Chapter 10
Senataxin, A Novel Helicase at the Interface of RNA Transcriptome Regulation and Neurobiology: From Normal Function to Pathological Roles in Motor Neuron Disease and Cerebellar Degeneration

Craig L. Bennett and Albert R. La Spada

Abstract Senataxin (SETX) is a DNA-RNA helicase whose C-terminal region shows homology to the helicase domain of the yeast protein Sen1p. Genetic discoveries have established the importance of SETX for neural function, as recessive mutations in the *SETX* gene cause Ataxia with Oculomotor Apraxia type 2 (AOA2) (OMIM: 606002), which is the third most common form of recessive ataxia, after Friedreich's ataxia and Ataxia-Telangiectasia. In addition, rare, dominant *SETX* mutations cause a juvenile-onset form of Amyotrophic Lateral Sclerosis (ALS), known as ALS4. SETX performs a number of RNA regulatory functions, including maintaining RNA transcriptome homeostasis. Over the last decade, altered RNA regulation and aberrant RNA-binding protein function have emerged as a central theme in motor neuron disease pathogenesis, with evidence suggesting that sporadic ALS disease pathology may overlap with the molecular pathology uncovered in familial ALS. Like other RNA processing proteins linked to ALS, the basis for SETX gain-of-function motor neuron toxicity remains ill-defined. Studies of yeast

C. L. Bennett
Department of Neurology, Duke University School of Medicine, Durham, NC, USA

Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA

Department of Cell Biology, Duke University School of Medicine, Durham, NC, USA

A. R. La Spada (✉)
Duke Center for Neurodegeneration & Neurotherapeutics, Duke University School of Medicine, Durham, NC, USA

Department of Neurology, Duke University School of Medicine, Durham, NC, USA

Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA

Department of Cell Biology, Duke University School of Medicine, Durham, NC, USA
e-mail: al.laspada@duke.edu

R. Sattler, C. J. Donnelly (eds.), *RNA Metabolism in Neurodegenerative Diseases*, Advances in Neurobiology 20,
https://doi.org/10.1007/978-3-319-89689-2_10

Sen1p and mammalian SETX protein have revealed a range of important RNA regulatory functions, including resolution of R-loops to permit transcription termination, and RNA splicing. Growing evidence suggests that *SETX* may represent an important genetic modifier locus for sporadic ALS. In cycling cells, SETX is found at nuclear foci during the S/G_2 cell-cycle transition phase, and may function at sites of collision between components of the replisome and transcription machinery. While we do not yet know which SETX activities are most critical to neurodegeneration, our evolving understanding of SETX function will undoubtedly be crucial for not only understanding the role of SETX in ALS and ataxia disease pathogenesis, but also for delineating the mechanistic biology of fundamentally important molecular processes in the cell.

Keywords Senataxin · Helicase · R-Loops · Nuclear exosome · RENT1 · IGHMBP2 · Sen1p · Exosc9 · Sumo · Nucleolus · tRNA

Senataxin (SETX) is now recognized as an important protein in the fields of molecular genetics and neurodegeneration. *SETX* gene mutations lead to two distinct neurological disorders, Ataxia with Oculomotor Apraxia type 2 (AOA2) and Amyotrophic Lateral Sclerosis type 4 (ALS4). AOA2 has uniformly early onset and leads to very severe disability, requiring life-long care [1]. ALS4 is moderate to severe with varying age of onset, but with its average onset at 17-years, ALS4 is considered a juvenile-onset form of familial ALS (FALS) [2]. The *SETX* gene has also attracted recent attention as a potential genetic modifier of sporadic ALS (SALS) [3].

The effort to define key functions of SETX continues as its roles in RNA processing and maintenance of genomic stability are now well established by the molecular genetics community. In early studies, the *SETX* gene was found to be ubiquitously expressed [4, 5], and many functional processes eventually attributed to SETX were originally described for its yeast orthologue, Sen1p. These SETX functions include: (1) RNA polymerase II (RNAP II) transcription termination; (2) the resolution of RNA/DNA hybrids, or R-Loops; (3) processing of noncoding RNAs and mRNA; (4) interaction with the nuclear exosome; and (5) the formation of replication stress-related foci during the S/G_2 transition phase. This last function suggests that SETX may be essential for cell cycling when long genes are being transcribed and RNAP II collides with the replisome. Other important roles for SETX that deserve attention include the regulation of the circadian rhythm genes, *Period* (PER) and *Cryptochrome* (CRY) [6]. Whether these functions are critical to neuron survival is unclear, but highlight the fact that the full spectrum of cellular processes for this helicase is extremely wide-ranging.

In this review, we attempt to clarify the many processes attributed to SETX, and evaluate if SETX gain-of-function toxicity impacts these functions and how this might contribute to motor neuron disease. Detailed proteomics studies have shown that human SETX, like Sen1p, has retained regulatory functions during gene transcription. Interactome analysis of purified TAP-tagged Sen1p identified the RNAP

II subunit Rpo21, along with subunits Rpb2 and Rpb4, as key interactors [7], confirming earlier yeast two-hybrid studies [8]. Furthermore, a number of RNAP I and RNAP III core subunits, elongation factors, and other key components were identified as protein interactors of Sen1p [7]. In contrast, when similar analyses were performed with SETX, no direct interaction with RNAP II was observed, but rather an enrichment within the chromatin fraction, and interaction with RNAP II-related core factors [7]. Importantly, these studies were undertaken with Flag/GFP-tagged SETX, used with HeLa cell stable integration, and at near endogenous levels of protein expression [7]. Hence, human SETX likely retains transcription-related functions, but the regulatory relationships may be quite different in comparison with budding yeast.

In regards to both SETX and Sen1p function, it is crucial to emphasize that these proteins are present at very low abundance; hence, overexpression can lead to aberrant cellular events. Gene duplication can uncover such sensitivity to protein levels as dosage-sensitivity [9, 10]. For example, increased levels of ataxin-2 can distort the cytoplasmic-to-nuclear ratio of TDP-43 and FUS proteins [11], both proteins that are critical in ALS disease. Sen1p is known to be maintained at very low levels in the cell, typically as low as 125 molecules per cell [12], which is much less than its known transcription termination partners, Nrd1p and Nab3p, present at ~19,000 and ~5800 molecules per cell respectively [13]. RNAP II is itself present in yeast at ~14,000 molecules/cell [14, 15]. In a detailed proteomics study using the human U2OS cell line, ~10,000 different proteins were quantified and found to span a concentration range of seven orders of magnitude up to 20,000,000 copies per cell [16]. From this study SETX was found to be in the lowest category of very low-abundant proteins at <500 molecules/cell [16]. Many SETX functional studies have been undertaken with tagged-constructs which will almost assuredly lead to cellular levels of recombinant SETX that are grossly elevated, which means that results of such studies must be interpreted with caution.

