Online 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 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-micron galactose-inducible TDP-43 yeast expression plasmids are 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-micron α–synuclein expression plasmid was as described [32]. Site-directed mutagenesis was performed with the QuickChange Multi kit (Stratagene). The ΔNLS-TDP-43-YFP construct was generated by mutating residues lysine 82, arginine 83, and lysine 84 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 [22]. Mammalian expression vectors were generated by shuttling TDP-43-YFP, ΔNLS-TDP-43-YFP, TDP-43-(5F→L)-YFP, or ΔNLS-TDP-43-(5F→L)-YFP from pDONR221 into pcDNA 3.2 (Invitrogen). The Ataxin-2 expression vectors (22Q, 31Q, 39Q) were generated by cloning human Ataxin-2 cDNA into pcDNA6/*myc*-His-A between the Bam HI and Xba I 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 OD_{600}, serially diluted and spotted onto synthetic solid media containing glucose or galactose lacking uracil, and were grown at 30°C for 2-3 d.

Yeast TDP-43 toxicity modifier screen

PBP1, the yeast ortholog of human Ataxin-2, was isolated in a high-throughput yeast transformation screen similar to previous screens [7,8]. 5,500 full-length yeast ORFs (Yeast FLEXGene collection, http://www.hip.harvard.edu/research/yeast_flexgene/) 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 in general similar, 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 temperature), or for time considerations (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 Ataxin-2 reagents $dAtx2^{X1}$ (a null allele of Ataxin-2) and $dAtx2^{EP3145}$ (an *EP* insertional allele of *dAtx2* that increases *dAtx2* expression when combined with a GAL4 driver line) have been previously described (gift of L Pallanck and T Satterfield) [19].

Immunoblots of fly tissue were performed following standard protocols [19]. Specifically, 10 fly heads (5 males and 5 females) were collected in 80μl 1X LDS sample buffer, with 1X sample reducing agent (Invitrogen), and ground by pestle. Lysates were incubated at 99°C for 5 min and then spun for 30 min at 20,000Xg to remove debris. Proteins were separated on 3-8% Tris-acetate gels or 4-12% Tris-glycine gels (Invitrogen), and transferred to nitrocellulose or PVDF membrane following standard protocols. Primary antibodies used were rabbit anti-TDP-43 (1:600; Proteintech), mouse anti-beta-galactosidase (1:1,000; Promega) and mouse anti-beta-tubulin (1:2000; E7, Developmental Studies Hybridoma Bank). Following incubation with 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 *dAtx2*, we used forward primer 5′-ACAATAGCAAGCGGAAAACC-3’ and reverse primer 5′-CCTTAGGCTGTTATCGTTGGA-3’. Ribosomal protein 49 (RP49) was used as the internal control.

Co-immunoprecipitation

For the yeast Co-IP, 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 to OD_{600} =1.2 – 1.5. Cells were then collected and washed twice with dH_2O and resuspended in 500 µl of 1X 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, St. Louis, MO). 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 Inc., Cambridge, MA) followed by 30 µl protein A agarose beads (Life Technologies Co.). After incubation at 4°C for 1 hour, 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-IP 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 48h, 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 RNase A for 15 minutes (Qiagen). Lysates were clarified by centrifugation and pre-cleared with Protein A agarose (Invitrogen). Immunoprecipitation was performed by incubating with α-GFP rabbit polyclonal antibody (1:750 dilution; Abcam) or α–Ataxin-2 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 3x with NP-40 lysis buffer and resuspended in 4X 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 5 min, then subjected to SDS/PAGE (4-12% gradient Bis-Tris, Invitrogen) and transferred to PVDF membrane (Invitrogen). Membranes were blocked 1 h in 5% non-fat dry milk at RT and then incubated O/N in primary antibody at 4°C. Membranes were washed 4x in PBS, then incubated in HRP-conjugated secondary antibody (1:5000) 1h, then washed 4x 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: α-GFP mouse polyclonal antibody (Roche), 1:1000; α-Ataxin-2 mouse antibody (BD), 1:500; α-Spinocerebellar Ataxin Type 3 mouse antibody, clone 1H9 (Millipore), 1:500; α-actin mouse monoclonal antibody, clone C4 (Millipore), 1:5000.

