ARTICLES

Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS

Andrew C. Elden[1]*, Hyung-Jun Kim[2]*, Michael P. Hart[1]*, Alice S. Chen-Plotkin[3,4]*, Brian S. Johnson[1], Xiaodong Fang[1], Maria Armakola[1], Felix Geser[3], Robert Greene[3], Min Min Lu[1], Arun Padmanabhan[1], Dana Clay-Falcone[3], Leo McCluskey[4], Lauren Elman[4], Denise Juhr[5], Peter J. Gruber[5], Udo Rüb[6], Georg Auburger[7], John Q. Trojanowski[3], Virginia M.-Y. Lee[3], Vivianna M. Van Deerlin[3], Nancy M. Bonini[2] & Aaron D. Gitler[1]

The causes of amyotrophic lateral sclerosis (ALS), a devastating human neurodegenerative disease, are poorly understood, although the protein TDP-43 has been suggested to have a critical role in disease pathogenesis. Here we show that ataxin 2 (ATXN2), a polyglutamine (polyQ) protein mutated in spinocerebellar ataxia type 2, is a potent modifier of TDP-43 toxicity in animal and cellular models. ATXN2 and TDP-43 associate in a complex that depends on RNA. In spinal cord neurons of ALS patients, ATXN2 is abnormally localized; likewise, TDP-43 shows mislocalization in spinocerebellar ataxia type 2. To assess the involvement of ATXN2 in ALS, we analysed the length of the polyQ repeat in the *ATXN2* gene in 915 ALS patients. We found that intermediate-length polyQ expansions (27–33 glutamines) in ATXN2 were significantly associated with ALS. These data establish *ATXN2* as a relatively common ALS susceptibility gene. Furthermore, these findings indicate that the TDP-43–ATXN2 interaction may be a promising target for therapeutic intervention in ALS and other TDP-43 proteinopathies.

ALS, also known as Lou Gehrig's disease, is a devastating adult-onset neurodegenerative disease with no cure[1]. The disease—the causes of which are unclear—is mostly sporadic but approximately 10% of cases have a first- or second-degree relative with ALS (familial ALS). Mutations in *SOD1*, encoding Cu/Zn superoxide dismutase, have been identified in ~20% of familial ALS cases[2], for an overall incidence of ~2%. Further ALS disease genes have been identified that are even rarer. Identifying new and potentially common genetic risk factors for ALS will accelerate understanding of the disease, aid the development of biomarkers and spur innovative new treatments.

Recently, the 43-kDa TAR DNA-binding protein (TDP-43) was identified as having a major involvement in sporadic and familial ALS. In 2006, TDP-43 was identified as the major disease protein in ubiquitinated cytoplasmic inclusions in neurons of patients with ALS and frontotemporal lobar degeneration with ubiquitinated inclusions[3]. Subsequently, mutations in the gene encoding TDP-43 (*TARDBP*) were found to be associated with some sporadic and familial cases of ALS and frontotemporal lobar degeneration with ubiquitinated inclusions[4,5], indicating a central role of TDP-43 in disease pathogenesis. TDP-43 is normally a nuclear protein but pathological inclusions contain cytoplasmic TDP-43 aggregates, indicating that altered subcellular localization of the protein may be critical to disease pathogenesis[6]. Little is known about how loss of one or more of the biological functions of TDP-43, or how a potential toxic gain-of-function activity, might contribute to neurodegenerative disease. Moreover, nothing is known about genetic modifiers of TDP-43 pathogenesis or how other factors that interact with TDP-43 contribute to the risk of developing ALS or the age of disease onset.

In an unbiased screen to define modifiers of TDP-43 toxicity in yeast, we identified the yeast orthologue of ataxin 2, Pbp1, as a potent, dose-sensitive modulator of TDP-43 toxicity, and extended the finding of a role for ataxin 2 across several model systems. We show that the two proteins associate in a complex and are mislocalized in ALS patient spinal cord neurons. Given that human *ATXN2* is a polyQ disease gene, we analysed the length of the polyQ repeat in over 900 sporadic and familial ALS patients. This revealed a significant association of *ATXN2* intermediate-length polyQ tract expansions with ALS (4.7% of cases). We propose that polyQ expansions in *ATXN2* are a new and potentially common risk factor for ALS. Furthermore, these findings indicate that the TDP-43–ATXN2 interaction may be a promising target for therapeutic intervention.

Pbp1 modifies TDP-43 toxicity in yeast

To gain insight into the mechanisms of TDP-43 pathogenesis, we used an unbiased genetic approach to identify genes that could suppress or enhance TDP-43 toxicity in yeast. Similar approaches have been used to discover modifiers of the Parkinson's disease protein α-synuclein[7,8]. We individually transformed 5,500 yeast genes, which comprise the yeast FLEXGene plasmid overexpression library[7] into a yeast strain expressing TDP-43. Thirteen genes were identified that suppressed and twenty-seven genes that enhanced TDP-43 toxicity when overexpressed (A.C.E. and A.D.G., unpublished data). The largest functional class enriched in the screen included RNA-binding proteins and proteins involved in RNA metabolism. We also identified kinases and proteases as potent modifiers of TDP-43 toxicity. Importantly, of 71 genes from this library that have been previously

[1]Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [2]Department of Biology, Howard Hughes Medical Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [3]Department of Pathology and Laboratory Medicine and Center for Neurodegenerative Disease Research, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [4]Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [5]The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [6]Institute of Clinical Neuroanatomy, Dr Senckenberg Anatomy, Goethe University, Frankfurt am Main D-60590, Germany. [7]Molecular Neurogenetics, Department of Neurology, Goethe University, Frankfurt am Main D-60528, Germany.
*These authors contributed equally to this work.

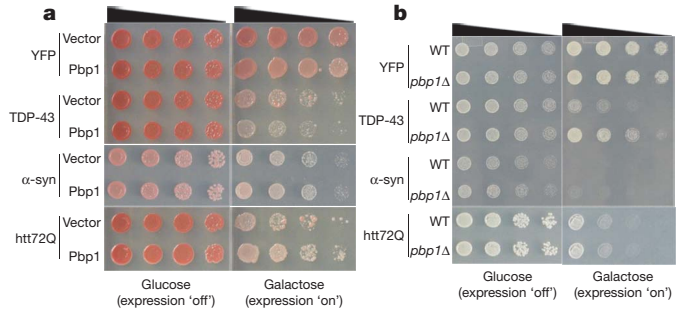


Figure 1 | Pbp1 is a dose-sensitive modifier of TDP-43 toxicity in yeast. **a**, Yeast spotting assays with yeast TDP-43 showing toxicity. Fivefold serial dilutions of yeast cells spotted onto glucose (expression repressed) or galactose (expression induced). Upregulation of *PBP1* enhances TDP-43 toxicity. Whereas Pbp1 has no effect on yeast viability when expressed with the control protein YFP, when co-expressed with TDP-43, it enhances toxicity. Enhancement is specific because Pbp1 does not affect the toxicity of a pathogenic huntingtin fragment (htt72Q) or α-synuclein (α-syn). **b**, Yeast spotting assays with yeast TDP-43 showing that *PBP1* deletion (*pbp1*Δ) suppresses TDP-43 toxicity. Whereas expression of TDP-43 from a plasmid in wild-type yeast (WT) was toxic, this was mitigated in *pbp1*Δ cells. The effect was specific because α-synuclein or htt72Q toxicity was not suppressed by *pbp1*Δ.

shown to modify α-synuclein toxicity in yeast[7,8], only one also affected TDP-43 toxicity, underscoring the specificity of the screen for TDP-43 biology.

One yeast gene identified that enhanced TDP-43 toxicity, *PBP1* (Pab1-binding protein 1), was notable as an orthologue of the human *ATXN2* gene, mutations in which cause the neurodegenerative disease spinocerebellar ataxia type 2 (SCA2). SCA2 is one of a heterogeneous group of 28 autosomal dominant hereditary ataxias[9] and is caused by polyQ tract expansions in *ATXN2*[10–13]. Interestingly, in SCA2, as in ALS, motor neurons are also known to degenerate, but these features typically occur later than the cerebellar degeneration. However, in select cases, the motor neuron features of SCA2 are sufficiently prominent to mimic an ALS presentation[14,15], indicating the potential for clinicopathological overlap. Although the precise functions of yeast Pbp1 and human ATXN2 are not fully understood, Pbp1 interacts with Pab1 to regulate mRNA polyadenylation and is involved in stress granule assembly[16]. Processing (P) bodies and stress granules have important roles in regulating translation, mRNA degradation and the subcellular localization of mRNAs[17]. Upregulation of Pbp1 enhanced TDP-43 toxicity in yeast (Fig. 1a), whereas Pbp1 loss of function suppressed toxicity (Fig. 1b), indicating that Pbp1 is a dose-sensitive modifier of TDP-43 toxicity. Up- or downregulation of Pbp1 did not enhance the toxicity of other human neurodegenerative disease proteins, such as α-synuclein or a mutant huntingtin fragment (Fig. 1a, b), demonstrating specificity of the Pbp1 interaction for TDP-43.

