fragile X syndrome (FXS)[261].

As an alternative approach, R-loops levels may be indirectly modulated by treatments that target proteins involved in R-loops biology. For instance, genomic instability caused by a widespread increase of R-loops due to loss of an R-loops suppressing protein could potentially be reverted by introduction of an alternative R-loops suppressor. Upregulation of genes encoding R-loops-specific helicases, such as senataxin, demonstrated that RNA-DNA helicase activity is responsible for R-loops resolution at transcription termination sites[248], potentially representing a promising strategy to treat pathology characterized by R-loops accumulation as noted in C9ORF72-associated diseases[262]. Similarly, modulation of ataxin could represent another therapeutic possibility, thanks to its R-loops regulatory role its involvement in numerous pathways related to R-loops, such as TDP-43 and FUS stress response pathway[254]. Overexpression of endonuclease RNase H1/H2 that specifically degrades RNA of RNA-dna:DNA hybrids could represent another strategy to reduce R-loops levels and consequently resolve DNA damage[263]. Finally, an alternative approach is to target the downstream effect of R-loops promoting DSB repair kinetics with chromatin-modulating drugs[262]. Furthermore, the S9.6 antibody offers new opportunities for research and for this it has been used in the development of biosensor systems[264], detection of miRNA targets[265], and as a key component of human papillomavirus (HPV) diagnostic kit (Qiagen). In front of these exciting prospect of

developing new targeted therapeutics for many human disorders, there are many studies for uncovering the role of R-loops in health and disease in recent years that have the necessity to individualise specific efficient treatments considering the ubiquitous nature of R-loops.

1.5 TDP-43 role in R-loops associated DNA damage

Different studies support the idea that RNA-binding proteins (RBPs), which are involved in different steps of mRNA life, from transcription to translation, can also affect genome stability programs[266], indicating that RNA metabolism and DNA repair pathways functionally intersect. There are several possible ways through which these two mechanisms can be connected, for example, RBPs action may affect the splicing profiles and levels of mRNAs of proteins involved in cellular response to DNA, since some splicing factors are included in a control strategy on splicing programs after DNA damage, modulating mRNA stability, monitoring cotranscriptional splicing and regulating the DDR after post-translational modifications[267]. In some cases, RNA-binding proteins may directly participate in the DDR, for example hnRNP G or SFPQ that are involved in the regulation of HR[268], or Ntr1, normally implicated in NHEJ[269], have been detected in DNA damage sites[270]. RNA-binding proteins can also prevent DNA damage by the regulation of hazardous R-loops' formation, which is controlled in particular in negatively super-coiled region by the physiological association of RBPs with the pre-mRNA molecule emerging from the transcriptional machinery RNA Pol II[271]. The nascent RNA molecule is sequestered in ribonucleoprotein (RNP) complexes and it is processed into mature mRNP particles, passing through multiple processing steps, including maturation of 5' and 3' ends and splicing, followed by shuttling to the cytoplasm. Lack of coordination among different RNA maturation steps, leads to the perturbation of the interaction between the nascent transcripts and the DNA template with deleterious effects on genome stability. DNA damage induces the activation of signalling pathways that target the expression and post-translational modification of RBPs involved in the metabolism of protein-coding transcripts. However, the mechanism through which this protein–RNA packaging influences the capacity to form R-loops is still under investigation, even if it is commonly assumed that binding to RBPs avoids R-loops formation[271]. The involvement of WT TDP-43 in non-homologous end joining (NHEJ)-mediated DSB repair has been reported by Hedge et al., showing a new role for this ribonucleoprotein as stabilizing factor in the scaffold for the recruitment of XRCC4-DNA ligase 4 complex. The mutation in TDP-43 affected the nuclear translocation of XRCC4-DNA ligase considering a reduction of their nuclear interaction due to an enhanced cytosolic mislocalization of mutated TDP-43. Given the elevated genome damage in the post mortem spinal cord tissue from an ALS patient with acquired TDP-43 Q331K mutation, they decided to investigate this particular missense point mutation in DNA damage response. From another study from the same

group, using conditional SH-SY5Y lines ectopically expressing wild-type (WT) or Q331K-mutant TDP-43, they confirmed that the increased cytosolic sequestration of the poly-ubiquitinated and aggregated form of mutant TDP-43 was correlated with increased genomic DNA strand breaks, activation of the DNA damage response factors phospho-ataxia-telangiectasia mutated (ATM), phospho-53BP1, γH2AX and neuronal apoptosis[272].

A study from Hill et al. confirmed the colocalization of WT TDP43 with active RNA pol II at sites of DNA damage along with the DNA damage repair protein, BRCA1, with γH2AX and phosphorylated RPA, while its depletion was associated with an excessive fork stalling. They also proved TDP-43 participation in the prevention/repair of R-loops associated DNA damage, using overexpression of RNase H as negative control for proving direct R-loops participation[183].

The activation of DNA damage pathway is interconnected with the signalling pathway that targets the expression and post-translational modification of RBPs, which in turn are involved in the metabolism of protein-coding transcripts. In the last ten years several studies showed that the DDR is important in regulating precursor or stem cell differentiation programs[273]. For example, the response to DSBs induces the development of vertebrate adaptive immune systems, based on programmed induction and subsequent damage repair during antigen receptor gene rearrangements to assemble a complete Ig gene via V(D)J recombination, and it even regulates a network of genes involved in proliferation, B-cell self-renewal, and cell differentiation[274]. The formation of DSBs, which is necessary for class switch recombination (CSR) and somatic hypermutation (SHM) required for the production of high-affinity antibodies of different isotypes, is characterized by the programmed formation of physiological R-loops[275] and deoxycytidine deamination mediated by AID[276]. Interestingly, unscheduled AID-mediated DSBs are implicated in cancer and neurodegeneration[277] but the mechanism which links these diseases to aberrant R-loops structures is not still totally clear.

In view of their association with large, still poorly characterized, multiprotein assemblies linked to genome stability, transcription, and pre-mRNA processing, RBPs and in particular TDP-43 may be crucial to be studied in DNA damage and in response to genotoxic stress in ALS. A better understanding of the function of these RNA-binding/processing proteins could also aid in identifying and generating new biomarkers in diseases that share these particular alterations in RNA metabolism such as ALS.

2 Aim of the work

TDP-43 aggregates are identified as a major component of the ubiquitinated neuronal cytoplasmic inclusions deposited in spinal motor neurons in ALS[88]. The pathogenesis of 97 % of both familial and sporadic forms of ALS are strictly associated to TDP-43 proteinopathies and the deposition of full-length and fragmented, ubiquitinated and hyperphosphorylated TDP-43 protein as detergent-resistant causes aggregates in the cytoplasm and clearing of TDP-43 from the nucleus[118]. Phosphorylated C-terminal fragments (CTFs) of TDP-43 are accumulated in aggresomes by proteic cleavage action of caspases. The cleavage is enhanced by C-terminal TDP-43 mutations[122,126], by cellular stress[107], and proteasomal inhibition[278]. Due to its sequestration in stress granules, misfolded mutated TDP-43 showed a decreased ability to exert its physiological function, such as transcription, mRNA splicing and transport[123,124] and miRNAs biogenesis[111,112], with a loss of function. Subsequent to this, the formation of oligomers and aggregates of TDP-43 in the cytoplasm may recruit native TDP-43 or other interactors proteins[279], constituting a gain of toxic function associated to neurodegeneration[120]. While the regulated aggregation of RBPs and RNA into stress granules is wholly reversible under physiological conditions[280,281], due to the persistence of conditions of prolonged or repeated neuronal damage, stress granules might act to "seed" the irreversible pathological transformation of proteins associated to ALS[173, 282, 283].

In addition to its RNA-binding activity, TDP-43 also binds to DNA[176,177]; however, until now, its possible role in DNA transactions is not so clear. Considering the significant accumulation of genomic damage that is consistently observed in multiple neurodegenerative diseases, recent studies from Hedge et al. investigated TDP-43's involvement in DDR. Researchers proved TDP-43 involvement in nonhomologous end joining (NHEJ), the major pathway for repair of DNA double-strand breaks (DSBs) in the postmitotic neurons[181], and they investigated how increased cytosolic sequestration of the poly-ubiquitinated and aggregated form of mutant TDP-43 is correlated with higher genomic DNA strand breaks, activation of the DNA damage response factors phospho-ataxia-telangiectasia mutated (ATM), phospho-53BP1, γH2AX and neuronal apoptosis in SH-SY5Y lines ectopically expressing wild-type (WT) or Q331K-mutant TDP-43[272].

The DDR in mammalian cells is a complex and highly orchestrated signalling process. In DNA DSB damage, ATM together with the Mre11/Rad50/NBS1 complex is the first responder. Activated ATM (autophosphorylation at serine 1981) phosphorylates a member of the H2A histone family H2AX bound to the DSB site, which then facilitates recruitment of other DNA damage repair (DDR) proteins in proximity of the DNA break[178]. Hill et al. proved that TDP-43 colocalizes with active RNA polymerase II at sites of DNA damage along with the DDR protein, BRCA1 and also participates in

the prevention and/or repair of R loop-associated DNA damage[183]. Apart from this, a clear vision of the mechanism through which TDP-43 is able to regulate the formation of genomic R-loops during DNA transcription in relation to ALS pathology is not well clarified until now.

From many studies, R-loops are becoming central molecular players in different neurological disorders and deserve attention for pathogenic theories and therapeutic strategies. Besides their presence is conserved ubiquitously in different conditions and in different organisms and involved in main biological processes such as bacterial plasmid replication, DNA transcription, mitochondrial and nuclear DNA replication, or immunoglobulin (Ig) class switching[194], R-loops' accumulation can also be deleterious to cells, promoting DNA damage and genome instability[230,231,232]. Currently, in neurodegenerative disorders two main mechanisms have been proposed to be responsible for non-physiological R-loops formation: first of all mutations in genes encoding factors playing a role in the R-loops generation and resolution pathways[284,285]: in second place the presence of pathogenic repeat expansion, which could be involved in thermodynamic stabilization of the structure of RNA-DNA hybrids that subsequently contribute to repeat instability and to the pathogenesis of expansion diseases[286,287]. Apart from these evidences, these findings and hypotheses lead to numerous other questions, like what is the relation of distinct mutations in disease-associated genes with aberrant R loops' formation and the consequences in the arising of neurodegenerative diseases. Despite the fact that these mutations occur in ubiquitously expressed genes, these diseases present primarily neurologic deficits, highlighting a particular sensitivity of the nervous system and motor neurons to defects in R-loops metabolism and genome stability, which may be related to an inefficient repair machinery[288]. The evidence that R-loops levels in neuronal cells during embryogenesis is crucial for the development of the central nervous system[289] underlines how much the role of R-loops in motor neuron biology is fundamental for understanding the selectivity of motor neuron death in neurodegenerative disorders.

Unfortunately, the study of the mechanisms that controls and prevents co-transcriptional R-loops formation is not well understood. So, the aim of this project was to investigate the role of TDP-43 mutation, overexpression and depletion on R-loops' formation and their impact on cellular processes involved in RNA modulation and regulation cooperating with resolving or suppressing RNA-DNA hybrid factors. The knowledge of this molecular mechanism will shed some light on R-loops pathways and could represent a step forward in the understanding of ALS and our ability to treat them.

3 Material and Methods

3.1 Cell culture

Lymphoblastoid cell lines (LCLs) immortalized with Epstein Barr Virus (EBV) were harvested with RPMI+20% of fetal bovine serum (FBS), 1% of penicillin/ streptomycin (100X) and 1% of L-glutamine (200mM). It was used LCLs derived from a healthy control (Ctrl LCLs), from an ALS patient carrier of a p.A382T TDP-43 mutation in the C-terminal domain of the protein (TDP-43 mut LCLs) and from a sporadic ALS patient (sALS LCLs).

SH-SY5Y (CRL-2266, American Type Culture Collection [ATCC]) and HeLa cells (CCL-2, American Type Culture Collection [ATCC]) were cultured in Dulbecco's modified Eagle medium (DMEM, Thermo Fisher Scientific) medium supplemented with 10% fetal bovine serum (FBS), 1% Penicillin-Streptomycin and 1% L-glutamine in a humidified chamber at 37◦C with 5% CO2.

Murine hybridoma cell lyne ATCC® HB-8730™ were grown in with DMEM +10% FBSm, 4 mM L-glutamine, 4500 mg/L glucose, 1 mM sodium pyruvate, and 1500 mg/L sodium bicarbonate.

3.2 Transformation of competent bacteria

Competent cells (TOP10) were thawed on ice for 10 minutes. 50ng of plasmid, dissolved in water, were added to the cells, gently mixed and left on ice for 30 minutes with regular mixing to resuspend the cells. The cell and DNA mix were then heat-shocked at 42°C for 90 seconds and then placed on ice for at least 2 minutes. 160ml of L-Broth was added and incubated with shaking at 37°C for 30 minutes. We plated cells at a correct confluency ($\geq 1\times10^8$ CFU/µg DNA) at 37°C ON on plates with selective agent (ampicillin, 100 µg/ml).

3.3 Preparation of plasmid DNA-Miniprep

After vector reconstitution, we followed a transformation protocol on efficient competent TOP10 E.Coli for high-efficiency cloning and plasmid propagation. We picked at least 4 independent colonies for each condition and we performed an extraction of high copy plasmid DNA using a plasmid miniprep kit (NucleoSpin® Plasmid NoLid, Macherey-Nagel GmbH & Co.).

3.4 SH-SY5Y stable transfection with GFP tagged TDP-43 vectors

SHSY5Y were stably transfected with GFP tagged plasmid coding for TDP-43 WT and with a GFP tagged plasmid coding for TDP-43 p.A382T mutation, using Fugene HD (Promega), with a Fugene:DNA ratio of 3:1. Plasmids harboring C-terminally GFP-tagged wild-type (PS100010 pCMV6-AC-GFP) or mutant human TDP-43 sequences (CW303334 Mutate ORF of RC210639 at nt position 1144, to nr change G to A,to achieve A382T, keep insert in PS100001) were purchased

from OriGene. Transfected SH-SY5Y were selected using 600 µg/ml geneticin (G418) and were clonally isolated. The presence of the A382T (c.1144G>A) mutation was confirmed by DNA sequencing performed by GENEWIZ (forward sequencing primer 5'-GGACTTTCCAAAATGTCGT-3'; reverse sequencing primer 5'- CCCACCAGCCTTGTCCTAAT-3').

3.5 TDP-43 silencing with SiCTRL and SiTDP-43 vectors in HeLa cell lines

For siRNA silencing subconfluent HeLa cells were transiently transfected using DharmaFect1 (Dharmacon) with siCTRL (D-001810-01-05) and siTDP-43 (L-012394-00-0005) using 20 nM from each siRNA SMARTpool for conditions. Cells were grown to densities of $5X10^4$ per 4-cm dish in 2 ml DMEM without antibiotics, they were transfected with DharmaFect according to the manufacturer's instructions and collected after 72h.

3.6 Silencing efficiency control

RNA was extracted from Hela silenced cells using RNeasy Mini Kit (Quiagen), in accordance with manufacturer' instructions. The efficiency of silencing was verified by qPCR on recovered cDNA using QuantiTect Reverse Transcription Kit (Quiagen), in accordance with manufacturer' instructions by techne TC312 thermal cycler. As housekeeping gene HPRT was used (HPRT forward sequencing primer 5'-GGACTAATTATGGACAGGAC-3', reverse sequencing primer 5'-TCCAGCAGGTCAGCAAAGAA-3'; TDP-43 forward sequencing primer 5'-GGCCTTTGCCTTTGTTACATTTG-3', reverse sequencing primer 5'-CAAAGAGACTGGCGCAATCTGA-3'), obtaining in media 80% of silenced Hela cells.

3.7 HeLa and SH-SY5Y transient transfection with RNaseH overexpressing plasmid

SHSY5Y and Hela were also transfected with the following plasmid: pEGFP, a vector expressing GFP (pEGFP-N2, Clontech), and pEGFP-M27, containing the GFP-fused RNase H1 lacking the first 26 amino acids responsible for its mitochondrial localization cloned into pEGFP. For each condition we used 2µg of plasmid DNA and we transfected by Lipofectamine 2000 (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions[208]. All assays were performed 72 h after siRNA transfection plus 24 h after plasmid transfection[290].

3.8 Stress granules induction and immunofluorescence in LCLs

Stress granules induction was performed using a treatment with 0.5 mM of sodium arsenite (Sigma) for 45 minutes on Ctrl, sALS and TDP-43 mut LCLs. As reported in literature, HRI (heme-regulated

initiation factor 2α kinase) is able to induce phosphorylation of eIF2α on Ser51 under stress triggered by arsenite treatment, inhibiting global protein translation through depletion of the eIF2–GTP–tRNA-met ternary complex. This mechanism allows the RBP TIA-1 to bind the 48S complex instead of the ternary complex and promotes polysome disassembly and the consequent recruitment of mRNAs to SGs[291]. For immunofluorescence analysis of LCLs, cells were fixed with two methods: one using formaldehyde and cold acetone and the other using cold methanol. After incubation in blocking solution (blocking buffer: PBS 1X, BSA 3%, Tween20 0,1%), Ctrl, sALS and TDP-43 mut LCLs were stained with the following primary antibody overnight (ON) at 4°C: rabbit **anti-TDP-43** antibody (1:100, Proteintech Europe), goat **anti-TIA** antibody (1:50, Santa Cruz Biotechnology). The next day after phosphate buffer saline (PBS 1X) washing, cells were stained either with anti-rabbit Alexa Fluor 488 (1:100, Sigma Aldrich) and anti-goat Alexa Fluor 647 (1:700, Sigma Aldrich) and a confocal laser microscope via z-stack acquisition was used for analysis (Leica DM IRBE, Leica Microsystems Srl).

3.9 S9.6 Purification

S9.6 antibody was purified from cell medium of confluent hybridoma cell line ATCC® HB-8730™ with Protein A/Protein G GraviTrap™ (GE Healthcare) using manufacturer' instructions. The medium from starved hybridoma cell line was collected and loaded in prepacked sepharose columns previously equilibrated with binding buffer (0.02 M sodium phosphate, pH 7.0). S9.6 antibody was captured with high specificity on protein A/G ligands in the gravity-flow columns and after a washing step with binding buffer is eluted through elution buffer (0.1 M glycine-HCl, pH 2.7). Different fractions of eluted antibody were collected, quantified through BCA proteic assay and stored at -80°C. As a safety measure to preserve the activity of acid-labile IgG, 1 M Tris-HCl, pH 9.0 was added to the tubes used for collecting antibody-containing fractions for maintaining the pH of the sample approximately neutral.

3.10 Immunofluorescence of R-loops and of R-loops related factors in in LCLs, SH-SY5Y and HeLa

For immunofluorescence analysis of LCLs, cells were fixed with two methods: one using formaldehyde and cold acetone and the other using cold methanol. After incubation in blocking solution (blocking buffer: PBS 1X, BSA 3%, Tween20 0,1%), Ctrl, sALS and TDP-43 mut LCLs were stained with the following primary antibody overnight (ON) at 4°C: murine derived hybridoma **anti-S9.6** antibody (1:500), rabbit **anti-TDP-43** antibody (1:100, Proteintech Europe). The next day after phosphate buffer saline (PBS 1X) washing, cells were stained either with anti-mouse Alexa

Fluor 594 (1:500, Sigma Aldrich), anti-rabbit Alexa Fluor 488 (1:100, Sigma Aldrich) and a confocal laser microscope via z-stack acquisition was used for analysis (Leica DM IRBE, Leica Microsystems Srl).

For S9.6 immunofluorescence analysis in SHSY5Y and Hela, cells were harvested from coverslips in 6-well plate and fixed with cold methanol for 10 minutes at -20°C.After two washes in PBS, cells were blocked in PBS-blocking solution (1X PBS+ 3%BSA+ 0,1%Tween 20) and incubated with the following primary antibody ON at 4°C: murine derived hybridoma **anti-S9.6** antibody (1:1000), rabbit **anti-nucleolin** antibody (1:2000. Abcam). The next day after phosphate buffer saline (PBS 1X) washing, cells were stained with anti-mouse Alexa Fluor 546 (1:1000, Sigma Aldrich) and anti-rabbit Alexa Fluor 647 (1:1000, Sigma Aldrich). The incubations with primary and secondary antibodies were performed in PBS-blocking solution. For γH2AX and FANCD2 immunofluorescence analysis in SHSY5Y and Hela, cells were harvested from coverslips in 6-well plate and incubated with a fixation solution (PFA 4%, TRITON 0,1%, PBS1X) for 10 minutes at room temperature (RT). After two washes in PBS cells were blocked in TBS-blocking solution (1X TBS+ 3%BSA+ 0,1%Tween 20) and incubated with the following primary antibody ON at 4°C: rabbit **anti- γH2AX** (1:1000, Abcam), mouse **anti-FANCD2** (1:1000, SantaCruz), rabbit **anti-H3S10P** (1:200, Millipore). In this case we used a TBS-blocking solution because phosphate presence in PBS may interfere with the immune reaction between the studied phosphorylated proteins protein and specific Ab. The next day after phosphate buffer saline (PBS 1X) washing, cells were stained either with anti-mouse Alexa Fluor 546 (1:1000, Sigma Aldrich) and anti-rabbit Alexa Fluor 647 (1:1000, Sigma Aldrich). The incubations with primary and secondary antibodies were performed in TBS-blocking solution. For avoiding false positive signal from dsRNA, SHSY5Y cells were treated with 40 U/ml RNaseIII (1 U/µl, Thermo Fisher Scientific) for 30 minutes at 37° using 1X RNase III Reaction Buffer. Antibody signal was detected on a Leica DM6000 microscope equipped with a DFC390 camera (Leica). Data acquisition was performed with LAS AF (Leica).

3.11 Data quantification and statistical analysis of IF

S9.6 raw integrated density of LCLs IF was calculated through ImageJ program for measure nuclear R-loops presence in these cells. R-loops signal in the nucleoplasm of SH-SY5Y and of HeLa cell lines was calculated using ImageJ program by the measure of S9.6 raw integrated density observed in the nucleus, positive to DAPI staining, subtracting the nucleolar contribution, detected by nucleolin antibody. γH2AX, FANCD2 and H3S10P number of foci was detected in IF of SH-SY5Y and of HeLa cell lines through ImageJ program, using a macro specific for recognition of foci set in the channel of the considered antibody. Then we calculated the foci count distribution in each condition,

obtaining the relative number of cells containing >10 foci in the nuclei. In each IF experiment we maintained the number of cells around 100 counted cells for improving their comparability. Statistical analysis was performed by Student t-test and by One-Way Analysis of Variance (ANOVA test) followed by post hoc comparison as a post-test (GraphPad Prism version 5, San Diego, CA, USA). Values were considered statistically significant when p values were < 0.05.

