

Rethinking ALS: The FUS about TDP-43

Clotilde Lagier-Tourenne¹ and Don W. Cleveland^{1,*}

¹Ludwig Institute for Cancer Research and Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA 92093-0670, USA

*Correspondence: dcleveland@ucsd.edu

DOI 10.1016/j.cell.2009.03.006

Mutations in TDP-43, a DNA/RNA-binding protein, cause an inherited form of the neurodegenerative disease amyotrophic lateral sclerosis (ALS). Two recent studies (Kwiatkowski et al., 2009; Vance et al., 2009) now report that mutations in FUS/TLS, another DNA/RNA-binding protein, also trigger premature degeneration of motor neurons. TDP-43 and FUS/TLS have striking structural and functional similarities, implicating alterations in RNA processing as a key event in ALS pathogenesis.

Amyotrophic lateral sclerosis (ALS) is an adult-onset neurodegenerative disorder in which premature loss of motor neurons leads to fatal paralysis with a typical disease course of 1 to 5 years. Most forms of ALS are sporadic, but ~10% of patients have an inherited familial form of the disease and a clear family history. Understanding of ALS pathogenesis began with the landmark discovery of dominant causative mutations in the gene encoding copper/zinc superoxide dismutase 1 (*SOD1*) in ~20% of familial ALS cases and ~1% of sporadic cases. Subsequently, other even rarer familial cases with atypical disease features were linked to mutations in several other genes.

Most efforts to understand ALS pathogenesis over the past 15 years have focused on mutations in the ubiquitously expressed *SOD1*. No consensus has yet emerged as to how *SOD1* mutations lead to selective premature death of motor neurons, except that damage within motor neurons expressing mutant *SOD1* drives disease onset, whereas damage within their glial cell neighbors expressing mutant *SOD1* accelerates disease progression (Yamanaka et al., 2008). Multiple toxicity pathways have been implicated including the ability of misfolded mutant *SOD1* to trigger aberrant mitochondrial function, endoplasmic reticulum stress pathways, axonal transport defects, or excessive production of extracellular superoxide radicals. Views about ALS pathogenesis

