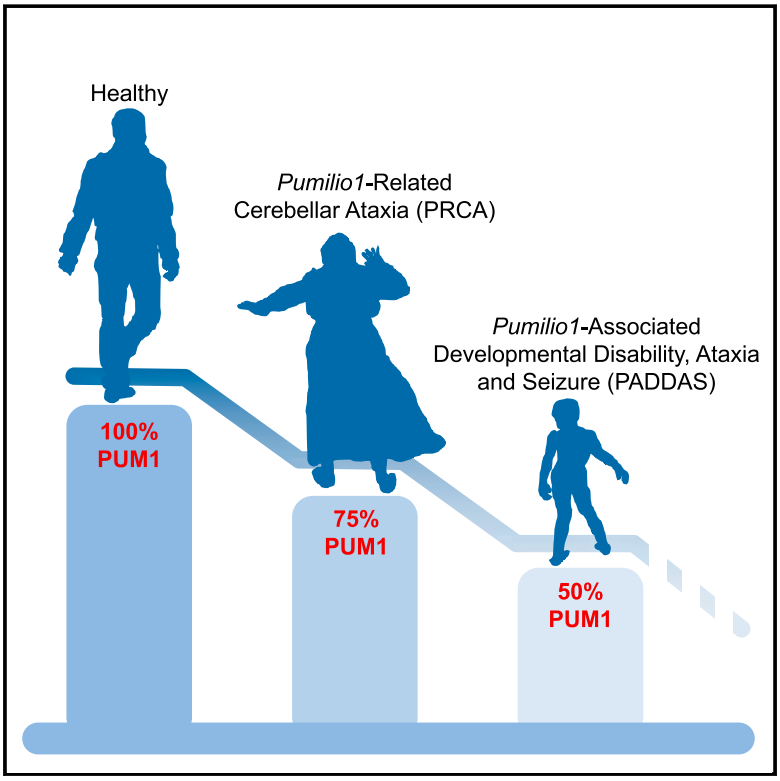
Article

Cell

A Mild *PUM1* Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures

Graphical Abstract



Highlights

- The brain is sensitive to levels of PUM1 and some of its targets
- *PUM1* haploinsufficiency causes developmental delay, ataxia, and other problems
- Mutations that reduce PUM1 levels by 25% are associated with adult-onset ataxia
- Regulators of disease-driving proteins are a pool of new candidate disease genes

Authors

Vincenzo A. Gennarino,
Elizabeth E. Palmer,
Laura M. McDonell, ..., Kym M. Boycott,
J. Lloyd Holder, Jr., Huda Y. Zoghbi

Correspondence

vag2138@cumc.columbia.edu (V.A.G.),
hzoghbi@bcm.edu (H.Y.Z.)

In Brief

Different dosages of an RNA-binding protein result in human neurological diseases of corresponding severities.

Gennarino et al., 2018, Cell *172*, 924–936
February 22, 2018
https://doi.org/10.1016/j.cell.2018.02.006

CellPress

Cell

Article

A Mild *PUM1* Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures

Vincenzo A. Gennarino,[1,2,23,*] Elizabeth E. Palmer,[3,4,5] Laura M. McDonell,[6] Li Wang,[1,2] Carolyn J. Adamski,[1,2,7] Amanda Koire,[8] Lauren See,[1] Chun-An Chen,[1,2] Christian P. Schaaf,[1,2] Jill A. Rosenfeld,[1] Jessica A. Panzer,[9,10,24] Ute Moog,[11] Shuang Hao,[2,12] Ann Bye,[3,4] Edwin P. Kirk,[3,4,13] Pawel Stankiewicz,[1,14] Amy M. Breman,[1,14] Arran McBride,[6] Tejaswi Kandula,[3,4] Holly A. Dubbs,[15] Rebecca Macintosh,[3] Michael Cardamone,[3,4] Ying Zhu,[13] Kevin Ying,[16] Kerith-Rae Dias,[16] Megan T. Cho,[17] Lindsay B. Henderson,[17] Berivan Baskin,[17] Paula Morris,[16] Jiang Tao,[16,18] Mark J. Cowley,[16,18] Marcel E. Dinger,[16,18] Tony Roscioli,[3,13,19] Oana Caluseriu,[20] Oksana Suchowersky,[20,21] Rani K. Sachdev,[3,4] Olivier Lichtarge,[1] Jianrong Tang,[2,12] Kym M. Boycott,[6] J. Lloyd Holder, Jr.,[2,12] and Huda Y. Zoghbi[1,2,7,12,22,25,*]