Atx2 modifies TDP-43 toxicity in *Drosophila*

To test the relevance of the ATXN2–TDP-43 genetic interaction in the nervous system we used *Drosophila*. A series of transgenic lines were generated that expressed wild-type human TDP-43 or an ALS-linked mutant form (Q331K) at varying levels. This yielded a set of lines with milder or more severe effects, enabling the detection of genetic interactions in a sensitive manner (Fig. 2). Directing expression of TDP-43 to the eye of *Drosophila* caused a progressive, age-dependent degeneration of the structure (Fig. 2a and Supplementary Fig. 2a). TDP-43 induced a markedly shortened lifespan when expressed in the nervous

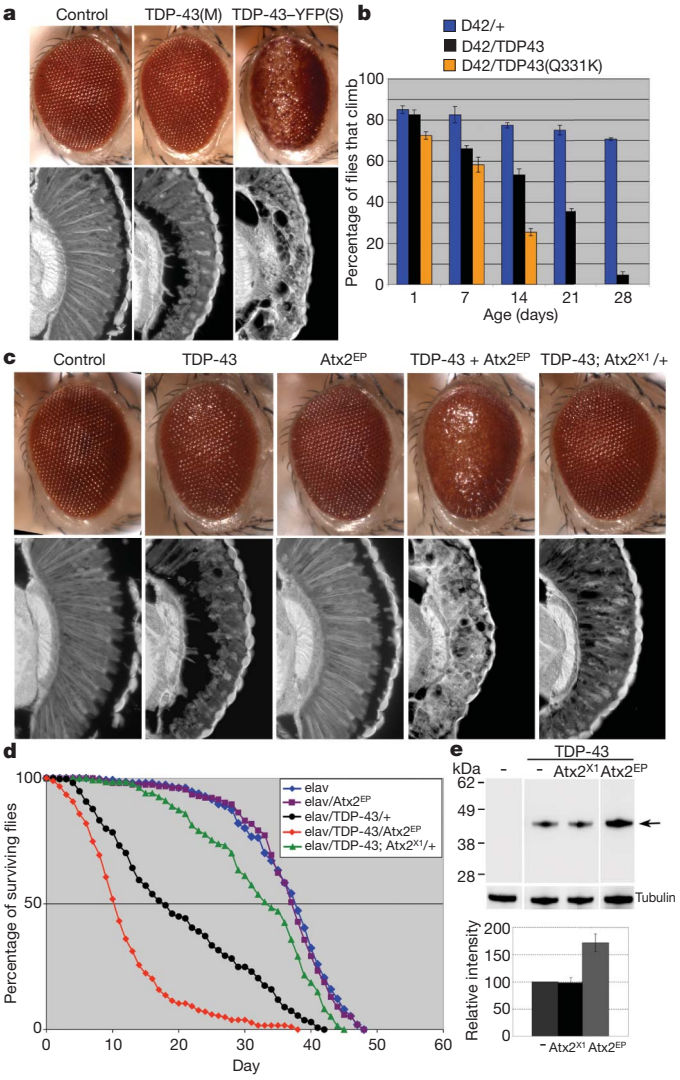


Figure 2 | Atx2 is a dose-sensitive modifier of TDP-43 toxicity in *Drosophila*. **a**, Expression of TDP-43 caused a dose-dependent disruption of retinal structure. Genotypes: *gmr-GAL4(YH3) in trans* to *UAS–GFP* (control), *UAS-TDP-43*(M) or *UAS-TDP-43–YFP*(S). (M) and (S) are moderate and strong TDP-43 expression, respectively (Methods and Supplementary Fig. 2). **b**, TDP-43 caused motility deficits when expressed in motor neurons. TDP-43(Q331K) caused a more severe loss of motility than the wild-type protein at the same level of expression (Supplementary Fig. 2). Genotypes: *D42-GAL4 in trans* to +, *UAS-TDP-43* or *UAS-TDP-43(Q331K)*. Error bars represent the mean ± 95% confidence interval of four climbing trials. **c**, Atx2 modulates TDP-43 toxicity. Flies expressing TDP-43 or Atx2 alone ($Atx2^{EP}$; Atx2 enhancer and promoter containing *P*-element (EP) insertion line $Atx2^{EP3145}$) have a mild effect on retinal structure. TDP-43 toxicity is more severe with upregulation of Atx2 (TDP-43 + $Atx2^{EP}$). TDP-43 toxicity is markedly mitigated on reduction of Atx2 (flies *in trans* to null allele $Atx2^{X1}$). See Methods for details of the genotypes. $Atx2^{X1}$/+ has wild-type retinal structure. **d**, Atx2 modulates the reduced lifespan conferred by TDP-43. Expression of TDP-43 in the nervous system reduces lifespan (black, compared to normal in blue). Upregulation of Atx2 causes more rapid death (red, compared to TDP-43 in black). Upregulation of Atx2 on its own has no effect (purple). Reduction of Atx2 significantly extends lifespan (green, compared to TDP-43 in black). Heterozygous loss of Atx2 on its own has no effect (not shown). Quantitative PCR with reverse transcription (RT–PCR) showed that TDP-43 expression has no effect on levels of *Atx2* transcript in $Atx2^{EP}$ flies. See Methods for details of the genotypes. Flies raised at 25 °C, lifespan performed at 29 °C. **e**, TDP-43 immunoblot on Atx2 modulation. Reduction of Atx2 has no effect on TDP-43 levels, whereas upregulation of Atx2 enhances TDP-43 protein levels. Atx2 has no effect on the transgene expression system and 2× TDP-43 does not produce the same effect as TDP-43+$Atx2^{EP}$ (Supplementary Figs 2 and 3). —, control lane. TDP-43 lanes show expression of TDP-43 alone (—) or with reduced Atx2 ($Atx2^{X1}$), or with upregulated Atx2 ($Atx2^{EP}$). Details of genotypes in Methods. Bottom, quantification of immunoblots from 3–5 independent experiments; normalized to tubulin. Data are presented as mean ± standard deviation of 3–5 independent immunoblots.

system (Fig. 2d) and led to a progressive loss of motility when directed to motor neurons (Fig. 2b). Notably, expression of an ALS-linked mutant TDP-43 had a more severe effect than wild-type TDP-43 (Fig. 2b and Supplementary Fig. 2b). Upregulation of the *Drosophila* homologue of ATXN2, Atx2, enhanced toxicity of TDP-43, resulting in more severe retinal degeneration (Fig. 2c) and a further reduction in lifespan (Fig. 2d). Atx2 upregulation also enhanced toxicity of an ALS-linked mutant TDP-43 (data not shown). The effect of Atx2 was markedly dose-dependent, as partial reduction of levels of Atx2 with a null allele mitigated TDP-43 toxicity in the eye (Fig. 2c) and markedly extended lifespan (Fig. 2d), indicating that toxicity of TDP-43 is extremely sensitive to the levels of Atx2. Atx2 enhanced TDP-43 protein accumulation when upregulated, although downregulation had no effect (Fig. 2e). The enhanced toxicity owing to Atx2 was not attributable simply to enhanced TDP-43 protein levels, because degeneration was markedly more severe on upregulation of Atx2 than with higher levels of TDP-43 only (Supplementary Fig. 3). The interaction was specific, as Atx2 had no effect on the expression of a control protein (β-galactosidase; Supplementary Fig. 4)[18,19]. Moreover, upregulation of the molecular chaperone Hsp70 did not suppress TDP-43 toxicity, as it does in models of Parkinson's disease and spinocerebellar ataxia type 3 (SCA3)[20,21], nor did upregulation of ataxin 3 (ATXN3) (the polyQ expansion of which is the basis of SCA3) have an effect (Supplementary Fig. 4). These data indicate that modulation of TDP-43 toxicity by Atx2 is conserved in the nervous system of *Drosophila*.