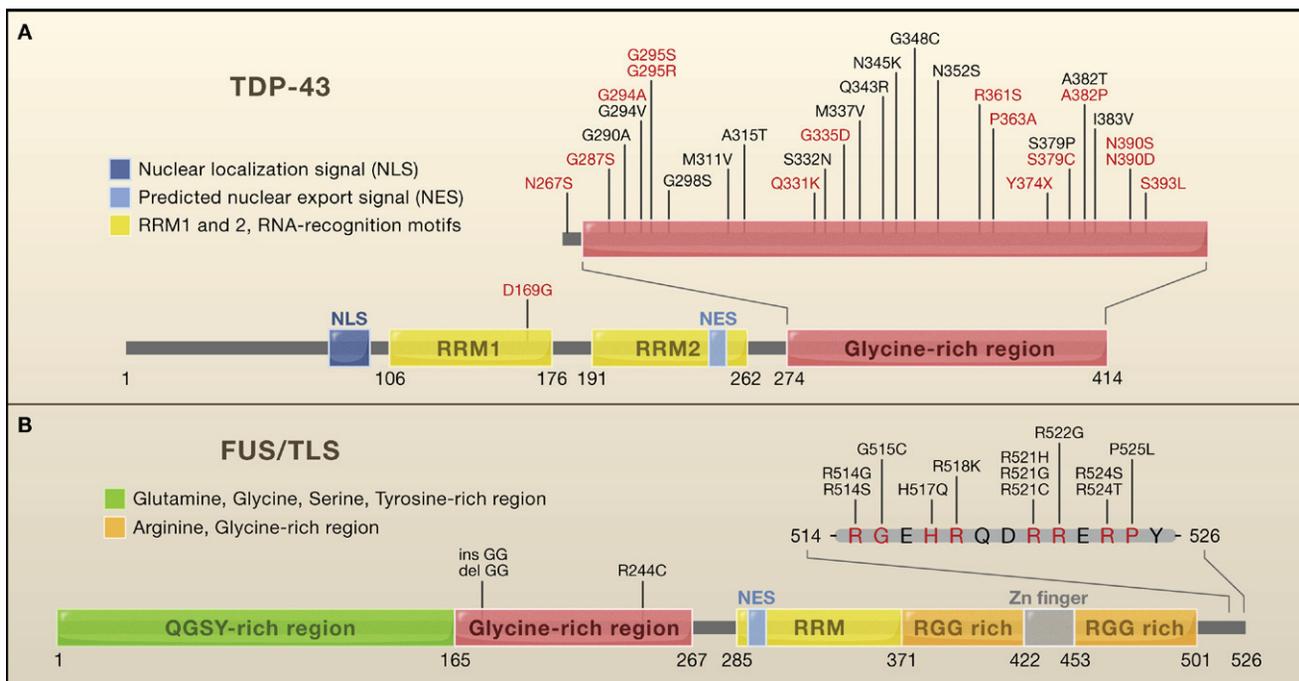


Figure 1. TDP-43 and FUS/TLS Mutations in ALS

(A) Thirty dominant mutations in TDP-43 have been identified in sporadic (red) and familial (black) ALS patients, with most lying in the C-terminal glycine-rich region of TDP-43. All are missense mutations, except for the truncating mutation TDP-43^{V374X}.

(B) Fourteen mutations in FUS/TLS have been identified in familial ALS cases, with most lying in the final 13 amino acids of this protein (R514S and G515S are found in cis). (Data compiled from Kwiatkowski et al., 2009; Vance et al., 2009; domains from <http://www.uniprot.org> and <http://www.cbs.dtu.dk/services/NetNES>.)

are now undergoing a seismic shift triggered by the recent discovery of mutations in a pair of DNA/RNA-binding proteins called TDP-43 (Gitcho et al., 2008; Sreedharan et al., 2008; Kabashi et al., 2008) and FUS/TLS (Kwiatkowski et al., 2009; Vance et al., 2009) (Figure 1) as causes of familial and sporadic forms of ALS.

TDP-43 Mutations in Familial and Sporadic ALS

This seismic shift in our understanding of ALS pathogenesis began with the identification (Arai et al., 2006; Neumann et al., 2006) of the 43 kDa TAR DNA-binding protein (TDP-43) as a major component of ubiquitinated protein aggregates found in many patients with sporadic ALS or the most common form of frontotemporal dementia called FTLD-U (frontotemporal lobar degeneration with ubiquitinated inclusions). In ALS and FTLD-U patients, TDP-43 immunoreactive inclusions are observed in the cytoplasm and nucleus of both neurons and glial cells. The brains and spinal cords of patients with TDP-43 proteinopathy present a biochemical signature that is characterized by abnormal hyperphosphorylation and ubiquitination of TDP-43 and the production of ~25 kDa C-terminal fragments that are missing their nuclear targeting domains (Arai et al., 2006; Neumann et al., 2006). TDP-43 is partly cleared from the nuclei of neurons containing cytoplasmic aggregates (Figure S1A available online) (e.g., Neumann et al., 2006; Van Deerlin et al., 2008) supporting the notion that pathogenesis of ALS in these cases may be driven, at least in part, by loss of normal TDP-43 function in the nucleus. Combined with a flurry of subsequent reports, TDP-43 inclusions are now recognized as a common characteristic of most ALS patients, with the striking exception of patients with familial ALS caused by *SOD1* mutations.

Although identification of TDP-43 aggregates proved to be a breakthrough, the pathology alone left it unclear whether aggregation of TDP-43 is a primary event in ALS pathogenesis or whether it is a byproduct of the disease process. Accumulation of intracellular or extracellular misfolded or misprocessed proteins in the central nervous system is a feature of many neurodegenerative conditions. Rare mutations have been found in the genes encoding misfolded proteins implicated in Alzheimer's disease, Parkinson's disease, the tauopathies, and prion diseases. Thus, the gene encoding TDP-43 on chromosome 1, *TARDBP*, constituted an excellent candidate for direct sequencing in search of disease-causing mutations in cohorts of patients with motor neuron disease or frontotemporal dementia.