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
[2]Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, Houston, TX 77030, USA
[3]Sydney Children's Hospital, Randwick, NSW 2031, Australia
[4]School of Women's and Children's Health, UNSW Medicine, The University of New South Wales, NSW 2031, Australia
[5]Genetics of Learning Disability Service, Waratah, NSW 2298, Australia
[6]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H 8L1, Canada
[7]Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX 77030, USA
[8]Program in Quantitative and Computational Biosciences, Baylor College of Medicine, Houston, TX 77030, USA
[9]Department of Pediatrics, Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[10]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[11]Institute of Human Genetics, Heidelberg University, Im Neuenheimer Feld 440, 69120 Heidelberg, Germany
[12]Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA
[13]Genetics Laboratory, NSW Health Pathology East Randwick, Sydney, NSW, Australia
[14]Baylor Genetics Laboratories, Baylor College of Medicine, Houston, TX 77030, USA
[15]Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[16]Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia
[17]GeneDx, 207 Perry Pkwy Gaithersburg, MD 20877, USA
[18]St. Vincent's Clinical School, University of New South Wales, Sydney, NSW 2010, Australia
[19]Neuroscience Research Australia and Prince of Wales Clinical School, University of New South Wales, Randwick, NSW 2031, Australia
[20]Department of Medical Genetics, University of Alberta, AB T6G 2H7, Canada
[21]Departments of Medicine (Neurology) and Pediatrics, University of Alberta, AB, Canada
[22]Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA
[23]Present address: Department of Genetics & Development, Columbia University Medical Center, New York, NY, 10032
[24]Now deceased
[25]Lead Contact
*Correspondence: vag2138@cumc.columbia.edu (V.A.G.), hzoghbi@bcm.edu (H.Y.Z.)
https://doi.org/10.1016/j.cell.2018.02.006

SUMMARY

Certain mutations can cause proteins to accumulate in neurons, leading to neurodegeneration. We recently showed, however, that upregulation of a wild-type protein, Ataxin1, caused by haploinsufficiency of its repressor, the RNA-binding protein *Pumilio1* (*PUM1*), also causes neurodegeneration in mice. We therefore searched for human patients with *PUM1* mutations. We identified eleven individuals with either *PUM1* deletions or *de novo* missense variants who suffer a developmental syndrome (*Pumilio1*-associated developmental disability, ataxia, and seizure; PADDAS). We also identified a milder missense mutation in a family with adult-onset ataxia with incomplete penetrance (*Pumilio1*-related cerebellar ataxia, PRCA). Studies in patient-derived cells revealed that the missense mutations reduced PUM1 protein levels by ~25% in the adult-onset cases and by ~50% in the infantile-onset cases; levels of known PUM1 targets increased accordingly. Changes in protein levels thus track with phenotypic severity, and identifying posttranscriptional modulators of protein expression should identify new candidate disease genes.

INTRODUCTION

Decades of human and mouse genetic studies have taught us that neurons are intolerant of significant alterations in protein abundance. Doubling a protein's levels, as in the case of chromosomal duplications, or halving it, as in the case of

haploinsufficiency, can lead to neurological disorders ranging from autism to Alzheimer's disease (Chartier-Harlin et al., 2004; La Cognata et al., 2017; Ramocki et al., 2009; Rovelet-Lecrux et al., 2006; Sleegers et al., 2006). That such large changes should prove detrimental to neurons is not surprising, but some studies have suggested that more modest alterations in protein levels can also be problematic (Gennarino et al., 2015a, 2015b; Xie et al., 2017). Yet the molecular mechanisms that allow the brain to achieve precise control of protein levels are not well understood, and not much research attention has been given to the possibility that changes in posttranscriptional regulation (Fukao and Fujiwara, 2017; Mata et al., 2005) might be connected to disease.