3.12 Flow Cytometry (FCM) analysis for S9.6

Ctrl, sALS and TDP-43 mut LCLs were incubated with a vitality dye for 15 minutes (Zombie Violet™ Fixable Viability Kit, BioLegend) for discriminating cells with regular vitality and cells with an apoptotic profile (Figure 9), thanks to its ability to bind cytoplasmatic proteins only in cells with a compromised membrane. Then, LCLs were incubated for 20 minutes with anti-CD19 antibody, which is positively expressed by B lymphocytes membrane. For RNA-DNA hybrids detection, cells were fixed and permeabilized for one hour using a kit based on saponin permeabilization (Fixation/Permeabilization Solution Kit, BD), chosen as the best method to permeabilize the membrane as described by the literature[291]. Cells were treated with 60 U/ml of ribonuclease H (RNase H, 5.000 units/ml, NEB) using RNase H buffer as negative control at 37°C for 1 hour. Using the same conditions, cells were treated with 100 µg/mL ribonuclease A (RNase A, 10 mg/mL, Thermofisher) in 0,4M NaCl buffer at 37°C for 1 hour, because this enzyme is able to digest of ssRNA at high salt concentration. In this last case, buffer at different salt concentration higher than 0.3 M were tested and it was chosen the one that was able to allow RNase A specific action against ssRNA and at the same time preserving LCLs vitality. At the end, cells were stained for one hour with conjugated S9.6 antibody (PE/R-Phycoerythrin Conjugation Kit, Abcam) and analysed by flow cytometry (BD FACS Canto II). Logarithmic amplification was used for all channels and FACSDIVA was used for the analysis.

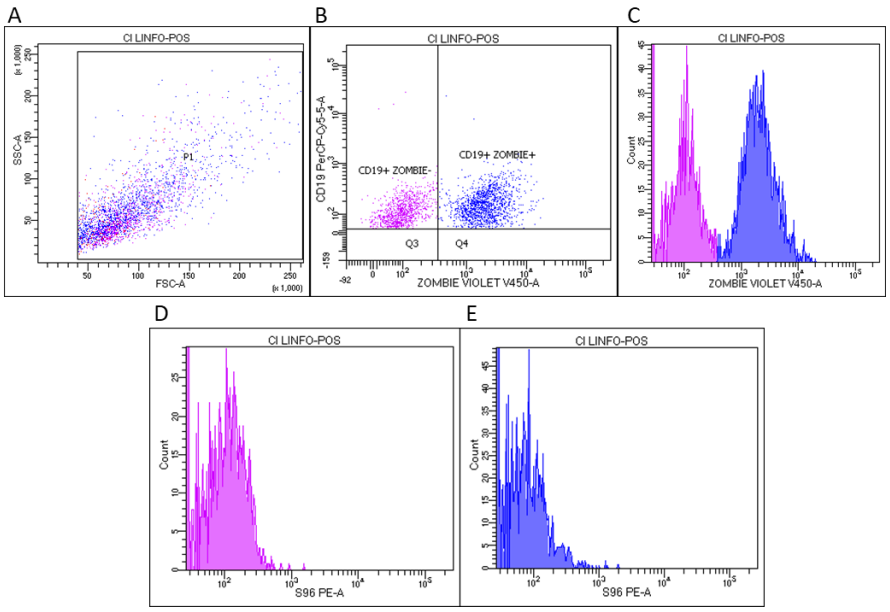

Figure 9. A) SSC-A and FSC-A channels allows the identification of Ctrl LCLs population in accordance with their cellular morphology. **B, C)** In this dot plot CD19+ Ctrl LCLs could be divided in Zombie+ (blue) and Zombie- (violet) with the correspondent pick in the beside histogram. **D, E)** For each individualized population is reported the histogram related to S9.6 mean fluorescence.

3.13 Flow Cytometry for analysis of transfection efficiency

$1*10^6$ cells from not transfected and transfected SH-SY5Y were permeabilized for 20 minutes at 4°C and washed in 1× BD Perm/Wash™ buffer, present in a Fixation/Permeabilization Solution Kit (Cat. No. 554714, BD Cytofix/Cytoperm™ Kit). Analysis was performed at flow cytometry (BD FACS Canto II), comparing mean intensity of the GFP signal detected using FITC channel.

3.14 Western Blotting

Cells were washed with ice-cold PBS. PBS was drained, and then lysis buffer (RIPA or Laemmli SDS PAGE buffer 1x) was added. The RIPA buffer (Radio Immuno Precipitation Assay buffer) was composed of 150 mM sodium chloride, 1.0% NP-40 or Triton X-100, 0.5% sodium deoxy-cholate, 0.1% SDS (sodium dodecyl sulphate) 50 mM Tris, pH 8.0.The lysate was incubated for 30 minutes at 4°C and then centrifuged in a microcentrifuge at 4°C. A BCA assay was performed (ThermoFisher). Bovine serum albumin (BSA) was used for the protein standard. The same amount of protein and volume was used for all the samples. They were diluted with 6x Laemmli buffer, boiled at 95°C, and loaded on the gel with 5 µL of Prosieve protein Ladder (Amersham). Proteins were fractionated by size on 10% polyacrylamide gels, transferred to a nitrocellulose membrane using a Trans-blot Turbo (BioRad, Italy) and blocked with blocking solution (5% non-fat dry milk in Tween-

20 Tris-Buffered Saline solution, TBS-T) for 1 hour. Membranes were incubated overnight with primary antibody in blocking solution at 4°C. Rabbit anti-TDP-43 antibody (1:100, Proteintech Europe) and mouse anti-GAPDH antibody (1:10.000, Abcam) as housekeeping gene were used. Immunoreactivity was detected using the donkey anti-rabbit or anti-mouse (GE Healthcare) secondary peroxidase-conjugated antibodies. Both incubations were performed using a blocking solution containing 5% of non-fat dry milk in TBS-T buffer (10 mM Tris-HCl, 100 mM NaCl, 0.1% Tween, pH 7.5). Bands were visualized using an enhanced chemiluminescence detection kit (ECL Advance, GE Healthcare, UK). For subsequent immunoreactions, primary and secondary antibodies were removed from the membrane with stripping solution incubated for 20 minutes (100 mM Glycine, 0.1% NP-40, 1% SDS pH 2.2). Densitometric analysis of the bands was performed using ImageJ software (National Institutes of Health, USA).

3.15 Immunoprecipitation coIP

Chromatin fraction (Chr fraction) and whole lysate fraction (Wl fraction) were extracted from Ctrl, sALS and TDP-43 mut LCLs ($2*10^7$ cells for each fraction) and immunoprecipitated with S9.6 antibody and mouse IgG conjugated using magnetic Protein G beads (Dynabeads™ Protein G for Immunoprecipitation, Thermofisher Scientific). The beads were previously washed and incubated O/N with sonication buffer (50 µM HEPES pH 7.8, 140 mM $NaCl_2$, 1mM EDTA, 1%Triton X-100, 0.1% Na-deoxycholate, 0.2% SDS, 1X PIC) +1%BSA at 4°C ON in agitation. For chromatin extraction, samples were resuspended in swelling buffer (25 µM HEPES pH 7.8, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5% NP-40, 1mM DTT, 1X PIC), then centrifuged at 5000 rpm for 5' and the pellet resuspended in sonication buffer and sonicated for 5cycles (30 ON/30 OFF). Both fractions were stained with previously blocked beads for 2h at 4°C in agitation, followed by incubation with 10 µg S9.6 antibody and respective murine IgG antibody (sc-2025, Santa Cruz) ON at the same conditions, recovering 10% of input from each sample. After washes, a part of the samples was eluted using elution buffer (50mM Tris pH 8.0, 1mM EDTA, 1% SDS) at 65°C for 10 minutes twice and resuspended in TE buffer for qPCR analysis. S9.6 immunoprecipitation was evaluated by qPCR in genomic sites enriched for R-loops' presence in these cells (NOP58 and ING3)[293]. Another part was boiled in 2X Laemmli Buffer at 80°C for 10 minutes and used for WB analysis. Input, S9.6 and IgG were loaded on a 10% SDS-PAGE and co-immunoprecipitation was evaluated by immunoblotting with rabbit **anti-TDP-43** (1:1000, Proteintech Europe), rabbit **anti-H3** (1:1000, Abcam), rabbit **anti-GAPDH** (1:10000, GeneTex).

3.16 DRIP-qPCR

DRIP was performed as described in the article released by Sanz L.A. and Chédin F. (2019)[295]. Briefly, cells were harvested from 100cm plaque and treated O/N at 37°C with TE buffer, 20% SDS and 20mg/ml Proteinase K. The next day, DNA was extracted carefully by phenol/chloroform extraction in phase lock tubes, precipitated with EtOH/sodium acetate, washed with 80% EtOH, and resuspended in TE. DNA was digested with a cocktail of restriction enzymes (Bsrg1, EcoR1, HindIII, SspI, XbaI) ON at 37˚C. For RNase H-treated samples, 10 µg of DNA was treated with RNase H for 5 hours at 37˚C. For both RNaseH treated and not RNaseH treated samples, 8 µg of DNA was bound with 20 µg of S9.6 antibody in 1X binding buffer (100 mM NaPO4 pH 7,1.4 M NaCl, 0.5% Triton X-100) overnight at 4˚C. After washing in 1X binding buffer, Protein A/G agarose beads (Thermo Fisher Scientific) were added for 2 hr at 4°C on rotation. Bound beads were washed twice in 1X binding buffer and elution was performed in elution buffer (50 mM Tris pH 8, 10 mM EDTA pH 8, 0.5% SDS, 20mg/ml Proteinase K) for 45 min at 55˚C. DNA was purified by standard methods described above.

qPCR was performed on 7500 Fast & 7500 Real-Time PCR System using SYBR-Green master mix (Biorad) following this program:

Cycle number	Duration	Temperature
1	30s	95°C
2-39	10s	95°C
	30s	60°C
40	Melting curve	

qPCR primers are the following:

APOE: forward sequencing primer 5'-CCCGACTGCGCTTCTCA-3', reverse sequencing primer 5'-GGGAGCCCTATAATTGGACAAGT-3';

RHOT2: forward sequencing primer 5'-GTGCCAGGCTGTATTGCTT-3', reverse sequencing primer 5'-GGGAAATGCAGACGTGTCAT-3':

RPL13: forward sequencing primer 5'-GCTTCCAGCACAGGACAGGTAT-3', reverse sequencing primer 5'-CACCCACTACCCGAGTTCAAG-3';

SNRPN: forward sequencing primer 5'-TGCCAGGAAGCCAAATGAGT-3', reverse sequencing primer 5'-TCCCTCTTGGCAACATCCA-3'.

4 Results

4.1. Characterization of R-loops in stable transfected GFP tagged TDP-43 WT and in TDP-43 p.A382T SH-SY5Y

4.1.1 Efficiency of stable transfection of a GFP tagged TDP-43 WT and a GFP tagged TDP-43 p.A382T vector in SH-SY5Y

For investigating the role of human ribonucleoprotein TDP-43 in the regulation of R-loops and in the control of DNA damage R-loops' dependent, we decided to use SH-SY5Y cell line since they have been widely used as a model for analysis of neuronal differentiation, metabolism, and function related to neurodegenerative and neuroadaptive processes, neurotoxicity, and neuroprotection[297]. These cells, which are a subclone of SK-N-SH isolated from a bone marrow biopsy of a female affected by neuroblastoma, have many biochemical and functional properties of neurons and can even be differentiated to a more mature neuron-like phenotype characterized by dopaminergic markers[298]. We perform SH-SY5Y stably transfected with a GFP tagged plasmid coding for TDP-43 p.A382T mutation compared to cells transfected with GFP tagged plasmid coding for TDP-43 WT. In particular, it was chosen p.A382T, a mutation on the C-terminal glycine-rich domain which is one of the most common presented in ALS patients affected by TDP-43 proteinopathies, that can influence TDP-43 functional role in different ways from its cytoplasmic localization, misfolding, aggregation, and cleavage to pathological posttranslational modifications such as phosphorylation and ubiquitination[299]. The efficiency of transfection was initially controlled by flow cytometry (FCM) for the selection of transfected SH-SY5Y, taking into account cellular vitality (Figure 1A and 1B).

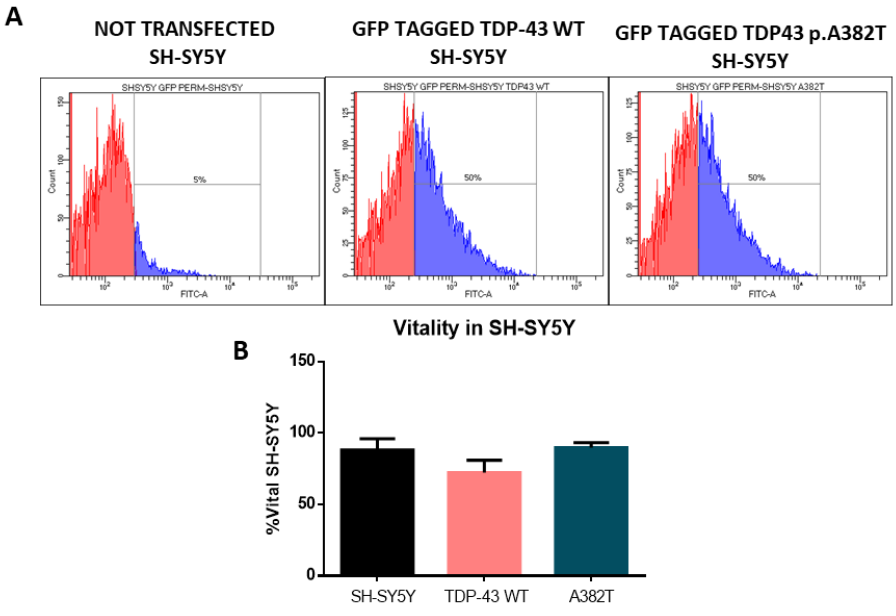

Figure 1. A) FCM analysis preformed in not transfected SH-SY5Y, TDP-43 WT GFP tagged SH-SY5Y and p.A382T GFP tagged SH-SY5Y showed the enrichment of 50% of GFP+ cellular population in transfected SH-SY5Y compared to not transfected ones, using the best conditions of selection **B)** The histogram showed the conservation in the percentage of vital SH-SY5Y at the same time analysed by FCM from a biological triplicate experiment.

Then, we performed immunofluorescence (IF) and western blot (WB) analysis in not transfected SH-SY5Y compared to the previously mentioned transfected SH-SY5Y, showed respectively in Figure 2A and 2B.

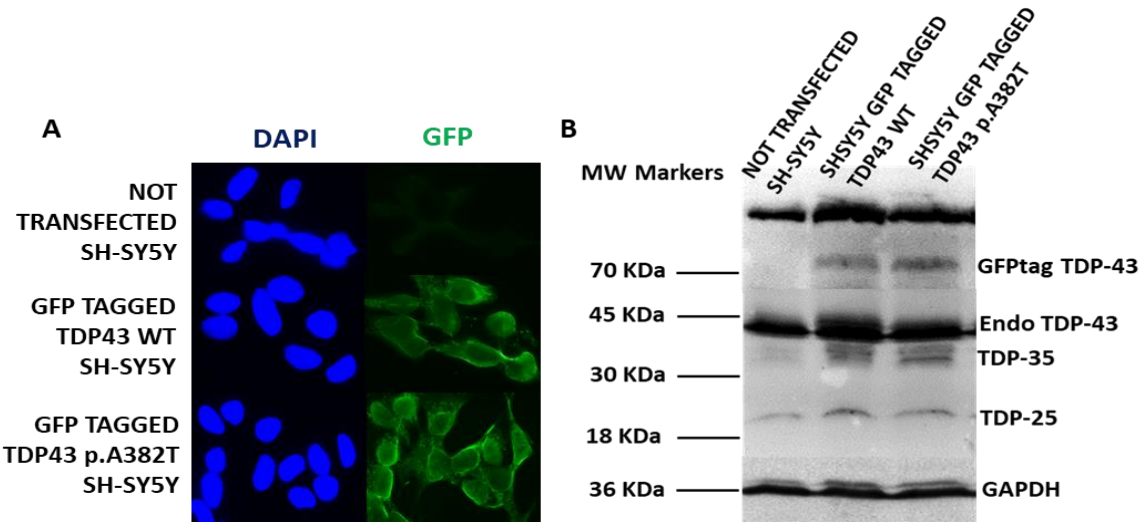

Figure 2. A) Detection of GFP signal by IF in not transfected SH-SY5Y, TDP-43 WT GFP tagged SH-SY5Y and p.A382T GFP tagged SH-SY5Y. Nuclei are stained in DAPI. **B)** Western blot analysis of not transfected SH-SY5Y, TDP-43 WT GFP tagged SH-SY5Y and p.A382T GFP tagged SH-SY5Y for TDP-43 GFP tagged TDP-43 (~70kDa). Signal from the endogeneous TDP-43 (43 kDa) is recorded in all samples. Two minor bands at 35 kDa and 25 kDa are observed in TDP-43 WT GFP tagged SH-SY5Y and in p.A382T GFP tagged SH-SY5Y.

From IF we could appreciate higher GFP signal in TDP-43 WT GFP tagged SH-SY5Y and in p.A382T GFP tagged SH-SY5Y, while a lower signal was detected in not transfected SH-SY5Y. WB analysis confirmed the transfection efficacy of TDP-43 WT GFP tagged SH-SY5Y and p.A382T GFP tagged SH-SY5Y by the presence of a major band of GFPtagTDP-43 (~70kDa), above signal from endogenous TDP-43 (43kDa). Immunoblot even showed that C-terminal fragments (CTFs) of TDP-43, localized with the bands at 35kDa and at 25kDa, were recovered in both transfected cell lines.

4.1.2 Mislocalization of TDP-43 in the cytoplasm of p.A382T TDP-43 SH-SY5Y

As previously reported, p.A382T is one of the most commonly occurring missense mutations, particularly present in Sardinia (Italy) where it is also associated with Parkinson disease (PD) and other degenerative parkinsonisms[156]. From a recent study, Mutihac R. et al. showed that both immortalized cells both primary neurons carrying p.A382T mutation were associated with a decrease of endoplasmic reticulum Ca^{2+} signalling and with an increased translocation of nuclear TDP-43 to cytoplasm after stress induction in endoplasmic reticulum (ER) causing elevated Bcl-2 levels and impaired Ca^{2+} release/storage[159]. In literature controversial results are reported about the mislocalization ability of p.A382T TDP-43, depending on the cellular model that is used. For example, Kabashi E. et al. showed that transient transfection of p.A382T TDP-43 in COS1 and Neuro2A cells induced its localization in the nuclei, but when expressed in motor neurons from dissociated spinal cord cultures and in zebrafish (Danio rerio) embryos were neurotoxic, concomitant with perinuclear localization and aggregation of TDP-43[125]. Nonaka T. et al. showed that the expression of TDP-43 C-terminal fragments (CTFs) fusion protein in SH-SY5Y cells results in the formation of abnormally phosphorylated and ubiquitinated inclusions that are similar to those found in FTLD-U and ALS; these inclusions were enhanced by the expression of GFP-tagged TDP-43 C-terminal fragments harboring pathogenic mutations such as p.A382T TDP-43[158]. Bossolasco P. et al. were able to differentiate motor neurons from iPSCs - derived from peripheral blood mononuclear cells (PBMC) of an ALS patient carrying a p.A382T TARDBP mutation, that anyway displayed a nuclear localization of TDP-43[300]. Orrù S. et al. didn't find any TDP-43 translocation in primary skin fibroblast cultures from an ALS patient with A382T mutation, but they showed that TDP-43 is still able to contribute to SGs formation through a regulatory effect on the G3BP1 core protein causing a significant reduction in the number of SGs per cell[164]. In order to investigate TDP-43 cellular localization, we performed IF in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y using an anti TDP-43 antibody able to recognize not only the native physiological form of the protein but even the CTFs as cleavage product of 25kDa and 35kDa in addition to the phosphorylated forms of TDP-43 (Figure 3). As reported by literature, these isoforms gain a highly neurotoxic function in the cell as they tend to form aggregates and inclusions in the cytoplasmic compartment sequestering the full-length form of TDP-43, which loses its physiological nuclear role[109].

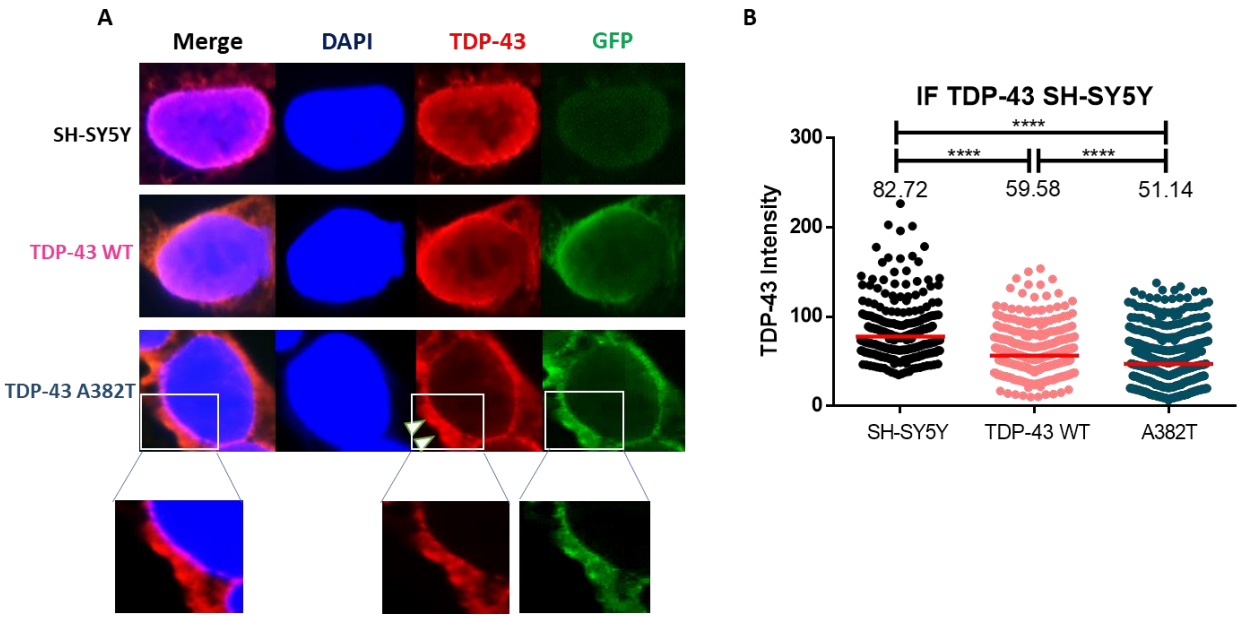

Figure 3. A) Detection of TDP-43 by IF in not transfected SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. Nuclei were stained with DAPI. An enlargement in GFP tagged p.A382T TDP-43 SH-SY5Y made more evident cytoplasmic TDP-43 inclusions. **B)** The slot blot shows the median of the TDP-43 signal intensity per nucleus and for each sample the mean of TDP-43 signal is reported. Around 300 cells from two independent experiments were considered. ****, $P < 0.0001$ (Mann-Whitney U test, two-tailed).