Starting in early 2008, dominant mutations in the *TARDBP* gene were reported by several groups as a primary cause of ALS (e.g., Corrado et al., 2009; Daoud et al., 2009; Gitcho et al., 2008; Kabashi et al., 2008; Sreedharan et al., 2008; Van Deerlin et al., 2008; for review see Banks et al., 2008). These studies collectively provided persuasive evidence that the aberrant form of TDP-43 can directly trigger neurodegeneration. A total of 30 different mutations are now known in 22 unrelated families (~3% of familial ALS cases) and in 29 sporadic cases of ALS (~1.5% of sporadic cases) (Figure 1A). Given that linkage of familial ALS to chromosome 1 had not been identified previously, key among these genetic efforts was a retrospective analysis of a large family in which direct sequencing of TDP-43 revealed a TDP-43 mutation M337V. This study (Sreedharan et al., 2008) identified linkage between the disease and only one genomic region—an

8.2 Mb region on chromosome 1p36 containing the *TARDBP* gene. Although unconventional, this approach provided strong support for the pathogenic effect of the TDP-43^{M337V} mutation.

Widely expressed and predominantly nuclear, TDP-43 is 414 amino acids long and is encoded by six exons. This protein contains two RNA-recognition motifs (RRM1 and 2) and a C-terminal glycine-rich region that may mediate interactions with other proteins. All but one of the mutations identified so far are localized in the C-terminal region encoded by exon 6 of *TARDBP* (Figure 1A). All of these mutations are dominantly inherited missense changes with the exception of a truncating mutation (Y374X) at the extreme C terminus of the protein (Daoud et al., 2009). The pathogenicity of these missense changes is strongly supported by several lines of evidence. First, they affect amino acids that are highly conserved throughout evolution. Second, they have not been found in large cohorts of control individuals who do not have ALS. Third, in familial ALS cases where DNA is available, the mutations segregate with the disease and no mutations have been found in unaffected family members (except those below the typical age of disease onset). This indicates a high degree of penetrance for *TARDBP* mutations in these families, although further studies are needed as mutations in TDP-43 have been identified in patients with apparent sporadic ALS. Collectively, the evidence is now overwhelming that aberrant TDP-43 can trigger ALS.

Patients with TDP-43 mutations develop typical ALS with some variability within families in the site and age of onset. Although ~50% of all ALS patients develop cognitive impairment of varying severity, only one patient carrying a TDP-43 mutation has been reported to develop cognitive deficits (Corrado et al., 2009). This is despite the presence of TDP-43 inclusions in neurons and glial cells within the spinal cords and throughout the brains of ALS patients (e.g., Neumann et al., 2006; Van Deerlin et al., 2008). Diffuse granular cytoplasmic staining of TDP-43 (which may represent an earlier stage of inclusion development) and nuclear clearing of TDP-43 have also been described in the spinal cords and brains of ALS patients carrying TDP-43 mutations (Figure S1). However, it is unclear whether TDP-43 mutations lead to motor neuron loss through a gain of one or more toxic properties or a loss of normal function arising from sequestration of the protein in nuclear or cytoplasmic inclusions and the corresponding disruption of its interactions with protein partners or RNA targets.

FUS/TLS Mutations in Familial ALS

The identification of TDP-43 in ALS pathogenesis fueled the discovery recently reported by Kwiatkowski et al. (2009) and Vance et al. (2009) in *Science* of additional ALS mutations in a gene encoding another DNA/RNA-binding protein called FUS (fused in sarcoma) or TLS (translocation in liposarcoma). Previous reports had identified linkage between chromosome 16 and a familial form of ALS, but the underlying mutations were not known. Based on the knowledge that TDP-43 is a DNA/RNA-binding protein, Vance et al. (2009) prioritized sequencing of genes within the linkage region identified in a large British family with familial ALS so as to target genes encoding DNA/RNA-binding proteins. This led to their identification of a dominant missense mutation (R521C) in the *FUS/TLS* gene. A survey of 197 familial ALS cases identified the same R521C mutation in four additional families, as well as two additional missense mutations in another four families.

Independently, Kwiatkowski et al. (2009) used a linkage study in an ALS family originating from the Cape Verde islands in which disease transmission was compatible with an autosomal recessive inheritance pattern. A region of homozygosity by descent shared by all affected members of this family was identified on chromosome 16. This region overlapped with the previously reported ALS locus and contained the *FUS/TLS* gene. Again, homing in on genes encoding DNA/RNA-binding proteins led to screening of the *FUS/TLS* locus for mutations, resulting in identification of a homozygous missense mutation (H517Q) in all affected members. Although three healthy siblings were also homozygous for this mutation, they were younger than the typical age of disease onset. None of the individuals heterozygous for the mutation developed ALS, confirming autosomal recessive inheritance of this particular mutation. Subsequent screening in 292 familial ALS cases identified 12 dominant mutations in 16 families including two large families previously shown to have linkage to chromosome 16 (Kwiatkowski et al., 2009). No *FUS/TLS* mutations were found in a survey of 293 patients with sporadic ALS.