10.1 SETX Mutations Cause Both Ataxia with Oculomotor Apraxia and Motor Neuron Disease

Causal links between SETX protein defects and neurological disease were first reported in 2004. We discovered *SETX* gene mutations as the cause of ALS4, a rare, dominantly inherited, juvenile onset form of ALS (OMIM: 602433). Importantly, all 49 affected members of an extended American pedigree were found to carry a L389S mutation [4]. Other sizable European pedigrees were found to segregate R2136H and T3I *SETX* mutations. The L389S mutation was subsequently found in Italian and Dutch pedigrees, confirming the pathogenicity of this mutation in ALS4 [17, 18]. The phenotype of ALS4 is unique compared with classical ALS due to a number of factors, including normal patient life expectancy due to the sparing of the respiratory musculature; absence of bulbar involvement; and the presentation of symmetrical atrophy and weakness [2].

Recessive *SETX* mutations were also reported in 2004 as the cause of a severe Ataxia with Oculomotor Apraxia—type 2 (AOA2; OMIM: 606002) [5]. While dominant mutations are rare, *SETX* recessive mutations are not nearly so rare. AOA2 is considered the third most common autosomal recessive cerebellar ataxia [1, 5], with the most common being Friedreich's ataxia, closely followed by Ataxia-Telangiectasia (A-T). One unique feature shared between AOA2 and A-T is elevated serum levels of alpha-feto-protein (AFP), which are ~9-fold higher in AOA2 patients than normals [19]. A-T is associated with unique DNA repair defects and extra-neurologic features, including a greatly increased cancer risk and immune defects (OMIM: 208900). With AOA2 patients, some groups have reported sensitivity to DNA-damaging agents in patient cells [20], yet others report normal sensitivity [21, 22]. Nonetheless, AOA2 patients do not display an increased cancer risk nor immunological abnormalities, and thus the above features form a differential diagnosis for A-T versus AOA2. AOA2 patients show non-cancer-related extra-neurologic features, such as ovarian failure [23], suggesting tissues other than just the central nervous system (CNS) are susceptible to SETX loss. At the molecular level, Moreira et al. initially reported 15 different *SETX* mutations, ten of which predict premature protein termination. Thus, parental carriers of AOA2 null mutations were carefully examined and found to harbor no neurological phenotypes [1, 5]. Importantly, this suggests that ALS4 dominant mutations possess toxic gain-of-function properties. Now that greater than 150 different *SETX* mutations have been identified to date [24], it is known that missense mutations cluster within either the helicase domain or the amino-terminal domain, confirming the critical nature of these two protein regions [25].

It is of note that within the human genome, only two human proteins exist with significant homology within the helicase domain to SETX: RENT1 (46% similarity) and IGHMBP2 (45% similarity) (Fig. 10.1). RENT1 (the yeast Upf1 orthologue) is an essential component of the nonsense-mediated RNA decay (NMD) pathway for degrading incorrectly spliced or stop codon-containing mRNAs [26], and *Rent1* null mice show embryonic lethality [27], attesting to the importance of this pathway for transcriptome stability. Over-expression of RENT1 (hUpf1) can significantly rescue toxicity in ALS/FTD cell models (TDP-43/FUS) that are likely induced by uncharacterized RNA dysfunction [28]. Recessive mutations of the *IGHMBP2* gene cause a severe spinal muscle atrophy with respiratory distress (SMARD) [29]. This disease has overlap with AOA2 and ALS4, which will be discussed below. Thus, these three human helicase homologs cause either embryonic lethality (Rent1)—due to failure of NMD surveillance of mRNA splicing errors, or specific neuronal vulnerability and neurodegeneration as loss-of-function mutations (SETX and IGHMBP2) or gain-of-function mutations (SETX). Given the known low cellular levels of SETX protein, one could speculate that trace levels of SETX, produced by alternate splicing, might prevent lethality in the human, akin to the minimal levels of normal survival motor neuron protein produced by the SMN2 gene in spinal muscular atrophy [30]. In regards to IGHMBP2, human mutations are known to be homozygous hypomorphic loss-of-function alleles [29], in agreement with the naturally occurring mouse model of SMARD, where ~20% of correctly spliced *Ighmbp2* mRNA is produced [31].

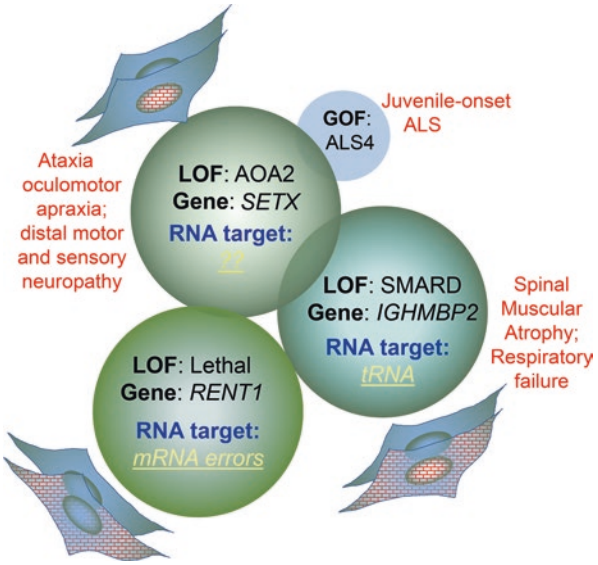

Fig. 10.1 SETX, RENT1, and IGHMBP2 form a helicase subfamily. While there are nearly 100 helicase proteins in the human genome, only two show significant homology to SETX. These three related homologues each perform critical roles in RNA processing, and are essential for survival or for normal central nervous system function. RENT1 is unique in that no human disease results from recessive mutation, likely due to embryonic lethality, and displays a cytosolic localization. The specific RNA targets linked to dysfunction or lethality are known for IGHMBP2 and RENT1, and are listed. In contrast, the specific RNA targets for SETX are unknown, and how altered RNA regulation results in AOA2 or ALS4 remains an open question. SETX resides in the nucleus, while IGHMBP2 localizes to both the nucleus and the cytosol, as shown