In GFP tagged TDP-43 WT SH-SY5Y, the ribonucleoprotein was localized in the perinuclear area compared to the not transfected SH-SY5Y, where its main localization was clearly visible inside the nuclear compartment. TDP-43 nuclear localization was significatively decreased in GFP tagged p.A382T TDP-43 SH-SY5Y compared both to not transfected SH-SY5Y and GFP tagged TDP-43 WT SH-SY5Y (****, $P < 0.0001$, Mann-Whitney U test, two-tailed), confirming that this mutation is linked to TDP-43 cytoplasmic myslocalization with formation of inclusions or aggregates (white arrows), as reported in literature[299,159]. In IF of transfected SH-SY5Y it can be noticed a colocalization between the signal detected by anti TDP-43 antibody in red and the signal detected by GFP tagged vectors in green, confirming the perinuclear distribution of wild type TDP-43 in GFP tagged TDP-43 WT SH-SY5Y and the translocation of mutated TDP-43 form in the cytoplasm of GFP tagged p.A382T TDP-43 SH-SY5Y. At this point, we wondered how the accumulation of both mutated and wild form of TDP-43 in the cytoplasmic compartment of p.A382T GFP tagged SH-SY5Y could affect TDP-43 nuclear physiological function investigating its role in the control and in the regulation of cotranscriptional R-loops.

4.1.3 R-loops accumulation and activation of DDR in stable transfected p.A382T TDP-43 SH-SY5Y

The characterization of R-loops' presence in SH-SY5Y neuroblastoma cell lines was initially investigated by IF using S9.6 antibody, which is able to recognize with an affinity of 0.6 nM RNA-DNA hybrids[301,191]. We performed IF preceded by methanol fixation on basal SH-SY5Y cells, on stable transfected GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. For discrimination of genomic R-loops, S9.6 signal in the nucleoplasm was measured as the total nuclear signal subtracting the nucleolar contribution[302, 192]. The nucleus and nucleolus were detected with DAPI and anti-nucleolin antibodies, respectively (Figure 4). Nucleolin is a nucleolar phosphoprotein that is highly expressed in proliferating cells, involved not only in ribosome biogenesis[303], but also in chromatin remodeling[304] (30, 31), transcription[305] (32–36), and apoptosis[306] (39) and G-quadruplex binding[307,308] (36–38). Wang I.X. et al. proved nucleolin ability to recognize nuclear RNA-DNA hybrids by pull-down assays followed by mass spectrometry, thanks to its K homology (KH) and helicase domains[309]. Haeusler A.R. et al. reported that nucleolin was able to specifically bind C9orf72 HRE G-quadruplexes containing R-loops in iPS cells derived from C9orf72 ALS patients[246], leading to RNA Pol II pausing, aborted transcripts formation and nucleolar stress due to nucleolin sequester[310]. As showed in figure 4, S9.6 immunostaining in SH-SY5Y cells confirmed colocalization of the nuclear RNA-DNA hybrids with nucleolin as protein attracted by R loops. All the samples were also transiently transfected with a mock control vector expressing GFP and with a plasmid overexpressing RNaseH1 enzyme, used as negative control for R-loops presence thanks to its ability to specifically degrade RNA-DNA hybrids. The effects of different ribonucleoleolytic treatments on R-loops formation have been widely described in literature. For example, Halász et al. tested the action of RNaseA enzyme on RNA-DNA hybrids depending on salt concentration[294] and Hartono S.R. et al. studied RNaseIII effect on S9.6 immunoprecipitation in DRIP-qPCR experiment[311]. Apart from this, RNaseH1 is the only accepted endonuclease able to pecifically hydrolyzes the phosphodiester bonds of RNA which is hybridized to DNA present in the RNA-DNA hybrids and so the one used for proving the real detection of R-loops[192].

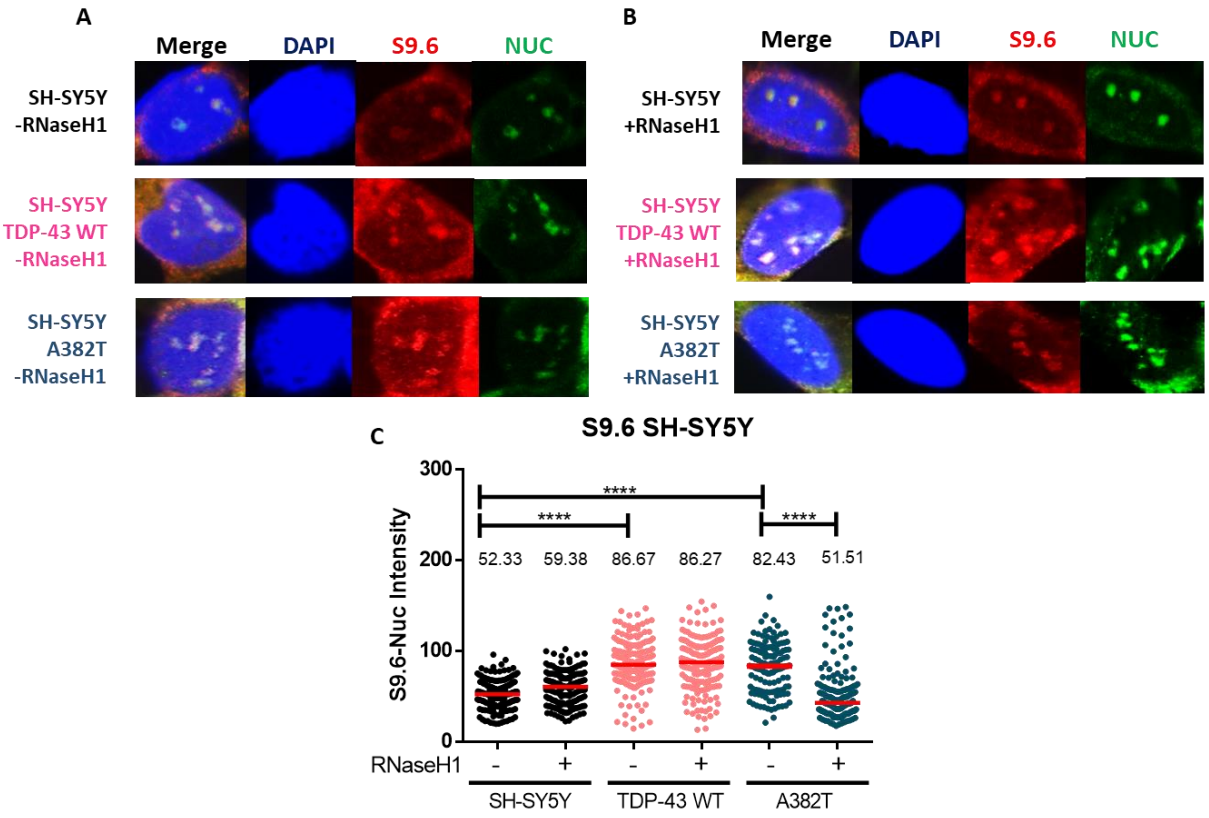

Figure 4. A and B) SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y were incubated with anti S9.6 antibody and anti-nucleolin antibody. Nuclei were stained with DAPI. Here samples without RNaseH1 overexpression are indicated with -RNaseH1 and samples with RNaseH1 overexpression are indicated with +RNaseH1. **C)** The slot blot shows the median of the S9.6 integrated raw intensity per nucleus after nucleolar signal removal. For each sample the median of S9.6 signal is reported. Around 300 cells from three independent experiments were considered. ****, $P < 0.0001$ (Mann-Whitney U test, two-tailed).

From three biological experiments, an increase of nucleolar S9.6 intensity was detected both in GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y compared to SH-SY5Y. However, RNaseH1 overexpression decreased S9.6 signal only in GFP tagged p.A382T TDP-43 SH-SY5Y cells, suggesting that only the mutation could induce S9.6 signal depending on R-loops presence (Figure 4B). At this point we wondered about the nature of the hybrids accumulated in the nucleus of GFP tagged TDP-43 WT SH-SY5Y. For this reason, we decided to repeat the same IF adding an extra treatment with RNaseIII enzyme, for its ability to degrade specifically dsRNAs considering that S9.6 antibody can also recognize dsRNAs with a binding affinity of 2.7 nM[312], as reported by Phillips et al. (Figure 5).

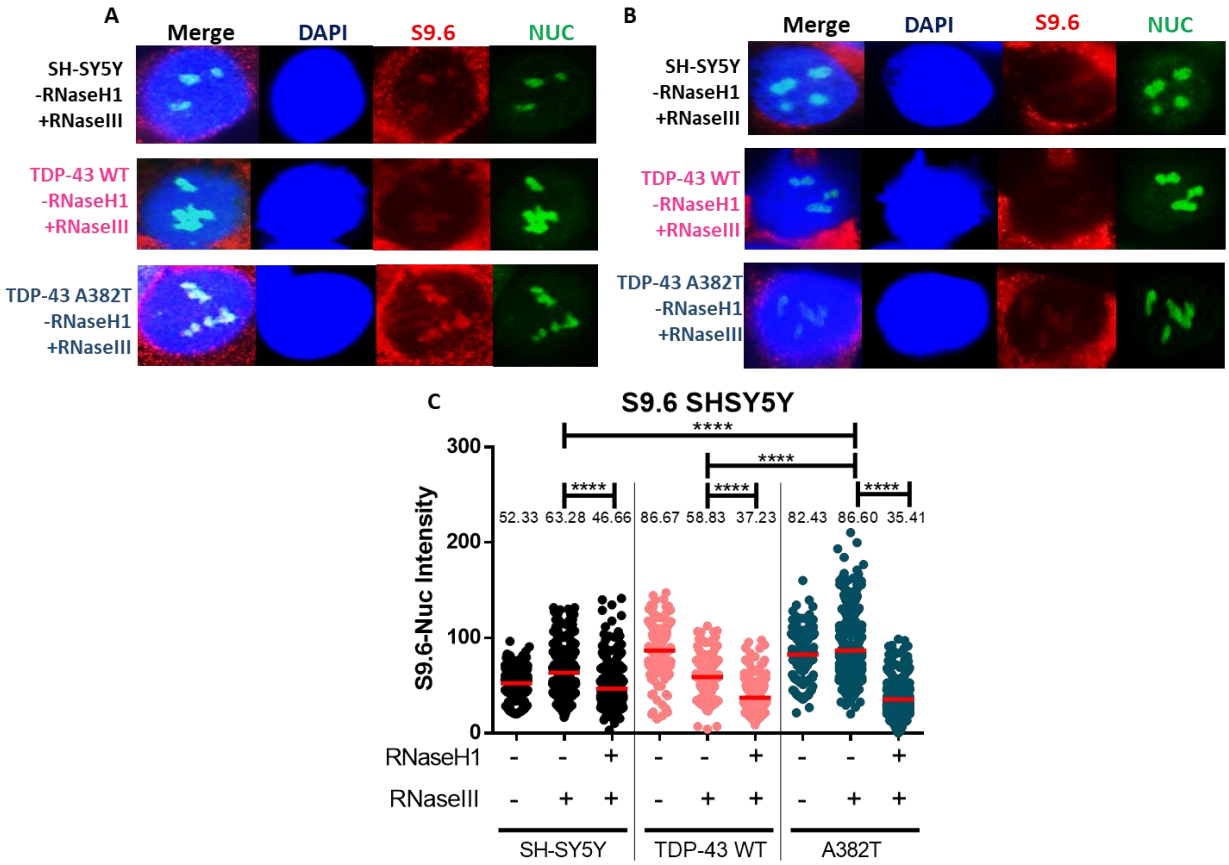

Figure 5. A and B) SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y were transiently transfected with a mock control plasmid and with RNaseH1 overexpressing plasmid, in addition with RNaseIII treatment were stained with anti S9.6 antibody and anti-nucleolin antibody. Nuclei were stained with DAPI. Here samples without RNaseH1 overexpression are indicated with -RNaseH1, samples with RNaseH1 overexpression are indicated with +RNaseH1 and samples with RNaseIII treatment are indicated with +RNaseIII. **C)** The slot blot shows the median of the S9.6 integrated raw intensity per nucleus after nucleolar signal removal and for each sample the mean of S9.6 signal is reported. Around 300 cells from three independent experiments were considered. ****, $P < 0.0001$ (Mann-Whitney U test, two-tailed).

RNaseIII treatment showed a several behaviors of analysed SH-SY5Y reflecting differential ability to process dsRNAs. In transient mock transfected SH-SY5Y RNaseIII action showed an increase in R-loops detection, which is confirmed by S9.6 signal decrease in condition of RNaseH1 overexpression. In GFP tagged p.A382T TDP-43 SH-SY5Y RNaseIII enzyme didn't modify the detected R-loops signal compared to untreated GFP tagged p.A382T TDP-43 SH-SY5Y, suggesting a connection between nuclear accumulation of genomic R-loops and p.A382T mutated TDP-43 impairment in their resolution. Moreover, in presence of RNaseIII an increase in S9.6 signal was detected in GFP tagged p.A382T TDP-43 SH-SY5Y compared both to transient mock transfected

SH-SY5Y and both to GFP tagged TDP-43 WT SH-SY5Y, showing that TDP-43 overexpression can induce the accumulation of different kind of hybrids, and not only of genomic R-loops. Possibly, the alteration of localization of TDP-43 registered in GFP tagged TDP43 WT vector compared to not transfected SH-SY5Y (Figure 3) could explain the increase in nuclear dsRNAs found in this sample. The decrease in S9.6 intensity observed in all the samples overexpressing RNaseH1 enzyme confirmed the dependence of detected signal from R-loops' presence and showed that the combined action of RNaseIII with RNaseH1 could be responsible of a greater clearance of observed genomic hybrids.

Once demonstrated R-loops' accumulation in GFP tagged p.A382T TDP-43 SH-SY5Y, we investigated the functional impact of RNA-DNA hybrids enrichment in the same cell line by detection of γH2AX foci as indirect measurement of DSBs, as it is known that R-loops accumulation is linked to DSBs' induction and genomic instability[190]. To determine the potential role of R-loops in TDP-43 mediated DSBs, as negative control it was investigated if RNase H1 overexpression was able to prevent γH2AX foci formation (Figure 6). In this case, cells were fixed with paraformaldehyde after permeabilization step with the detergent Triton X-100, allowing the stain of overexpressed RNase H1 only in nucleus and nucleoli, which is fundamental for R-loops elimination.

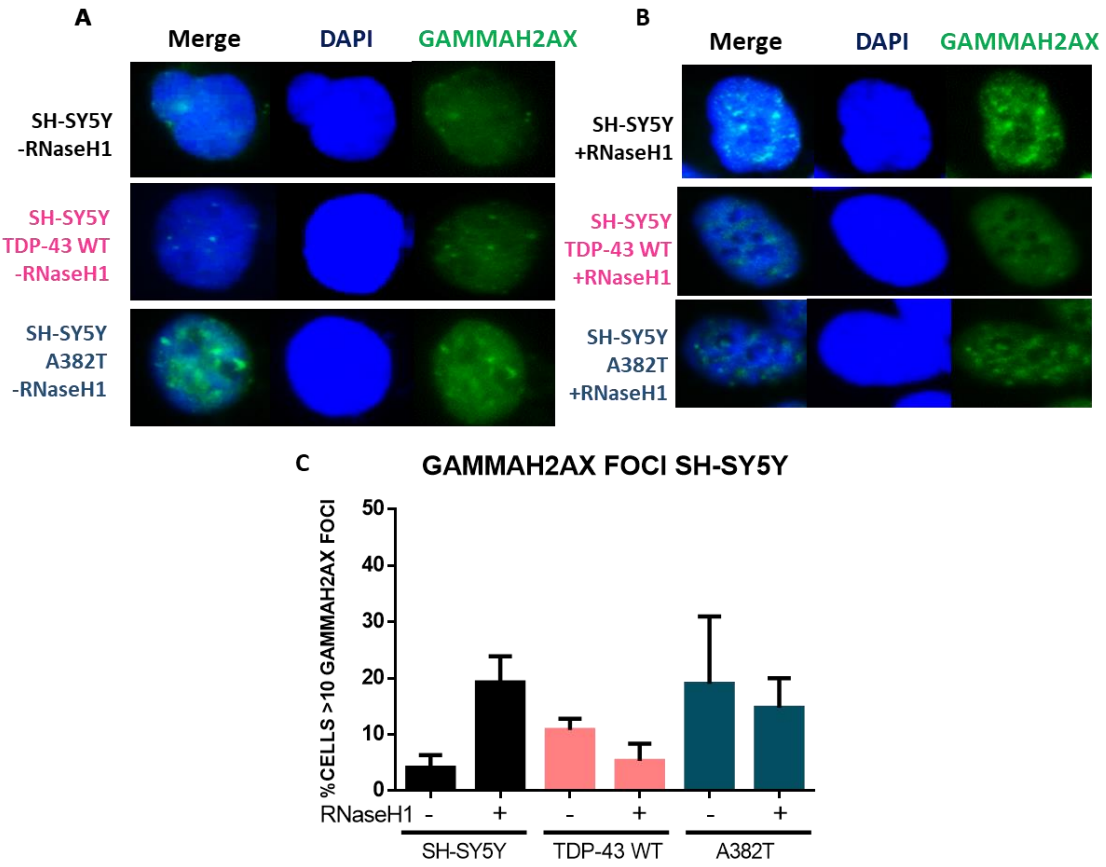

Figure 6. A and B) Detection of γH2AX foci by IF in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. Cells were also transiently transfected with a mock control plasmid and with RNaseH1 overexpressing plasmid. Nuclei were stained with DAPI. Here samples without RNaseH1 overexpression are indicated with -RNaseH1 and samples with RNaseH1 overexpression are indicated with +RNaseH1. **C)** The graph shows the quantification of the relative amount of cells in percentage containing >10 γH2AX foci in each case. More than 100 cells were counted in each of the three experiments. Data represent mean ± SEM from three independent experiments.

By quantifying the number of cells with more than 10 γH2AX foci, an accumulation in foci count was detected in GFP tagged p.A382T TDP-43 SH-SY5Y compared to basal SH-SY5Y and GFP TDP-43 WT SH-SY5Y (Figure 6A). These data were not significative due to high variability recorded in TDP-43 mutated SH-SY5Y. Importantly, γH2AX foci formation was reduced by RNase H1 overexpression, confirming that their formation in p.A382T GFP tagged SH-SY5Y and TDP-43 WT GFP tagged SH-SY5Y was R loops-dependent. The increase of foci count observed in transiently mock transfected SH-SY5Y overexpressing RNaseH1 may be related to the fact that transient transfection could affect in some way the physiological DNA integrity in basal cells.

It is very well known that the Fanconi Anemia pathway is involved in the interstrand crosslinks (ICLs) repair pathway[313], and that FANCD2 associated to BRCA1 plays a relevant role in removing R-loops via the replication of R-loop-containing regions[301]. Moreover, García-Rubio M.L. et al. using FANCD2-/- cells demonstrated the involvement of the FA pathway to remove R loops or R loop-associated DNA damage, including presumably RNA-DNA ICLs[292]. Recently it has been proposed that R-loops lead to formation and to accumulation of a chromatin condensation marker H3S10P, which in turn could cause replication fork stalling, transcription or replication collisions, and, ultimately, DSBs in the newly synthesized DNA[226]. For this reason, we also investigated the accumulation of FANCD2 and H3S10P foci in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. All the samples were also transiently transfected with a control plasmid and with a plasmid overexpressing RNaseH1 enzyme, (Figure 7).

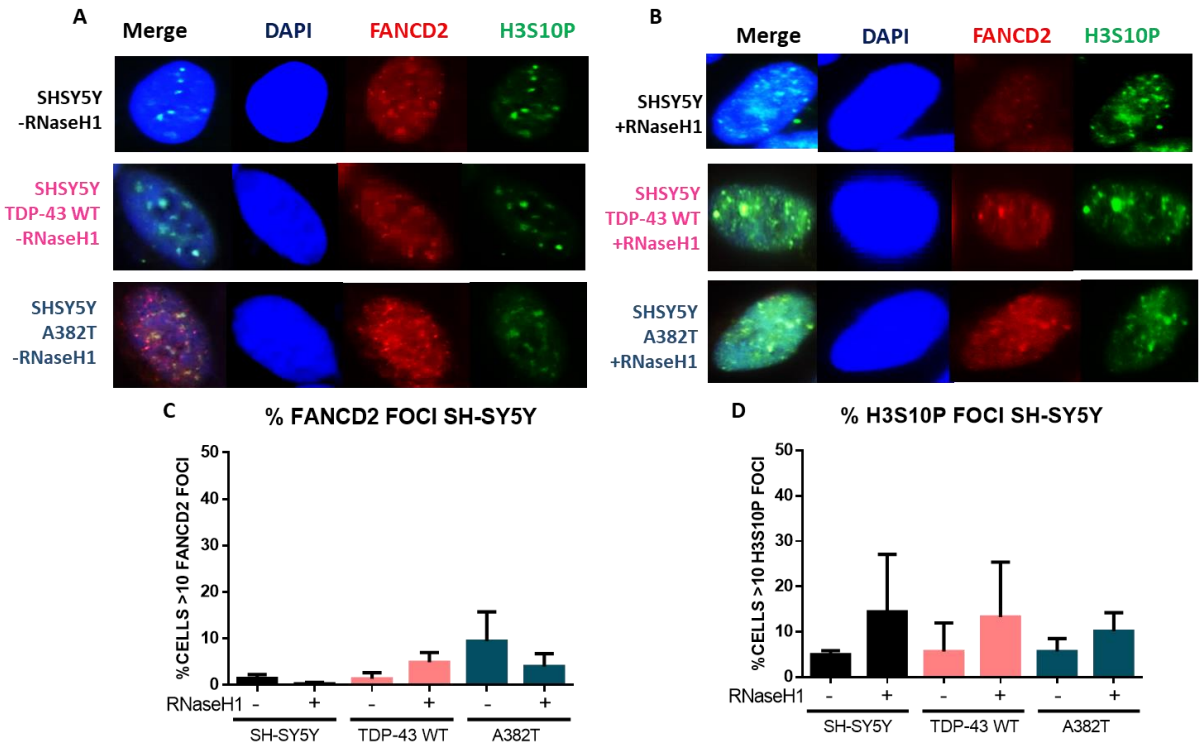

Figure 7. A and B) Detection of FANCD2 and H3S10P foci in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y by IF. Cells were also transiently mock transfected with a control plasmid and with RNaseH1 overexpressing plasmid. Nuclei were stained with DAPI. Here samples without RNaseH1 overexpression are indicated with -RNaseH1 and samples with RNaseH1 overexpression are indicated with +RNaseH1. **C and D)** The two histograms show the quantification of the relative amount of cells in percentage containing >10 FANCD2 and H3S10P foci in each case. More than 100 cells were counted in each of the three experiments. Data represent mean ± SEM from three independent experiments.