Combining the efforts of both teams, *FUS/TLS* mutations were detected in ~4% of familial ALS (~0.4% of all ALS). As is the case for TDP-43 mutations, all patients developed classical ALS with no cognitive deficits. Except for the recessive mutation in the family of Cape Verdean origin, the inheritance pattern is dominant (albeit with an incomplete penetrance reported for the R521G mutation) (Kwiatkowski et al., 2009).

The *FUS/TLS* protein is 526 amino acids long and is encoded by 15 exons. It is characterized by an N-terminal domain enriched in glutamine, glycine, serine, and tyrosine residues (QGSY region), a glycine-rich region, an RNA-recognition motif (RRM), multiple RGG repeats implicated in RNA binding, a C-terminal zinc finger motif, and a highly conserved extreme C-terminal region (Figure 1B). The vast majority of ALS-linked mutations are clustered in the extreme C terminus, with reports of mutations in all five arginine residues in this region. All mutations are missense changes except for two, both of which are located in the glycine-rich region and correspond to an insertion or a deletion of two glycines in a 10 glycine-long tract.

Like TDP-43, *FUS/TLS* is almost ubiquitously expressed. It is mainly localized in the nucleus, but cytoplasmic accumulation has been detected in most cell types. Analysis of the brains and spinal cords of ALS patients with *FUS/TLS* mutations revealed normal staining of *FUS/TLS* in the nuclei of many neurons and glial cells but aggregates of *FUS/TLS* in the cytoplasm of neurons (Kwiatkowski et al., 2009; Vance et al., 2009) (Figure S1C). It has not been reported whether *FUS/TLS* inclusions are also present in the cytoplasm of glial cells. Cell fractionation experiments after expression of tagged wild-type or mutant *FUS/TLS* confirmed an increase in the cytoplasmic accumulation of this mutant protein (Kwiatkowski et al., 2009; Vance et al., 2009).

A very curious aspect of mutant TDP-43 pathology is its partial clearance from the nucleus of either neuronal or glial cells when there are aggregates of TDP-43 in the cytoplasm. In a minority of neurons from ALS patients with *FUS/TLS* mutations (Figure S1C) or cells transfected to express fluorescently tagged mutant *FUS/TLS*^{R521G}, a uniquely cytoplasmic pattern of *FUS/TLS* aggregates has been reported. Cytoplasmic inclusions containing the *FUS/TLS* protein are absent in normal individuals, in ALS patients

with *SOD1* mutations, and in sporadic ALS patients who presumably are positive for TDP-43 aggregates. Importantly, TDP-43-positive inclusions are absent in ALS patients with *FUS/TLS* mutations, implying that neurodegenerative processes driven by *FUS/TLS* mutations are independent of TDP-43 aggregation (Vance et al., 2009). It will now be essential to assess *FUS/TLS* accumulation and localization in ALS patients with TDP-43 mutations, as well as in patients with other neurodegenerative diseases, especially those with mislocalized TDP-43.

TDP-43 and *FUS/TLS* in Gene Regulation

The precise roles of TDP-43 and *FUS/TLS* have not been fully elucidated, but both are multifunctional proteins that have been implicated in several steps of gene expression regulation including transcription, RNA splicing, RNA transport, and translation (for review see Buratti and Baralle, 2008; Janknecht, 2005). They might also be involved in the processing of microRNAs, and *FUS/TLS* may play a role in the maintenance of genomic integrity. Both proteins contain RNA-binding motifs and are structurally close to a family of heterogeneous ribonucleoproteins (hnRNPs). Indeed, *FUS/TLS* is sometimes referred to as hnRNP P2. Consistently, both TDP-43 and *FUS/TLS* directly bind to RNA, as well as to single- and double-stranded DNA.