Immunofluorescence

HEK293T cells or patient derived lymphoblastoid cell lines were washed in PBS and fixed in 4% paraformaldehyde 15 min, then washed in 1X PBS 4x. Cells were blocked for 1h in blocking solution (2% Fetal Bovine Serum, 0.02% Triton X-100, 1X PBS), and then incubated 1h in primary antibody at RT. Cells were then washed 3x in PBS, then incubated with secondary antibody 1h RT. Cells were then washed with blocking solution and mounted in Vectashield mounting media with DAPI (Vector). Antibodies used were: α-Ataxin-2 mouse antibody (BD), 1:50; α-TDP-43 rabbit antibody (Proteintech Group), 1:200; Cy-3 conjugated α-mouse IgG (Jackson Immunoresearch), 1:500; and Cy-2 conjugated α-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 minutes. After washing with 0.1% PBST and blocking with 10% goat serum, 0.5% PBST for 30-60 minutes at 25°C. Sections were incubated with mouse anti-Ataxin-2

(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 hour at 25°C. After washing with 0.1% PBST, sections were then incubated with Vectastain ABC (Vector Laboratories) for 45 minutes. After washing with 0.1% PBST followed by 0.1M Tris (pH 7.5) and 0.3M NaCl. Peroxidase activity was then detected with DAB (Sigma). Detailed immunohistochemistry protocols are available at http://www.uphs.upenn.edu/mcrc.

Patient-derived lymphoblastoid cell culture and Ataxin-2 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 Ataxin-2 stability, protein synthesis was inhibited by treating cells with cycloheximide (0.5 μM) for 0, 16 or 24 hours. 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 4X sample buffer and subjected to SDS/PAGE followed by immunoblotting for Ataxin-2 and β−Actin.

Patient-derived lymphoblastoid cell heat shock experiment

Lymphoblastoid cell lines were heat shocked at 44°C for 1 hour, then fixed and immunostained as mentioned above. 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 IL 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.

Ataxin-2 trinucleotide repeat size determination in ALS patients and controls

Details on the patients and controls included in this study are compiled in Table S2. 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. 454 ALS samples from Coriell were distributed in 96-well plates NDPT103, NDPT026, NDPT025, NDPT100, NDPT030, and NDPT106. 100 additional ALS cases unselected for family history and a subsequent cohort of 23 FALS 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, 80/123 CNDR ALS samples were neuropathologically confirmed to have ALS pathology with TDP-43 immunopositivity, while the remainder was 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% FALS. Clinical details were collected from 65/100 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/absence of family history, and ALS functional rating scale score (ALS-FRS) at the time of initial neurological evaluation. Mutations in *SOD1* or *TARDBP* were excluded from the 23 additional CNDR cases of FALS. 286 neurologically normal control samples from Coriell were distributed in 96-well plates 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 82 neurologically normal control samples were obtained from the Children's Hospital of Philadelphia under an IRB approved protocol.

We amplified Ataxin-2 CAG repeats from individual samples by polymerase chain reaction (PCR). PCR primers used for amplification were designed to amplify the CAG repeat region of human Ataxin-2 (bp 442-598). The 5' primer was SCA2-Anew: 5' - CCC CGC CCG GCG TGC GAG CCG GTG TAT G - 3'. The 3' primer was SCA2-B: 5' - CGG GCT TGC GGA CAT TGG - 3'. PCR cycles were as follows: 2 min 94°C, 35 cycles (1 min 94°C, 1 min 60°C, 1 min 72°C), and 5 min 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 Ataxin-2 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). 32 samples with repeat expansions were verified by independent PCR as above, 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.

Most of both our cases and controls were North American individuals self-described as Caucasian. However, some of these, while 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). 21 of 529 ALS cases had Ataxin-2 expansions ≥27 while 9 of 560 controls had them (4.0% vs. 1.6%, p=0.02), similar results to what we find with the entire cohort, giving us confidence that the association of Ataxin-2 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 Ataxin-2 repeats after ascertainment that distributions met assumptions of normality. Two-tailed Fisher's exact tests were used to compare gender and presence/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 Ataxin-2 repeats and ALS, and odds ratios were calculated, under an intermediate-length Ataxin-2 repeat-dominant model.

References for Online Methods

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 U S A* **97**, 1589-94 (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 U S A* **105**, 6439-44 (2008).
32. Outeiro, T.F. & Lindquist, S. Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* **302**, 1772-5 (2003).
33. Li, L.B., Yu, Z., Teng, X. & Bonini, N.M. RNA toxicity is a component of ataxin-3 degeneration in Drosophila. *Nature* (2008).