We became interested in the posttranscriptional regulation of ATAXIN1 (ATXN1), the protein whose mutation causes spinocerebellar ataxia type 1 (SCA1), after discovering that transgenic mice and flies overexpressing ATXN1 develop neurological degeneration reminiscent of SCA1 (Burright et al., 1995; Fernandez-Funez et al., 2000). We found that ATXN1 is a target of PUMILIO1 (PUM1), a member of the PUMILIO/FBF (PUF) RNA-binding protein (RBP) family that is evolutionarily conserved from plants to humans (Gennarino et al., 2015b). The PUF protein family is characterized by a highly conserved RNA-binding *Pumilio* homology domain (PUM-HD) consisting of eight tandem repeats that regulate specific mRNA targets posttranscriptionally, with or without the contribution of the microRNA machinery (Kedde et al., 2010; Miles et al., 2012). Different studies have shown that Pum1 is an essential regulator of spermatogenesis in mice (Chen et al., 2012) and critical for differentiation of embryonic stem cells (Leeb et al., 2014; Spassov and Jurecic, 2003), cell-cycle control (Kedde et al., 2010; Miles et al., 2012), and genomic stability (Lee et al., 2016). *Drosophila pum* mutants show sterility, behavioral defects, and neuronal hyperexcitability (Schweers et al., 2002), and it was recently reported that RNAi-mediated knockdown of *pum* exacerbates seizures in flies bearing a gain-of-function mutation in the *Drosophila* voltage-gated sodium channel (Lin et al., 2017).

The importance of PUM1 function in the mammalian brain, however, was not apparent until we found that *Pum1*-heterozygous mice develop a progressive ataxia and Purkinje cell degeneration that resembles what we see in SCA1 mice (Watase et al., 2002), largely due to an increase in Atxn1 RNA and protein levels of about 30%–40% in the cerebellum (Gennarino et al., 2015b). *Pum1* null mice experience a roughly 50%–60% rise in Atxn1 levels and develop a more severe phenotype, with hyperactivity, developmental delay, smaller body size and weight, infertility, and shortened lifespan (Chen et al., 2012; Gennarino et al., 2015b). Removing a copy of *Pum1* in SCA1 mice accelerates their disease progression, whereas removing a copy of *Atxn1* in *Pum1*-heterozygous mice rescues the cerebellar phenotypes by normalizing wild-type (WT) Atxn1 levels (Gennarino et al., 2015b). Other features of the *Pum1* nulls, such as hyperactivity, are not rescued after normalizing Atxn1 levels in *Pum1* mutant mice, confirming the contributions of other Pum1 targets to the null phenotype.

These data led us to hypothesize that *PUM1* insufficiency, as would be caused by heterozygous loss-of-function (LoF) mutations or genomic deletions, would cause neurological disease in humans as well. The Exome Aggregation Consortium (ExAC) database (Lek et al., 2016) gives a probability of LoF intolerance (pLI) of 1.00 and a missense constraint metric (MCM) score of 4.59 for *PUM1*, which further suggests that LoF of the gene could be pathogenic in humans. Here, we describe fifteen individuals with either severe early-onset disease due to complete loss of one *PUM1* allele or milder, late-onset disease with partial loss of function in one *PUM1* allele. These data not only demonstrate the importance of *PUM1* for human neurological development and function, but also suggest that the class of RBPs should be investigated for involvement in neurodegenerative and neurodevelopmental disorders.

RESULTS

PUM1 Deletions Are Associated with Syndromic Developmental Delay

We sought patients with copy-number variations (CNVs) of *PUM1* from public databases and from a cohort of 52,000 patients who underwent clinical chromosome microarray analyses for neurodevelopmental disorders (see STAR Methods for more details). We found no duplication events spanning *PUM1* in affected individuals or *PUM1*-spanning chromosomal deletions among 41,345 healthy controls ($p = 0.0092$, one-tailed Chi square test without Yates correction; see STAR Methods), but we did identify 9 patients with heterozygous deletions ranging in size from 0.6 to 5.60 Mb, where the minimal region of overlap included only *PUM1* (Figure 1A).

All nine subjects were documented to have developmental delay; eight had language delay, seven have intellectual disability, six have ataxia, and three were noted to have seizures (Tables 1 and S1). Much more comprehensive information was available for subject 9, who had been previously reported as having a microdeletion of 1p35.2 (Wilson et al., 2015) (Table S1).

To evaluate all the genes encompassed by the nine deletions and assess the likelihood of *PUM1* pathogenicity, we interrogated the ExAC database again and used three additional tools: pcGERP (Petrovski et al., 2015), EvoTol (Rackham et al., 2015), and residual variation intolerance score (RVIS) (Petrovski et al., 2013). Out of all the deleted genes, *PUM1* had the highest probability to be pathogenic (Table S2).