TDP-43 was initially proposed to repress transcription by binding to the TAR DNA sequence of human immunodeficiency virus type-1 and to the mouse SP-10 gene promoter, but little is known about the mechanisms and selectivity of transcriptional repression. The normal function of *FUS/TLS* has been studied more extensively following its identification as a fusion protein generated by chromosomal translocations in human cancers. It is a member of the TET protein family that also includes the Ewing's sarcoma protein and the TATA-binding protein-associated factor (TAF_{II}68). Wild-type *FUS/TLS* associates with both general and more specialized factors, presumptively influencing transcription initiation. Indeed, *FUS/TLS* interacts with several nuclear hormone receptors and with gene-specific transcription factors. It also associates with the general transcriptional machinery, interacting with RNA polymerase II and the TFIID complex.

Recently, a very interesting and unexpected mechanism of transcriptional regulation was described for *FUS/TLS* (Wang et al., 2008). In response to DNA damage, *FUS/TLS* is recruited by sense and antisense noncoding RNAs transcribed in the 5' regulatory region of the gene encoding cyclin D1. Then *FUS/TLS* binds and inhibits CREB-binding protein and p300 histone acetyltransferase activities leading to the repression of *cyclin D1* transcription (Wang et al., 2008). This provides a direct link between the RNA-binding properties of *FUS/TLS* and a role in transcriptional regulation. Moreover, this kind of regulation might be more general in light of four recent reports in *Science* demonstrating that production of short sense and antisense noncoding RNAs upstream of the active transcription start site occurs in other contexts (e.g., Core et al., 2008).

TDP-43 and *FUS/TLS* in RNA Splicing and Localization

Beyond transcription, TDP-43 and *FUS/TLS* have been implicated in RNA maturation and splicing. Only a few of their respective RNA targets have been identified, and a comprehensive map

of their RNA targets is a crucial next goal. Recent technologies using high-throughput sequencing have demonstrated that a single RNA-binding protein can affect many alternatively spliced transcripts (Licatalosi et al., 2008). Such approaches will be necessary to understand the role of TDP-43 and FUS/TLS in neurodegeneration. Indeed, observation of a widespread mRNA splicing defect in diseases characterized by aggregation of TDP-43 or FUS/TLS would reinforce the crucial role of splicing regulation in neuronal integrity and potentially could identify candidate genes whose altered splicing is central to ALS pathogenesis. It should not be overlooked that TDP-43 and FUS/TLS also may be involved in microRNA processing as both have been found by mass spectrometry to associate with Droscha (Gregory et al., 2004).

Despite their enrichment in the nucleus and potential roles in nuclear RNA maturation, TDP-43 and FUS/TLS shuttle between the nucleus and cytosol. In addition, both are found in granules associated with RNA transport in neurons, with translocation to dendritic spines following different neuronal stimuli (e.g., Fujii et al., 2005). Moreover, abnormal spine morphology is observed in cultured neurons from FUS/TLS knockout mice (Fujii et al., 2005). These results suggest that both proteins could play a role in the modulation of neuronal plasticity or other properties by altering mRNA transport and local RNA translation in neurons.

Beyond ALS: The Next Steps

The identification of causative TDP-43 and FUS/TLS mutations, along with TDP-43 pathology in most cases of ALS, represents a dramatic shift in our understanding of this disease. Now we need to define the normal roles of TDP-43 and FUS/TLS and to determine whether mutant forms of these proteins and their abnormal aggregation lead to general or specific alterations in gene expression. Both cellular and animal models will be essential to define the link between mutations in TDP-43 and FUS/TLS and disease.

The lessons to be learned for TDP-43 and FUS/TLS in triggering neurodegenerative disease will not be unique to ALS. TDP-43 aggregation is present in most sporadic and familial FTL-D-U patients including those with mutations in progranulin and valosin-containing protein. Moreover, abnormal TDP-43 inclusions have been reported for several other neurodegenerative conditions, including in ~30% of Alzheimer's disease patients (reviewed in Banks et al., 2008). Wild-type FUS/TLS was recently identified as a major component of polyQ aggregates in cellular models of spinal cerebellar ataxia type 3 and Huntington's disease (Doi et al., 2008). The latter observation was confirmed by the finding of intranuclear inclusions in neurons from Huntington's disease patients, provoking the proposal that the protein binds directly to polyQ aggregates at an early stage of disease (Doi et al., 2008).