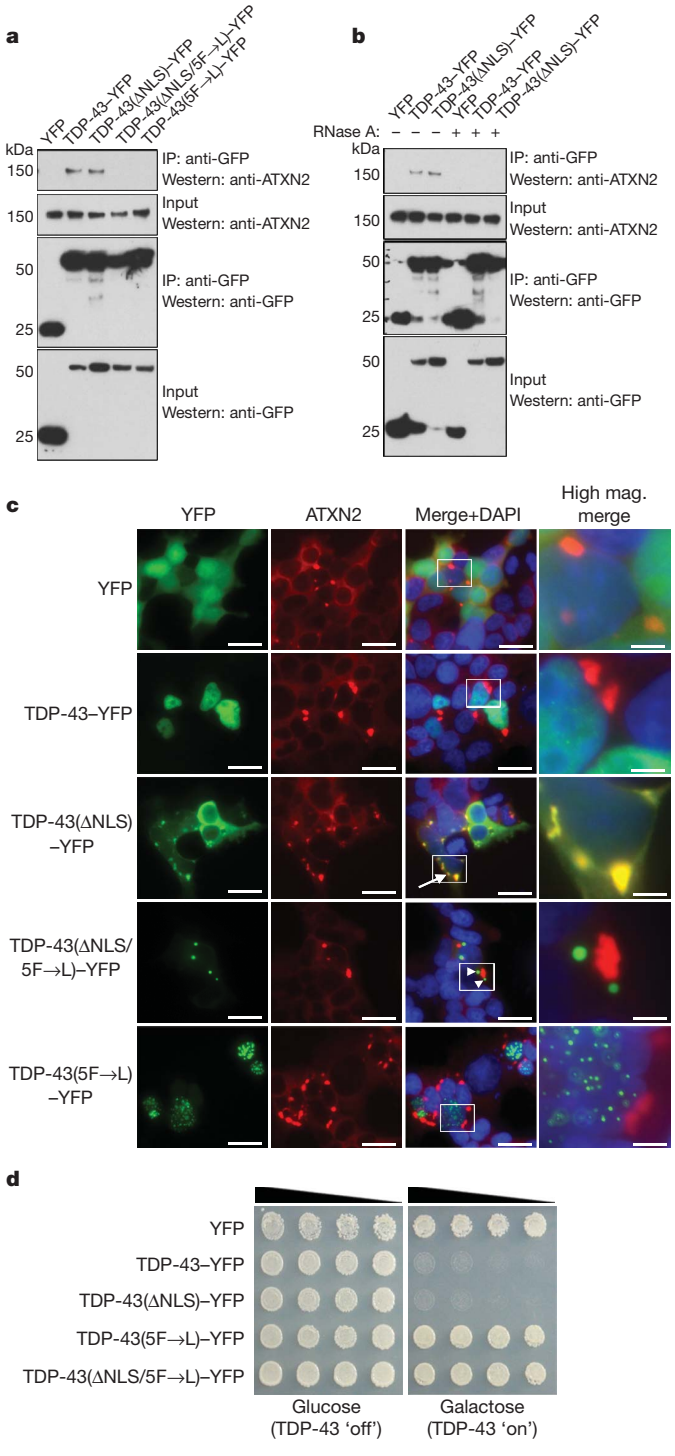


Figure 3 | ATXN2 and TDP-43 interact in a manner dependent on RNA. **a**, TDP-43 and ATXN2 associate in mammalian cells in a manner dependent on the RRMs. HEK293T cells were transfected with plasmids encoding YFP, TDP-43–YFP, TDP-43(ΔNLS)–YFP (NLS mutant localizes to cytoplasm), TDP-43(ΔNLS/5F→L)–YFP (NLS mutant + RNA-binding mutant), or TDP-43(5F→L)–YFP (RNA-binding mutant). Protein was immunoprecipitated (IP) with anti-GFP antibody and then subjected to immunoblotting with anti-ataxin-2 to detect endogenous ATXN2. Whereas TDP-43 and TDP-43(ΔNLS) both interact with ATXN2, RNA-binding mutant versions do not. **b**, Co-immunoprecipitation in HEK293T cells as in **a**, but now with lysates treated with RNase. The interaction between ATXN2 and TDP-43 seen normally was abolished on RNase treatment. **c**, HEK293T cells transfected with YFP-tagged wild-type and mutant TDP-43 constructs then immunostained for endogenous ATXN2. Normally, ATXN2 is localized to the cytoplasm forming occasional cytoplasmic accumulations. TDP-43 localized to the nucleus in a diffuse pattern. TDP-43(ΔNLS) localized to the cytoplasm where it occasionally formed cytoplasmic aggregates; these aggregates always co-localized with ATXN2 cytoplasmic accumulations (arrow). Abolishing the ability of TDP-43 to interact with RNA (TDP-43(ΔNLS/5F→L) or TDP-43(5F→L)) eliminated ATXN2 co-localization (white arrowheads pointing at green TDP-43 accmulations; see high magnification merge). TDP-43(5F→L)–YFP was restricted to the nucleus where it formed multiple foci. Scale bar is 2.5 μm for merge panels and 0.5 μm for high magnification merge panels (high mag. merge). **d**, Yeast spotting assays for TDP-43 toxicity. Whereas wild-type and TDP-43(ΔNLS) constructs are toxic, mutations of TDP-43 that prevent RNA binding abolish toxicity.

TDP-43 and ATXN2 interactions

Given the marked effects of Pbp1 and Atx2 on TDP-43 toxicity in yeast and *Drosophila*, we determined whether the two proteins could physically interact in yeast and human cells. Yeast cells were transformed with yellow fluorescent protein (YFP)-tagged TDP-43 and either cyan fluorescent protein (CFP)-tagged Pbp1 or CFP alone, and the localization of the proteins visualized by fluorescence microscopy. Upregulation of Pbp1 increased the number of TDP-43–YFP foci present in yeast cells compared to situations with TDP-43–YFP alone[22] (A.C.E. and A.D.G., unpublished observations). These studies showed that Pbp1 localized to TDP-43 cytoplasmic accumulations (Supplementary Fig. 5). To determine whether Pbp1 and TDP-43 could associate in the same protein complex, we performed immunoprecipitation assays with an antibody directed against the Pbp1 epitope tag followed by immunoblotting to detect TDP-43. These studies confirmed the ability of TDP-43 to interact with Pbp1 in the same complex (Supplementary Fig. 5). To determine whether this interaction was conserved in human cells, HEK293T cells were transfected with YFP-tagged TDP-43 or YFP alone. TDP-43–YFP, but not YFP alone, immunoprecipitated endogenous human ATXN2 (Fig. 3a). This interaction was specific to ATXN2 because ATXN3 did not co-immunoprecipitate with TDP-43 (Supplementary Fig. 6a). Finally, an ALS-linked TDP-43 mutant (Q331K) also interacted with ATXN2 (Supplementary Fig. 6c). These data indicate that ATXN2 and TDP-43 can physically interact in both yeast and human cells.

Because both TDP-43 and ATXN2 are involved in RNA metabolism[22,23], we considered that RNA binding may be important for

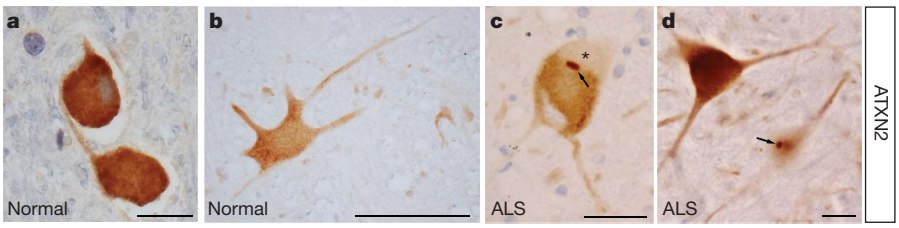
the TDP-43–ATXN2 interaction. TDP-43 is an RNA-recognition-motif (RRM)-containing protein with highly conserved RNP-1 and RNP-2 consensus motifs in each RRM. Within these motifs, specific aromatic residues have been shown to be necessary for RNA-base-stacking interactions, and mutation of these residues (Phe to Leu) reduces the ability of TDP-43 to bind RNA *in vitro*[23]. To address the importance of RNA binding by TDP-43, we mutated all five of these residues to generate a TDP-43(5F→L)–YFP protein, and determined the effect of this on the interaction with ATXN2. Mutation of the RRM domain abolished the ability of TDP-43 to interact with ATXN2 (Fig. 3a), indicating that RNA probably serves as a bridge between the two proteins. To rule out potential effects of these RRM mutations on protein–protein interactions or TDP-43 stability, we performed several further control experiments. We tested single, double and triple mutants within the RRMs of TDP-43 and found that even single mutations markedly diminished the interaction with ATXN2 (Supplementary Fig. 6b), with double and triple mutations abolishing the interaction completely (Supplementary Fig. 6c). To demonstrate further a role for RNA in mediating the TDP-43–ATXN2 interaction and exclude potential deleterious effects of point mutations on TDP-43 folding or stability, we performed the immunoprecipitation with full-length wild-type protein in the presence of RNase. RNase treatment abolished the interaction between TDP-43 and ATXN2 (Fig. 3b). Finally, HEK293T cells were transfected with YFP-tagged wild-type and RRM-domain mutant TDP-43 constructs and immunostained for endogenous ATXN2 (Fig. 3c). Consistent with previous reports, ATXN2 was predominantly localized to the cytosol and occasionally formed punctate cytoplasmic accumulations[24]. TDP-43–YFP appeared mostly in the nucleus, although it has been reported to shuttle back and forth to the cytoplasm[6]. However, a form of TDP-43 with a mutated nuclear localization signal (NLS; TDP-43(ΔNLS)–YFP) to restrict TDP-43 to the cytosol, which is its presumed pathogenic localization in disease[6], remained in the cytosol where it occasionally formed aggregates; these aggregates always co-localized with ATXN2 (Fig. 3c). Notably, mutating the RRM domains of TDP-43 in the context of the ΔNLS construct, despite also resulting in TDP-43 aggregation in the cytosol, never led to co-localization with ATXN2 (Fig. 3c). In addition to blocking the interaction between TDP-43 and ATXN2, mutating the RRMs of TDP-43 also completely eliminated TDP-43 toxicity in yeast (Fig. 3d). Taken together, these data indicate that TDP-43 and ATXN2 can, although perhaps transiently, interact in a complex in the cytoplasm—the site of toxic function of TDP-43 in disease—and that this interaction probably depends on RNA binding. In a broader sense, these results also highlight that cytoplasmic aggregation per se is not sufficient for toxicity but that an RNA-binding component is also involved, which provides some mechanistic insight into TDP-43 pathogenesis.