We also interrogated three databases for phenotypic associations: Online Mendelian Inheritance in Man (OMIM, https://www.omim.org/), Mouse Genome Informatics (MGI) (Blake et al., 2017; Eppig et al., 2017), and DisGeNET (a platform integrating information on gene-disease association from several public data sources and literature) (Piñero et al., 2017). A genome-wide association study cited in DisGeNET suggested that *FABP3* may be associated with intellectual disability and schizophrenia (Table S2), but the deletion in subject 4, who is noted to have an intellectual disability, does not include this gene (Figure 1A). No other genes in the deleted regions have been associated with neurodevelopmental abnormalities (Table S2), and the minimal shared deletion includes only PUM1.

Because none of the deletions included only *PUM1*, however, and the deletion of multiple genes at once can have an additive effect beyond the direct activities of any single gene, we searched for additional patients with mutations in this gene.

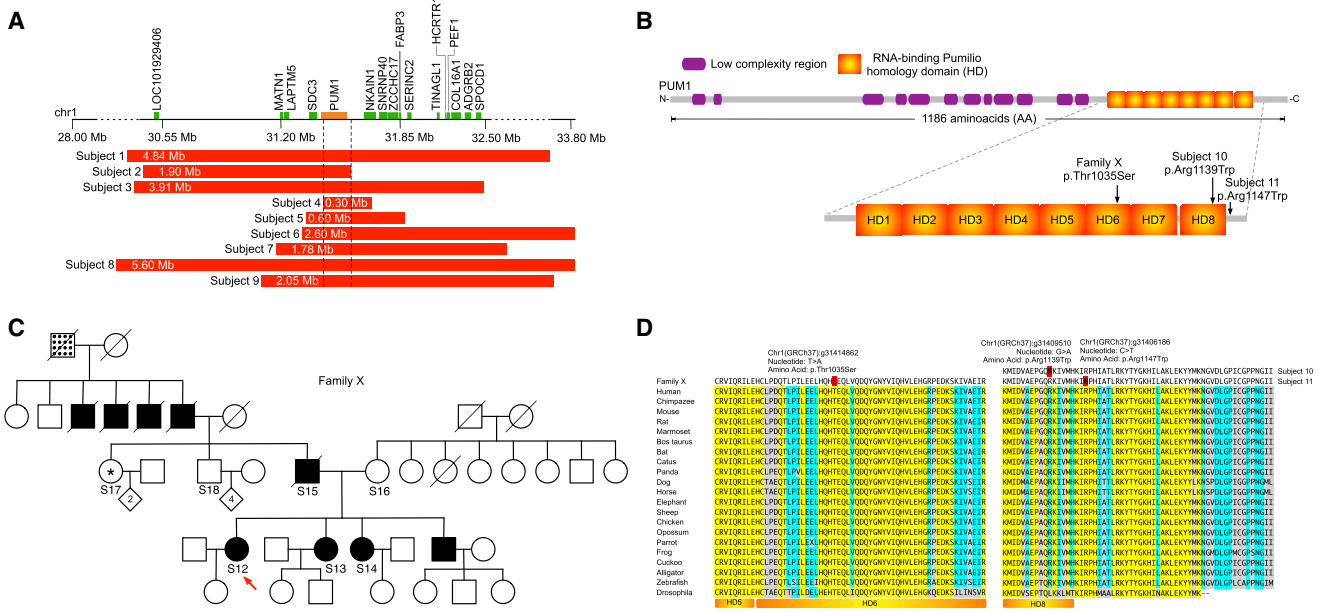


Figure 1. Deletions and Mutations in *PUM1* Identified in Early- and Late-Onset Diseases

(A) Deletions spanning *PUM1* on chromosome 1p35.2 (shown in red) were identified in nine patients with developmental disability. Dashed lines indicate the minimal region spanning *PUM1*.

(B) Schematic of the PUM1 protein. Low-complexity regions are shown as purple boxes and PUM1 homology domains (HDs) are shown as orange boxes. Locations of the PUM1 mutations in subjects 10 and 11 and family X are indicated.

(C) Pedigree shows autosomal dominant inheritance of adult-onset ataxia in family X. White and black denote unaffected and affected individuals, respectively; squares indicate males and circles indicate females; diamonds and numbers indicate the respective offspring; a line through the box indicates that the individual is deceased. Subjects 12–18, who have been sequenced, are numbered in the order in which they were identified; DNA was not available from affected individuals without a subject number. Subject 17 (asterisk) carries the familial mutation but does not have reported ataxia. The square box with dots (the great-grandfather of the proband) is a deceased individual who began using a walker in his 30s or 40s. The red arrow indicates the proband (subject 12).