In the IF experiments we observed a higher percentage of cells with more than 10 FANCD2 foci in GFP tagged p.A382T TDP-43 SH-SY5Y compared to transiently mock transfected SH-SY5Y and GFP tagged TDP-43 WT SH-SY5Y. FANCD2 foci count accumulation was R-loops dependent only in GFP tagged p.A382T TDP-43 SH-SY5Y since the decrease in FANCD2 foci count was only observed with RNaseH1 overexpression. Instead, the percentage of cells with more than 10 H3S10P foci remained unvariate in the three samples and no correlation with R-loops was found. The model of correlation between R-loops and H3S10P proposed by Castellano-Pozo et al. has been proved in yeast, nematodes, in R-loop-accumulating THOC1 and senataxin-depleted HeLa cells, with a particular enrichment in G1-selected cells[226]. It would be interesting to further investigate the mechanisms of H3S10P accumulation linked to R-loops regulation, blocking SH-SY5Y in a precise phase of the cell cycle.

4.1.4 DRIP-qPCR shows R-loops accumulation in RPL13 gene of p.A382T TDP-43 SH-SY5Y

TDP-43 role in prevention or resolution of R-loops' accumulation in SH-SY5Y neuroblastoma cells was also investigated through DRIP-qPCR, which is based on RNA-DNA hybrids immunoprecipitation recognized by S9.6 antibody. We decided to investigate the effect of TDP-43 overexpression and of p.A382T TDP-43 mutation on R-loops presence choosing three human genes APOE, RPL13A and RHOT2 as they are indicated in literature as target regions prone to form R loops[301,314,315], while SNRPN gene was used as negative control for lacking accumulation of R-loops[314,315] (Figure 8). Following DRIP-qPCR protocol, negative control for R-loops presence in each locus was represented by treatment with RNaseH1 enzyme, produced in recombinant Escherichia Coli and purified by proprietary chromatographic techniques.

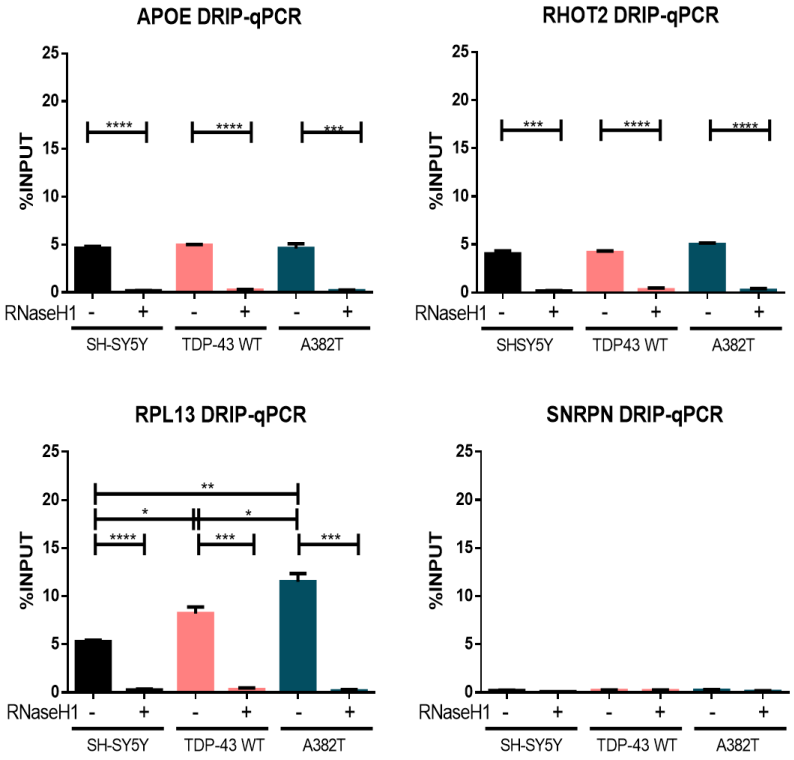

Figure 8. DRIP-qPCR using the anti RNA-DNA hybrids S9.6 monoclonal antibody at APOE, RPL13A, RHOT2 genes and SNRPN as negative control region are shown in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. Pre-immunoprecipitated samples were untreated (-) or treated (+) with RNaseH1, as indicated. Signal values of RNA-DNA hybrids immunoprecipitated in each region were normalized to input values (10% of each sample). Data represent mean ± SEM from three independent experiments. *, P < 0,0393; **, P < 0,0021; *** P < 0,0003; ****, P < 0,0001 (Upaired t test, two-tailed)

From the results of DRIP-qPCR we obtained a significative R-loops accumulation in RPL13 gene in GFP tagged p.A382T TDP-43 SH-SY5Y compared to not transfected SH-SY5Y and GFP tagged TDP-43 WT SH-SY5Y. Importantly, when the samples were treated with RNaseH1, that digested the RNA moiety of RNA-DNA hybrids, the levels of R-loops dramatically decreased, confirming that indeed the signal detected was specific for RNA-DNA hybrids. A very substantial decrease of signal was also recorded in SNRPN locus represented as negative control gene. The amount of R-loop presence was measured as a function of input DNA, that for each sample was 10% of the entire amount. RPL13 gene encodes for a ribosomal protein that is involved in pathways of viral mRNA translation and of rRNA processing and that is implicated in TDP-43 cytoplasmic interactome cluster fundamental for RNA metabolism regulation[184]. Recently, Tank E.M. et al. used Bru-seq and BruChase-seq genome-wide analysis for showing the connection of TDP-43 deposition in ALS patients' derived iPSCs with destabilization in ribosomal transcripts and with increase of cytoplasmic ribosome proteins including RPL13 as compensatory response for preservation in cell's capacity to synthesize proteins. They also proved that transcripts demonstrating altered stability in ALS derived iPSCs were highly enriched in motifs recognized by RBPs that are involved in the formation of cytoplasmic stress granules[316].

For ensuring R-loop sequences specific precipitation, Zhang et al. suggested to include an RNaseA preclearing step prior to the immunoprecipitation (IP) reaction[317], but König F. et al. proved that this treatment could result in an underestimation of R-loop regions throughout the genome, especially if short R-loop stretches are studied[318]. In accordance with this, we performed a DRIP-qPCR experiment in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y adding a RNaseIII treatment for degradation of dsRNAs (Figure 9).

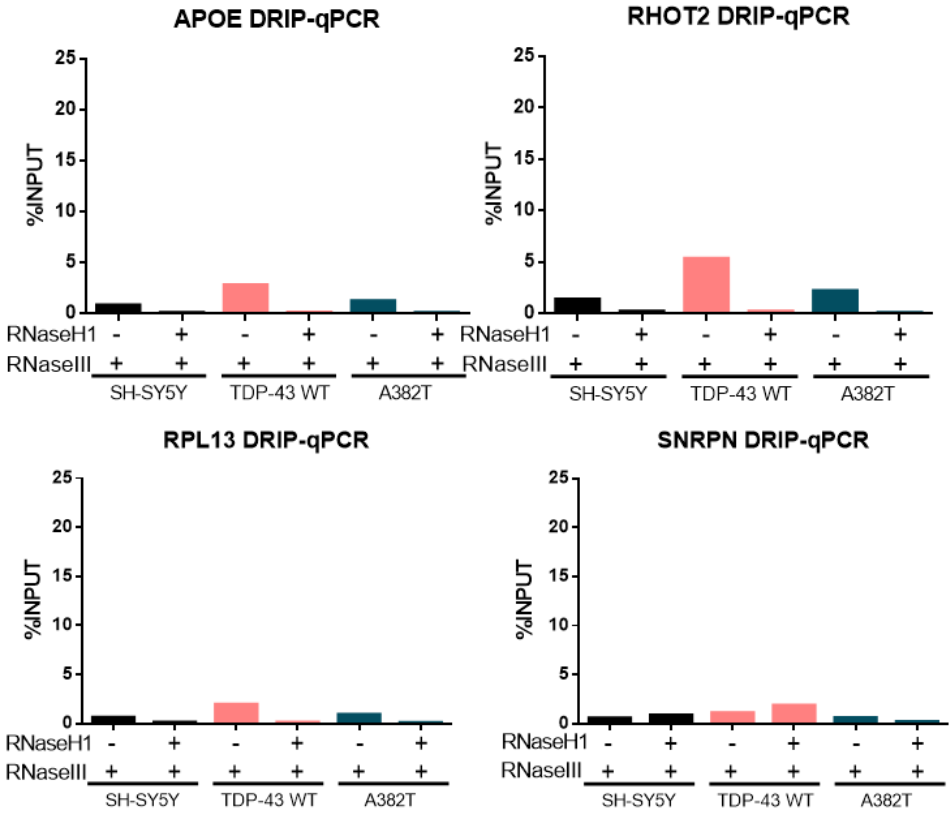

Figure 9. DRIP-qPCR using the anti RNA-DNA hybrids monoclonal antibody, S9.6 at APOE, RPL13A, RHOT2, SNRPN genes are shown in basal SH-SY5Y, GFP tagged TDP-43 WT SH-SY5Y and GFP tagged p.A382T TDP-43 SH-SY5Y. Pre-immunoprecipitated samples were untreated (-) or treated (+) with RNaseH1 with an additional RNaseIII treatment, as indicated. Signal values of RNA-DNA hybrids immunoprecipitated in each region were normalized to input values (10% of each sample). Data represent mean from one experiment.

With this experiment we proved that the addition of RNaseIII treatment to DRIP-qPCR decreased R-loops related signal to a very low level in all the samples and did not discriminate a cell line from the other with an action comparable to RNaseA treatment described by the work from König F. et al.[318].

4.2 Characterization of R-loops in siTDP-43 HeLa cells shows conservation of TDP-43 role in R-loops regulation

4.2.1 Accumulation of R-loops and activation of DDR in siTDP-43 HeLa cells

Considering the obtained results from *in vitro* SH-SY5Y cellular model, we demonstrated that TDP-43 function linked to regulation of genomic R-loops was affected due to the loss of its nuclear physiological role. P.A382T TDP-43 mutation impaired protein conformational and aggregative property inducing a mislocalization of both mutated and WT form in cellular cytoplasm (Figure 3) as partially suggested in the literature[159], leading to the formation of inclusions with CTFs and causing an accumulation of DNA damage R-loops dependent (Figure 6). For reproducing TDP-43 nuclear depletion described in GFP tagged p.A382T TDP-43 SH-SY5Y (Figure 3), we decided to silence TDP-43 in HeLa cells and to perform the same analysis based on R-loops characterization to check if the role of TDP-43 in hybrids' regulation was conserved in the two cell lines. Ayala Y.M. proved an important regulatory role of TDP-43 in essential metabolic processes also in siTDP-43 HeLa showing that the loss of TDP-43 resulted in dysmorphic nuclear shape, misregulation of the cell cycle, apoptosis, increase in cyclin-dependent kinase 6 (Cdk6) protein and transcript levels[319]. In the future we would be interested to further investigate this aspect also in silenced TDP-43 SH-SY5Y.

The efficiency of silencing in HeLa cells was checked controlling the elimination of TDP-43 transcript using Real Time PCR, that showed 80% decrease of relative TDP-43 mRNA in SiTDP-43 HeLa compared to the control both without both with RNaseH1 overexpression (Figure 10).

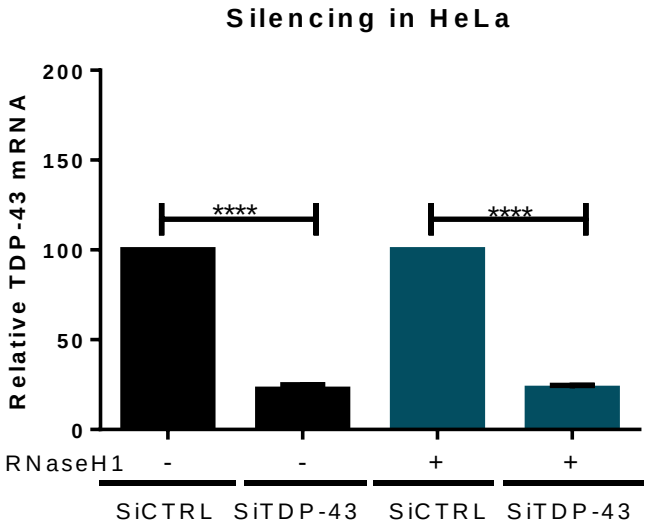

Figure 10. Relative TDP-43 mRNA was analysed by qPCR in silenced HeLa compared to the housekeeping control transcript HPRT obtaining the reduction of 80% of TDP-43 mRNA in siTDP-

43 HeLa compared to siCTRL HeLa both without both with RNaseH1 overexpression. Data represent mean ± SEM from three independent experiments. ****, P < 0,0001 (Upaired t test, two-tailed)

IF experiments for analyzing R-loops accumulation was performed using S9.6 antibody on HeLa transfected with a SiControl (SiCTRL) and a SiTDP-43 (SiTDP-43) oligonucleotides, after methanol fixation. In this case SiCTRL is a negative control of silencing process because it presented a sequence that did not target any gene product. For discrimination of genomic R-loops, S9.6 signal in the nucleoplasm was measured as the total nuclear signal subtracting the nucleolar contribution as previously described[292,302]. (Figure 11). As negative control for R-loops presence, cells were transfected with a control plasmid and a plasmid overexpressing RNaseH1 enzyme.

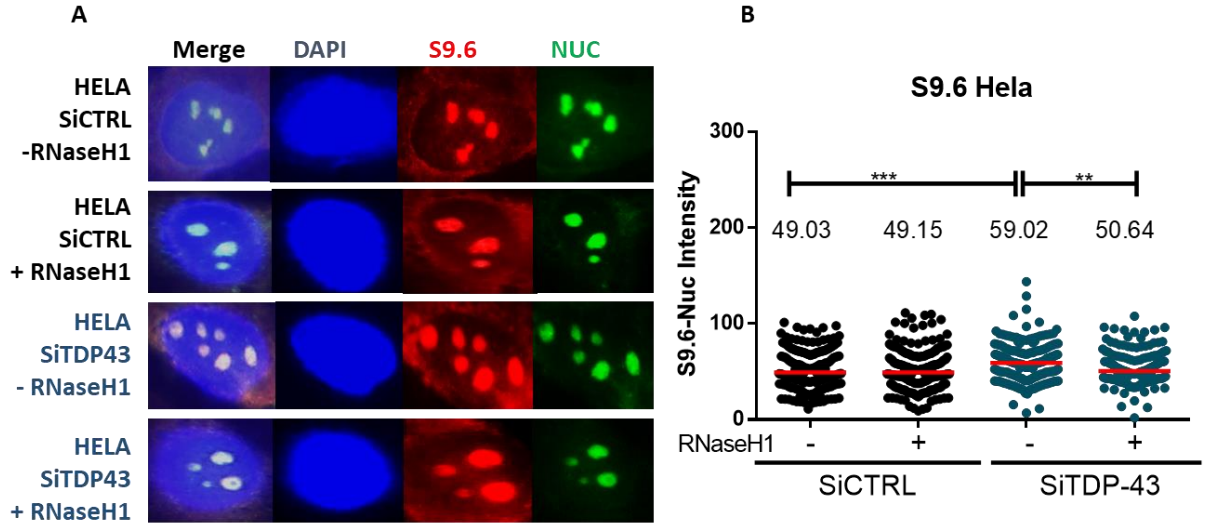

Figure 11. A) SiCTRL and SiTDP-43 HeLa were incubated with antiS9.6 antibody and anti-nucleolin antibody. They were also transiently transfected with a control plasmid (- RNaseH1) and with RNaseH1 overexpressing plasmid (+ RNaseH1), as indicated. Nuclei were stained with DAPI. **B)** The slot blot shows the median of the S9.6 signal intensity per nucleus after nucleolar signal removal and for each sample the mean of S9.6 signal is also reported. Around 300 cells from three independent experiments were considered. ***, P < 0,0002; **, P < 0,0012 (Mann-Whitney U test, two-tailed).

A significative slight induction of R-loops presence was observed in SiTDP-43 HeLa in comparison to SiCTRL HeLa and the decrease after RNaseH1 overexpression recorded in the same cell line confirmed that the recorded signal is R-loops dependent. Then, we performed an IF on SiCTRL and SiTDP-43 HeLa to investigate DDR activation connected to DNA DSBs accumulation using an anti-γH2AX antibody (Figure 12).

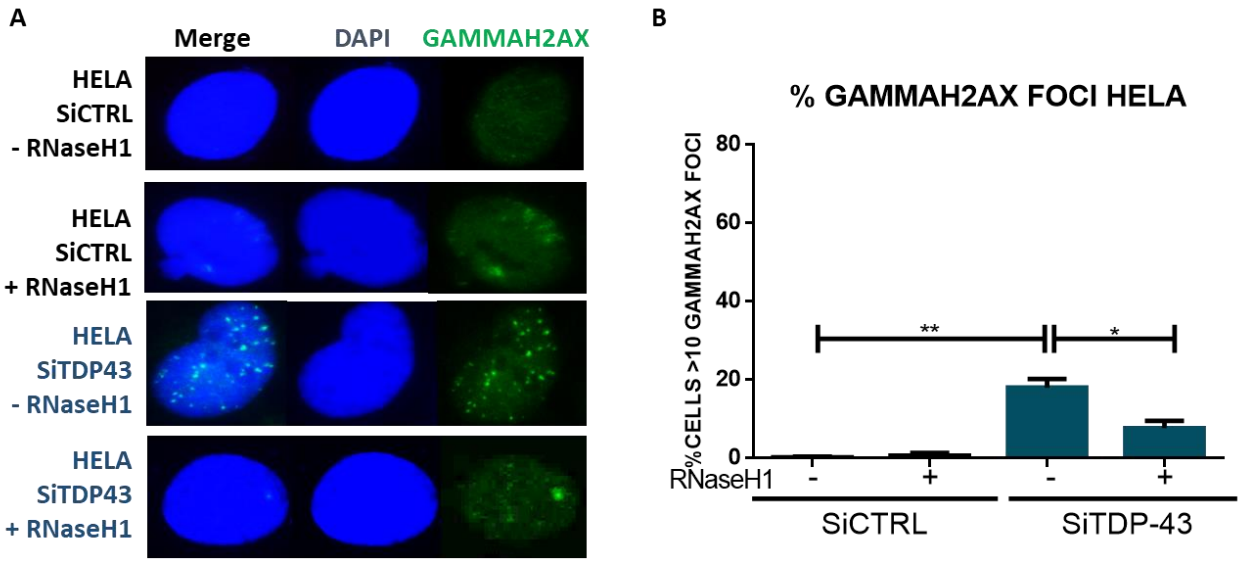

Figure 12. A) Detection of γH2AX foci by IF in SiCTRL and SiTDP-43 HeLa. Cells were also transiently transfected with a control plasmid (- RNaseH1) and with RNaseH1 overexpressing plasmid (+ RNaseH1), as indicated. Nuclei were stained with DAPI. **B)** The histogram shows the quantification of the relative amount of cells in percentage containing >10 γH2AX foci in each case. More than 100 cells were counted in each of the three experiments. Data represent mean ± SEM from three independent experiment. **, P = 0,0012 and *, P = 0,0228 (Upaired t test, two-tailed).

γH2AX IF showed a significative increase in the percentage of cells that present more than 10 γH2AX foci in the nucleus of SiTDP-43 HeLa compared to SiCTRL HeLa. Moreover, the decrease of γH2AX foci after RNaseH1 overexpression recorded in SiTDP-43 HeLa confirmed that the recorded signal is R-loops dependent. These data are also confirmed by Hill et al., who proved that TDP-43 colocalizes with active RNA polymerase II and with the DNA damage repair protein, BRCA1, and its depletion leads to increased R loop-associated DNA damage[183].

At the end, we investigated the accumulation of FANCD2 foci and H3S10P foci in -RNaseH1 and +RNaseH1 SiCTRL and SiTDP-43 HeLa because, as previously reported, these factors are closely connected with R-loops regulation[226, 292] (Figure 13).

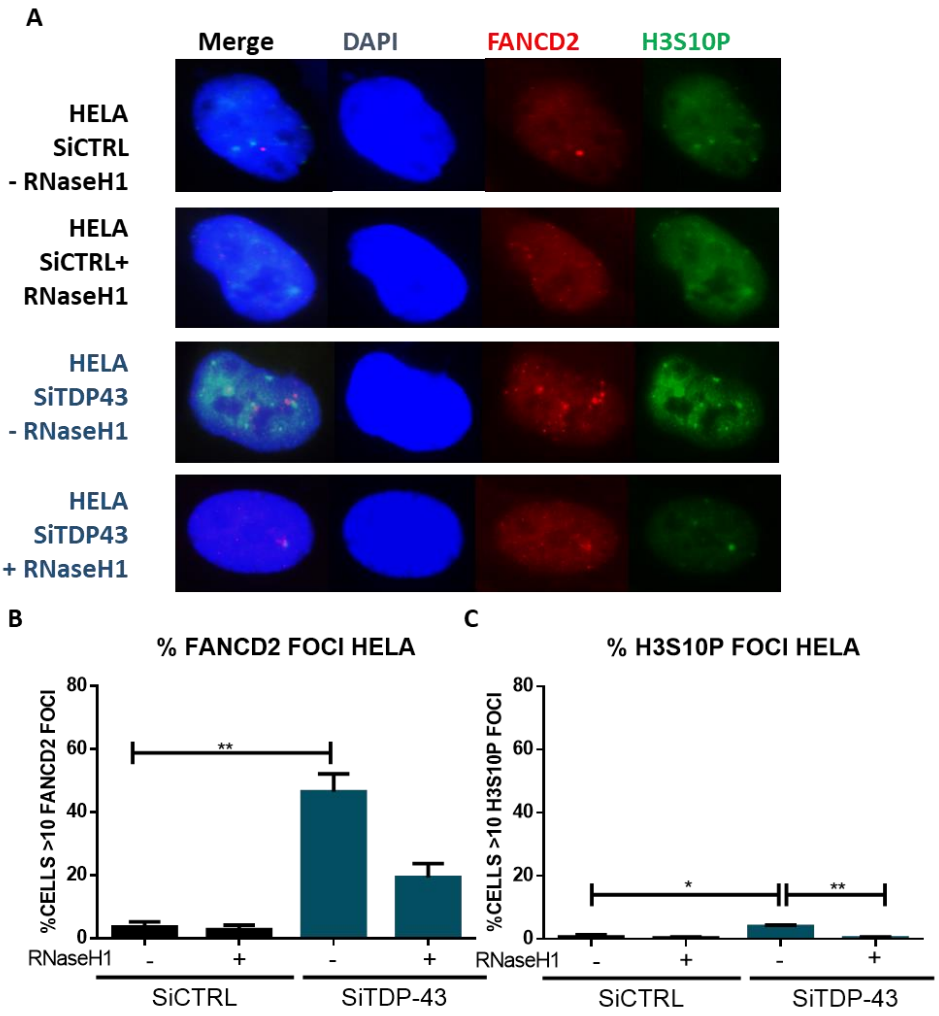

Figure 13. A) Detection of FANCD2 and H3S10P foci in SiCTRL and SiTDP-43 HeLa. Cells were also transiently transfected with a control plasmid (- RNaseH1) and with RNaseH1 overexpressing plasmid (+ RNaseH1), as indicated by IF. Nuclei were stained with DAPI. **B and C)** The two histograms show the quantification of the relative amount of cells in percentage containing >10 FANCD2 and H3S10P foci in each case. More than 100 cells were counted in each of the three experiments. Data represent mean ± SEM from three independent experiments. *, P< 0,037**, P = 0,00175 (Upaired t test, two-tailed).