10.2 SETX Function: Similarities and Differences with Yeast Sen1p

Soon after *SETX* mutations were discovered, it was noted that SETX and Sen1p shared a highly conserved C-terminal helicase domain. Researchers hypothesized that key functions attributed to Sen1p would likely be retained by human SETX. The *Sen1* gene was named for its suspected function as a splicing endonuclease, but was thereafter found to function primarily in processing a diverse class of noncoding RNAs (ncRNAs) [32, 33]. In further studies, Sen1p was shown to process intron-containing tRNA precursors [34], rRNA precursors [33], 3′-extended forms of some small nucleolar RNAs (snoRNAs), and to a lesser extent, small nuclear RNAs (snRNAs) [33, 35, 36]. Sen1p directly interacts with the RNA-binding proteins Nrd1 and Nab3 [36, 37], and was confirmed as a functioning component of the NRD complex in transcription termination of RNAP II ncRNA transcripts [38, 39]. Taken together, years of dedicated yeast research portray a complex picture of RNA processing, which is essential to yeast survival and directed primarily to the regulation of ncRNAs, as well as a central role in RNAP II transcription termination that may be unique to yeast.

Early SETX work yielded divergent and contradictory views as to what might be its key disease-linked protein function [40]. Readers new to the field could easily be left confused by perusing summaries of the very broad array of proposed SETX functions. An example of one early discrepancy was whether cells lacking SETX were sensitive to DNA-damaging agents. A second function attributed to SETX is a generalized regulation of RNAP II transcription termination. While some reports strongly favor a role for SETX in mediating Xrn2-dependent transcription termination via the formation of R-loops (which are DNA-RNA hybrids formed during the process of transcription) [41], in much the same way as Sen1p, others have concluded that decreased Xrn2 or SETX levels yield only marginal effects on the regulation of transcription termination [42]. However, it is clear that for certain genes and biological processes, SETX is central to transcription termination regulation, as SETX has been shown to control the cyclic expression of the circadian rhythm genes PER and CRY [6]. This work has yielded a model in which recruitment of PER complexes to the elongating polymerase at *Per* and *Cry* termination sites inhibits SETX action, impeding RNAP II release and thereby repressing transcription re-initiation. A third SETX function, reported to be critical for neuron survival, is the resolution of RNAP II-mediated R-loops more generally (independent of termination), which will be discussed later. Fourth, a role for SETX in directing incomplete RNA transcripts to the nuclear exosome has been found in cycling cells when the DNA polymerase machinery collides with active RNAP II transcription.

10.3 R-Loop Resolution Is Not Defective in Setx Null Mice

Another avenue to address the question of which SETX RNA processing functions are most crucial to neuron survival is to create animal models. *Setx* knock-out mice, generated by gene targeted removal of exon four, resulted in near complete loss of SETX protein [43]. However, for a range of reasons possibly including differences in neuroanatomy and lifespan between mice and humans, *Setx* null mice show neither ataxia nor cerebellar degeneration, preventing the possibility to characterize mechanisms of neuron cell death. Failure to recapitulate human recessive ataxias in mice is not without precedent [44], yet a range of critical in vivo studies have been examined in these mice nonetheless. In post-mitotic neurons, R-loops were resolved normally in the cerebellum or brain of mice lacking SETX (similar to wild-type mice), as R-loops could not be detected in wild-type controls or SETX knock-outs. Furthermore, there was no evidence of cells undergoing apoptosis in these tissues in SETX null mice [45]. The authors postulated that the "*major clinical neurodegenerative phenotype seen in AOA2 patients is more likely to be due to a more general defect in RNA processing …rather than a failure to resolve R-loops*". Recent studies from the Libri lab indicate that Sen1p has relatively low processivity on RNA [46]. This is relevant to R-loop removal, R-loops that form in mammals are believed to be very long (>1 kb in humans) [47]. Thus SETX, with its likely low processivity (based upon the Sen1p findings), would not be able to unwind such long structures [46].

10.4 SETX Is SUMO-Modified and Regulates RNAP II Transcription

The multiple RNA processing functions identified for Sen1p undoubtedly require the coordinated efforts of both the amino-terminal protein interaction domain for trafficking, as well as the carboxy-terminal helicase domain for RNA/DNA interaction and processing. In yeast, truncation of the amino-terminal region of Sen1p prevented its proper localization to the nucleolus, though only the helicase domain is required for survival [32]. To better define critical Sen1p protein-binding partners, a yeast two-hybrid (Y2H) screen was employed with follow-up co-immunoprecipitation validation. Only the first 565 residues of Sen1p were required for the identified interactions, again supporting the hypothesis of a crucial function for the amino-terminal domain. Specifically, Sen1p was shown to bind with Rpo21p, the large subunit of RNAP II; Rad2p, a deoxyribonuclease required in DNA repair; and Rnt1p, an endoribonuclease required for RNA maturation [8].

Interestingly, we noted that the Sen1p and SETX protein interaction domains are not conserved at the primary amino acid level [25]. The SETX orthologue in marine vertebrates, such as zebrafish, shows conservation within this domain, but not with fly [25]. However, based upon a hypothesis of functional conservation, we reasoned that a human-specific Y2H screen with the first 650 residues of SETX may identify overlap with Sen1p interactors. The results did bear out some overlap, but to a lesser degree than was expected [48], and was confirmed by a second independent SETX Y2H screen [49]. Our screen also included ALS4 mutants to potentially identify gain-of-function interactors [48], and revealed several important interactor groups that were confirmed with alternate techniques and in other independent proteomics screens [7, 50]. The key interactor groups were: (1) SETX self-interaction or dimerization; (2) critical Sumo/Ubiquitin posttranslational modification; and (3) DNA/RNA-binding proteins, including the exosome component 9 protein (Exosc9).

SETX amino-terminal domain self-interaction and dimerization were validated by purification, size exclusion chromatography, protein cross-linking, and Western blot analysis [48] (Fig. 10.2). Importantly, as ALS is one of a number of conditions in which protein aggregation may drive disease pathogenesis [51], we examined this further. Using techniques including targeted mammalian two-hybrid (M2H) analysis, we found that the SETX mutants L389S (ALS4) and W305C (AOA2) can still engage in self-interaction, and do not lead to excessive aggregation. SETX's ability to dimerize may thus set it apart from Sen1p. Additionally, we found five SETX interactors representing proteins in either the SUMO protein trafficking cascade or the ubiquitin protein degradation pathway. The key interaction with Exosc9 was shown to require SETX SUMOylation [49]. Dramatically, when the exosome components Exosc9 or Exosc10 were depleted by targeted siRNA, this yielded significant co-depletion of SETX [49].