(D) Protein alignment and comparison of the affected *PUM1* residues compared to 21 organisms from human to *Drosophila melanogaster*. Different colors highlight degree of conservation: yellow for full conservation, light blue if conserved in all but one organism, and gray if more than one organism does not share the same amino acid. The human PUM1 amino acid sequence is used here as the reference protein.

See also Figure S1 and Tables S1, S2, and S3.

Severe *PUM1* Missense Mutations Are Also Associated with Syndromic Developmental Delay

We soon identified two individuals (subjects 10 and 11) with different *de novo* missense variants in *PUM1* with early-onset ataxia and developmental delay (Figures 1B and 1D and Tables 1 and S1).

Subject 10 is a 9-year-old girl who was born small for her age and remained so (she is currently in the 5th centile). She manifested chorea, gait ataxia, and fine-motor incoordination by the age of 5 years (Table 1 and Movie S1). She has dysarthria and spasticity upon waking that improve over the course of the day. Magnetic resonance imaging (MRI) of the brain revealed no brain malformations (Figure 2) at the age of 6 years. Whole-exome sequencing of this subject and her parents revealed a heterozygous *de novo* substitution from arginine 1139 to tryptophan, Chr1(GRCh37):g31409510 G>A (p.Arg1139Trp, transcript NM_001020658.1), in the highly conserved RNA-binding PUM-HD (Zamore et al., 1997). This amino acid change is located in the eighth repeat of the PUM-HD (Figures 1B, 1D and S1A). The p.Arg1139Trp variant was not found in the ExAC database or in BRAVO, a whole-genome sequencing consortium that include functional variants for 62,785 healthy individuals (https://bravo.sph.umich.edu/freeze5/hg38/). Therefore, this is not a common benign variant. Moreover, p.Arg1139Trp is a nonconservative amino acid substitution; because these residues differ in polarity, charge, and size, the mutation is likely to affect secondary protein structure. The *in silico* predictors SIFT (Kumar et al., 2009) and PolyPhen2 (Adzhubei et al., 2010) indicate this variant should be pathogenic (Table S1). This subject also harbored a *de novo* nonsense variant of uncertain significance in the *XPR1* gene, in which missense variants have been implicated in late-onset primary familial brain calcification (PFBC) (Anheim et al., 2016; Legati et al., 2015; Moura and Oliveira, 2015). Subject 10 had congenital, not late-onset, disease, but only time will tell if she develops PFBC later in life.

Subject 11 is a 9-year-old girl who developed generalized epilepsy starting at the age of 5 months that worsened over time. The seizures were not controllable by polypharmacy (carbamazepine, phenobarbitone, lamotrigine, levetiracetum, clobazam, oxcarbazepine, cannabidiol, and combinations) or the ketogenic diet; her electroencephalogram (EEG) findings were consistent with infantile-onset epileptic encephalopathy (Table 1 and Figure S2). She developed early-onset, progressive ataxia with hypotonicity in her lower limbs (Table 1 and Movie

Table 1. Summary of Molecular and Clinical Symptoms of 15 Individuals with *PUM1* Mutations

	PADDAS Syndrome											PRCA Syndrome			
Feature	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Subj. 9[a]	Subj. 10	Subj. 11	Subj. 12	Subj. 13	Subj. 14	Subj.15
Gender	–	–	male	–	–	–	male	female	female	female	female	female	female	female	male
Current age	–	–	16 years	–	–	–	2 years	6 months	7 years	9 years	9 years	59 years	58 years	52 years	81 years (deceased)
Age of onset	–	–	–	–	–	–	–	–	–	<5 years	5 months	early 30s	early 30s	early 40s	early 50s
Chr1 (hg19):	28751378-33588455	29600988-31598923	28716929-32629424	31442430-31720099	31284806-31872758	31239605-33825029	31113947-32897001	28743173-34340430	31091243-33142346	31409510	31406186	31414862	31414862	31414862	31414862
Size	4.84 Mb	1.9 Mb	3.91 Mb	0.3 Mb	0.6 Mb	2.6 Mb	1.78 Mb	5.6 Mb	2.05 Mb	1 nt	1 nt	1 nt	1 nt	1 nt	1 nt
Nucleotide change	deletion	deletion	deletion	deletion	deletion	deletion	deletion	deletion	deletion	G > A	G > A	T > A	T > A	T > A	T > A
AA change	–	–	–	–	–	–	–	–	–	R1139W	R1147W	T1035S	T1035S	T1035S	T1035S
Type	deletion	deletion	deletion	deletion	deletion	deletion	deletion	deletion	deletion	missense	missense	missense	missense	missense	missense
Inheritance	–	–	*de novo*	–	–	–	–	*de novo*	*de novo*	*de novo*	*de novo*	AD	AD	AD	AD
Developmental delay	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no
Intellectual disability	yes	–	yes	yes	yes	yes	–	yes	yes	no	yes	no	no	no	no
Seizures	yes	–	no	yes	–	–	yes	no	–	–	yes	no	no	no	no
Ataxia	yes	yes	no	yes	–	–	yes	yes	yes	yes	yes	yes	yes	yes	yes
Progressive	–	–	yes	–	–	–	–	–	–	no	yes	yes	yes	yes	yes