IF showed a significative increase in the percentage of cells that present more than 10 FANCD2 foci and H3S10P foci in the nucleus of SiTDP-43 HeLa compared to SiCTRL HeLa. Moreover, the decrease of FANCD2 foci and H3S10P foci, after RNaseH1 overexpression, registered in siTDP-43 HeLa confirmed that the recorded signal is R-loops dependent. We obtained a higher activation of FANCD2 pathway in siTDP-43 HeLa cells compared to GFP tagged p.A382T TDP-43 SH-SY5Y (Figure 7B), suggesting that in cervical cancer cell line this is the main mechanism involved in resolution of genomic R-loops.

4.2.2 DRIP-qPCR shows R-loops accumulation in genomic regions of siTDP-43 HeLa cells

DRIP-qPCR was also performed on SiCTRL HeLa and siTDP-43 HeLa for confirming the previous obtained results based on the involvement of TDP-43 role in prevention or resolution of R-loops' accumulation in human cells using S9.6 antibody. R-Loops enrichment was evaluated on the genes described above in paragraph 1.4 as reported in the literature[190,312] (Figure 14).

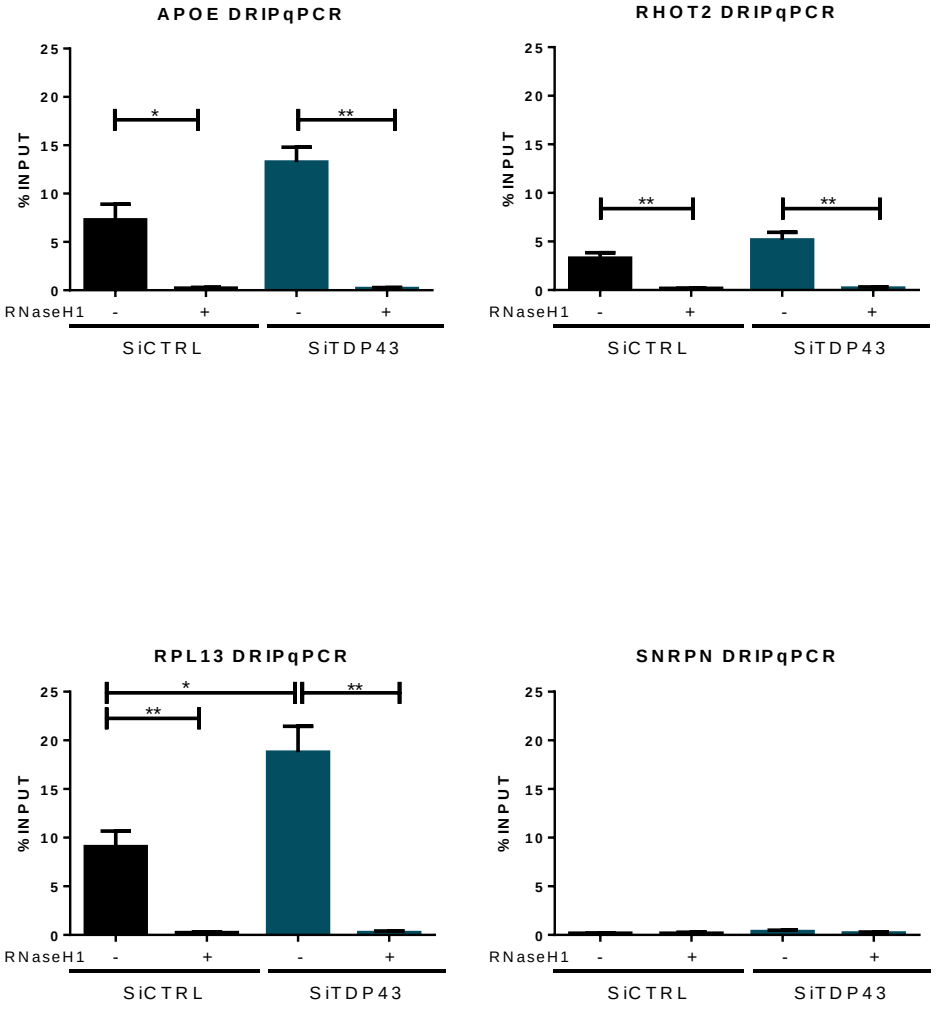

Figure 14. DRIP-qPCR using the anti RNA-DNA hybrids S9.6 monoclonal antibody, at APOE, RPL13A, RHOT2 genes and SNRPN as negative control region are shown in SiCTRL HeLa and siTDP-43 HeLa. Pre-immunoprecipitated samples were untreated (-) or treated (+) with RNaseH1, as indicated. Signal values of RNA-DNA hybrids immunoprecipitated in each region were normalized to input values (10% of each sample). Data represent mean ± SEM from three independent experiments. *, P <0,0353, **, P < 0,0005 (Upaired t test, two-tailed)

From results obtained using DRIP-qPCR, R-loops accumulation was detected in the three genes used as positive control for their presence with a significative increase in particular in RPL13 gene of siTDP-43 HeLa compared to SiCTRL HeLa. Importantly, when the samples were treated with RNase H, thanks to its ability to eliminate RNA moiety of RNA-DNA hybrids, R-loops level dramatically decreased, confirming that the detected signal was RNA-DNA hybrids specific. The amount of R-loops presence was measured as described above.

From R-loops characterization performed in HeLa cells, we can conclude that TDP-43 role is directly involved in genomic RNA-DNA hybrids regulation through a conservative process among human cells, confirming its importance in control of genetic instability and DNA damage associated to R-loops accumulation. Comparing R-loops level found in p.A382T TDP-43 SH-SY5Y model to the ones observed in silenced Hela cell line we could appreciate a higher hybrids' accumulation in the neuroblastoma cells. This result underlined that both p.A382T missense mutation had higher detrimental effects on R-loops' resolution despite of TDP-43 silencing both that SH-SY5Y, used as ALS disease model, presented a particular affection in the elimination of accumulated R-loops compared to HeLa.

4.3 R-loops and stress granules characterization in EBV LCLS

4.3.1 Accumulation of R-loops in p.A382T TDP-43 mutated LCLs

Considering that ALS patients with TDP-43 mutations are rare we decided to use lymphoblastoid cell lines (LCLs), i.e. B cells derived from whole blood immortalized with Epstein Barr Virus (EBV), from a TDP-43 mutated patient (TDP-43 mut LCLs), an healthy control (Ctrl LCLs) and a sporadic ALS patient (sALS LCLs) in order to confirm the role of TDP-43 in R-loops removal in a model derived from patients. Co-localization of R-loops with the ribonucleoprotein TDP-43 was studied by immunofluorescence (IF) in CTRL, sALS and TDP-43 mut LCLs using two fixation methods, methanol and paraformaldehyde. As reported in Figure 15B, the presence of R-loops co-localization with TDP-43 in the perinuclear area of TDP-43 mut LCLs in comparison to CTRL LCLs and sALS LCLs was particularly appreciable using paraformaldehyde than methanol fixation method (Figure 15A). R-loops quantification in the slot blot analysis of figure 15 confirmed significative R-loops accumulation in TDP-43 mut LCLs using both fixation methods. Paraformaldehyde fixation guaranteed a decreased detection level of TDP-43 signal in the nucleus of TDP-43 mut LCLs in comparison to CTRL LCLs and sALS LCLs, which is in accordance with the mislocalization of the mutated protein in the cellular cytoplasm, causing the loss of its nuclear physiological function[149].

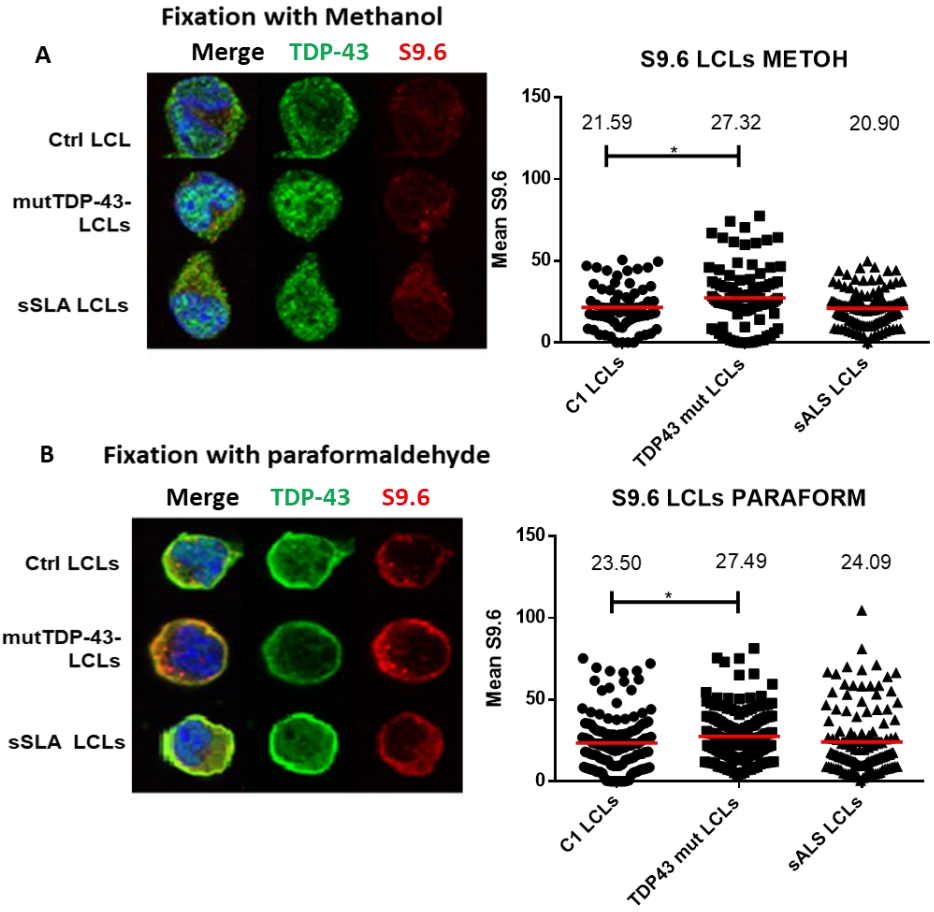

Figure 15. A) Immunofluorescence of CTRL, sALS and TDP-43mut LCLs using an anti-TDP-43 antibody in green and an anti-S9.6 antibody in red against R-loops after methanol fixation. Nuclei are stained with DAPI in blue. The slot blot shows the median of the S9.6 signal intensity per nucleus and for each sample the mean of S9.6 signal is also reported. Around 100 cells from two independent experiments were considered. *, P = 0,0382 (Mann-Whitney U test, two-tailed). **B)** Immunofluorescence of CTRL, sALS and TDP-43mut LCLs using an anti-TDP-43 antibody in green and an anti-S9.6 antibody in red against R-loops after paraformaldehyde fixation. Nuclei are stained with DAPI in blue. The slot blot shows the median of the S9.6 signal intensity per nucleus and for each sample the mean of S9.6 signal is also reported. Around 100 cells from two independent experiments were considered. *, P = 0,0345 (Mann-Whitney U test, two-tailed).

R-loops presence was then investigated in LCLs by flow cytometry analysis. A vitality dye (zombie violet dye) able to label cytoplasmatic proteins only in cells with a compromised membrane and fundamental for discrimination between alive and apoptotic cells was used. In these experiments vital cells were around 60% and 40% accounted for the apoptotic LCLs, (53,67 ± 6,386 of vital cells vs 41,67 ± 5,925 of apoptotic cells in CTRL LCLs; 72,67 ± 9,615 of vital cells vs 26,33 ± 9,615 of apoptotic cells in TDP43mut LCLs and of 62,67 ± 4,256 vital cells vs 34,33 ± 3,844 of apoptotic cells

in sALS LCLs for three biological experiments). For the detection of cotranscriptional R-loops as genomic three-stranded nucleic acid structures, permeabilization of the cellular membrane was necessary. For this reason, different permeabilization substances such as Tween20 and saponin were tested, both preceded by paraformaldehyde fixation. Results showed that saponin was the less effective treatment on LCLs vitality, recorded in a biological triplicate of CTRL LCLs 72,83 ± 1,014 of vital cells vs 26,50 ± 0,7638 of apoptotic cells, instead in the same conditions Tween20 lead to 21,67 ± 2,028 of vital cells vs 72,67 ± 9,615 of apoptotic cells. Thus, cells were treated with and without RNaseH1 and analysed for R-loops formation. An increased amount of R-loops was found in TDP-43 mut LCLs (orange) in comparison to Ctrl LCLs (blue) as reported in Figure 16A and the S9.6 signal in TDP-43 mut LCLs was proved as R loop dependent because it was reverted by the action of RNaseH1 enzyme (Figure 16.B, green peak). In a biological triplicate experiment, S9.6 mean fluorescence of TDP-43 mut LCLs was found significantly increased compared to Ctrl LCLs and sALS LCLs, in absence (-) and in presence (+) of RNaseH1 as shown in the histogram of figure 16C.

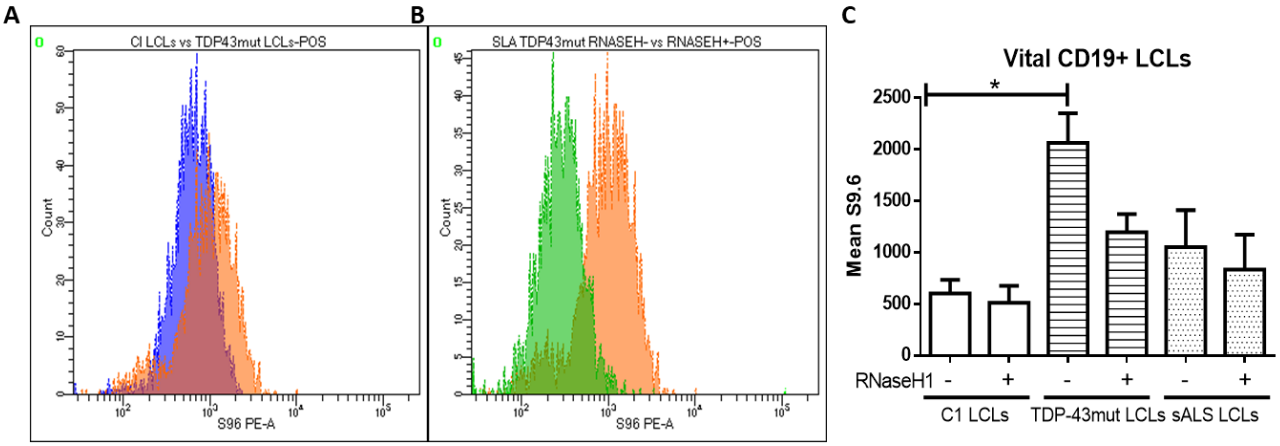

Figure 16. A) Flow cytometry plot reports the amount of R-loops in Ctrl LCLs (blue) in comparison to TDP-43 mut LCLs (orange). **B)** The RNaseH1 action on TDP-43 mut LCLs decreases the presence of R-loops signal (orange peak vs green peak). **C)** Histogram of S9.6 mean fluorescence of Ctrl, sALS and TDP-43 mut LCLs in presence (+) and in absence (-) of RNaseH1. The value represented is the mean ± SEM of three biological experiments. ANOVA, Newman-Keuls Multiple Comparison Test, *p <0,05.

As proved by Phillips et al., S9.6 antibody can bind RNA duplexes with an approximately fivefold reduced affinity compared to RNA-DNA hybrids[320] and considering its ability in recognizing different kind of hybrids with different specificity the use of control enzyme becomes necessary for the future experiments. Considering that RNase H was not able to totally revert the S9.6 signal in all

LCLs samples, we decided to remove competing free RNA for improving the specificity of the RNA-DNA hybrid signal. As reported by Zhang and colleagues, S9.6 antibody is able to detect even the presence of ssRNAs annealed as duplex RNAs and as adventitious RNA-DNA hybrids[317], thus treatment with RNaseA enzyme is fundamental for avoiding this effect. So, R-loops detection was repeated with the addition of RNaseA enzyme that, in presence of high salt concentration (higher than 300 mM), is able to degrade selectively ssRNAs (Figure 17). For the preservation of cellular vitality around the values established in the previous experiment, NaCl concentration was set to 0,4M.

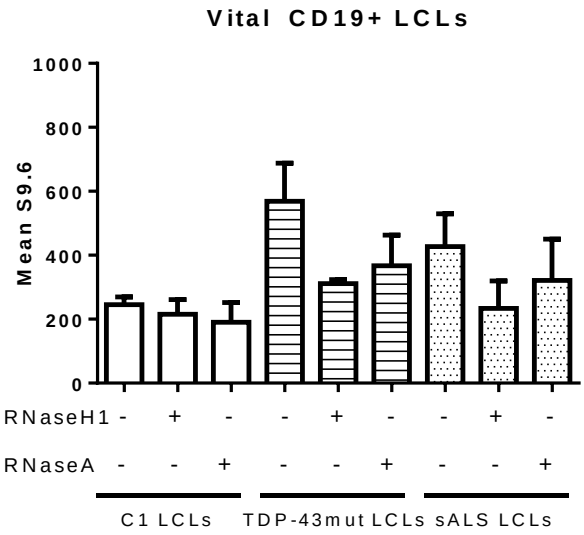

Figure 17. Histogram of mean ± SEM of mean fluorescence of S9.6 in Ctrl, sALS and TDP-43 mut LCLs of three biological experiments in normal conditions, with RNaseH1 and RNaseA treatments, as indicated.

From the results reported in Figure 16, we confirmed the previous observations, i.e. a higher R-loops' accumulation in TDP43 mut LCLs and a recovery of S9.6 signal with RNaseH1 treatment both in sALS and TDP-43 mut LCLs at similar level of the Ctrl LCLs, representing a backgroung signal. In comparison with RNaseH1, RNaseA treatment did not vary the recovered R-loops' signal in the three cell lines, confirming the ability of S9.6 antibody to recognize not specifically R-loops but also other molecular structure[320] which tend to accumulate in LCLs. For this reason, in these cells the use of supplementary treatments besides RNaseH1 could give clearer signal unrelated specifically to accumulation of R-loops.

4.3.2 R-loops interacts less with TDP-43 at chromatin level in p.A382T TDP-43 mutated LCLs

Co-immunoprecipitation (coIP) between S9.6 antibody and TDP-43 on both chromatin (Chr) fraction and whole lysate (Wl) fraction of the same cell lines was run. As control of the S9.6 immunoprecipitation, DNA was extracted from the same samples followed by a qPCR in R-loops' enriched genes ING3 and NOP58, as genes prone to form R-loops in this tissue as reported by Halász et al.[294] (Figure 18).

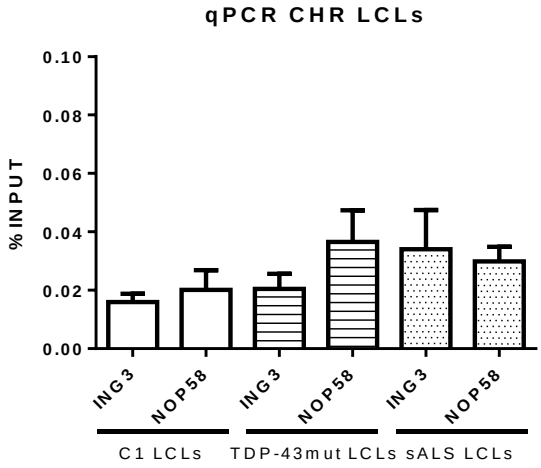

Figure 18. The histogram quantifies mean ± SEM from qPCR in ING3 and NOP58 genes from three biological experiments of CTRL LCLs, sALS LCLs and mutTDP-43-LCLs p.A382T.

The interaction between R-loops and TDP-43 protein was then tested by western blot. In Figure 19, the Chr fraction and the Wl fraction from CTRL LCLs, sALS LCLs and mutTDP-43-LCLs p.A382T are shown.

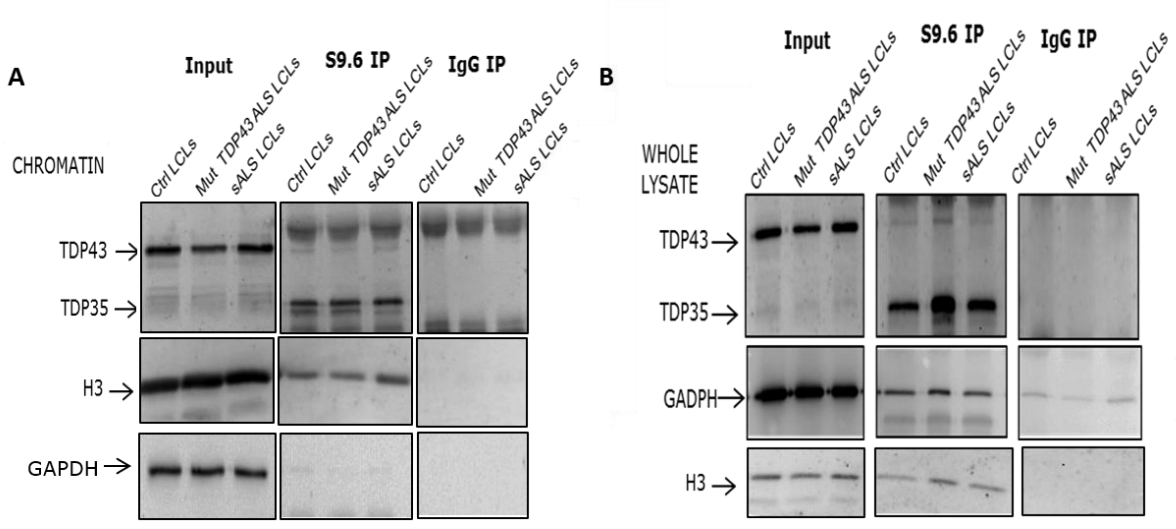

Figure 19. Co-immunoprecipitation between S9.6 and TDP-43 in chromatin of CTRL, TDP-43 mut and sALS LCLs. Input, S9.6 IP and IgG IP of chromatin fraction were loaded on a 10% SDS-PAGE

and then immunoblotted with TDP-43, H3 as nuclear normalizator and GAPDH and cytosolic normalizator. S9.6 binding was tested by qPCR.