As for the overlap of SETX and Sen1p interactors, several interesting distinctions were noted. The three major Sen1p interactors, Rpo21p, Rad2p, and Rnt1p, were not detected by our Y2H screen. Similarly, but this time in mammalian cells using targeted M2H analysis, we did not detect a direct interaction between SETX and the human

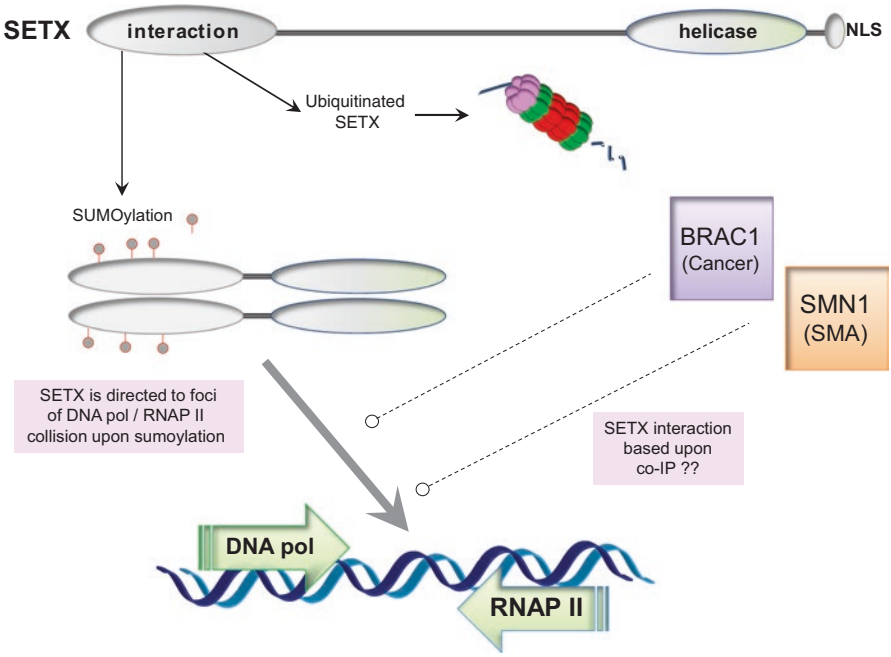

Fig. 10.2 SETX protein domain organization, proposed functions, and protein interactions. The SETX protein is ~303 kDa in molecular mass, and possesses just three known domains: the amino-terminal protein interaction domain, the carboxy-terminal helicase domain, and a nuclear localization signal (NLS) domain. SETX contains amino-terminal sequences that are targets of ubiquitination utilized to degrade SETX protein via the proteasome. The SETX amino-terminal domain possesses regions required for dimerization and for SUMOylation (gray ball and red stem). It is thought that SETX needs to be SUMOylated to direct it to sites of collision between the DNA polymerase-containing replisome and the RNA polymerase-containing transcription machinery. Other key proteins of interest that have been linked with SETX include BRCA1 and SMN1

orthologues to these three proteins (unpublished data). In another proteomics study of Sen1p, tandem affinity purification/mass spectrometry analysis defined this protein as a general transcription factor based upon interactions with RNAP I and RNAP III subunits, as well as with the classic mRNA polymerase complex, RNAP II [7]. This result is consistent with the types of ncRNAs that Sen1p has been previously shown to regulate. Alternatively, proteomics analysis of full-length, GFP/Flag-tagged SETX indicated a general association with the RNAP II complex. A direct interaction was not shown with the core RNAP II subunits (RPB1, RPB2, and RPB3). This suggests that while Sen1p may interact directly with RNAP I, II and III subunits, SETX likely interacts with RNAP II subunits via intermediary associations as periodically directed.

One can conclude that both SETX and Sen1p contain amino acid sequences within the relatively large 500–600 amino-terminal region that are targets for Sumo and Ubiquitin-mediated regulation (Fig. 10.2). For Sen1p, it has been clearly demonstrated that this region is required for signaling its degradation via the ubiquitin proteasome system to maintain low cellular protein levels [52]. Our Y2H screen identified

ubiquitin pathway proteins, Ubc9 and UBC with the amino-terminal region as bait, suggesting similar regulation.

10.5 SETX and Sen1p: A Convincing Role in Connecting RNAP II and the Exosome

Characterizing a specific role for SETX in RNAP II termination has not been convincing, despite some initial reports [50]. A role for Sen1p in this process has been well characterized for ncRNAs, which employ a distinctive mechanism specific for these transcripts in yeast, and not likely used in higher eukaryotes (as noted above). For example, termination of the elongated snoRNA precursors relies upon different machinery than the cleavage and polyadenylation mechanism used for mRNA termination. Rather, ncRNA-specific processing relies on the NRD complex containing the Nrd1 and Nab3 RNA-binding proteins in association with Sen1p [36]. In this case, termination occurs downstream of tetranucleotide motifs, which form binding sites for Nrd1 and Nab3 on the nascent RNA [53]. The Nrd1–Nab3–Sen1p complex, which directly interacts with the RNAP II Carboxy-Terminal Domain (CTD), also directly interacts with the nuclear exosome [54]. Thus, Sen1p as part of the NRD complex forms a bridge between the RNAP II and the exosome to aid in the termination of ncRNA transcripts, such as snoRNAs. In these cases, transcription termination is coupled to 3′–5′ exonuclease trimming by the TRAMP–exosome complex [54]. Extrapolation of the Sen1p termination process to mammals is not readily possible, as the RNA-binding protein Nab3, a critical protein bridging the interaction of Nrd1 and Sen1, has no human homologue [55], and Nab3 has no known role in poly-A-dependent termination [36]. In higher eukaryotes, RNAP II utilizes alternate means of transcription termination [56], and previous studies confirm that SETX is not required for snRNA termination [50].