Clinical information on subjects 1–8 was taken from clinical databases (see STAR Methods for more detail).

Mb, megabase; nt, nucleotide; –, not known. R, arginine; W, tryptophan; S, serine; T, threonine. Subj., subject; AD, autosomal dominant inheritance.See also Tables S1 and S3 and Movies S1, S2, S3, and S4.[a]Subject 9's clinical details were obtained from Wilson BT., et al., 2015 (Wilson et al., 2015) and Decipher (Firth et al., 2009).

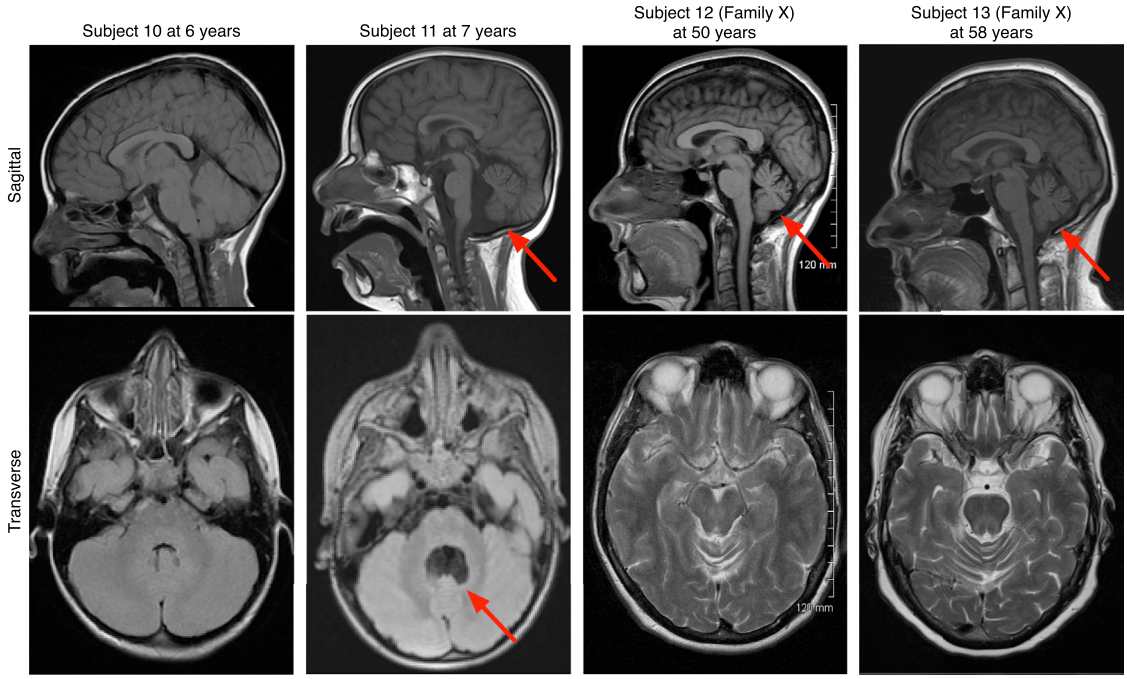


Figure 2. Magnetic Resonance Imaging for Subjects 10, 11, 12, and 13

Representative sagittal and transverse MRI images show normal imaging for subject 10, an enlarged fourth ventricle with elevation and shortening of the vermis in subject 11, and cerebellar atrophy in subjects 12 and 13 (family X).

See also Figure S2.