In the chromatin fraction of TDP-43 mut LCLs, TDP-43 had less interaction with R-loops. This may lead to the conclusion that at the chromatin level full length TDP-43 was not able to interact with R-loops maybe due to its sequestration to the cytosolic compartment in the cell. Performing the same experiment on the whole lysate fraction of LCLs, we observed the recovery of the interaction signal confirming the ability of S9.6 antibody to recognize different kind of hybrids that can accumulate in the cytoplasm of LCLs. In both cases of the chromatin and of the whole lysate, it is showed a high interaction of s9.6 with the truncated form TDP-35 in TDP-43 mut LCLs in comparison with Ctrl LCLs and sALS LCLs. These preliminary data in LCLs demonstrated the ability of TDP-43 to be involved in genomic R-loops regulation in the nucleus. We wondered if R-loops accumulation could also be involved in the cytoplasm and in particular in stress granules. Increased R-loop accumulation in fact leads to genomic instability and aberrant transcript accumulations, which in turn promotes SG formation as reported by Salvi and Mekhail. (2015)[236].

4.3.3 Hybrids presence in stress granules of p.A382T TDP-43 mutated LCLs

CTRL, sALS and TDP-43 mut LCLs were fixed both with paraformaldehyde and methanol and then treated with arsenite to induce the formation of SG[321, 322]. Using the two methods of fixation, a higher amount of SG was visible in the perinuclear area of TDP-43 mut LCLs in comparison with the other two cell lines and were strictly linked to a visible co-localization with RNA-DNA hybrids, confirming their mutual association and reciprocal enrichment due to inert RNA transcripts accumulation (Figure 20).

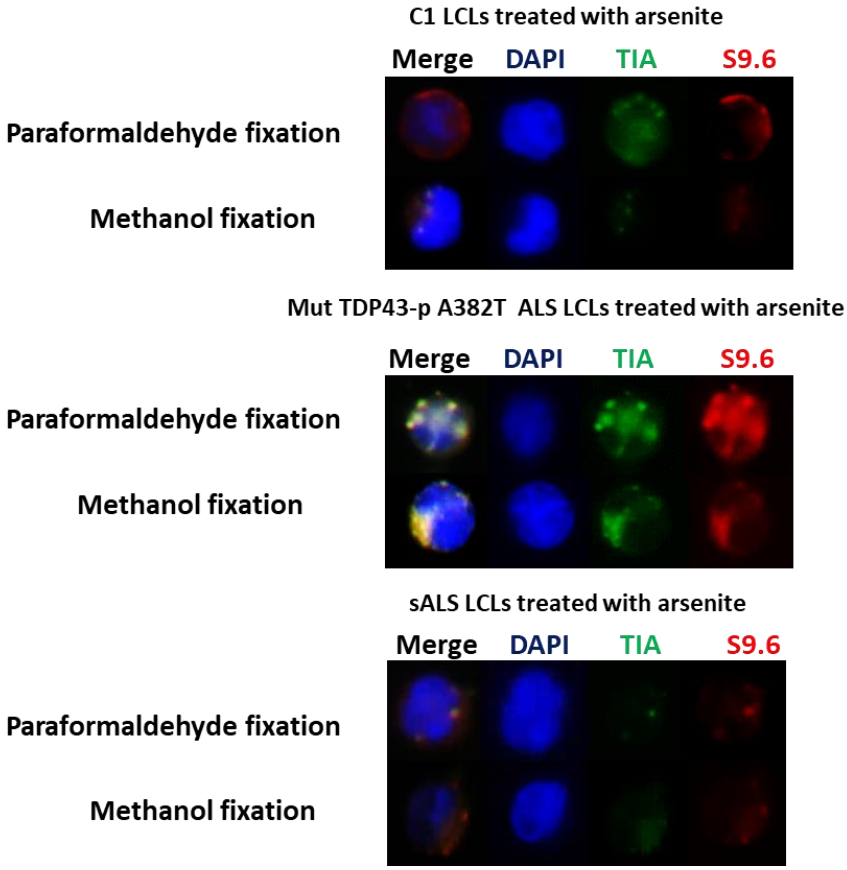

Figure 20. SGs co-localize with R-loops in TDP43 mut LCLs, compared to CTRL and sALS LCLs. Cells were treated with arsenite and then stained with DAPI for nuclear staining in blue, TIA for SG marker represented in green and S9.6 for R-loops represented in red.

Considering the obtained results from the previous experiments, we proved that R-loops formation is regulated and controlled by TDP-43 to not only in transfected SH-SY5Y and silenced HeLa but also in ALS derived LCLs. TDP-43 missense mutation p.A382T impaired its role in R-loops resolution leading to the accumulation of aberrant transcripts and hybrids that are trapped in persistent RNA-processing foci such as stress granules present in the cytoplasmic compartment of TDP-43 mutated LCLs. So, from one side the sequester of the misfolded and mutated form of TDP-43 in stress granules could cause a loss of its nuclear physiological function and from the other the formation of TDP-43 aggregates in the cytoplasm could recruit native TDP-43 or other interactors proteins constituting a gain of toxic function associated to neurodegeneration.

5 Discussion and Conclusions

Amyotrophic lateral sclerosis ALS represented one of the most common adult-onset motor neuron disease characterized by the rapidly progressive neurodegeneration of upper and lower motor neurons from the spinal cord, brain stem, and motor cortex causing muscle weakness and eventual respiratory failure[1]. Like other neurodegenerative diseases, ALS disease raised due to a complex interplay between multiple pathogenic cellular mechanisms which may not be mutually exclusive[22,23] engaging factors involved in several molecular pathways[34] and also environmental predisposing agents[17]. Despite the multifactorial nature of this disease, sporadic and familial ALS are clinically and pathologically similar, suggesting a common pathogenesis. TDP-43 was shown to be a major disease protein in ubiquitinated inclusions characteristic of high majority of ALS patients, where it is redistributed from its physiological nuclear location to cytoplasmic aggresomes called stress granules, together with caspase cleaved phosphorylated C terminal fragments CTFs, represented by TDP-35 and TDP-25[118]. Moreover, the identification of TDP-43 mutations in affected ALS patients suggested an important role for this protein in the pathogenesis of ALS disease[158]. TDP-43 is a highly conserved and ubiquitously expressed member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family of proteins[95] able to shuttle to cellular cytoplasm as regulatory factor in all steps in mRNA life cycle such as splicing, transcription, translation, mRNA transport and stabilization[111,112]. TDP-43 proteinopathy is based on a molecular mechanism widely described in literature and considered as pathogenic for ALS development. The loss of TDP-43 nuclear function due to sequestration in cytoplasmic stress granules and gain of proteinopathic transformation of mutant TDP-43 associated to an altered proteic structure could act to "seed" the irreversible pathological transformation of proteins associated to ALS, exacerbated by the persistence of conditions of prolonged or repeated neuronal damage[173, 282, 283]. Besides its role in RNA maturation, they are increasing researches on TDP-43 possible involvement in DNA transactions thanks to its DNA binding ability[176,177] and its involvement in DNA damage response DDR, considering the significant accumulation of genomic damage consistently observed in ALS. Researchers proved TDP-43 involvement in nonhomologous end joining (NHEJ), as stabilizing factor in the scaffold for the recruitment of XRCC4-DNA ligase 4 complex in postmitotic neurons[181] and showed that TDP-43 depletion lead to increased sensitivity to various forms of transcription-associated DNA damage such as R-loop-associated damage[272]. Moreover, it has been proved that mutations in low complexity domains associated with the loss of TDP-43 important nuclear function is known to be associated with excessively stable R-loops, stalled transcription machinery and stable open DNA[183]. So, the aim of this project was based on the investigation of TDP-43 role in R loops' formation and on the study of p.A382T TDP-43 mutation

functional effects in the alteration of TDP-43 nuclear function, using different cellular model such as neuroblastoma cell line SH-SY5Y, HeLa cervical cancer cells and lymphoblastoid cell line LCLs derived from ALS patients. Confirming evidences reported in literature, we firstly proved that p.A382T TDP-43 missense mutation, pathogenic for ALS, caused an increased nuclear TDP-43 translocation into cellular cytoplasm p.A382T TDP-43 SH-SY5Y detected by TDP-43 immunofluorescence. Mislocalization of mutated p.A382T TDP-43 could sequester full length TDP-43 form in cytoplasmic inclusions avoiding the performance of its physiological nuclear function and could cause a gain of neurotoxic cellular function fundamental in neurodegeneration process in ALS patients[317]. S9.6 immunofluorescence performed on the neuronal model SH-SY5Y proved that p.A382T mutation affected TDP-43 physiological regulative role associated with R-loops' formation showing a higher detection of genomic RNA-DNA hybrids in GFP tagged p.A382T TDP-43 SH-SY5Y. This result was even more confirmed by reversion of signal in condition of RNaseH1 overexpression. The lack of S9.6 signal decrease in GFP tagged TDP-43 WT SH-SY5Y overexpressing RNaseH1 proved that in these neuroblastoma cells we didn't detect a real R-loops accumulation. The use of RNaseIII treatment used specifically for dsRNAs degradation proved that the previously obtained signal in TDP-43 WT GFP tagged SH-SY5Y was due to dsRNA hybrids accumulation and this was related to the ability of S9.6 antibody of recognizing also these molecular structures with lower binging affinity[312]. DRIP-qPCR experiment showed in GFP tagged p.A382T TDP-43 SH-SY5Y the recovery of a significative increase in R-loops accumulation, particularly in RPL13 gene, that encodes for a ribosomal protein involved in pathways of viral mRNA translation and of rRNA processing. R-loops related impairment of RPL13 gene could lead to alterations of protein homeostasis as well as modifications of RNA metabolism that are now recognized as major players in the pathogenesis of ALS[323], considering that RPL13 is implicated in TDP-43 cytoplasmic interactome cluster fundamental for RNA translation regulation[184]. R-loops accumulation in p.A382T GFP tagged SH-SY5Y induced the activation of DNA damage response due to double strand breaks DSBs detected by presence of γH2AX foci which dependence from R-loops was confirmed by foci decrease with RNaseH1 overexpression. In p.A382T GFP tagged SH-SY5Y we also observed also an increase R-loops dependent of FANCD2 foci which is involved in the interstrand crosslinks (ICLs) repair pathway[313] and in R-loops' resolution as component of Fanconi Anemia complementation group[292,301]. Prpar Mihevc S. et al. showed that TDP-43 aggregation had a comparable effect to TDP-43 knockdown by identifying a common set of proteins that were differentially expressed in a similar manner in these two different conditions[324]. Considering the loss of the nuclear localization of both p.A382T mutated TDP-43 and of WT TDP-43 and their translocation in cytoplasmic aggregates of p.A382T TDP-43 SH-SY5Y, we decided to investigate R-loops development in TDP-43 silenced

HeLa cells. In these cells we obtained a significative increase of R-loops signal by S9.6 immunofluorescence, that was confirmed by DRIP-qPCR with an important R-loops' accumulation in RPL13 gene. We confirmed the work from Hill et al.[183] showing that the absence of nuclear TDP-43 gave rise to DNA damage response R-loops' dependent, observing a significative increase of DSBs by γH2AX foci count totally reverted in conditions of RNaseH1 overexpression. We also recovered an elevated accumulation of FANCD2 and H3S10P as R-loops associated factors in SiTDP-43 HeLa that was suppressed in conditions of RNaseH1 overexpression. From these experiments we can conclude that in HeLa cell line TDP-43 silencing was responsible of loss in the regulation of R-loops, confirming the importance of this RBP in the prevention and in the resolution of these genomic RNA-DNA hybrids. Anyway, higher level of R-loops was recovered in GFP tagged p.A382T TDP-43 SH-SY5Y, proving that TDP-43 mutation had a very detrimental effect on hybrids' resolution and that cell lines respond in a different way to these events. Even if both cell lines were able to activate DDR, we could appreciate that the activation of FANCD2 pathways played a leading role in the cervical cancer cell line. This could suggest a preferential activation in HeLa cells of homologous recombination (HR) pathway for repair of DNA damages, which instead in SH-SY5Y could be solved by NHEJ mechanism. This could confirm TDP-43 involvement in NHEJ process showed by Hegde et al.[175] and previously described. To better study the contribution of TDP-43 in R-loops regulation in ALS, preliminary experiments were performed on lymphoblastoid cell lines from a healthy control (CTRL LCLs), a sporadic ALS patient (sALS LCLs) and a p.A382T TDP-43 mutated patient (p.A382T mut TDP4-3 LCLs). We found an increased level of R-loops in p.A382T TDP-43 mut LCLs co-localizing with TDP-43 in its perinuclear area, confirming the results previously described in SH-SY5Y. In p.A382T TDP-43 mut LCLs it was even observed SGs formation as well as a relation between SG and accumulation of R-loops or RNA molecules, which could cooperate to cell neurodegeneration[236]. Chromatin fraction of p.A382T TDP-43 mut LCLs showed an important interaction between R-loops and TDP-43, in particular with the truncated form at 35kDa. Instead the mutated cell line showed a wicker binding of TDP-43 to R-loops, underlining TDP-43 less presence in the nucleus and less capacity to bind and resolve R-loops with this mutation. This was confirmed by the recovery a higher binding signal between R-loops and TDP-43 in whole lysate fraction of the same cell line. Upon cleavage by caspases at intrinsic caspase cleavage sites, TDP-43 can generates C-terminal fragments (CTF) such as 35 kDa fragments (TDP-35), that due to the lack of nuclear localization signal (NLS), mislocalizes in the cytoplasm, where may associate with RNA-DNA hybrids forming cytoplasmic inclusions[325]. Nishimoto et al. investigated the biochemical properties of these inclusions, reporting that they exhibit properties of SGs cellular structures that package

mRNA and RNA-binding proteins during cell stress, and proposed that truncated TDP-35 fragments facilitate SGs assembly and promotes SGs formation[326].

From these results we can conclude that p.A382T missense mutation altering conformational properties of TDP-43 induced its cytoplasmic mislocalization and its sequestration in cytoplasmic aggregates together with WT form and with CTFs avoiding the performance of its physiological nuclear function involved in the prevention and in the resolution of R-loops formation. This was observed in p.A382T GFP tagged SH-SY5Y, in p.A382T TDP-43 mut LCLs and with less extent in SiTDP-43 HeLa, showing that as reported by Prpar Mihevc S. et al.[324] TDP-43 knockdown had a comparable effect with TDP-43 aggregation, but with more detrimental result in this latter case. TDP-43 has been reported to co-localize with the major siRNA pathway components, Dicer and Argonaute, in both cell culture and human patient tissue, but their interaction is inhibited by SGs formation in response to cellular stressors or by overexpression of ALS-linked genes including human TDP-43[327]. According to this, nuclear dsRNAs accumulation observed in SH-SY5Y overexpressing TDP-43 could be associated to inhibition of Dicer processing function responsible for the loss of maturation of pre-miRNAs, presented as dsRNAs hairpin structure, in mature miRNAs[327]. This signature was detected in both sporadic and familial ALS spinal column motor neurons as a dramatic global reduction in mature miRNAs in comparison to control tissue[327]. Response to R-loops accumulation was studied under different aspects in the used cellular models. Both silenced HeLa and transfected SH-SY5Y were able to activate DDR by phosphorylation of H2AX used as biomarker of DNA DSBs. Selection of HR pathway in SiTDP-43 HeLa could be associated with higher involvement of FANCD2 R-loops resolutive agent in condition of TDP-43 silencing. Lower FANCD2 detection in SH-SY5Y could be associated with choose of NHEJ as preferential mechanism for DNA damage repair, confirming TDP-43 function in this latter pathway as proved by Hegde et al.[181]. At least in response to R-loops accumulation LCLs showed SGs formation as safe “storage and sorting stations” for RNA binding proteins, translationally stalled mRNAs and arrested pre-initiation complexes. In LCLs a very important role in R-loops regulation and in SGs development seemed to be represented by TDP-35 CTF. The presence of TDP-35 was also observed both in total proteic extract of p.A382T GFP tagged SH-SY5Y and of TDP-43 WT GFP tagged SH-SY5Y, confirming the research by Wobst H.J. et al.[328].

For this reason, it should be interesting to further investigate TDP-35 role in R-loops regulation and in SGs formation in these transfected SH-SY5Y and finally to confirm in silenced TDP-43 SH-SY5Y the data previously obtained in SiTDP-43 HeLa. Further studies are needed to understand the affinity of TDP-43 to different kind of RNA species (dsRNA, ssRNA, RNA-DNA hybrids) and if different mutations can give the same results. Overall this work of thesis has demonstrated a new role of TDP-

43 in RNA-DNA hybrids removal dependent on Fanconi Anemia pathway, which was never described in relation to this RNA binding protein, before.

6 References

1. Robberecht W. & Philips T. The changing scene of amyotrophic lateral sclerosis *Nat Rev Neurosci*., **14,** 248–264 (2013).
2. Joffroy J. & M. C. A. Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arch Physiol Neurol Path*., **2,** 744 (1869).
3. Ingre C. *et al*. Risk factors for amyotrophic lateral sclerosis. *Clin Epidemiol*., **7,** 181–93 (2015).
4. DeLoach A. *et al.* A retrospective review of the progress in amyotrophic lateral sclerosis drug discovery over the last decade and a look at the latest strategies. *Expert Opin. Drug Discov*. **10**, 1099–1118 (2015).
5. Petrov D. *et al.* ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment? *Front Aging Neurosci*. **22**;9:68 (2017).
6. Rusmini P. et al. The Role of the Heat Shock Protein B8 (HSPB8) in Motoneuron Diseases. *Front Mol Neurosci*. **10**:176 (2017).
7. Tsaytler P. *et al.* Selective inhibition of a regulatory subunit of protein phosphatase 1 restores proteostasis. *Science* **332**, 91-94 (2011).
8. Johnston C. A. *et al*. Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *J. Neurol*. **253**, 1642-1643 (2006).
9. Logroscino G. and Piccininni M. Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference. *Neuroepidemiology*. **52**(1-2):93-103 (2019).
10. Mehta P. *et al.* Prevalence of Amyotrophic Lateral Sclerosis - United States, 2014. *MMWR Morb Mortal Wkly Rep.* **23**;67(7):216-218 (2018).
11. Al-Chalabi A. *et al.* Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nat. Rev. Neurol*. **13**, 96-104 (2017).
12. Nguyen H.P. ALS Genes in the Genomic Era and their Implications for FTD. *Trends Genet.*;**34**(6):404-423 (2018)
13. Renton A. E. *et al*. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268. (2011).

14. Chio, A. *et al*. Genetic counselling in ALS: facts, uncertainties and clinical suggestions. J Neurol Neurosurg Psychiatry 85, 478-85 (2014).
15. Gordon P. H. Amyotrophic Lateral Sclerosis: An update for 2013 Clinical Features, Pathophysiology, Management and Therapeutic Trials. *Aging Dis*, **4,** 295–310 (2013).
16. Al-Chalabi A. *et al.* An estimate of amyotrophic lateral sclerosis heritability using twin data. *J Neurol Neurosurg Psychiatry*, **81,** 1324–6 (2010).
17. Hanby M.F. *et al.* The risk to relatives of patients with sporadic amyotrophic lateral sclerosis. *Brain*, **134,** 3454–7 (2011).
18. Ajroud-Driss S. & Siddique T. Sporadic and hereditary amyotrophic lateral sclerosis (ALS). *Biochim Biophys Acta.* **1852,** 679–684 (2015).
19. Andersen P. M. & Al-Chalabi A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat Rev Neurol*., **7,** 603–615 (2011).
20. Bosco D. A. *et al.* Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nat Neurosci,* **13,** 1396–1403 (2010).
21. Saberi S. *et al.* Neuropathology of Amyotrophic Lateral Sclerosis and Its Variants. *Neurol. Clin*. **33**, 855-876 (2015).
22. Shaw P.J., Molecular and cellular pathways of neurodegeneration in motor neurone disease. *J Neurol Neurosurg Psychiatry*, **76,**1046-1057(2005)
23. Cozzolino M. *et al.* Amyotrophic lateral sclerosis: from current developments in the laboratory to clinical implications. *Antiox Redox Signal,* **10,** 405-443, (2008).
24. Kipnis J. *et al.* Dual effect of CD4+CD25+ regulatory T cells in neurodegeneration: A dialogue with microglia. *Proc Natl Acad Sci U S A*, **101,** 14663–14669 (2004).
25. Renton A. E. *et al.* State of play in amyotrophic lateral sclerosis genetics. *Nature Neuroscience*, **17,** 17–23 (2014).
26. Conicella A. E., Zerze G. H., Mittal J. et al. ALS mutations disrupt phase separation mediated by alpha-helical structure in the TDP-43 low-complexity C-terminal domain. *Structure* **24**, 1537–1549 (2016).
27. Wang, W. Y. *et al*. Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. *Nat Neurosci*. **16**, 1383-91 (2013).
28. Fang, X. *et al*. The NEK1 interactor, C21ORF2, is required for efficient DNA damage repair. *Acta Biochim Biophys Sin*. **47**, 834-41 (2015).
29. Corcia, P. *et al.* Molecular imaging of microglial activation in amyotrophic lateral sclerosis. *PLoS One.* **7**, e52941 (2012).
30. Ferrante R. J. *et al.* Evidence of increased oxidative damage in both sporadic and familial

amyotrophic lateral sclerosis. *Journal of Neurochemistry*, **69**, 2064-2074 (1997).

31. Atkin J.D. *et al.* Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. *Neurobiol Dis*, **30**, 400-407 (2008).
32. Saxena S. *et al.* A role for motoneuron subtype-selective ER stress in disease manifestations of FALS mice. Nat. Neurosci. *Nat Neurosci.*, **12**, 627-636 (2009).
33. Vande Velde C. *et al.* Misfolded SOD1 associated with motor neuron mitochondria alters mitochondrial shape and distribution prior to clinical onset. *PLoS One*. **6**, e22031 (2011).
34. Hayashi Y. *et al.* SOD1 in neurotoxicity and its controversial roles in SOD1. *Adv Biol Regu.*, **60,** 95–104, (2016).
35. Brotherton T.E. *et al*. Cellular toxicity of mutant SOD1 protein is linked to an easily soluble, non-aggregated form in vitro. *Neurobiol Dis*, **49**, 49-56 (2013).
36. Tadic V. *et al*. The ER mitochondria calcium cycle and ER stress response as therapeutic targets in amyotrophic lateral sclerosis. *Front Cell Neurosci*., **8**, 147 (2014).
37. Weiduschat N. *et al.* Motor cortex glutathione deficit in ALS measured in vivo with the J-editing technique. *Neurosci Lett.*, **570,** 102–7 (2014).
38. Carrì M. T. *et al.* M. Oxidative stress and mitochondrial damage: importance in non-SOD1 ALS. *Front Cell Neurosci.*, **9,** 41 (2015).
39. Shodai A. *et al.* Aberrant assembly of RNA recognition motif 1 links to pathogenic conversion of TAR DNA-binding protein of 43 kDa (TDP-43). *J Biol Chem.*, **288,** 14886–905 (2013)
40. Parker S. J. *et al.* Endogenous TDP-43 localized to stress granules can subsequently form protein aggregates. *Neurochem Int.*, **60,** 415–24 (2012).
41. Vance C. *et al.* ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. *Hum Mol Genet*., **22,** 2676–88 (2013).
42. Zhang P. *et al*. Chronic optogenetic induction of stress granules is cytotoxic and reveals the evolution of ALS-FTD pathology. *Elife*. **8**. pii: e39578 (2019).
43. Chen L., Liu B. Relationships between Stress Granules, Oxidative Stress, and Neurodegenerative Diseases. *Oxid Med Cell Longev*. **2017**:1809592 (2017).
44. Colombrita C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem.*, **111,** 1051–61 (2009).
45. Gasset-Rosa F. *et al.* Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. *Neuron*. **102**(2):339-357.e7 (2019).