ATXN2 localization perturbed in ALS

The genetic interactions between TDP-43 and Pbp1/Atx2 in yeast and *Drosophila*, and the physical association in yeast and mammalian cells, indicated that ATXN2 might show abnormal localization in human disease. To address this, we examined ATXN2 localization in spinal cord neurons from six ALS patients and three neurologically normal controls (Fig. 4). Normally, ATXN2 is localized in a diffuse or fine-granular pattern throughout the cytoplasm of spinal cord neurons (Fig. 4a, b). However, in ALS spinal cord neurons, ATXN2 localization was altered, showing more distinct cytoplasmic accumulations (27% of ALS spinal cord neurons versus 5% of control neurons; Fig. 4c, d, see arrows). Interestingly, we did not observe a significant difference in ATXN2 localization in spinal cord neurons from ALS patients with normal or expanded ATXN2 polyQ lengths (Fig. 4e, cases 1–3 versus cases 4–6; see later). Importantly, mislocalization of ATXN2 in ALS appeared specific because ATXN3 localization was unaffected (Supplementary Fig. 7). These studies indicate that ATXN2 localization is altered in spinal cord neurons of ALS patients (see Supplementary Results for further analysis of ATXN2 localization in frontotemporal lobar degeneration with ubiquitinated inclusions and TDP-43 analysis in SCA2).

ATXN2 polyQ expansions in ALS

These genetic, biochemical and neuropathological interactions between ATXN2 and TDP-43 raised the possibility that mutations in *ATXN2* could have a causative role in ALS. The ATXN2 polyQ tract length, although variable, is most frequently 22–23, with expansions of >34 causing SCA2[10–13]. However, the variable nature of the ATXN2 repeat indicated a mechanism by which such mutations in ATXN2 could be linked to ALS: we proposed that intermediate-length repeat expansions greater than 23 but below the threshold



e

Diagnosis and case no.	Total motor neurons counted	ATXN2 diffuse/ fine granular staining	ATXN2 large accumulations	Percentage of neurons with large accumulations
Normal 1	31	29	2	6.5
Normal 2	55	54	1	1.8
Normal 3	17	16	1	5.9
ALS 1	52	41	10	19.2
ALS 2	40	25	15	37.5
ALS 3	28	24	4	14.3
ALS 4	13	7	6	46.2
ALS 5	61	49	12	19.7
ALS 6	15	11	4	26.7

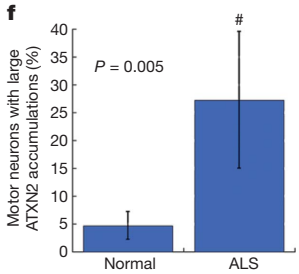


Figure 4 | ATXN2 localization is perturbed in ALS patient neurons. **a**–**d**, Immunostaining for ATXN2 in spinal cord. **a**, **b**, In control spinal cord neurons, ATXN2 is localized throughout the cytoplasm in a diffuse pattern (also Supplementary Fig. 11). **c**, **d**, In ALS spinal cord neurons, ATXN2 was present in distinct cytoplasmic accumulations (arrows). In some cases, ATXN2-positive accumulations were adjacent to clearings indicative of TDP-43 aggregates (* in **c**). **e**, **f**, Quantification of large accumulations of ATXN2 in control (normal) versus ALS spinal cord neurons. In ALS patients, 27.2 ± 12.3% of spinal cord neurons had large accumulations of ATXN2 compared to 4.7 ± 2.6% of control neurons. ALS patients with normal (ALS 1–3) and intermediate-length (ALS 4–6) *ATXN2* polyQ repeats (see Fig. 5) were included and the ATXN2 pathology was not significantly different. An ALS case with both a *SOD1* mutation and an *ATXN2* polyQ expansion (27 glutamines) was included (case ALS 4). Data are presented as mean ± standard deviation. Scale bars, 1.25 μm for **a**; 5 μm for **b**–**d**.

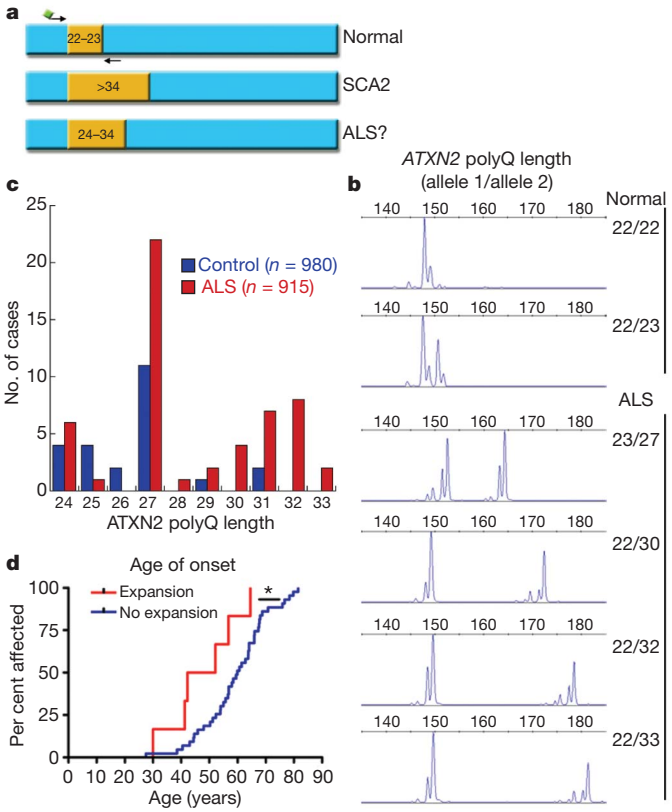


Figure 5 | Intermediate-length *ATXN2* polyQ expansions linked to ALS. **a**, The *ATXN2* gene contains a trinucleotide repeat encoding polyQ. The repeat length is normally 22–23 glutamines. Expansions of >34 cause SCA2[12]. We proposed that intermediate-length polyQ expansions (for example, 24–34) could be linked to ALS. The *ATXN2* polyQ length was defined by Genescan analysis of ALS cases and neurologically normal controls (for details, see Table 1 and Methods). **b**, Representative examples of Genescan analysis of polyQ lengths from control and ALS cases. **c**, The distribution of *ATXN2* polyQ repeat lengths in ALS and control cases. PolyQ lengths ≥27 are significantly enriched in ALS versus controls. Furthermore, polyQ lengths >31 were never observed in our controls but we found nine ALS patients above this threshold. **d**, In a selected cohort of ALS patients ($n = 65$), those with *ATXN2* polyQ expansions showed a significantly lower age of onset (compared by survival analysis).*, $P = 0.01$.

for SCA2 (for example, 24–34) may be associated with ALS (Fig. 5a and Supplementary Fig. 1). To test this, the *ATXN2* polyQ repeat length was defined in genomic DNA from 915 individuals diagnosed with ALS and 980 neurologically normal controls (Fig. 5b). We found that 24 of 980 control cases (2.4%) harboured a single intermediate-length *ATXN2* allele, whereas 50 of 915 ALS cases (5.5%) possessed one allele with an intermediate-length *ATXN2* repeat (mean repeat length, 28; range 24–33; $P = 8.0 \times 10^{-4}$; odds ratio, 2.3 with a 95% confidence interval of 1.41–3.76). *ATXN2* polyQ repeat lengths of 22 and 23 are the most common alleles; however, our findings and other studies have reported slightly longer repeat lengths in some control individuals[10,11,13]. Receiver operating characteristic (ROC) analysis of our data set showed that a cutoff of ≥27 polyQ repeats in ATXN2 provided the greatest sensitivity and specificity for discriminating normal versus ALS subjects. By using this cutoff (≥27 glutamines), the association with ALS is even stronger ($P = 3.6 \times 10^{-5}$; odds ratio, 2.8 with 95% confidence interval of 1.54–5.12; Table 1). Notably, in 980 neurologically normal controls, we identified only three individuals with expansions of >28 (all three were under age 60), whereas 23 ALS cases fell in this range (Fig. 5c). Moreover, we never observed an *ATXN2* polyQ repeat length greater than 31 in controls, whereas nine ALS patients harboured repeat lengths of 32 or 33 (Fig. 5c). Thus, intermediate-length ATXN2 polyQ repeat expansions are significantly associated with ALS.