S2). In addition, she suffers global developmental delay, cortical visual impairment, stereotypic hand-clasping, scoliosis, facial dysmorphia, and low bone mineral density in the neck of femur (four standard deviations below normal for age) (Tables 1 and S1). MRI revealed an enlarged fourth ventricle with elevation and shortening of the cerebellar vermis (Figure 2). Recently, she has been noted to have mild idiopathic intracranial hypertension with bilateral papilledema that improves with lumbar punctures. She was enrolled with her parents in a trio exome-sequencing project (Figures S1B and S1C) and was found to be heterozygous for a *de novo* missense variant adjacent to the eighth PUM-HD repeat just outside the PUM-HD region, Chr1(GRCh37):g31406186 G>A, a nonconservative amino acid substitution from arginine 1147 to tryptophan (p.Arg1147Trp, transcript NM_001020658.1) (Figures 1B and 1D). This variant affects a highly conserved residue and was not present in ExAC or BRAVO; it is predicted to be pathogenic by SIFT and PolyPhen2 (Table S1).

Based on the phenotype of subjects 10 and 11 and the areas of overlap with subjects 1–9, we refer to this disease as *Pumilio1*-associated developmental disability, ataxia, and seizure syndrome (PADDAS).

A *PUM1* Missense Mutation Associated with Adult-Onset Ataxia

A rare missense variant was identified in *PUM1* in a family with cerebellar ataxia (family X) (Figure 1C and Table 1). Previous testing in this family had ruled out triplet repeat expansions in SCA 1, 2, 3, 6, 7, 8, and 17. Exome sequencing identified a heterozygous missense variant, Chr1(GRCh37):g.31414862 T>A, resulting in an amino acid change from threonine 1035 to serine (p.Thr1035Ser, transcript NM_001020658.1) located within PUM1-HD 6 (Figures 1C, 1D, S1D, and S1E) in two affected sisters (subjects 12 and 13) but not their unaffected mother (subject 16); this finding was confirmed by Sanger sequencing. Segregation analysis using Sanger sequencing of DNA from the third affected sister (subject 14), her affected, deceased father (subject 15), and his 80-year-old sister (subject 17, who carries the variant but is reported to be unaffected) and brother (subject 18, who does not carry the variant and is reported to be unaffected) indicates an autosomal-dominant mode of inheritance with incomplete penetrance. We refer to this condition as adult-onset syndrome *Pumilio1*-related cerebellar ataxia (PRCA).

Subjects 12, 13, and 14 are currently 59, 58, and 52 years of age, respectively (Figure 1C and Table 1), and developed an adult-onset, slowly progressive cerebellar ataxia in their 30s or 40s; MRI studies of subjects 12 and 13 at 50 and 58 years of age, respectively (Figure 2), revealed mild vermian atrophy compared to normal MRI images taken 8 and 11 years previously. Their father, subject 15, developed ataxia in his 50s and died at the age of 81 years. Subject 15's uncles, as well as his own father, were reported by the family to have adult-onset ataxia and needed walkers, but there is no available clinical documentation on these individuals. The cerebellar degeneration in the affected members of family X is (or was) characterized by gait ataxia, dysmetria, dysarthria, and in some cases, diplopia (Tables 1 and S3; see also Movie S3 for subject 12 and Movie S4 for subject 13). The affected individuals have no co-morbidities.