46. Peters O. M. *et al.* Emerging mechanisms of molecular pathology in ALS. *J Clin Invest.*, **125,** 1767–1779 (2015).
47. Lee Y. B. *et al.* Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.*, **5,** 1178–86 (2013).
48. Donnelly C. J. *et al.* RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron*, **80,** 415–28 (2013).
49. Sareen D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med*., **5,** 208ra149 (2013).
50. Maurel C. *et al.* Causative Genes in Amyotrophic Lateral Sclerosis and Protein Degradation Pathways: a Link to Neurodegeneration. *Mol Neurobiol.***55**(8):6480-6499 (2018).
51. Gamerdinger M. *et al.* Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J*. **28**(7):889-901 (2009).
52. Rubinsztein DC, Mariño G, Kroemer G. Autophagy and aging. *Cell* **146**(5):682-95 (2011).
53. Tomaru U. *et al*. Decreased proteasomal activity causes age-related phenotypes and promotes the development of metabolic abnormalities. *Am J Pathol.* **180**(3):963-72 (2012).
54. Munch C. & Bertolotti A. Self-propagation and transmission of misfolded mutant SOD1: prion or prion-like phenomenon? *Cell Cycle*, **10,** 1711 (2011).
55. Grad L. I. *et al.* Intercellular propagated misfolding of wild-type Cu/Zn superoxide dismutase occurs via exosome-dependent and -independent mechanisms. *Proc Natl Acad Sci U S A.*, **111,** 3620–5 (2014).
56. Gitler A. D. & Shorter J. RNA-binding proteins with prion-like domains in ALS and FTLD-U. *Prion*, **5,** 179–87 (2011)
57. Kim H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*, **95,** 467–73 (2013).
58. Mori K. *et al.* hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol*., **125,** 413–23 (2013).
59. Li Y. R. *et al*. Stress granules as crucibles of ALS pathogenesis. *J Cell Biol*., **01,** 361–72 (2013).
60. Harrison A.F. and Shorter J. RNA-binding proteins with prion-like domains in health and disease. *Biochem J.* **7**;474(8):1417-1438 (2017).
61. Popoli M. *et al.* The stressed synapse: the impact of stress and glucocorticoids on glutamate transmission. **13,** 22–37 (2012).

62. Rothstein, J.D. *et al*. Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis. *N Engl J Med*., **326**, 1464-1468 (1992).
63. Howland, D.S. *et al*. Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proc Natl Acad Sci U S A*, **99**, 1604-1609 (2002).
64. Morel L. *et a*l. Neuronal exosomal miRNA-dependent translational regulation of astroglial glutamate transporter GLT1. *J Biol Chem.,* **288**, 7105-7116 (2013).
65. Li K. *et al.* GLT1 overexpression in SOD1(G93A) mouse cervical spinal cord does not preserve diaphragm function or extend disease. *Neurobiol Dis.,* **78**, 12-23 (2015).
66. Milanese M. *et al*. Abnormal exocytotic release of glutamate in a mouse model of amyotrophic lateral sclerosis. *J Neurochem*., **116**, 1028-1042 (2011).
67. Mesci P. *et al*. System xC- is a mediator of microglial function and its deletion slows symptoms in amyotrophic lateral sclerosis mice. *Brain*, **138**, 53.-68 (2015).
68. Van Damme P. *et al.* Astrocytes regulate GluR2 expression in motor neurons and their vulnerability to excitotoxicity. *Proc Natl Acad Sci U S A*., **104**, 14825-14830 (2007).
69. Staats K. A. *et al.* Neuronal overexpression of IP3 receptor 2 is detrimental in mutant SOD1 mice. *Biochem Biophys Res Commun*., **429**, 210-213 (2012).
70. Rodolfo C. *et al.* Proteomic analysis of mitochondrial dysfunction in neurodegenerative diseases. *Expert Rev Proteomics.,* **7,** 519–42 (2010).
71. Damiano M. *et al.* Neural mitochondrial Ca^{2+} capacity impairment precedes the onset of motor symptoms in G93A Cu/Zn-superoxide dismutase mutant mice. *Nat Rev Neurosci*., **96**, 1349-1361 (2006).
72. Liu J. *et al*. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron,* **43**, 5-17 (2004).
73. Israelson A. *et al*. Misfolded mutant SOD1 directly inhibits VDAC1 conductance in a mouse model of inherited ALS. *Neuron,* **67**, 575-587 (2010).
74. Jiang Z., Wang W. *et al.* Mitochondrial dynamic abnormalities in amyotrophic lateral sclerosis. *Transl Neurodegener*. **29**;4:14 (2015).
75. Pansarasa O. *et al.* Lymphoblastoid cell lines as a model to understand amyotrophic lateral sclerosis disease mechanisms. *Dis Model Mech.* **26**;11(3) (2018).
76. Jaronen M., Goldsteins G. *et al.* ER stress and unfolded protein response in amyotrophic lateral sclerosis-a controversial role of protein disulphide isomerase. *Front Cell Neurosci*. **2**;8:402 (2014).
77. Kikuchi H. *et al*. Spinal cord endoplasmic reticulum stress associated with a microsomal

accumulation of mutant superoxide dismutase-1 in an ALS model. *Proc Natl Acad Sci U S A*., **103**, 6025-6030 (2006).

78. Zoghbi H.Y. Silencing misbehaving proteins. *Nat Genet*., **37**, 1302-1303 (2005).
79. Filezac de L'Etang *et al*. Marinesco-Sjogren syndrome protein SIL1 regulates motor neuron subtype-selective ER stress in ALS. *Nat Genet*., **18**, 227-238.
80. Homma K. *et al*. SOD1 as a molecular switch for initiating the homeostatic ER stress response under zinc deficiency. *Mol Cell*., **52**, 75-86 (2013).
81. Atkin J.D. *et al*. Mutant SOD1 inhibits ER-Golgi transport in amyotrophic lateral sclerosis. *J Neurochem*., **129**, 190-204 (2014).
82. Lyon M.S., Wosiski-Kuhn M. Inflammation, Immunity, and amyotrophic lateral sclerosis: I. Etiology and pathology. *Muscle Nerve*. **59**(1):10-22 (2019).
83. Nagai M. *et al.* Astrocytes expressing ALS-linked mutated SOD1 release factors selectively toxic to motor neurons. *Nat Neurosci*., **10,** 615–22 (2007).
84. Mantovani S. *et al.* Immune system alterations in sporadic amyotrophic lateral sclerosis patients suggest an ongoing neuroinflammatory process. *J Neuroimmunol*., **210,** 73–9 (2009).
85. Cereda C. et al. TNF and sTNFR1/2 plasma levels in ALS patients. *J Neuroimmunol*. **194**(1-2):123-31 (2008).
86. Kipnis J. *et al.* Dual effect of CD4+CD25+ regulatory T cells in neurodegeneration: A dialogue with microglia. *Proc Natl Acad Sci U S A*, **101,** 14663–14669 (2004).
87. Henkel J. S. *et al.* Regulatory T-lymphocytes mediate amyotrophic lateral sclerosis progression and survival. *EMBO Mol Med.*, **5**, 64–79 (2013).
88. Ou S. H., Wu F., Harrich D., García-Martínez L. F., e*t al.* Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J. Virol*. **69**, 3584–3596 (1995).
89. Sproviero D. *et al*. Leukocyte Derived Microvesicles as Disease Progression Biomarkers in Slow Progressing Amyotrophic Lateral Sclerosis Patients. *Front Neurosci*. **15**;13:344 (2019).
90. Al-Chalabi A. *et al.* The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol*. **124**, 339-352 (2012).
91. Arai T. *et al*. Tdp-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commu* **351**:602-11(2006).
92. Mackenzie I.R. *et al.* TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal

dementia. *Lancet Neurol*. **9**(10):995-1007 (2010).

93. Buratti E., and Baralle F. E. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J.Biol. Chem*. **276**, 36337–36343 (2001).
94. Krecic A. M., and Swanson M. S. hnRNP complexes: Composition, structure, and function. *Curr. Opin. Cell Biol*. **11**, 363–371 (1999).
95. Martinez-Contreras R., Cloutier P., Shkreta L. *et al.* hnRNP proteins and splicing control. *Adv. Exp. Med. Biol*. **623**, 123–147 (2007).
96. Sephton C. F., Good S. K., Atkin S., Dewey C. M. *et al.* TDP-43 is a developmentally regulated protein essential for early embryonic development. *J. Biol. Chem.* **285**, 6826–6834 (2010).
97. Geuens T., Bouhy D. and Timmerman V. The hnRNP family: insights into their role in health and disease. Hum. Genet. 135, 851–867 (2016).
98. Lukavsky P. J., Daujotyte D., Tollervey J. R. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. Nat. Struct. Mol. Biol. 20, 1443–1449 (2013).
99. Ayala Y. M., De Conti L., Avendaño-Vázquez S. E. *et al.* TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288 (2011).
100. Ayala Y. M., Zago P., D'ambrogio A. *et al.* Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci*. **121**, 3778–3785 (2008).
101. Wang W., Wang L., Lu J. *et al.* The inhibition of TDP-43 mitochondrial localization blocks its neuronal toxicity. *Nat. Med*. **22**, 869–878 (2016).
102. Mompeán M., Romano V., Pantoja-Uceda D. *et al.* The TDP-43 N-terminal domain structure at high resolution. *FEBS J*. **283**, 1242–1260 (2016).
103. Jiang L.-L., Zhao J., Yin X.-F. Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Sci. Rep*. **6**:23928 (2016).
104. Prasad A., Bharathi V., Sivalingam V. *et al.* Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Front Mol Neurosci*. **14**;12:25 (2019).
105. Qin H., Lim L. Z., Wei Y. *et al.* TDP-43N terminus encodes a novel ubiquitin-like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18619–18624 (2014).
106. Zhang Y. J., Caulfield T., Xu Y. F. *et al*. The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum.Mol. Genet*. **22**,

3112–3122 (2013).

107. Wang, Y. T. *et al.* The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J. Biol. Chem*. **288**, 9049–9057 (2013)

108. Santamaria N., Alhothali M., Alfonso M. H. *et al*. Intrinsic disorder in proteins involved in amyotrophic lateral sclerosis. *Cell. Mol. Life Sci*. **74**, 1297–1318 (2017).

109. Zhang Y. J., Xu Y. F., Cook C. *et al*. Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7607–7612 (2009).

110. Colombrita C. *et al.* TDP-43 and FUS RNA-binding proteins bind distinct sets of cytoplasmic messenger RNAs and differently regulate their posttranscriptional fate in motoneuron-like cells. *J. Biol. Chem*. **287**, 15635–15647 (2012).

111. Ling S. C., Albuquerque C. P.,Han J. S. *et al*. ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13318–13323 (2010).

112. Kawahara Y. And Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of theDrosha andDicer complexes. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 3347–3352 (2012).

113. Polymenidou M., Lagier-Tourenne C., Hutt K. R. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).

114. Tollervey J. R., Curk T., Rogelj B. et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci*. **14**, 452–458 (2011).

115. McDonald K. K., Aulas A., Destroismaisons L., Pickles et al. TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum. Mol. Genet*. **20**, 1400–1410 (2011).

116. Bentmann E., Haass C. and Dormann D. Stress granules in neurodegeneration–lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. *FEBS J*. **280**, 4348–4370 (2013).

117. Ratti A. , Buratti E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J Neurochem. Aug* **138,** 1:95-111 (2016).

118. Neumann M., Sampathu D.M., Kwong L.K. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**:130–133 (2006).

119. Buratti E. TDP-43 post-translational modifications in health and disease. *Expert Opin Ther Targets*. **3**:279-293 (2018).

120. Yang C., Tan W., Whittle C. *et al.* The C-terminal TDP-43 fragments have a high

aggregation propensity and harm neurons by a dominant-negative mechanism. *PLoS One* **5**:e15878 (2010).

121. Renton A.E., Chio A., Traynor B.J. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci* **17**:17–23 (2014).
122. Sreedharan J., Blair I.P., Tripathi V.B. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science 319*:1668–1672 (2008).
123. Alami N.H., Smith R.B., Carrasco M.A. *et al.* Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* **81**:536–543 (2014).
124. Liu-Yesucevitz L., Lin A.Y., Ebata A. *et al.* ALS-linked mutations enlarge TDP-43-enriched neuronal RNA granules in the dendritic arbor. *J Neurosci* **34**:4167–4174 (2014).
125. Kabashi E., Lin L., Tradewell M.L. *et al.* Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum Mol Genet* **19**:671–683 (2010).
126. Rutherford N.J., Zhang Y.J., Baker M. *et al*. Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet* **4**:e1000193 (2008).
127. Kabashi E., Valdmanis P.N., Dion P. *et al*. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet **40**:572–574 (2008).
128. Serio A., Bilican B., Barmada S.J. *et al.* Astrocyte pathology and the absence of non-cell autonomy in an induced pluripotent stem cell model of TDP-43 proteinopathy. *Proc Natl Acad Sci U S A* **110**:4697–4702 (2013).
129. Han J.H., Yu T.H., Ryu H.H. *et al.* ALS/FTLD-linked TDP-43 regulates neurite morphology and cell survival in differentiated neurons. *Exp Cell Res* **319**:1998–2005 (2013).
130. Johnson B.S., Snead D., Lee J.J. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J Biol Chem* **284**:20329–20339 (2009).
131. Budini M., Romano V., Avendano-Vazquez S.E. *et al*. Role of selected mutations in the Q/N rich region of TDP-43 in EGFP-12xQ/N-induced aggregate formation. *Brain Res* **1462**:139–150 (2012).
132. Watanabe S., Kaneko K., Yamanaka K. Accelerated disease onset with stabilized familial amyotrophic lateral sclerosis (ALS)-linked mutant TDP-43 proteins. *J Biol Chem* **288**:3641–3654 (2013).
133. Wu L.S., Cheng W.C., Shen C.K. Similar dose-dependence of motor neuron cell death caused by wild type human TDP-43 and mutants with ALS-associated amino acid

substitutions. *J Biomed Sci* **20**:33 (2013)

134. Wang X., Fan H., Ying Z. *et al.* Degradation of TDP-43 and its pathogenic form by autophagy and the ubiquitinproteasome system. *Neurosci Lett* **469**:112–116 (2010).
135. Zhang Y.J., Xu Y.F., Cook C. *et al.* Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proc Natl Acad Sci US A* **106**:7607–7612 (2009).
136. Fecto F., Yan J., Vemula S.P. *et al.* SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol* **68**:1440–1446 (2012).
137. Watts G.D., Wymer J., Kovach M.J. *et al.* Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* **36**:377–381 (2004).
138. Deng H.X., Chen W., Hong S.T. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**:211–215 (2011).
139. De Jesus-Hernandez M., Mackenzie I.R., Boeve B.F. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**:245–256 (2011).
140. Renton A.E., Majounie E., Waite A. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**:257–268 (2011).
141. Troakes C., Maekawa S., Wijesekera L. *et al.* An MND/ALS phenotype associated with C9orf72 repeat expansion: abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline. *Neuropathology* **32**:505–514 (2012).
142. Brettschneider J., Van Deerlin V.M., Robinson J.L. *et al.* Pattern of ubiquilin pathology in ALS and FTLD indicates presence of C9ORF72 hexanucleotide expansion. Acta *Neuropathol* **123**: 825–839 (2012).
143. Ash P.E., Bieniek K.F., Gendron T.F. *et al.* Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**:639–646 (2013).
144. Mori K., Weng S.M., Arzberger T. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**:1335–1338 (2013).
145. Johnson J.O., Pioro E.P., Boehringer A. *et al.* Mutations in theMatrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci* **17**:664–666 (2014).
146. Kim H.J., Kim N.C.,Wang Y.D. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**:467–473 (2013).
147. Lagier-Tourenne C., Polymenidou M., Hutt K.R. *et al*. Divergent roles of ALS-linked

proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci* **15**:1488–1497 (2012).

148. Honda D., Ishigaki S., Iguchi Y. *et al*. The ALS/FTLD-related RNAbinding proteins TDP-43 and FUS have common downstream RNA targets in cortical neurons. *FEBS Open Bio* **4**:1–10 (2013).
149. Scotter E.L., Chen H.J., Shaw C.E. TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets. *Neurotherapeutics*. **12**(2):352-63 (2015).
150. Johnson B. S., Snead D., Lee J. J. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *The Journal of Biological Chemistry*, **284**(30), 20329e20339 (2009).
151. Buratti E. Functional Significance of TDP-43 Mutations in Disease. *Adv Genet*. **91**:1-53. (2015).
152. Mitsuzawa S., Akiyama T., Nishiyama A. TARDBP p.G376D mutation, found in rapid progressive familial ALS, induces mislocalization of TDP-43. *eNeurologicalSci*. **12**;11:20-22 (2018).
153. Corrado L., Ratti A., Gellera C., Buratti *et al.* High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Hum. Mutat.*, **30**, 688–694 (2009).
154. Millecamps S., Salachas F., Cazeneuve C. *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. *J. Med. Genet*., **47**, 554–560 (2010).
155. Chiò A., Borghero G., Pugliatti M. *et al*. Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. *Arch. Neurol*., **68**, 594–598 (2011).
156. Orrù S., Manolakos E., Orrù N. *et al.* High frequency of the TARDBP p.Ala382Thr mutation in Sardinian patients with amyotrophic lateral sclerosis. *Clin Genet*., **81**, 172–178 (2012).
157. Floris G., Borghero G., Cannas A. *et al*. Clinical phenotypes and radiological findings in frontotemporal dementia related to TARDBP mutations. *J. Neurol*., **262**, 375–384 (2015).
158. Nonaka T., Kametani F., Arai T. *et al*. Truncation and pathogenic mutations facilitate the formation of intracellular aggregates of TDP-43. *Hum. Mol. Genet*., **18**, 3353–3364 (2009).
159. Mutihac R., Alegre-Abarrategui J., Gordon D. *et al.* TARDBP pathogenic mutations increase cytoplasmic translocation of TDP-43 and cause reduction of endoplasmic reticulum Ca^{2+} signaling inmotor neurons. *Neurobiol. Dis*., **75**, 64–77 (2015).
160. Watanabe S., Kaneko K. and Yamanaka K. Accelerated disease onset with stabilized

familial amyotrophic lateral sclerosis (ALS)-linked mutant TDP-43 proteins. *J. Biol. Chem*., **288**, 3641–3654 (2013).

161. Araki W., Minegishi S., Motoki K. *et al.* Disease-associated mutations of TDP-43 promote turnover of the protein through the proteasomal pathway. *Mol. Neurobiol*., **50**, 1049–1058 (2014).
162. Sabatelli M., Zollino M., Conte A. *et al.* Primary fibroblasts cultures reveal TDP-43 abnormalities in amyotrophic lateral sclerosis patients with and without SOD1 mutations. *Neurobiol. Aging*, **36**, 2005.e5–2005.e13 (2015).
163. Dewey C.M., Cenik B., Sephton C.F. TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol Cell Biol*. **31**(5):1098-108 (2011).
164. Orrù S., Coni P., Floris A. Reduced stress granule formation and cell death in fibroblasts with the A382T mutation of TARDBP gene: evidence for loss of TDP-43 nuclear function. *Hum Mol Genet*. **5**;25(20):4473-4483 (2016).
165. Aulas A., Caron G., Gkogkas C.G. *et al*. G3BP1 promotes stress-induced RNA granule interactions to preserve polyadenylated mRNA. *J. Cell Biol*., **209**, 73–84 (2015).
166. Gilks N., Kedersha N., Ayodele M. *et al*. Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* **15**:5383–5398 (2004).
167. Tourriere H., Chebli K., Zekri L. *et al*. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol*, **160**:823–831 (2003).
168. Kedersha N., Anderson P. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem Soc Trans*, **30**:963–969 (2002).
169. McDonald K.K., Aulas A., Destroismaisons L. *et al*. TAR DNA-binding protein 43 (TDP-43 regulates stress granule dynamics via differential regulation of G3BP and TIA-1. *Hum Mol Genet*, **20**:1400–1410 (2011).
170. Colombrita C., Zennaro E., Fallini C. *et al*. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem*, **111**:1051–1061 (2009).
171. Aulas A., Stabile S., Vande Velde C. Endogenous TDP-43, but not FUS, contributes to stress granule assembly via G3BP. *Mol Neurodegener*. **24**;7:54 (2012).
172. Anderson P., Kedersha N. Stress granules. *Curr Biol* **19**:R397–R398 (2009).
173. Dormann D., and Haass C. TDP-43 and FUS: a nuclear affair. *Trends Neurosci*. **34**, 339–348 (2011).
174. Bose J.K., Huang C.C., Shen C.K. Regulation of autophagy by neuropathological protein TDP-43. *J Biol Chem* **286**:44441–44448 (2011).
175. Swarup V., Phaneuf D., Dupré N. *et al.* Deregulation of TDP-43 in amyotrophic lateral

sclerosis triggers nuclear factor κB-mediated pathogenic pathways. *J Exp Med* **208**:2429–2447 (2011).

176. Kuo P.H., Chiang C.H., Wang Y.T. *et al.* The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res* **42**:4712–4722 (2014).
177. Lalmansingh A.S., Urekar C.J., Reddi P.P. TDP-43 is a transcriptional repressor: The testis-specific mouse acrv1 gene is a TDP-43 target in vivo. *J Biol Chem* **286**:10970–10982 (2011).
178. Shiloh Y., Ziv Y. The ATM protein kinase: Regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* **14**:197–210 (2013).
179. Barnes D.E. Non-homologous end joining as a mechanism of DNA repair. *Curr Biol* **11**:R455–R457 (2001).
180. Lieber M.R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**:181–211 (2010).
181. Mitra J., Guerrero E.N., Hegde P.M. Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. Proc Natl Acad Sci U S A. pii: 201818415. (2019)
182. Yu Z, Fan D., Gui B. *et al.* Neurodegeneration-associated TDP-43 interacts with fragile X mental retardation protein (FMRP)/Staufen (STAU1) and regulates SIRT1 expression in neuronal cells. *J Biol Chem* **287**:22560–22572 (2012).
183. Hill S.J., Mordes D.A., Cameron L.A. *et al*. Two familial ALS proteins function in prevention/repair of transcription-associated DNA damage. *Proc Natl Acad Sci USA* **113**:E7701–E7709 (2016).
184. Freibaum B.D., Chitta R.K., High A.A. *et al*. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J Proteome Res* **9**:1104–1120 (2010).
185. Guerrero E.N., Mitra J., Wang H. Amyotrophic lateral sclerosis-associated TDP-43 mutation Q331K prevents nuclear translocation of XRCC4-DNA ligase 4 complex and is linked to genome damage-mediated neuronal apoptosis. *Hum Mol Genet*. pii: ddz062 (2019).
186. Sreedharan J., Blair I.P., Tripathi V.B. *et al*. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*, **319**, 1668–1672 (2008).
187. Lim L., Wei, Y. Lu Y. and Song J. ALS-causing mutations significantly perturb the self-assembly and interaction with nucleic acid of the intrinsically disordered prion-like domain

of TDP-43. *PLoS Biol*., **14**, e1002338 (2016).