For a subset of the ALS cases ($n = 65$; screened negative for mutations in *SOD1*, *TARDBP* and *FUS* (also known as *TLS*)), extensive clinical details have been assessed. We interrogated the clinical characteristics in this cohort and compared the ALS cases with ($n = 8$) and without ($n = 57$) intermediate-length *ATXN2* repeats (≥24 glutamines). This analysis suggested that the age of onset was significantly earlier in ALS patients in this cohort with intermediate-length *ATXN2* repeats (mean age 47.8 versus 59.4 in ALS without intermediate-length *ATXN2* repeat expansion, $P = 0.01$; Fig. 5d), although it will be important to extend this analysis to a larger ALS population. Taken together, these data indicate a strong link between intermediate-length polyQ expansions in *ATXN2* and ALS. In addition, *ATXN2*-associated ALS cases may have distinct characteristics, such as an earlier age of onset (see also clinical anecdote in Supplementary Text).

Effects of ATXN2 polyQ expansions

To provide insight into how an intermediate-length ATXN2 polyQ repeat could enhance pathogenesis we considered whether the expanded polyQ repeat might increase ATXN2 stability and/or affect its degradation, as is the case for other polyQ proteins[25]. Therefore, we analysed ATXN2 protein levels in patient-derived lymphoblastoid cells from ALS cases harbouring intermediate-length polyQ expansions, ALS cases with normal-range repeat lengths, and controls. These studies showed that whereas the steady-state levels of ATXN2 were comparable, cycloheximide treatment, which blocks new protein synthesis, revealed an increase in stability (or decreased degradation) of ATXN2 in cells with intermediate-length polyQ repeats (Supplementary Fig. 8). Thus, intermediate-length repeats could increase ATXN2 stability or inhibit its degradation, which could result in an increase in the effective concentration of ATXN2. This may further promote TDP-43 pathology beyond the interactions of ATXN2 harbouring normal repeat lengths.

We next examined the effect of polyQ repeat expansions in ATXN2 on its interaction with TDP-43. We transfected HEK293T cells with YFP-tagged TDP-43 or YFP alone along with ATXN2 containing 22, 31, or 39 glutamines (Supplementary Figs 8 and 9). As before, TDP-43–YFP, but not YFP alone, immunoprecipitated endogenous ATXN2 (Supplementary Fig. 8c, long exposure). Notably, this interaction was enhanced by longer polyQ lengths; ATXN2 with 39 glutamines immunoprecipitated with TDP-43–YFP more robustly than ATXN2 with 22 glutamines (Supplementary Fig. 8). These data indicate that polyQ expansions in ATXN2 might enhance its interaction with TDP-43.

Finally, although we were unable to detect a physical interaction by co-immunoprecipitation between endogenous TDP-43 and ATXN2, with or without polyQ expansions, in ALS-patient-derived lymphoblastoid cells (data not shown), we did detect a functional interaction. Both ATXN2 and TDP-43 have been shown to re-localize to stress granules, sites of RNA processing, under various stress situations (heat shock, oxidative stress)[26–28]. Under normal conditions, TDP-43 was localized to the nucleus and ATXN2 to the cytoplasm in both control cells and cells harbouring ATXN2 polyQ repeat expansions (Supplementary

Table 1 | Increased frequency of intermediate-length *ATXN2* polyQ repeat expansions in ALS

Subjects	Total	≤26 repeats	27–33 repeats	Percentage of 27–33 repeats	*P* value	OR (95% CI)
ALS	915	872	43	4.7%	3.6×10^{-5}	2.80 (1.54–5.12)
Neurologically normal	980	966	14	1.4%		

OR, odds ratio; CI, confidence interval

Fig. 8d). However, following 1-h heat shock at 44 °C, ATXN2 coalesced into multiple discreet cytoplasmic foci (Supplementary Fig. 8d). By blinded analysis, in the same situation, there was a significant increase in the number of cells with TDP-43 mislocalization to the cytoplasm in polyQ-expanded *ATXN2* cells compared to controls (53% versus 23%, $P = 0.02$; *ATXN2*-expanded cell lines, $n = 4$; control cell lines, $n = 3$; Supplementary Fig. 8). Although further analysis is required in disease-relevant cells, these data indicate a mechanism by which intermediate-length *ATXN2* polyQ repeats might confer genetic risk for ALS: by making TDP-43 more prone to mislocalize from the nucleus to cytoplasm under situations of stress.

Discussion

We present evidence for intermediate-length polyQ expansions in the *ATXN2* gene as a potentially common genetic contributor to ALS. This finding extends from a simple modifier screen in yeast for genes with activity that affects TDP-43 toxicity. Confirmation of these studies in *Drosophila* and human cells followed by biochemical analysis in yeast and human cells revealed that ATXN2 and TDP-43 can associate in a complex, and that this interaction depends on RNA. Furthermore, ATXN2 is abnormally localized in ALS-patient motor neurons, and TDP-43 pathology characterizes SCA2. Whereas long polyQ expansions in *ATXN2* are the cause of SCA2, our studies show that intermediate-length polyQ expansions of 27–33 are associated with ALS, with a frequency of 4.7% in cases unselected for family history. These findings indicate that intermediate-length polyQ expansions in *ATXN2* may be the most common genetic risk factor for ALS defined to date (Supplementary Fig. 1; see also Supplementary Discussion).

The identification of a novel and potentially common ALS disease gene from a simple yeast screen, leveraged by the more complex model organism *Drosophila*, underscores the power of yeast and fly as model systems for gaining insight into human disease pathogenesis. There is no cure for ALS and currently the only treatment is riluzole, which slows disease progression by only 3 months[29]. The identification of pathological interactions between ATXN2 and TDP-43, together with the strong genetic association of *ATXN2* intermediate-length polyQ expansions and ALS, will empower the development of new therapies for this devastating disease.

METHODS SUMMARY

The yeast TDP-43 modifier screen was performed in a similar manner to previous screens[7,8]. Transgenic flies expressing human TDP-43 were generated by standard techniques using the pUAST vector. Co-immunoprecipitation in yeast and mammalian cells (HEK293T) was performed using standard techniques. Lymphoblastoid cell lines were obtained from patients with ALS or unaffected normal controls (Coriell) and cultured according to instructions from Coriell. Detailed immunohistochemistry protocols are available at http://www.med.upenn.edu/mcrc/histology_core/. For *ATXN2* polyQ-repeat size determination, we amplified *ATXN2* CAG repeats from individual samples by polymerase chain reaction (PCR). The 5′ primer was SCA2-Anew: 5′-CCCCGCCCGGCGTGCG AGCCGGTGTATG-3′. The 3′ primer was SCA2-B: 5′-CGGGCTTGCGG ACATTGG-3′. PCR cycles were as follows: 2 min at 94 °C, 35 cycles (1 min at 94 °C, 1 min at 60 °C, 1 min at 72 °C), and 5 min at 72 °C. We determined *ATXN2* CAG repeat lengths by capillary electrophoresis, incorporating 6FAM fluorophor into PCR products in 5′ SCA2-Anew primer. PCR products were mixed with Liz-500 size standard (Applied Biosystems) and processed for size determination on an ABI3730 sequencer. Repeat sizes were determined with GeneMapper 4.0 software (Applied Biosystems). Select PCR products were sequenced to confirm repeat expansions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 November 2009; accepted 28 June 2010.

