Review

Taking a risk: a therapeutic focus on ataxin-2 in amyotrophic lateral sclerosis?

Dianne M.A. van den Heuvel[1,2*], Oliver Harschnitz[1,2*], Leonard H. van den Berg[2], and R. Jeroen Pasterkamp[1]

[1] Department of Translational Neuroscience, UMC Utrecht Brain Center Rudolf Magnus, Utrecht, 3584 CG, The Netherlands
[2] Department of Neurology and Neurosurgery, UMC Utrecht Brain Center Rudolf Magnus, Utrecht, 3584 CG, The Netherlands

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease caused by the loss of lower and upper motor neurons leading to progressive muscle weakness and respiratory insufficiency. No treatment is currently available to cure ALS. Recent progress has led to the identification of several novel genetic determinants of this disease, including repeat expansions in the ataxin-2 (*ATXN2*) gene. Ataxin-2 is mislocalized in ALS patients and represents a relatively common susceptibility gene in ALS, making it a promising therapeutic target. In this review, we summarize genetic and pathological data implicating ataxin-2 in ALS, discuss potential disease mechanisms linked to altered ataxin-2 localization or function, and propose potential strategies for therapeutic intervention in ALS based on ataxin-2.

Amyotrophic lateral sclerosis (ALS)

ALS is a fatal adult-onset neurodegenerative disease characterized by progressive muscle weakness, muscle atrophy, and fasciculations (see Glossary) caused by the loss of motor neurons in the motor cortex, brainstem, and spinal cord. Progressive paralysis leads to death due to respiratory failure on average within 3 years after symptom onset. The risk of developing ALS peaks between the ages of 50 and 75 years and the overall lifetime risk of developing the disease is 1:400 [1,2]. Treatment options for ALS remain limited. The only effective drug, riluzole, extends median survival by a few months, emphasizing the need for novel therapeutic agents and strategies [3].

Traditionally, ALS is divided into familial ALS and sporadic ALS. Although approximately 5–10% of patients have a family history of ALS (fALS) [4], most patients have no family history and are referred to as sporadic ALS (sALS) patients. The majority of fALS kindreds show a Mendelian pattern of inheritance and are clinically indistinguishable from sALS. To date, approximately 15 genes have been identified as causative for classical fALS and a similar, although rapidly increasing, number of genes have been linked to sALS (Box 1). Mutations in the superoxide dismutase (*SOD1*) gene were identified as the first cause of ALS in 1993 and account for ~20% of all fALS cases [5]. Two other genes with a relatively high incidence in ALS are TAR DNA-binding protein 43 (*TARDBP*), which encodes for TDP-43, and fused in sarcoma/translated in liposarcoma (*FUS*/*TLS*). Mutations in *TARDBP* are responsible for ~5% of fALS and ~1% of apparently sALS, whereas mutations in *FUS* are causative of ~4% and ~1% of fALS and sALS, respectively [6]. In 2011, a hexanucleotide repeat expansion in chromosome 9 open reading frame 72 (*C9ORF72*) was identified in 40–50% of all fALS and

Glossary

Clustered regularly interspaced short palindromic repeats (CRISPRs): RNA-based arrays containing multiple short direct repeats (24–48 base pairs) to guide the cleavage of a specific DNA sequence by a CRISPR-associated (Cas) nuclease.
Fasciculations: small, local, involuntary muscle contractions or twitches.
Frontotemporal dementia (FTD): a clinical syndrome characterized by changes in social behavior, personality, and language. The disease, which shows genetic overlap with ALS, is a result of relatively selective degeneration of the frontal and temporal lobes.
Huntington's disease (HD): a neurodegenerative disease caused by polyQ expansions in the huntingtin gene (*HTT*) and clinically characterized by involuntary movements, cognitive decline, and psychiatric problems.
Inclusions: nuclear or cytoplasmic aggregates of proteins or other molecules.
Induced pluripotent stem cells (iPSCs): a type of pluripotent stem cell generated from somatic cell types through expression of specific genes, typically *OCT4, SOX2, c-MYC,* and *KLF4,* that induce a pluripotent state.
Like Sm/Like Sm associated domain (Lsm domain): RNA-binding domain shared by proteins involved in RNA processes, such as mRNA splicing, post-transcriptional modification, and RNA decay.
Messenger ribonucleoprotein particles (mRNPs): highly dynamic granules consisting of RNA and RNA-binding proteins that serve in the transport of mRNAs to their site of translation.
Muscle atrophy: decrease and wasting away of muscle mass.
PolyQ diseases: neurodegenerative disorders caused by trinucleotide repeat expansions in the affected gene, where the repeated codon is CAG/CAA encoding for glutamine.
Stress granules (SGs): cytoplasmic structures that recruit stalled protein–RNA complexes as a cellular response to environmental stressors.
Transcription activator-like effector nucleases (TALENs): restriction enzymes consisting of TALE proteins which function as the DNA-binding domains and contain repeats called repeat variable diresidues (RVDs) and a FokI cleavage domain. RVDs are made up out of 33–25 amino acid repeats, each recognizing a single base pair.