With regard to transcription, despite likely divergence between the Sen1p and SETX regulation, there is significant evidence to suggest SETX has retained a role for linking RNAP II to the nuclear exosome. As noted above, some degree of SETX co-depletion occurs when major components of the exosome, Exosc9 (Rrp6) and Exosc10 (Rrp45), are depleted. Thus, a model can be proposed that SETX needs to dimerize, and then be SUMOylated as a requirement for its interaction with Exosc9 (and the nuclear exosome) [49], and that transcription-related DNA damage directs the lowly abundant SETX to the exosome in response to such transcription pausing (Fig. 10.3).

10.6 SETX Localization and Function in Cycling Cells

SETX is a large 303 kDa protein that localizes to the nucleus in unsynchronized cell lines. But with different antibodies, several investigators have observed SETX clearly in the nucleolus. Initially, this was not unexpected, as yeast Sen1p was

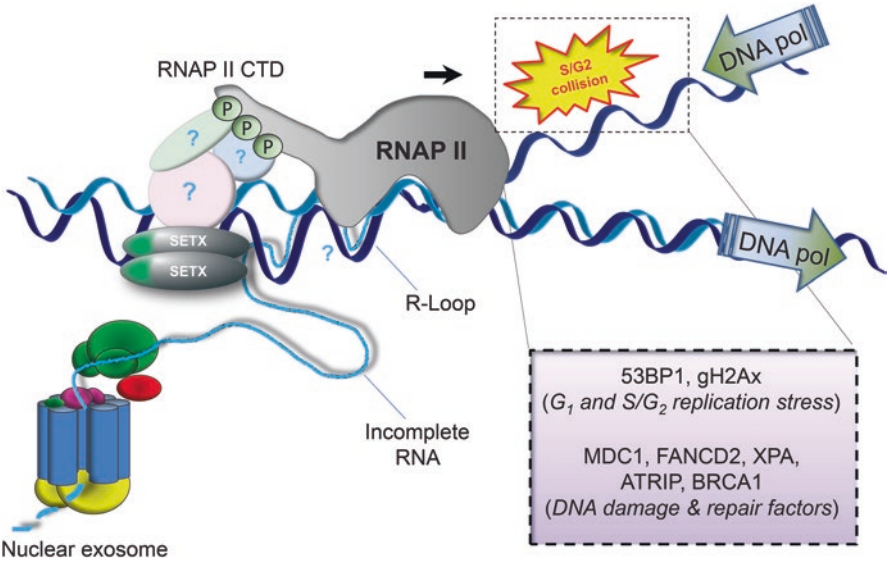

Fig. 10.3 SETX mediates RNA processing and degradation under specific circumstances. The SETX protein interacts with RNAP II via unknown intermediates, but only under certain circumstances, as SETX protein levels are exceedingly low, and SETX may require SUMOylation to direct it to specific foci. SETX foci form at times of replication stress and colocalize with markers such as 53BP1 and γH2Ax. Typical markers of DNA damage and repair factors are found in SETX foci at sites of DNA polymerase and RNAP II collision. According to this model, one likely SETX function is RNA processing via directing incompletely transcribed RNAs to the nuclear exosome for degradation

found to play a major role in the nucleolus, processing rRNA precursors and snoRNAs [33]. Sen1p was also required to maintain the normal crescent shape of the yeast nucleolus, and the temperature sensitive mutant, *Sen1-1*, caused mislocalization of nucleolar proteins, fibrillarin, and Ssb1 [32]. Our initial studies with an affinity-purified SETX antibody revealed strong colocalization with fibrillarin in the nucleolus [57], but further studies are ongoing. A recent publication looking at the possibility of SETX mislocalization found near complete localization to the nucleolus in both control ($SETX^{+/+}$) and patient ($SETX^{R332W/fs}$) fibroblasts [58]. Despite these findings, immunocytochemistry analysis of tagged or endogenous SETX mostly shows a general nuclear localization.

A more detailed analysis of nuclear foci during the cell cycle was recently undertaken with GFP-tagged SETX. These investigators used double thymidine block to synchronize cells and automated wide-field microscopy to visualize SETX dynamic localization. They found that SETX foci were indeed present in the nucleolus at S-phase periodically, but as cells progressed into G_2-phase, SETX became distributed throughout the nucleoplasm [7]. In new studies, evidence was found to link the functions of SETX in potentially directing RNA to the exosome. How does this occur? It should be noted that transcription of large genes can take longer than the

replication phase of the cell-cycle, such that the transcription and replication machinery may collide [59, 60], and SETX has been placed at the sites of these collisions (Fig. 10.3). Upon this backdrop, a range of studies were undertaken to define SETX nuclear foci, in response to phase transitions and drug treatments. In the nucleus, SETX distinct foci were found to be strongest during DNA replication or the S/G_2 phase. When cells were treated with aphidicolin to retard the replication fork, which is a form of replication stress, a two-fold increase in the number of SETX foci resulted [7]. These foci were perfectly colocalized with 53BP1 and γH2AX, markers of spontaneous DNA lesions and transcriptionally active nuclear bodies that form at fragile sites during replication [61] (Fig. 10.3). Then, after treatment of cycling cells with α-amanitin to inhibit RNAP II-mediated transcription, a significant reduction in SETX foci occurred, supporting the idea that coalescence of these distinct foci is dependent on RNAP II transcription. Such SETX-Exosc9 targeted interactions may represent one of the most pivotal roles of SETX, namely to bring functioning exosomes to sites of transcription—replication fork collisions, an interaction that is suggested to depend upon SUMO-2 and SUMO-3 SETX modification [49].

10.7 Lessons from the SETX Homologue IGHMBP2: Role of tRNA Regulation

What is missing from the study of SETX in neurological disease is a smoking gun pointing the way to the RNA pathways that are most affected. Interestingly, the study of SMARD was in a similar quandary, lacking knowledge of the affected RNA pathways; however, recent work has yielded a mechanistic understanding. In 2001, the *IGHMBP*2 gene was identified as the molecular basis of SMARD, when key mutations in six families were reported [29]. SMARD is clinically distinct from SMA, but the IGHMBP2 protein, like SMN1, colocalizes with the RNA-processing machinery in both the cytosol and the nucleus [29]. The mouse model for SMARD is a spontaneous mutant discovered at The Jackson Laboratory known as *nmd* (for neuromuscular degeneration). An important clue to disease mechanism was provided with the discovery that the *nmd* phenotype is suppressed in a semi-dominant fashion by the presence of a modifier region on mouse Chromosome 13 from strain CAST/EiJ [31]. The critical region for rescue is limited to just 166 kb, defined by a BAC clone which contains several tRNA genes, including five $tRNA^{Tyr}$ genes, one $tRNA^{Ala}$ gene, and activator of basal transcription 1 (Abt1) [62]. The *nmd* mice are characterized by motor neuron degeneration with axonal loss leading to neurogenic muscle atrophy and death at 8–12 weeks of age. The phenotypic rescue of *nmd* mice by the Chromosome 13 modifier is dramatic with ventral nerve roots showing completely normal axonal morphology and density at 6–7 weeks of age [63]. Finally, IGHMBP2 has been shown to physically associate with tRNAs, and in particular with $tRNA^{Tyr}$ and the tRNA transcription factor TFIIIC220 [62].