1. Cleveland, D. W. & Rothstein, J. D. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. *Nature Rev. Neurosci.* **2,** 806–819 (2001).
2. Valentine, J. S. & Hart, P. J. Misfolded CuZnSOD and amyotrophic lateral sclerosis. *Proc. Natl Acad. Sci. USA* **100,** 3617–3622 (2003).
3. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314,** 130–133 (2006).
4. 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.* **18,** R156–R162 (2009).
5. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: the FUS about TDP-43. *Cell* **136,** 1001–1004 (2009).
6. Winton, M. J. *et al.* Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *J. Biol. Chem.* **283,** 13302–13309 (2008).
7. Cooper, A. A. *et al.* α-Synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* **313,** 324–328 (2006).
8. Gitler, A. D. *et al.* α-Synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nature Genet.* **41,** 308–315 (2009).
9. Orr, H. T. & Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30,** 575–621 (2007).
10. Imbert, G. *et al.* Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nature Genet.* **14,** 285–291 (1996).
11. Lorenzetti, D., Bohlega, S. & Zoghbi, H. Y. The expansion of the CAG repeat in ataxin-2 is a frequent cause of autosomal dominant spinocerebellar ataxia. *Neurology* **49,** 1009–1013 (1997).
12. Pulst, S. M. *et al.* Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nature Genet.* **14,** 269–276 (1996).
13. Sanpei, K. *et al.* Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nature Genet.* **14,** 277–284 (1996).
14. Infante, J. *et al.* Spinocerebellar ataxia type 2 with levodopa-responsive parkinsonism culminating in motor neuron disease. *Mov. Disord.* **19,** 848–852 (2004).
15. Nanetti, L. *et al.* Rare association of motor neuron disease and spinocerebellar ataxia type 2 (SCA2): a new case and review of the literature. *J. Neurol.* **256,** 1926–1928 (2009).
16. Mangus, D. A., Amrani, N. & Jacobson, A. Pbp1p, a factor interacting with *Saccharomyces cerevisiae* poly(A)-binding protein, regulates polyadenylation. *Mol. Cell. Biol.* **18,** 7383–7396 (1998).
17. Buchan, J. R., Muhlrad, D. & Parker, R. P bodies promote stress granule assembly in *Saccharomyces cerevisiae. J. Cell Biol.* **183,** 441–455 (2008).
18. Al-Ramahi, I. *et al.* dAtaxin-2 mediates expanded Ataxin-1-induced neurodegeneration in a *Drosophila* model of SCA1. *PLoS Genet.* **3,** e234 (2007).
19. Lessing, D. & Bonini, N. M. Polyglutamine genes interact to modulate the severity and progression of neurodegeneration in *Drosophila. PLoS Biol.* **6,** e29 (2008).
20. Auluck, P. K., Chan, H. Y., Trojanowski, J. Q., Lee, V. M. & Bonini, N. M. Chaperone suppression of α-synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science* **295,** 865–868 (2002).
21. Warrick, J. M. *et al.* Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. *Nature Genet.* **23,** 425–428 (1999).
22. Johnson, B. S. *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).
23. Buratti, E. & 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).
24. Huynh, D. P., Yang, H. T., Vakharia, H., Nguyen, D. & Pulst, S. M. Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Hum. Mol. Genet.* **12,** 1485–1496 (2003).
25. Venkatraman, P., Wetzel, R., Tanaka, M., Nukina, N. & Goldberg, A. L. Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. *Mol. Cell* **14,** 95–104 (2004).
26. Colombrita, C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111,** 1051–1061 (2009).
27. Nonhoff, U. *et al.* Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Mol. Biol. Cell* **18,** 1385–1396 (2007).
28. Freibaum, B. D., Chitta, R. K., High, A. A. & Taylor, J. P. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J. Proteome Res.* **9,** 1104–1120 (2010).
29. Jackson, M., Llado, J. & Rothstein, J. D. Therapeutic developments in the treatment of amyotrophic lateral sclerosis. *Expert Opin. Investig. Drugs* **11,** 1343–1364 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported in part by a Pilot grant from the University of Pennsylvania Institute on Aging (A.D.G.), an NIH Director's New Innovator Award 1DP2OD004417-01 (A.D.G.), 1R01NS065317-01 (A.D.G.), P01 AG-09215 (N.M.B.), AG-10124 (J.Q.T., V.M.V.D.) and AG-17586, (V.M.-Y.L., V.M.V.D.). A.D.G. is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trusts. A.S.C.-P. is supported by a Burroughs Wellcome Fund Career Award and NIH K08 AG-033101-01. N.M.B. is an Investigator of the Howard Hughes Medical Institute. U.R. has support from the Deutsche Heredo-Ataxie Gesellschaft (DHAG) and Autosomal Dominant Cerebellar Ataxia

(ADCA) Vereniging Nederland, G.A. from the European Integrated Project on Spinocerebellar Ataxias (EuroSCA) and the Deutsche Forschungsgemeinschaft (DFG) (AU96/11-1). We acknowledge W. den Dunnen and E. Brunt for autopsy tissue and M. Babl for technical assistance. We thank J. Epstein, J. Shorter, A. Cashmore and members of the Gitler laboratory for comments on the manuscript and discussions. We are grateful for the dedication of the patients and their families and for their invaluable contributions to this research.

Author Contributions N.M.B. and A.D.G. are co-senior authors. M.P.H., H.-J.K., A.S.C.-P., F.G., U.R., G.A., J.Q.T., V.M.-Y.L., V.M.V.D., N.M.B. and A.D.G. designed the experiments. A.C.E., H.-J.K., M.P.H., B.S.J., X.F., M.A., R.G., M.M.L., U.R. and A.D.G. performed the research. D.J. and P.J.G. provided the reagents. D.C. and V.M.V.D. collected the clinical data. A.S.C.-P., L.E. and L.M. assessed clinical characteristics. M.P.H., H.-J.K., A.S.C.-P., F.G., A.P., L.E., L.M., U.R., G.A., J.Q.T., V.M.-Y.L., V.M.V.D., N.M.B. and A.D.G. analysed and interpreted data. N.M.B. and A.D.G. wrote the paper with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.D.G. (gitler@mail.med.upenn.edu) or N.M.B. (nbonini@sas.upenn.edu).

METHODS

Yeast strains and media. The strain used in the modifier screen was TDP-43, *MATa can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 ade2-1* pAG303Gal-TDP-43. The huntingtin and α-synuclein strains are as described[7,30]. The *pbp1*Δ strain was obtained by replacing the *PBP1* coding region with a KanMX4 cassette in the BY4741 strain background. Colony PCR was used to verify correct gene disruption. Strains were manipulated and media prepared using standard techniques.

Plasmids. The CEN and 2-μm galactose-inducible TDP-43 yeast expression plasmids are as described[31]. The *PBP1* expression plasmid was constructed by shuttling *PBP1* from the Gateway entry vector (pDONR221) into pBY011, a CEN, URA3, galactose-inducible yeast expression plasmid[7]. The CFP-tagged Pbp1 construct was made by shuttling *PBP1* into pAG413GPD-CFP-ccdB, a CEN, HIS3, constitutive promoter yeast expression plasmid. The 2-μm α-synuclein expression plasmid was as described[32]. Site-directed mutagenesis was performed with the QuickChange multi kit (Stratagene). The TDP-43(ΔNLS)–YFP construct was generated by mutating residues L82, R83 and L84 to alanine[6]. The TDP-43(5F→L), TDP-43(F194L)–YFP, TDP-43(F229L)–YFP, TDP-43(F147,149L)–YFP and TDP-43(F194,229,231L)–YFP constructs were generated by mutating phenylalanine residues 147, 149, 194, 229, and 231 to leucine[23]. Mammalian expression vectors were generated by shuttling TDP-43–YFP, TDP-43(ΔNLS)–YFP, TDP-43(5F→L)–YFP, or TDP-43(ΔNLS/5F→L)–YFP from pDONR221 into pcDNA 3.2 (Invitrogen). The *ATXN2* expression vectors (22, 31, 39 polyQ repeats) were generated by cloning human *ATXN2* complementary DNA (courtesy of S. Pulst) into pcDNA6/*myc*-His-A between the BamHI and XbaI cloning sites.

Yeast transformation and spotting assays. The PEG/lithium acetate method was used to transform yeast with plasmid DNA. For spotting assays, yeast cells were grown overnight at 30 °C in liquid media containing raffinose (SRaf/−Ura) until log or mid-log phase. Cultures were then normalized for cell number, serially diluted and spotted onto synthetic solid media containing glucose or galactose lacking uracil, and were grown at 30 °C for 2–3 days.

Yeast TDP-43 toxicity modifier screen. *PBP1*, the yeast orthologue of human *ATXN2*, was isolated in a high-throughput yeast transformation screen similar to previous screens[7,8]. Five-thousand and five-hundred full-length yeast open reading frames (yeast FLEXGene collection, http://plasmid.med.harvard.edu/PLASMID/Home.jsp) were transformed into a strain expressing TDP-43 integrated at the *HIS3* locus. A standard lithium acetate transformation protocol was modified for automation and used by employing a BIOROBOT Rapidplate 96-well pipettor (Qiagen). Transformants were grown overnight in synthetic-deficient media lacking uracil (SD/−Ura) with glucose. Overnight cultures were inoculated into fresh SD/−Ura media with raffinose and allowed to reach stationary phase. The cells were spotted onto SD/−Ura + glucose and SD/−Ura + galactose agar plates. Modifiers were identified on galactose plates after 2–3 days of growth at 30 °C.