Discovery of the involvement of TDP-43 and FUS/TLS in ALS and other neurodegenerative diseases reinforces the role of altered RNA processing in neurodegeneration. Earlier well-established examples of altered RNA processing in neurodegeneration include errors in RNA metabolism due to loss of the SMN (survival of motor neuron) protein in spinal muscular atrophy and of FMRP in fragile-X mental retardation. In addition, an RNA gain-of-function mechanism has been implicated in a set of diseases including the myotonic dystrophies, where a transcript with an abnormal repeat expansion alters the function and local-

ization of alternative splicing regulators. The emerging TDP-43 and FUS/TLS stories add considerable support to the proposal that defects in RNA processing play a central role in neurodegenerative diseases.

Supplemental Data

Supplemental Data contain one Figure and can be found with this article online at [http://www.cell.com/supplemental/S0092-8674\(09\)00263-3](http://www.cell.com/supplemental/S0092-8674(09)00263-3).

REFERENCES

- Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., et al. (2006). *Biochem. Biophys. Res. Commun.* 351, 602–611.
- Banks, G.T., Kuta, A., Isaacs, A.M., and Fisher, E.M. (2008). *Mamm. Genome* 19, 299–305.
- Buratti, E., and Baralle, F.E. (2008). *Front. Biosci.* 13, 867–878.
- Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). *Science* 322, 1845–1848.
- Corrado, L., Ratti, A., Gellera, C., Buratti, E., Castellotti, B., Carlomagno, Y., Ticozzi, N., Mazzini, L., Testa, L., Taroni, F., et al. (2009). *Hum. Mut.* Published online February 17, 2009. 10.1002/humu.20950.
- Daoud, H., Valdmanis, P.N., Kabashi, E., Dion, P., Dupre, N., Camu, W., Meisinger, V., and Rouleau, G.A. (2009). *J. Med. Genet.* 46, 112–114.
- Doi, H., Okamura, K., Bauer, P.O., Furukawa, Y., Shimizu, H., Kurosawa, M., Machida, Y., Miyazaki, H., Mitsui, K., Kuroiwa, Y., et al. (2008). *J. Biol. Chem.* 283, 6489–6500.
- Fujii, R., Okabe, S., Urushido, T., Inoue, K., Yoshimura, A., Tachibana, T., Nishikawa, T., Hicks, G.G., and Takumi, T. (2005). *Curr. Biol.* 15, 587–593.
- Gitcho, M.A., Baloh, R.H., Chakraverty, S., Mayo, K., Norton, J.B., Levitch, D., Hatanpaa, K.J., White, C.L., 3rd, Bigio, E.H., Caselli, R., et al. (2008). *Ann. Neurol.* 63, 535–538.
- Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). *Nature* 432, 235–240.
- Janknecht, R. (2005). *Gene* 363, 1–14.
- Kabashi, E., Valdmanis, P.N., Dion, P., Spiegelman, D., McConkey, B.J., Vande Velde, C., Bouchard, J.P., Lacomblez, L., Pochigaeva, K., Salachas, F., et al. (2008). *Nat. Genet.* 40, 572–574.
- Kwiatkowski, T.J., Bosco, J.D., LeClerc, A.D., Tamrazian, E., Van den Berg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., et al. (2009). *Science* 323, 1205–1208.
- Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). *Nature* 456, 464–469.
- Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., et al. (2006). *Science* 314, 130–133.
- Sreedharan, J., Blair, I.P., Tripathi, V.B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E., et al. (2008). *Science* 319, 1668–1672.
- Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D.M., Wright, P., et al. (2009). *Science* 323, 1208–1211.
- Van Deerlin, V.M., Leverenz, J.B., Bekris, L.M., Bird, T.D., Yuan, W., Elman, L.B., Clay, D., Wood, E.M., Chen-Plotkin, A.S., Martinez-Lage, M., et al. (2008). *Lancet Neurol.* 7, 409–416.
- Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., and Kurokawa, R. (2008). *Nature* 454, 126–130.
- Yamanaka, K., Chun, S.J., Boillee, S., Fujimori, N., Yamashita, H., Gutmann, D.H., Misawa, H., Takahashi, R., and Cleveland, D.W. (2008). *Nat. Neurosci.* 11, 251–253.