In 2013, further investigations revealed that aberrant tRNA processing can lead to neurodegeneration. The first mammalian RNA kinase to be identified was CLP1, and kinase-dead mice for this protein (*Clp1*$^{K/K}$) were generated [64]. On several genetic backgrounds, *Clp1*$^{K/K}$ homozygous pups were nonviable, but on the CBA/J background, mice survived to ~23 weeks of age. *Clp1*$^{K/K}$ mice display loss of spinal motor neurons associated with axonal degeneration in the peripheral nerves and denervation of neuromuscular junctions and respiratory failure [64]. Transgenic studies demonstrated that CLP1 functions in motor neurons, and that reduced CLP1 activity results in the accumulation of a novel set of small RNA fragments, derived from aberrant processing of pre-tRNA^{Tyr}. In 2014, a CLP1 R140H homozygous missense mutation was reported in five unrelated human families [65]. These patients suffered severe motor-sensory defects, cortical dysgenesis, and microcephaly. Biochemically, these presumed hypomorphic mutations lead to a loss of CLP1 interaction with the tRNA splicing endonuclease complex, greatly reduced pre-tRNA cleavage activity, and accumulation of linear tRNA introns [65].

Many other examples of tRNA biogenesis dysfunction leading to neurodegeneration exist. For brevity, we name just two: (1) an editing-defective tRNA synthetase causes protein misfolding and neurodegeneration in the *sticky* mouse [66]; and (2) a mutation of a CNS-specific tRNA causes neurodegeneration induced by ribosome stalling [67]. These examples serve to support the story of tRNA processing dysfunction in neurological disease and lend further credence to the mechanistic understanding of SMARD caused by IGHMBP2 recessive loss-of-function mutations. Similarly, methods and approaches that will reveal key insights into the most critical SETX RNA processing pathways for neuron health and survival are needed. While the lack of neurological phenotypes in *Setx* knock-out mice prevents the identification of similar modifier effects, other methodologies are likely to emerge to provide similar insight into RNA processing maintenance in neurodegeneration phenotypes caused by loss of SETX function.

10.8 SETX Gain-of-Function Motor Neuron Toxicity in ALS4 and Its Possible Role in Sporadic ALS

Here, we have considered two unique neurodegenerative disorders, AOA2 and ALS4, which represent the genotype/phenotype spectrum resulting from *SETX* mutation, and sought to underscore which RNA processing functions are most relevant to neurodegeneration. We began by recognizing that many functions attributed to SETX were extrapolated from its yeast orthologue Sen1p, which had been thoroughly studied long before *SETX* mutations were first discovered. Notably, Sen1p homology to SETX is restricted to an ~500 amino acid carboxy-terminal helicase domain, with no other regions of the large 303 kDa SETX protein conserved. The SETX amino-terminal protein interaction domain is divergent at the sequence level, but appears functionally conserved (with new protein interactions). Upon detailed examination, not all functions of yeast Sen1p were retained by

mammalian SETX, which functions in a multicellular organism where cell cycling regulation has become much more elaborate. The well-characterized function of *general* transcription termination in Sen1p is not likely conserved in SETX; instead, SETX regulation of transcription termination is restricted to specific genes and cellular pathways, including interestingly circadian rhythm control. Whether disruption of this pathway is relevant to motor neuron health and ALS neurodegeneration remains to be studied. Insofar as future research is concerned, it is important to recognize that Sen1p and SETX proteins are present at exceedingly low levels in the cell. Thus, studies with Sen1p or SETX which utilize massive transient over-expression will likely generate results that are not physiologically relevant.

One key SETX function, conserved from Sen1p, is direct engagement with the nuclear exosome. Two studies demonstrate that SETX interacts with Exosc9 [48, 49], and is regulated by the SUMO and ubiquitin cascade pathways. One group has demonstrated that it is this SUMO-2/3 modification at the amino-terminus that is specifically required for interaction with the exosome and that co-depletion of SETX occurs with either Exosc9 or Exosc10 knock-down [49]. SETX was shown to be present in specific nuclear foci during S/G2-phase human synchronized cells coincident with collision of the DNA replication machinery and the RNA transcriptome [7]. These SETX foci were described as representing replication stress, and at these foci, the SETX interaction with the nuclear exosome was specifically present and enriched [49]. SETX thus appears to play a key role in directing incomplete RNA transcripts to the exosome for degradation (Fig. 10.3). The connection between SETX and exosome regulation deserves further consideration as a possible explanation for how SETX gain-of-function toxicity results in motor neuron disease. Interestingly, recessive loss-of-function mutations in Exosc3 yield infantile-onset motor neuron disease in human pontocerebellar hypoplasia with spinal muscular atrophy type 1B (PCH1B; OMIM 614678) [68], and Exosc3 interacts with matrin-3 [69], a known ALS gene. Furthermore, recessive loss-of-function mutations in Exosc8 yield infantile-onset motor neuron disease in human pontocerebellar hypoplasia with spinal muscular atrophy type 1C (PCH1C; OMIM 616081). These inherited motor neuron degeneration phenotypes highlight that alterations of exosome function are particularly poorly tolerated in cerebellar and motor neurons, two CNS regions where altered SETX function results in neuronal demise. However, another point to consider is that neurons are not cycling cells; hence, the role of SETX in resolving collisions between the replication machinery and the RNAP II transcription complex could actually play out in non-neuronal cells. As glia comprise the bulk of CNS cells, it seems reasonable to propose that neuron demise in ALS4 and AOA2 could be the result of a non-cell-autonomous process occurring in astrocytes or another non-neural CNS cell type.