188. Wang H., Guo W., Mitra J. *et al*. Mutant FUS causes DNA ligation defects to inhibit oxidative damage repair in amyotrophic lateral sclerosis. *Nat. Commun*., **9**, 3683 (2018).
189. Reaban M.E., Lebowitz J., Griffin J.A. Transcription induces the formation of a stable RNA.DNA hybrid in the immunoglobulin a switch region. *J Biol Chem* **269**: 21850–21857 (1994).
190. Li X., Manley J.L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122**: 365–378 (2005).
191. Ginno P.A., Lott P.L., Christensen H.C. et al. Rloop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* **45**: 814–825 (2012).
192. Ginno P.A., Lim Y.W., Lott P.L. *et al.* GC skew at the 59 and 39 ends of human genes links R-loops formation to epigenetic regulation and transcription termination. *Genome Res* **23**: 1590–1600 (2013).
193. Sugimoto N., Nakano S., Katoh M. *et al.* Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **34**: 11211–11216 (1995).
194. Aguilera A., Garcia-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell* **46**: 115–124 (2012).
195. Westover K.D., Bushnell D.A., Kornberg R.D. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**: 481–489 (2004).
196. Roy D., Lieber M.R. G clustering is important for the initiation of transcription induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter. *Mol Cell Biol* **29**: 3124–3133 (2009).
197. Roy D., Zhang Z., Lu Z. *et al.* Competition between the RNA transcript and the nontemplate DNA strand during R-loops formation in vitro: a nick can serve as a strong R-loops initiation site. *Mol Cell Biol* **30**: 146–159 (2010).
198. Roberts R.W., Crothers D.M. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* **258**: 1463–1466 (1992).
199. Shaw N.N., Arya D.P. Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* **90**: 1026–1039 (2008).
200. Masukata H., Tomizawa J. Effects of point mutations on formation and structure of the RNA primer for ColE1 DNA replication. *Cell* **36**: 513–522 (1984).
201. Baker T.A., Kornberg A. Transcriptional activation of initiation of replication from the E. coli chromosomal origin: an RNA–DNA hybrid near oriC. *Cell* **55**: 113–123 (1988).
202. Carles-Kinch K., Kreuzer K.N. RNA–DNA hybrid formation at a bacteriophage T4

replication origin. *J Mol Biol* **266**: 915–926 (1997).

203. Xu B., Clayton D.A. RNA–DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA–DNA hybrids serving as primers. *EMBO J* **15**: 3135–3143 (1996).
204. Yu K., Chedin F., Hsieh C.L. *et al.* R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* **4**: 442–451 (2003).
205. Duquette A., Roddier K., McNabb-Baltar J. *et al*. Mutations in senataxin responsible for Quebec cluster of ataxia with neuropathy. *Ann Neurol* **57**: 408–414 (2005).
206. Bhatia V., Herrera-Moyano E., Aguilera A. *et al.* The Role of Replication-Associated Repair Factors on R-loops. *Genes* **8**:7-171 (2017).
207. Cerritelli S.M., Crouch R.J. Ribonuclease H: the enzymes in eukaryotes. *FEBS J* **276**: 1494–1505 (2009).
208. Cerritelli S.M., Frolova E.G., Feng C. *et al*. Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol Cell* **11**: 807–815 (2003).
209. Mischo H.E., Gomez-Gonzalez B., Grzechnik P. *et al.* Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell* **41**: 21–32 (2011).
210. Skourti-Stathaki K., Proudfoot N.J., Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* **42**: 794–805 (2011).
211. Chakraborty P., Grosse F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair* **10**: 654–665 (2011).
212. Becherel O.J., Yeo A.J., Stellati A. *et al.* Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet* **9**: e1003435 (2013).
213. Yu K., Chedin F., Hsieh C.L. *et al*. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* **4**: 442–451 (2003).
214. McKinnon P.J. DNA repair deficiency and neurological disease. *Nat Rev Neurosci* **10**: 100–112 (2009).
215. James P.A., Talbot K. The molecular genetics of non-ALS motor neuron diseases. *Biochim Biophys Acta* **1762**: 986– 1000 (2006).
216. Palau F., Espinos C. Autosomal recessive cerebellar ataxias. *Orphanet J Rare Dis* **1**: 47 (2006).
217. El Hage A., French S.L., Beyer A.L. *et al*. Loss of topoisomerase I leads to R-loops-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev* **24**: 1546–

1558 (2010).

218. Yang Y., McBride K.M., Hensley S. *et al.* Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol Cell* **53**: 484–497 (2014).
219. Dominguez-Sanchez M.S., Barroso S., Gomez-Gonzalez B. *et al*. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet* **7**: e1002386 (2011).
220. Paulsen R.D., Soni D.V., Wollman R. *et al.* A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell* **35**: 228–239 (2009).
221. Baumann P., Benson F.E., West S.C. Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell* **87**: 757–766 (1996).
222. Wahba L., Gore S.K., Koshland D. The homologous recombination machinery modulates the formation of RNA– DNA hybrids and associated chromosome instability. *eLife* **2**: e00505 (2013).
223. Chen Y.-H., Keegan S., Kahli M. *et al.* Transcription shapes DNA replication initiation and termination in human cells. Nat. Struct. *Mol. Biol*. **26**, 67-77 (2019).
224. Helmrich A., Ballarino M., Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell* **44**: 966–977 (2011).
225. Debatisse M., Le Tallec B., Letessier A. *et al.* Common fragile sites: mechanisms of instability revisited. *Trends Genet* **28**: 22–32 (2012).
226. Castellano-Pozo M., Santos-Pereira J.M., Rondon A.G. *et al*. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol Cell* **52**: 583–590 (2013).
227. Madireddy A., Kosiyatrakul S.T., Boisvert R.A. *et al*. FANCD2 facilitates replication through common fragile sites. *Mol. Cell* **64**, 388-404 (2016).
228. Minocherhomji S., Ying S., Bjerregaard V.A. *et al*. Replication stress activates DNA repair synthesis in mitosis. *Nature* **528**, 286-290 (2015).
229. Chan K.L., Palmai-Pallag T., Ying S. *et al*. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat. Cell Biol*. **11**, 753-760 (2009).
230. Santos-Pereira J.M., Aguilera A. R loops: newmodulators of genome dynamics and function. *Nat Rev Genet* **16**:583–597 (2015).
231. D'Alessandro G., d'Adda di Fagagna F. Transcription and DNA damage: holding hands or crossing swords? *J Mol Biol* **429**:3215–3229 (2017).
232. Richard P., Manley J.L. R loops and links to human disease. *J Mol Biol* **429**:3168–3180

(2017).

233. Duquette M.L., Handa P., Vincent J.A. *et al.* Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* **18**:1618–1629 (2004).
234. Skourti-Stathaki K., Proudfoot N.J., Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* **42**: 794–805 (2011).
235. Byrd A.K., Zybailov B.L., Maddukuri L. Evidence That G-quadruplex DNA Accumulates in the Cytoplasm and Participates in Stress Granule Assembly in Response to Oxidative Stress. *J Biol Chem*. **291(34)**:18041-57 (2016).
236. Salvi J.S., Mekhail K. R-loops highlight the nucleus in ALS. *Nucleus* **6**:23–29 (2015).
237. Mischo H.E., Gomez-Gonzalez B., Grzechnik P. *et al.* Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell* **41**: 21–32 (2011).
238. Suraweera A., Becherel O.J., Chen P. *et al*. Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *J Cell Biol* **177**: 969–979 (2007).
239. Yuce O., West S.C. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. *Mol Cell Biol* **33**: 406–417 (2013).
240. Becherel O.J., Yeo A.J., Stellati A. *et al*. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet* **9**: e1003435 (2013).
241. Vantaggiato C., Bondioni S., Airoldi G. *et al.* Senataxin modulates neurite growth through fibroblast growth factor 8 signalling. *Brain* **134**: 1808–1828 (2011).
242. Yeo A.J., Becherel O.J., Luff J.E. *et al*. R-loops in Proliferating Cells but Not in the Brain: Implications for AOA2 and Other Autosomal Recessive Ataxias. *PLoS ONE* **9**: e90219 (2014).
243. Powell W.T., Coulson R.L., Gonzales M.L. *et al*. R-loops formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proc Natl Acad Sci U S A* **110**:13938–13943 (2013).
244. Haeusler A.R., Donnelly C.J., Periz G. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**:195-200 (2014).
245. Fratta P., Mizielinska S., Nicoll A.J. *et al*. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes.

Sci Rep **2**:1016 (2012).

246. Haeusler A.R., Donnelly C.J., Periz G. *et al*. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**:195-200 (2014).
247. Wang J., Haeusler A.R., Simko E.A. Emerging role of rna-dna hybrids in c9orf72-linked neurodegeneration. *Cell Cycle* **14:**526–532 (2015).
248. Groh M., Gromak N. Out of balance: R-loops in human disease. *PLoS Genet* **10**:e1004630 (2014).
249. Liu E.Y., Russ J., Wu K. *et al.* C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/ FTD. *Acta Neuropathol* **128**:525–541 (2014).
250. Russ J., Liu E.Y., Wu K. *et al*. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathol* **129**:39–52 (2015).
251. Salvi J.S., Chan J.N., Szafranski K. *et al*. Roles for Pbp1 and Caloric Restriction in Genome and Lifespan Maintenance via Suppression of RNA-DNA Hybrids. *Dev Cell* **30**:177-91 (2014).
252. Farg M.A., Soo K.Y., Warraich S.T. *et al*. Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. *Hum Mol Genet* **22**:717-28 (2013).
253. Elden A.C., Kim H.J., Hart M.P. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**:1069-75 (2010).
254. Becker L.A., Huang B., Bieri G. *et al*. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature* **544**:367–371 (2017).
255. Van Blitterswijk M.,Mullen B., Heckman M.G. *et al*. Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. *Neurobiol Aging* **35**:2421.e13–2421.e17 (2014).
256. Shaw N.N., Arya D.P. Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* **90**: 1026–1039 (2008).
257. Shaw N.N., Xi H., Arya D.P. Molecular recognition of a DNA:RNA hybrid: sub-nanomolar binding by a neomycin-methidium conjugate. *Bioorg Med Chem Lett* **18**: 4142–4145 (2008).
258. Sordet O., Redon C.E., Guirouilh-Barbat J. *et al*. Ataxia telangiectasia mutated activation by transcription- and topoisomerase Iinduced DNA double-strand breaks. *EMBO Rep* **10**: 887–893 (2009).
259. Powell W.T., Coulson R.L., Gonzales M.L. *et al*. Rloop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. *Proc Natl Acad Sci U S A* **110**: 13938–13943 (2013).

260. McIvor E.I., Polak U., Napierala M. New insights into repeat instability: role of RNA/DNA hybrids. *RNA Biol* **7**: 551–558 (2010).
261. Colak D., Zaninovic N., Cohen M.S. *et al.* Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* **343(80):**1002–1005 (2014).
262. Walker C., Herranz-Martin S., Karyka E. *et al.* C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nat Neurosci* **20**:1225–1235 (2017).
263. Jangi M., Fleet C., Cullen P. *et al.* SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. *Proc Natl Acad Sci* **114**:E2347–E2356 (2017).
264. Sipova H., Zhang S., Dudley A.M. *et al*. Surface plasmon resonance biosensor for rapid label-free detection of microribonucleic acid at subfemtomole level. *Anal Chem* **82**: 10110–10115 (2010).
265. Qavi A.J., Kindt J.T., Gleeson M.A. *et al.* Anti-DNA:RNA antibodies and silicon photonic microring resonators: increased sensitivity for multiplexed microRNA detection. *Anal Chem* **83**: 5949–5956 (2011).
266. Matsuoka S., Ballif B. A., Smogorzewska A. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
267. Paulsen R.D.,Soni D.V.,Wollman R. *et al*. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol. Cell* **35**, 228–239 (2009).
268. Rajesh C.,Baker D.K.,Pierce A. J. *et al*. The splicing-factor related protein SFPQ/PSF interacts with RAD51D and is necessary for homology-directed repair and sister chromatid cohesion. *Nucleic Acids Res*. **39**, 132–145 (2011).
269. Sancar A.,Lindsey-Boltz L.A., Unsal-Kaçmaz K. *et al*. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* **73**, 39–85 (2004).
270. Adamson B.,Smogorzewska A.,Sigoillot F.D. *et al*. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat. Cell Biol.* **14**, 318–328 (2012).
271. Huertas P. and Aguilera A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711–721 (2003).
272. Guerrero E.N., Mitra J., Wang H. *et al.* Amyotrophic lateral sclerosis-associated TDP-43

mutation Q331K prevents nuclear translocation of XRCC4-DNA ligase 4 complex and is linked to genome damage-mediated neuronal apoptosis. *Hum Mol Genet*. pii: ddz062 (2019).

273. Sherman M.H.,Bassing C.H., and Teitell M.A. Regulation of cell differentiation by the DNA damage response. *Trends Cell Biol.* **21**, 312–319 .(2011).
274. Sherman M.H.,Kuraishy A.I.,Deshpande C. *et al.* AID- induced genotoxic stress promotes B cell differentiation in the germinal center via ATM and LKB1 signaling. *Mol. Cell* **39**, 873–885 (2010).
275. Roy D.,Yu K.,and Lieber M.R. Mechanism of R-loops formation at immunoglobulin class switch sequences. *Mol. Cell. Biol.* **28**, 50–60 (2008).
276. Chaudhuri J.,Basu U.,Zarrin A., *et al.* Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* **94**, 157–214 (2007).
277. Park S. R. Activation induced cytidine deaminase in B cell immunity and cancers. *Immune Netw.* 12, 230–239 (2012).
278. Huang C.C., Bose J.K., Majumder P. *et al.* Metabolism and mismetabolism of the neuropathological signature protein TDP-43. J Cell Sci 127:3024–3038 (2014).
279. Tsuiji H., Iguchi Y., Furuya A. *et al.* Spliceosome integrity is defective in the motor neuron diseases ALS and SMA. EMBO Mol Med 5:221–234 (2013).
280. Wang I.F., Guo B.S., Liu Y.C. *et al.* Autophagy activators rescue and alleviate pathogenesis of a mouse model with proteinopathies of the TAR DNA-binding protein 43. *Proc Natl Acad Sci U S A* **109**: 15024–15029 (2012).
281. Udan-Johns M., Bengoechea R., Bell S. *et al*. Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. *Hum Mol Genet* **23**:157–170 (2014).
282. King O.D., Gitler A.D., Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* **1462**:61–80 (2012).
283. Kim H.J., Raphael A.R., La Dow E.S. *et al.* Therapeutic modulation of eIF2alpha phosphorylation rescues TDP-43 toxicity in amyotrophiclateral sclerosis disease models. *Nat Genet* **46**:152–160 (2014).
284. Chen Y.Z., Bennett C.L., Huynh H.M. *et al.* DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet* **74**:1128–1135 (2004).
285. Groh M., Albulescu L.O., Cristini A. *et al.* Senataxin: genome guardian at the interface of transcription and neurodegeneration *J Mol Biol* **429**:3181–3195 (2017).
286. Freudenreich C.H. R-loops: targets for nuclease cleavage and repeat instability. *Curr Genet*

*:***1**-6 (2018)

287. Reddy K., Schmidt M.H.M., Geist J.M. *et al*., Processing of double R-loops in (CAG)•(CTG) and C 9 o r f 72 (GGGGCC)•(GGCCCC) repeats causes instability. *Nucleic Acids Res* **42**:10473–10487 (2014).
288. Jangi M., Fleet C., Cullen P. *et al*., SMN deficiency in severe models of spinal muscular atrophy causes widespread intron retention and DNA damage. *Proc Natl Acad Sci* **114**:E2347–E2356 (2017).
289. Sorrells S., Nik S., Casey M. *et al*., Spliceosomal components protect embryonic neurons from R-loops-mediated DNA damage and apoptosis. *Dis Model Mech* **11**:dmm.031583 (2018).
290. Dominguez-Sanchez M.S., Barroso S., Gomez-Gonzalez B. *et al*. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet* **7(12):**e1002386. (2011)
291. Scheffler J.M., Schiefermeier N., Huber L.A. Mild fixation and permeabilization protocol for preserving structures of endosomes, focal adhesions, and actin filaments during immunofluorescence analysis. *Methods Enzymol*. **535**:93-102 (2014).
292. García-Rubio M.L., Pérez-Calero C., Barroso S.I. The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. *PLoS Genet.* **11(11):**e1005674 (2015).
293. Vanderweyde T. *et al*. Role of stress granules and RNA-binding proteins in neurodegeneration: a mini-review. *Gerontology*. **59**(6):524-33 (2013).
294. Halász L., Karányi Z, Boros-Oláh B. *et al*. RNA-DNA Hybrid (R-loops) Immunoprecipitation Mapping: An Analytical Workflow to Evaluate Inherent Biases. *Genome Research* **27**:6, 1063–1073 (2017).
295. Sanz L.A., Chédin F. High-resolution, strand-specific R-loops mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. *Nat Protoc*. **14(6):**1734-1755 (2019).
296. Perego M.G.L. *et al.* R-Loops in Motor Neuron Diseases. *Mol Neurobiol*. **56**(4):2579-2589 (2019).
297. Yang M.H., Chen K.C., Chiang P.W. *et al.* Proteomic Profiling of Neuroblastoma Cells Adhesion on Hyaluronic Acid-Based Surface for Neural Tissue Engineering. *Biomed Res Int.* **2016**:1917394 (2016).
298. Krishna A., Biryukov M., Trefois C. *et al.* Systems genomics evaluation of the SH-SY5Y neuroblastoma cell line as a model for Parkinson's disease *BMC Genomics*, **15**:1154, (2014).

299. Pesiridis G.S., Lee V.M., Trojanowski J.Q. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. **15**;18(R2), R156-62 (2009).
300. Bossolasco P., Sassone F., Gumina V. Motor neuron differentiation of iPSCs obtained from peripheral blood of a mutant TARDBP ALS patient. *Stem Cell Res*. **30**:61-68 (2018).
301. Bhatia V., Barroso S.I., García-Rubio M.L. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**: 362-5 (2014).
302. Mersaoui S.Y., Yu Z., Coulombe Y. *et al.* Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. *EMBO J.* **21**:e100986 (2019).
303. Ginisty H. *et al*. Structure and functions of nucleolin. *J. Cell Sci*. **112**, 761–772 (1999)
304. Angelov D. *et al.* Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes *EMBO J*. **25**, 1669–1679 (2006)
305. Yang T. H. *et al.* Purification and characterization of nucleolin and its identification as a transcription repressor. *Mol. Cell. Biol*. **14**, 6068–6074 (1994).
306. He T. C. *et al*. Identification of c-MYC as a target of the APC pathway.Science 281,1509–1512 (1998).
307. Dempsey L. A. *et al*. G4 DNA Binding by LR1 and Its Subunits, Nucleolin and hnRNP D, A Role for G-G pairing in Immunoglobulin Switch RecombinationJ. *Biol. Chem*. **274**, 1066–1071 (1999)
308. Hanakahi L. A. *et al*. High Affinity Interactions of Nucleolin with G-G-paired rDNA J. *Biol. Chem*. **274**,15908–15912(1999).
309. Wang I.X. *et al*. Human proteins that interact with RNA/DNA hybrids. *Genome Res*. **28**(9):1405-1414 (2018).
310. Tsoi H. *et al*. CAG expansion induces nucleolar stress in polyglutamine diseases. *Proc Natl Acad Sci U S A* **109**(33): 13428-13433 (2012).
311. Hartono S.R. *et al*. The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *J Mol Biol*. **2**;430(3):272-284 (2018).
312. Phillips-Cremins J.E., Sauria M.E.G., Sanyal A. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell*. **153**:1281–1295 (2013).
313. Kottemann M.C., Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature*. **493**: 356–63 doi (2013).
314. Ginno PA, Lim YW, Lott PL, Korf I, Chedin F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination.

Genome Res. **23**: 1590–600 (2013).

315. Herrera-Moyano E., Mergui X., Garcia-Rubio M.L. *et al*. The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts. *Genes & development*. **28**: 735–48 (2014).
316. Tank E.M. *et al.* Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat Commun*. **20**;9(1):2845 (2018).
317. Zhang Z.Z., Pannunzio N.R., Hsieh C.L., Yu K., Lieber M.R. and Michael R. Complexities due to single-stranded RNA during antibody detection of genomic rna:dna hybrids, *BMC Research Notes* **8**, 127 (2015)
318. König F., Schubert T., Längst G. The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. PLoS One 12(6):e0178875 (2017).
319. Ayala Y.M. et al., TDP-43 regulates retinoblastoma protein phosphorylation through the repression of cyclin-dependent kinase 6 expression. *Proc Natl Acad Sci U S A*. **11**;105(10):3785-9 (2008).
320. Phillips D.D., Garboczi D.N., Singh K., Hu Z., Leppla S.H., Leysath C.E. The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. *J. Mol. Recognit*.**26**:376–381 (2013).
321. Boque-Sastre R., Soler M., Guil S. Detection and Characterization of R loop Structures, *Methods Mol Biol*. **1543**, 231-24 (2017)
322. Mazroui R. *et al.* Inhibition of Ribosome Recruitment Induces Stress Granule Formation Independently of Eukaryotic Initiation Factor 2α Phosphorylation, *Mol Biol Cell* **17**(10), 4212-9, (2006)
323. Cestra G. *et al*., Control of mRNA Translation in ALS Proteinopathy. *Front Mol Neurosci*. **23**;10:85 (2017).
324. Prpar Mihevc S. *et al.* TDP-43 aggregation mirrors TDP-43 knockdown, affecting the expression levels of a common set of proteins. *Sci Rep*. **26**;6:33996 (2016).
325. Crippa V. *et al.* The chaperone HSPB8 reduces the accumulation of truncated TDP-43 species in cells and protects against TDP-43-mediated toxicity. *Hum Mol Genet*. **15**;25(18), 3908-3924 (2016)
326. Nishimoto Y., Ito D., Yagi T., Nihei Y., Tsunoda Y., Suzuki N. Characterization of alternative isoforms and inclusion body of the TAR DNA-binding protein-43, *J Biol Chem*, **285**, 608 – 619 (2010).
327. Emde A. *et al*. Dysregulated miRNA biogenesis downstream of cellular stress and ALS-causing mutations: a new mechanism for ALS. *The EMBO journal*. **34**(21):2633–51

(2015).

328. Wobst H.J. *et al*. Truncation of the TAR DNA-binding protein 43 is not a prerequisite for cytoplasmic relocalization, and is suppressed by caspase inhibition and by introduction of the A90V sequence variant. *PLoS One.* **16**;12(5):e0177181 (2017).