***Drosophila* experiments.** Transgenic flies expressing human TDP-43, TDP-43–YFP and TDP-43(Q331K) were generated by standard techniques using the pUAST vector. Characterization of these lines for protein expression levels and toxicity indicated a tight correlation between expression level and effect. As is typical of fly *UAS*-transgenic lines, the effects of any single insertion were consistent among all the flies of that cross. The effects of untagged and YFP-tagged TDP-43 proteins were similar in general, but the YFP-tagged proteins were expressed at a higher steady-state level, which caused more severe effects. Most experiments were performed at 25 °C, although select experiments were performed at 29 °C for a stronger effect (the GAL4/UAS system drives expression more strongly at higher temperatures), or for time considerations (the lifespan of flies is shorter at 29 °C). For lifespan analysis, flies were raised at 25 °C, and lifespan assessed at 29 °C. Climbing and lifespan analyses were performed as described[33]. Fly *Atx2* reagents $Atx2^{X1}$ (a null allele of *Atx2*) and $Atx2^{EP3145}$ (an *EP* insertional allele of *Atx2* that increases *Atx2* expression when combined with a GAL4 driver line) have previously been described (gift of L. Pallanck and T. Satterfield)[19].

Immunoblots of fly tissue were performed following standard protocols[19]. Specifically, ten fly heads (five males and five females) were collected in 80 μl of LDS sample buffer with sample reducing agent (Invitrogen), and ground by pestle. Lysates were incubated at 99 °C for 5 min and then spun for 30 min at 20,000*g* to remove debris. Proteins were separated on 3–8% Tris-acetate gels or 4–12% Tris-glycine gels (Invitrogen), and transferred to nitrocellulose or polyvinylidene difluoride (PVDF) membrane following standard protocols. Primary antibodies used were rabbit anti-TDP-43 (1:600; Proteintech), mouse anti-β-galactosidase (1:1,000; Promega) and mouse anti-β-tubulin (1:2,000; E7, Developmental Studies Hybridoma Bank). Following incubation with horseradish peroxidase (HRP)-coupled secondary antibodies (goat anti-mouse or goat anti-rabbit, 1:2,000; Chemicon), blots were visualized using a chemiluminescent detection kit (Amersham ECL Plus; GE Healthcare).

Quantitative real-time PCR was performed in triplicate on each cDNA sample as described previously[19] with an Applied Biosystems 7500 Fast system. To amplify *Atx2*, we used forward primer 5′-ACAATAGCAAGCGGAAAACC-3′ and reverse primer 5′-CCTTAGGCTGTTATCGTTGGA-3′. Ribosomal protein 49 (RP49) was used as the internal control.

Fly genotypes for Fig. 2c are as follows. Control is *gmr-GAL4(YH3)/ UAS–GFP*. TDP-43 is *UAS-TDP-43*(M)/+; *gmr-GAL4(YH3)*/+. $Atx2^{EP}$ is *gmr-GAL4(YH3)*/ $Atx2^{EP3145}$. TDP43+$Atx2^{EP}$ is *UAS-TDP-43*(M)/+; *gmr-GAL4(YH3)*/ $Atx2^{EP3145}$. TDP-43; $Atx2^{X1}$/+ is *UAS-TDP-43*(M)/+; *gmr-GAL4(YH3)*/ $Atx2^{X1}$. Fly genotypes for Fig. 2d are as follows. elav is $elav^{C155}$/*Y*. elav/$Atx2^{EP}$ is $elav^{C155}$/*Y*; +/ $Atx2^{EP3145}$. elav/TDP-43/+ is $elav^{C155}$/*Y; UAS-TDP-43*(M)/+. elav/TDP-43/$Atx2^{EP}$ is $elav^{C155}$/*Y; UAS-TDP-43*(M)/+; +/ $Atx2^{EP3145}$. elav/TDP-43/$Atx2^{X1}$ is $elav^{C155}$/*Y; UAS-TDP-43*(M)/+; +/ $Atx2^{X1}$. Fly genotypes for Fig. 2e are as follows. — is *gmr-GAL4(YH3)*/+. TDP-43 alone is *UAS-TDP-43(M)*/+; *gmr-GAL4(YH3)*/+. TDP-43 with $Atx2^{X1}$ is *UAS-TDP-43*(M)/+; *gmr-GAL4(YH3)*/ $Atx2^{X1}$. TDP-43 with $Atx2^{EP}$ is *UAS-TDP-43*(M)/+; *gmr-GAL4(YH3)*/ $Atx2^{EP3145}$.

Co-immunoprecipitation. For the yeast co-immunoprecipitation, yeast cells were transformed with CFP alone and untagged TDP-43 (pAG413GPD-CFP and pAG416Gal-TDP-43) or CFP–Pbp1 and untagged TDP-43 (pAG413GPD-CFP-Pbp1 and pAG416Gal-TDP-43). Double-transformants were selected on synthetic media containing glucose and lacking histidine and uracil (SD/−His/−Ura). Transformants were grown in 25 ml of SRaf/−His/−Ura media at 30 °C overnight to saturation, diluted into 250 ml SGal/−His/−Ura to OD_{600} 0.4 and induced for 8 h at 30 °C and then normalized for cell number. Cells were then collected and washed twice with distilled H_2O and resuspended in 500 μl of PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) containing 0.1% NP-40 and a cocktail of protease and phosphatase inhibitors for yeast (Sigma). Cells were disrupted with glass beads in a bead beater. Lysates were clarified by centrifugation and supernatants normalized for protein concentration. Immunoprecipitation was performed by incubating lysates with 2.5 μl rabbit polyclonal GFP antibody (Abcam; crossreacts with CFP and YFP proteins) followed by 30 μl protein A agarose beads (Life Technologies). After incubation at 4 °C for 1 h, beads were washed three times each with 500 μl lysis buffer without protease and phosphatase inhibitors, and proteins bound to the beads were eluted with 25 μl SDS-sample buffer and subjected to SDS–PAGE followed by immunoblotting.

Co-immunoprecipitation in mammalian cells was performed as follows. HEK293T cells were transfected with TDP-43–YFP fusion constructs using FuGene 6 (Roche) according to the manufacturer's instructions. After 48 h, cells were washed with PBS, trypsinized and collected by centrifugation. Cells were washed in ice-cold PBS containing protease inhibitor cocktail (Roche) then lysed in NP-40 lysis buffer (150 mM NaCl, 50 mM Tris, pH 8.0, 1% NP-40 and protease inhibitor). In the case of RNA digestion, lysates were treated with 200 μg ml^{-1} RNase A for 15 min (Qiagen). Lysates were clarified by centrifugation and precleared with protein A agarose (Invitrogen). Immunoprecipitation was performed by incubating with anti-GFP rabbit polyclonal antibody (1:750 dilution; Abcam) or anti-ATXN2 mouse monoclonal antibody (1:500 dilution; BD Transduction Laboratories) for 2 h, then protein A agarose beads (50 μl) for 1 h. The beads were washed three times with NP-40 lysis buffer and resuspended in 4× SDS sample buffer (40% glycerol, 240 mM Tris HCL pH 6.8, 8% SDS, 0.04% bromophenol blue, 5% β-mercaptoethanol).

Immunoblotting of yeast and human cells. Lysates were boiled for 5 min, then subjected to SDS–PAGE (4–12% gradient Bis-Tris, Invitrogen) and transferred to PVDF membrane (Invitrogen). Membranes were blocked for 1 h in 5% non-fat dry milk at room temperature (22 °C) and then incubated overnight (16–18 h) in primary antibody at 4 °C. Membranes were washed four times in PBS, then incubated in HRP-conjugated secondary antibody (1:5,000) for 1 h, then washed four times in PBST (PBS + 0.1% Tween20). Proteins were detected with Immobilon Western Chemiluminescent HRP Substrate (Millipore) and visualized on Biomax MR film (Kodak). Primary antibodies were: anti-GFP mouse polyclonal antibody (Roche), 1:1,000; anti-ataxin-2 mouse antibody (BD), 1:500; anti-spinocerebellar-ataxin-type-3 mouse antibody, clone 1H9 (Millipore), 1:500; anti-actin mouse monoclonal antibody, clone C4 (Millipore), 1:5,000.