Corresponding author: Pasterkamp, R.J. (r.j.pasterkamp@umcutrecht.nl).
Keywords: ataxin-2; amyotrophic lateral sclerosis; neurodegeneration; TDP-43; FUS; antisense oligonucleotide therapy.
[*] These authors contributed equally to this work.

1471-4914/$ – see front matter
 http://dx.doi.org/10.1016/j.molmed.2013.09.001

Box 1. Amyotrophic lateral sclerosis (ALS) proteins

For many years, mutations in copper–zinc superoxide dismutase 1 (*SOD1*) were the only known genetic cause of ALS [5]. However, recently, there has been considerable progress in the identification of additional genetic causes of ALS. Mutations in various genes have been shown to cause fALS, and the most prevalent mutations are those in fused in sarcoma/translated in liposarcoma (*FUS/TLS*) [93,94], TAR DNA-binding protein (*TARDBP*) [95,96], and chromosome 9 open reading frame 72 (*C9ORF72*) [7,8].

SOD1 is an abundant copper/zinc enzyme that protects cells from oxidative damage by metabolizing superoxide radicals. The exact mechanism through which SOD1 mutations lead to ALS pathology is unknown, but SOD1 function is related to cellular mechanisms such as oxidative stress and mitochondrial dysfunction, which are characteristic for ALS pathogenesis [97]. Furthermore, mutant SOD1 forms cellular aggregates that may indirectly affect cell function. TDP-43 and FUS have numerous roles in RNA processes including RNA transcription, mRNA transport, nucleocytoplasmic shuttling, RNA splicing, miRNA processing, and stress granule dynamics [98]. Most of the ALS-linked mutations in TDP-43 are localized to its C-terminal domain, which is important for TDP-43 aggregation [99]. The majority of ALS-associated FUS mutations occur in the nuclear localization signal of FUS, thereby impairing nuclear transport of the protein [100]. Both mutant TDP-43 and FUS form cytoplasmic aggregates leading to depletion of the protein from other cellular sites. Furthermore, TDP-43 is present in aggregates in the majority of ALS patients, even in those that do not carry *TDP-43* mutations. At present, an intronic hexanucleotide repeat expansion in *C9ORF72* is the most prevalent cause of ALS [7,8]. The function of the C9ORF72 protein is unknown, but it shows homology to differentially expressed in normal and neoplasia (DENN), which is a GDP/GTP exchange factor for Rab GTPases [101]. More rare mutations in fALS include those in angiogenin (*ANG*) [102], senataxin (*SETX*) [103], ubiquilin 2 (*UBQLN2*) [104], and vesicle-associated membrane protein-associated protein B (*VAPB*) [105].

The majority of patients suffer from sALS, which is thought to be a complex disorder caused by interplay between environmental exposure and genetic factors. Genes implicated in sALS are, for example, charged multivesicular body protein 2b (*CHMP2B*) [106], unc-13 homolog A (*UNC13A*) [107], non-imprinted in Prader–Willi/Angelman syndrome 1 (*NIPA1*) [9], hemochromatosis (*HFE*), peripherin (*PRPH*) [108], and survival motor neuron (SMN) [109].

~20% of sALS patients, making it the most prevalent cause of ALS identified so far [7,8].