A final important point to consider when seeking an explanation for SETX neurotoxicity is that SETX belongs to a group of just three homologous proteins, the other two being IGHMBP2 and RENT1. Of this trio, RENT1 is specifically implicated in NMD [70], and its role in NMD appears critical, as loss of function of RENT1 leads to embryonic lethality in mice, with no known human disease correlate. The importance of this helicase protein for RNA toxicity in neurons is suggested by its ability to rescue TDP-43 and FUS ALS-linked cellular pathology [28].

The other member of the trio is IGHMBP2, recessive mutations of which cause SMARD. In this disorder, key processing events for tRNA appear to be the responsible RNA pathway affected. As SMARD is related to autosomal spinal muscular atrophy (SMA), the theme of altered RNA function is reinforced, as SMN protein is essential to spliceosomal snRNP biogenesis and thus the integrity of RNA splicing, and therefore has become a model for understanding RNA dysfunction in neurodegeneration [71]. RNA-binding proteins, such as TDP-43, have also been centrally implicated in ALS, but understanding the role of TDP-43 in motor neuron neurodegeneration is proving to be challenging. There may be multiple dominant, recessive, and toxic mechanisms at play throughout the disease process. However, an intriguing theory based upon loss of function, which necessarily occurs with nuclear clearance of TDP-43, is that of impaired repression of non-conserved cryptic exons [72]. SETX gain-of-function mutations cause ALS4, which is a rare disease, and while several dominant mutations have been linked to ALS4, by far the most penetrant mutation to study is the L389S substitution. Clues to disease mechanism based upon SETX L389S toxic gain-of-function await the description of new mouse models that have been produced and are being characterized. Furthermore, based upon independent SALS exome sequencing reports [3, 73–75], SETX is emerging as a common target for mutation, especially in SALS patients who carry mutations in established pathogenic genes, including C9orf72 repeat expansion carriers [75]. These observations support the hypothesis that these recently discovered disease-linked polymorphisms in SETX could be modifiers of SALS. Hence, future research into SETX normal function and altered action upon gain-of-function mutation holds great potential for advancing our understanding of not just ALS4 motor neuron disease but also for much more common sporadic ALS as well.

Acknowledgments Our SETX research is supported by a grant from the Robert Packard Center for ALS Research at the Johns Hopkins School of Medicine.

References

1. Le Ber I, et al. Frequency and phenotypic spectrum of ataxia with oculomotor apraxia 2: a clinical and genetic study in 18 patients. Brain. 2004;127:759–67.
2. Rabin BA, et al. Autosomal dominant juvenile amyotrophic lateral sclerosis. Brain. 1999;122:1539–50.
3. Cady J, et al. Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. Ann Neurol. 2015;77:100–13.
4. Chen YZ, et al. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). Am J Hum Genet. 2004;74:1128–35.
5. Moreira MC, et al. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. Nat Genet. 2004;36:225–7.
6. Padmanabhan K, Robles MS, Westerling T, Weitz CJ. Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. Science. 2012;337:599–602.

7. Yuce O, West SC. 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. 2013;33:406–17.
8. Ursic D, Chinchilla K, Finkel JS, Culbertson MR. Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. Nucleic Acids Res. 2004;32:2441–52.
9. Lupski JR. Charcot-Marie-tooth polyneuropathy: duplication, gene dosage, and genetic heterogeneity. Pediatr Res. 1999;45:159–65.
10. Warner LE, Roa BB, Lupski JR. Absence of PMP22 coding region mutations in CMT1A duplication patients: further evidence supporting gene dosage as a mechanism for Charcot-Marie-Tooth disease type 1A. Hum Mutat. 1996;8:362.
11. Nihei Y, Ito D, Suzuki N. Roles of ataxin-2 in pathological cascades mediated by TAR DNA-binding protein 43 (TDP-43) and Fused in Sarcoma (FUS). J Biol Chem. 2012;287:41310–23.
12. Kim HD, Choe J, Seo YS. The sen1(+) gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. Biochemistry. 1999;38:14697–710.
13. Ghaemmaghami S, et al. Global analysis of protein expression in yeast. Nature. 2003;425:737–41.
14. Borggrefe T, Davis R, Bareket-Samish A, Kornberg RD. Quantitation of the RNA polymerase II transcription machinery in yeast. J Biol Chem. 2001;276:47150–3.
15. Svejstrup JQ, et al. Evidence for a mediator cycle at the initiation of transcription. Proc Natl Acad Sci U S A. 1997;94:6075–8.
16. Beck M, et al. The quantitative proteome of a human cell line. Mol Syst Biol. 2011;7:549.
17. Avemaria F, et al. Mutation in the senataxin gene found in a patient affected by familial ALS with juvenile onset and slow progression. Amyotroph Lateral Scler. 2011;12:228–30.
18. Rudnik-Schoneborn S, Arning L, Epplen JT, Zerres K. SETX gene mutation in a family diagnosed autosomal dominant proximal spinal muscular atrophy. Neuromuscul Disord. 2012;22:258–62.
19. Anheim M, et al. Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. Brain. 2009;132:2688–98.
20. Suraweera A, et al. Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. J Cell Biol. 2007;177:969–79.
21. Vantaggiato C, et al. Novel SETX variants in a patient with ataxia, neuropathy, and oculomotor apraxia are associated with normal sensitivity to oxidative DNA damaging agents. Brain Dev. 2014;36:682–9.
22. De Amicis A, et al. Role of senataxin in DNA damage and telomeric stability. DNA Repair. 2011;10:199–209.
23. Lynch DR, Braastad CD, Nagan N. Ovarian failure in ataxia with oculomotor apraxia type 2. Am J Med Genet A. 2007;143:1775–7.
24. Vance C, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science. 2009;323:1208–11.
25. Bennett CL, La Spada AR. Unwinding the role of senataxin in neurodegeneration. Discov Med. 2015;19:127–36.
26. Weng Y, Czaplinski K, Peltz SW. Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. Mol Cell Biol. 1996;16:5477–90.
27. Medghalchi SM, et al. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum Mol Genet. 2001;10:99–105.
28. Barmada SJ, et al. Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. Proc Natl Acad Sci U S A. 2015;112:7821–6.
29. Grohmann K, et al. Mutations in the gene encoding immunoglobulin mu-binding protein 2 cause spinal muscular atrophy with respiratory distress type 1. Nat Genet. 2001;29:75–7.
30. Lefebvre S, Burglen L, Frezal J, Munnich A, Melki J. The role of the SMN gene in proximal spinal muscular atrophy. Hum Mol Genet. 1998;7:1531–6.