Immunofluorescence. HEK293T cells or patient-derived lymphoblastoid cell lines were washed in PBS and fixed in 4% paraformaldehyde for 15 min, then washed in PBS four times. Cells were blocked for 1 h in blocking solution (2% fetal bovine serum, 0.02% Triton X-100, PBS), and then incubated for 1 h in primary antibody at room temperature. Cells were then washed three times in PBS, then incubated with secondary antibody for 1 h at room temperature (22 °C). Cells were then washed with blocking solution and mounted in Vectashield mounting media with DAPI (Vector). Antibodies used were: anti-ataxin-2 mouse antibody (BD), 1:50; anti-TDP-43 rabbit antibody (Proteintech Group), 1:200; Cy-3 conjugated anti-mouse IgG (Jackson Immunoresearch), 1:500; and Cy-2 conjugated anti-rabbit IgG (Jackson Immunoresearch), 1:500. Cells were visualized by light microscopy.

Immunohistochemistry. SCA2 patient brain tissue was embedded in polyethylene glycol and cut into 100-μm thick serial sections. All other sections were

deparaffinized before pretreatment using heat antigen retrieval with Bull's Eye Decloaker (BioCare Medical). Endogenous peroxidase was then blocked with 3% hydrogen peroxide in PBS for 10 min. After washing with 0.1% PBST and blocking with 10% goat serum, 0.5% PBST for 30–60 min at 25 °C, sections were incubated with mouse anti-ATXN2 (1:500; BD Biosciences) or rabbit anti-TDP-43 (1:500; Proteintech Group) in 0.1% PBST overnight at 4 °C. After washing with 0.1% PBST, sections were incubated with biotinylated goat anti-mouse or rabbit IgG (1:200; Vector Laboratories) for 1 h at 25 °C. After washing with 0.1% PBST, sections were then incubated with Vectastain ABC (Vector Laboratories) for 45 min. After washing with 0.1% PBST followed by 0.1 M Tris (pH 7.5) and 0.3 M NaCl, peroxidase activity was detected with 3,3′-diaminobenzidine (DAB; Sigma). Detailed immunohistochemistry protocols are available at http://www.med.upenn.edu/mcrc/histology_core/.

Patient-derived lymphoblastoid cell culture and ATXN2 protein stability. Lymphoblastoid cell lines were obtained from patients with ALS or unaffected normal controls (Coriell) and cultured in RPMI 1640 medium supplemented with 2 mM L-glutamine, 15% fetal bovine serum, penicillin and streptomycin. To assess ATXN2 stability, protein synthesis was inhibited by treating cells with cycloheximide (0.5 μM) for 0, 16 or 24 h. Cells were washed once in PBS and lysed in ice-cold NP-40 lysis buffer containing protease inhibitors. Lysates were cleared by centrifugation and then suspended in 4× sample buffer and subjected to SDS–PAGE followed by immunoblotting for ATXN2 and β-actin.

Patient-derived lymphoblastoid cell heat-shock experiment. Lymphoblastoid cell lines were heat shocked at 44 °C for 1 h, then fixed and immunostained as described earlier. The percentage of cells with abnormal TDP-43 localization was quantified by counting >65 cells for each cell line ($n = 3$ for normal polyQ, $n = 4$ for intermediate-length polyQ) in three separate experiments. Cells lacking detectable TDP-43 immunoreactivity in the nucleus using DAPI counterstain were counted and divided by the total number of cells counted.

***ATXN2* trinucleotide repeat size determination in ALS patients and controls.** Details of the patients and controls included in this study are compiled in Supplementary Table 2. Genomic DNA from human ALS patients was obtained from the Coriell Institute for Medical Research (Coriell) or the Center for Neurodegenerative Disease Research (CNDR) at the University of Pennsylvania. Five-hundred and seventy-six ALS samples from Coriell were distributed in 96-well plates NDPT103, NDPT026, NDPT025, NDPT100, NDPT030 and NDPT106. Three-hundred and thirty-nine additional ALS cases unselected for family history and a subsequent cohort of 23 familial ALS cases were obtained from the CNDR. To our knowledge, none of the ALS cases belonged to the same pedigree. CNDR ALS samples were verified to meet El Escorial criteria for definite or probable ALS. In addition, 80 of 123 CNDR ALS samples were neuropathologically confirmed to have ALS pathology with TDP-43 immunopositivity, while the remainder were from living patients. Among the original CNDR cohort of 100 cases unselected for family history, 13 cases (13%) are known to have a first- or second-degree relative with ALS, in line with published estimates of ~10% familial ALS. Clinical details were collected from 65 of 100 CNDR ALS cases unselected for family history via chart review by a neurologist; these details included age of onset, age of death, disease duration, gender, presence or absence of family history, and ALS functional rating scale score (ALS-FRS) at the time of initial neurological evaluation. Six-hundred and forty-two neurologically normal control samples from Coriell were distributed in 96-well plates NDPT084, NDPT090, NDPT093, NDPT094, NDPT095, NDPT096, NDPT098 and NDPT099. One control sample (ND12820 from plate NDPT096) was excluded because of a documented family history of motor neuron disease; sibling ND12819 was diagnosed with progressive bulbar palsy. An additional 163 neurologically normal control samples were obtained from the Children's Hospital of Philadelphia under an institutional review board approved protocol. An additional one-hundred and seventy-five DNA samples from cognitively normal individuals >60 years of age were obtained from the National Cell Repository for Alzheimer's Disease (NCRAD).

We amplified *ATXN2* CAG repeats from individual samples by PCR. PCR primers used for amplification were designed to amplify the CAG repeat region of human *ATXN2* (base pairs 442–598). The 5′ primer was SCA2-Anew: 5′-CCCCGCCCGGCGTGCGAGCCGGTGTATG-3′. The 3′ primer was SCA2-B: 5′-CGGGCTTGCGGACATTGG-3′. PCR cycles were as follows: 2 min at 94 °C, 35 cycles (1 min at 94 °C, 1 min at 60 °C, 1 min at 72 °C), and 5 min at 72 °C. Initially, PCR products were resolved on a 2% agarose gel by electrophoresis, amplicons purified and cloned into the PCRII TA vector (Invitrogen), and repeat lengths were determined by DNA sequencing. Subsequently, for large-scale analysis of *ATXN2* CAG repeat lengths, a capillary electrophoresis approach was used, incorporating the 6FAM fluorophor into the PCR products in the 5′ SCA2-Anew primer. PCR products were mixed with Liz-500 size standard (Applied Biosystems) and were processed for size determination on an ABI3730 sequencer. The sizes of the repeats were determined with GeneMapper 4.0 software (Applied Biosystems). Thirty-two samples with repeat expansions were verified by independent PCR as described earlier, followed by resolution on a 4% agarose gel to confirm relative lengths, and also by capillary electrophoresis. To further confirm repeat expansions, amplicons from 21 of 32 samples were cloned and sequenced.

The majority of both our cases and controls were North American individuals self-described as Caucasian. However, some of these, although North American, were not Caucasian. And for some of our other cases and controls, ethnicity information was not available. We therefore performed a secondary analysis of only those cases and controls that we were certain were self-described as Caucasian from North America (560 controls and 529 ALS cases). Twenty-one of 529 ALS cases had *ATXN2* expansions of ≥27 whereas 9 of 560 controls had them (4.0% versus 1.6%, $P = 0.02$). This result is similar to findings with the entire cohort, giving us confidence that the association of *ATXN2* expanded polyQ repeats with ALS is not due to population stratification.

Statistical analyses. Two-tailed *t*-tests were used to compare age of onset, ALS-FRS at the time of initial neurological evaluation and age of death in ALS with and without intermediate-length *ATXN2* repeats after ascertainment that distributions met assumptions of normality. Two-tailed Fisher's exact tests were used to compare gender and presence or absence of family history between the two groups. Age of onset and duration compared with survival analysis. For all tests, percentages and statistical testing were calculated based only on the cases for which relevant clinical data were available.

Two-tailed Fisher's exact tests were used to evaluate genetic association between intermediate-length *ATXN2* repeats and ALS, and odds ratios were calculated under an intermediate-length *ATXN2*-repeat-dominant model.

30. Krobitsch, S. & Lindquist, S. Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. *Proc. Natl Acad. Sci. USA* **97,** 1589–1594 (2000).
31. Johnson, B. S., McCaffery, J. M., Lindquist, S. & Gitler, A. D. A yeast TDP-43 proteinopathy model: exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proc. Natl Acad. Sci. USA* **105,** 6439–6444 (2008).
32. Outeiro, T. F. & Lindquist, S. Yeast cells provide insight into α-synuclein biology and pathobiology. *Science* **302,** 1772–1775 (2003).
33. Li, L. B., Yu, Z., Teng, X. & Bonini, N. M. RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*. *Nature* **453,** 1107–1111 (2008).