Although this knowledge of the genetic basis of ALS has significantly improved our understanding of ALS pathogenesis, it does not explain the variability in disease penetrance, severity, and age of onset observed in patients harboring the same genetic defect. Also, variation in disease duration is large, with some patients dying within months after onset and others surviving for more than two decades. Furthermore, large differences in survival and age at disease onset exist even between individuals from one family, in whom ALS is caused by the same mutation. Together, these differences between individual patients suggest the existence of disease-modifying factors. Recent studies have discovered genetic susceptibility factors that may explain disease variability and may serve as therapeutic targets for larger cohorts of patients, for example, non-imprinted in Prader–Willi/Angelman syndrome 1 (*NIPA1*) and unc-13 homolog A (*UNC13A*), which are both associated with shorter median survival in ALS [9,10]. Here, we review another recently identified risk factor for ALS: ataxin-2 (*ATXN2*). Ataxin-2 is a ubiquitously expressed cytoplasmic protein that interacts with the ALS proteins FUS and TDP-43 and modifies their cellular toxicity [11–14]. In this review, we first provide an overview of genetic data linking repeat expansions in *ATXN2* to ALS. We then discuss the normal biological functions of ataxin-2 followed by potential disease mechanisms linked to altered function and localization of this protein in ALS. Finally, we highlight and discuss putative therapeutic strategies focused on ataxin-2 aiming to slow down disease progression and to increase patient survival.

Intermediate *ATXN2* polyQ expansions are a risk factor for ALS

In healthy individuals, the ataxin-2 polyglutamine (polyQ) repeat sequence, which is present in the most N-terminal part of the protein, is 22–23 repeats long [15,16]. At the DNA level, CAG repeats in *ATXN2* are interspersed with CAA codons (both encoding for glutamine) to form the most common repeat length of 22 repeats as follows: (CAG)8-CAA(CAG)4CAA(CAG)8 [17]. The first disease linked to polyQ repeat expansions in *ATXN2* was spinal cerebellar ataxia type 2 (SCA2) (Box 2) [18]. Sequencing revealed that the most frequent *ATXN2* disease allele in SCA2 consists of 37 uninterrupted CAG repeats [17,19,20]. In rare cases, repeat lengths of 32–33 CAG repeats were found to give rise to late-onset disease [21,22].

Expansion of the *ATXN2* polyQ repeat length in ALS was first reported in a study looking for toxic modifiers of TDP-43 [11]. Intermediate-length polyQ repeats (24–33) were significantly associated with ALS, with an ideal cut-off set at ≥27 repeats [11]. This association between intermediate polyQ repeats in *ATXN2* and ALS has been further established in larger European cohorts, in which the association is strongest at a cut-off of ≥29 repeats, which has been established through meta-analysis of the original Northern American and European cohorts combined [23,24]. Therefore, the specific cut-off for *ATXN2* repeat length and risk for ALS may vary among different human populations. To assess whether polyQ expansion in ALS was specific to *ATXN2* or could also be observed in other genes, seven other polyQ disease genes were analyzed in a Northern American cohort for their association with ALS. No associations were found between ALS and ataxin-1 (*ATXN1*), ataxin-3 (*ATXN3*), ataxin-6 (*ATXN6*), ataxin-7 (*ATXN7*), TATA-binding protein (*TBP*), atrophin-1 (*ATN1*), and huntingtin (*HTT*) [25]. For *HTT*, these results were replicated in a larger European cohort, in which no association with polyQ repeat length and ALS was found [26], as was also the case for *ATXN3* [27]. However, another study focusing on a European cohort recently found that both *ATXN1* (≥29 repeats) and *ATXN2* intermediate polyQ expansions are independently associated with an increased risk for ALS [28]. It is possible that this difference between studies reflects variation among populations, as has been reported for the cut-off for *ATXN2* repeat length and the risk for ALS.

A wide variety of neurological disorders are caused by repeat expansions in different genes. Given the fact that several diseases are linked to (intermediate) repeat expansions in *ATXN2* (e.g., SCA2, SCA2–parkinsonism, and ALS), other neurodegenerative disorders were studied to examine whether these are also associated with *ATXN2* (see Table 1 for overview). In a large study, an