31. Cox GA, Mahaffey CL, Frankel WN. Identification of the mouse neuromuscular degeneration gene and mapping of a second site suppressor allele. Neuron. 1998;21:1327–37.
32. Ursic D, DeMarini DJ, Culbertson MR. Inactivation of the yeast Sen1 protein affects the localization of nucleolar proteins. Mol Gen Genet. 1995;249:571–84.
33. Ursic D, Himmel KL, Gurley KA, Webb F, Culbertson MR. The yeast SEN1 gene is required for the processing of diverse RNA classes. Nucleic Acids Res. 1997;25:4778–85.
34. Winey M, Culbertson MR. Mutations affecting the tRNA-splicing endonuclease activity of *Saccharomyces cerevisiae*. Genetics. 1988;118:609–17.
35. Rasmussen TP, Culbertson MR. The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nucleolar RNAs in *Saccharomyces cerevisiae*. Mol Cell Biol. 1998;18:6885–96.
36. Steinmetz EJ, Conrad NK, Brow DA, Corden JL. RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature. 2001;413:327–31.
37. Steinmetz EJ, Brow DA. Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol Cell Biol. 1996;16:6993–7003.
38. Rondon AG, Mischo HE, Kawauchi J, Proudfoot NJ. Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. Mol Cell. 2009;36:88–98.
39. Steinmetz EJ, et al. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell. 2006;24:735–46.
40. Reynolds JJ, Stewart GS. A single strand that links multiple neuropathologies in human disease. Brain. 2013;136:14–27.
41. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell. 2011;42:794–805.
42. Banerjee A, Sammarco MC, Ditch S, Wang J, Grabczyk E. A novel tandem reporter quantifies RNA polymerase II termination in mammalian cells. PLoS One. 2009;4:e6193.
43. Becherel OJ, et al. Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. PLoS Genet. 2013;9:e1003435.
44. Elson A, et al. Pleiotropic defects in ataxia-telangiectasia protein-deficient mice. Proc Natl Acad Sci U S A. 1996;93:13084–9.
45. Yeo AJ, et al. R-loops in proliferating cells but not in the brain: implications for AOA2 and other autosomal recessive ataxias. PLoS One. 2014;9:e90219.
46. Han Z, Libri D, Porrua O. Biochemical characterization of the helicase Sen1 provides new insights into the mechanisms of non-coding transcription termination. Nucleic Acids Res. 2017;45:1355–70.
47. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol. 2003;4:442–51.
48. Bennett CL, et al. Protein interaction analysis of senataxin and the ALS4 L389S mutant yields insights into senataxin post-translational modification and uncovers mutant-specific binding with a brain cytoplasmic RNA-encoded peptide. PLoS One. 2013;8:e78837.
49. Richard P, Feng S, Manley JL. A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. Genes Dev. 2013;27:2227–32.
50. Suraweera A, et al. Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. Hum Mol Genet. 2009;18:3384–96.
51. Deng HX, et al. Amyotrophic lateral sclerosis and structural defects in Cu,Zn superoxide dismutase. Science. 1993;261:1047–51.
52. DeMarini DJ, et al. The yeast SEN3 gene encodes a regulatory subunit of the 26S proteasome complex required for ubiquitin-dependent protein degradation in vivo. Mol Cell Biol. 1995;15:6311–21.

53. Carroll KL, Pradhan DA, Granek JA, Clarke ND, Corden JL. Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol. 2004;24:6241–52.
54. Vasiljeva L, Buratowski S. Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. Mol Cell. 2006;21:239–48.
55. Kuehner JN, Pearson EL, Moore C. Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol. 2011;12:283–94.
56. Richard P, Manley JL. Transcription termination by nuclear RNA polymerases. Genes Dev. 2009;23:1247–69.
57. Chen YZ, et al. Senataxin, the yeast Sen1p orthologue: characterization of a unique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. Neurobiol Dis. 2006;23:97–108.
58. Roda RH, Rinaldi C, Singh R, Schindler AB, Blackstone C. Ataxia with oculomotor apraxia type 2 fibroblasts exhibit increased susceptibility to oxidative DNA damage. J Clin Neurosci. 2014;21:1627–31.
59. Helmrich A, Ballarino M, Nudler E, Tora L. Transcription-replication encounters, consequences and genomic instability. Nat Struct Mol Biol. 2013;20:412–8.
60. Helmrich A, Ballarino M, Tora L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol Cell. 2011;44:966–77.
61. Lukas C, et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat Cell Biol. 2011;13:243–53.
62. de Planell-Saguer M, Schroeder DG, Rodicio MC, Cox GA, Mourelatos Z. Biochemical and genetic evidence for a role of IGHMBP2 in the translational machinery. Hum Mol Genet. 2009;18:2115–26.
63. Maddatu TP, Garvey SM, Schroeder DG, Hampton TG, Cox GA. Transgenic rescue of neurogenic atrophy in the nmd mouse reveals a role for Ighmbp2 in dilated cardiomyopathy. Hum Mol Genet. 2004;13:1105–15.
64. Hanada T, et al. CLP1 links tRNA metabolism to progressive motor-neuron loss. Nature. 2013;495:474–80.
65. Karaca E, et al. Human CLP1 mutations alter tRNA biogenesis, affecting both peripheral and central nervous system function. Cell. 2014;157:636–50.
66. Lee JW, et al. Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. Nature. 2006;443:50–5.
67. Ishimura R, et al. RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. Science. 2014;345:455–9.
68. Wan J, et al. Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. Nat Genet. 2012;44:704–8.
69. Erazo A, Goff SP. Nuclear matrix protein Matrin 3 is a regulator of ZAP-mediated retroviral restriction. Retrovirology. 2015;12:57.
70. Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. Hum Mol Genet. 1999;8:1893–900.
71. Li DK, Tisdale S, Lotti F, Pellizzoni L. SMN control of RNP assembly: from post-transcriptional gene regulation to motor neuron disease. Semin Cell Dev Biol. 2014;32:22–9.
72. Ling JP, Pletnikova O, Troncoso JC, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science. 2015;349:650–5.
73. Cirulli ET, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015;347:1436–41.
74. Couthouis J, Raphael AR, Daneshjou R, Gitler AD. Targeted exon capture and sequencing in sporadic amyotrophic lateral sclerosis. PLoS Genet. 2014;10:e1004704.
75. Kenna KP, et al. Delineating the genetic heterogeneity of ALS using targeted high-throughput sequencing. J Med Genet. 2013;50:776–83.