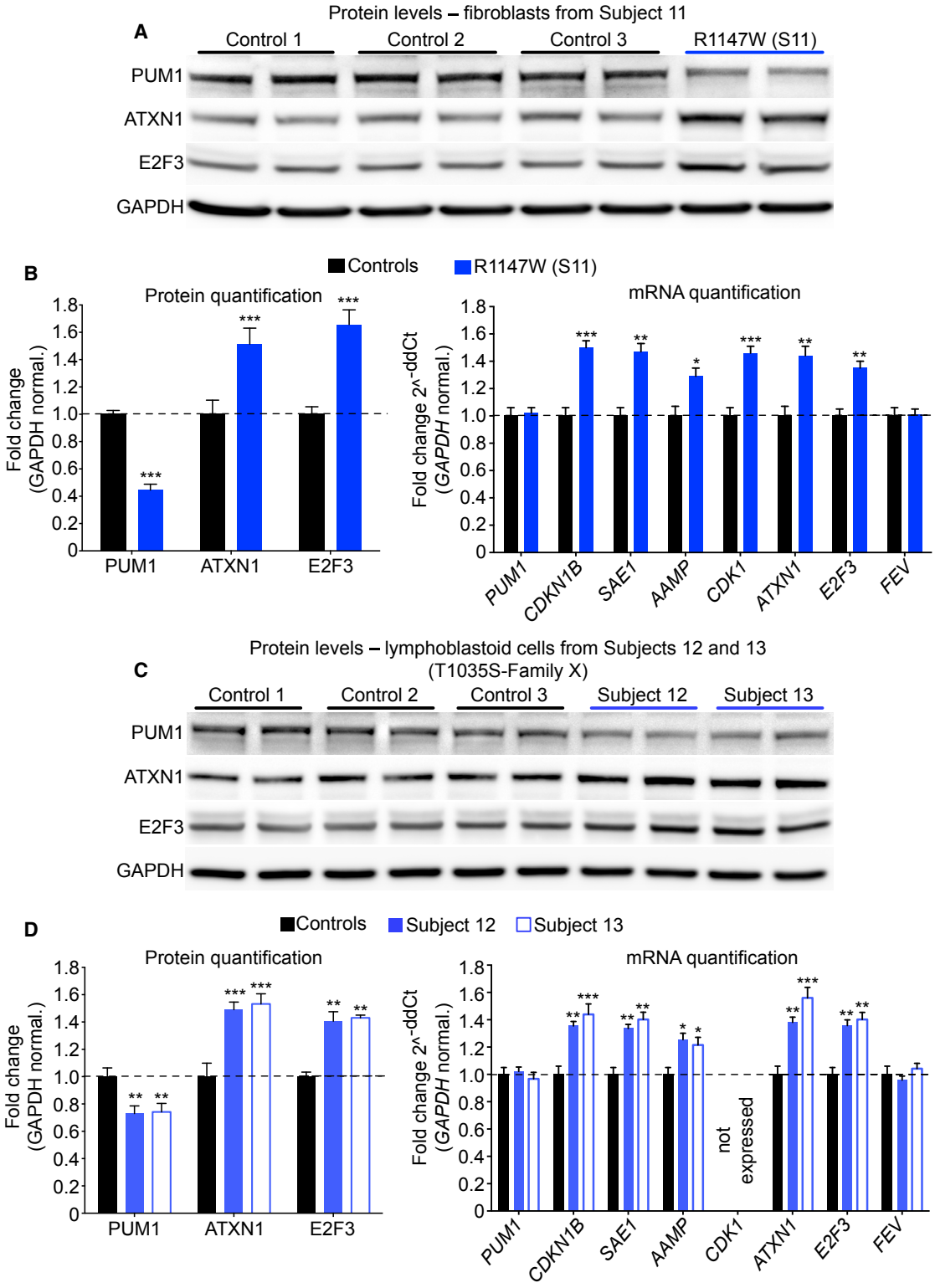


Figure 3. Missense Mutations Decrease PUM1 Stability in Patient-Derived Cells

(A and B) (A) Representative western blot and (B) (left) quantification of protein levels of PUM1 and its targets in patient-derived fibroblast cells from subject 11 (PADDAS) compared to three age-matched fibroblast control cell lines. PUM1 levels are about 50%–60% lower than in healthy controls. (Right) RNA

(legend continued on next page)

Thr1035 is highly conserved, and interestingly, the family's Thr1035Ser variant appeared once among 121,296 alleles (i.e., 60,706 unrelated individuals without neurological disease) in the ExAC database (8.244e−06), but not in BRAVO. Given the late onset of the disease, this single individual does not necessarily present counterevidence to the pathogenicity of the mutation, since someone who contributes to the database when apparently healthy could develop disease later in life (Tarailo-Graovac et al., 2017); this individual might, alternatively, be nonpenetrant. Moreover, that the same mutation could present symptoms in an individual as early as their 30s or as late as their 50s suggests that there are factors that influence penetrance.

The *PUM1* Missense Mutations Diminish PUM1 Protein Stability and Increase Levels of PUM1 Targets

To evaluate the molecular effects of the identified missense mutations, we sought cell lines from patients. We were able to obtain a cultured fibroblast cell line from subject 11 (early-onset PADDAS) and lymphoblastoid cells from subjects 12 and 13 (adult-onset PRCA).

The R1147W mutation (subject 11, juvenile-onset disease) appears to markedly reduce PUM1 protein stability. Western blot analysis showed that, compared with fibroblast cell lines from three age-matched control individuals, subject 11's fibroblasts had only ~43% of WT PUM1 levels. Protein levels of ATXN1 and E2F3, two well-known PUM1 targets (Gennarino et al., 2015b; Miles et al., 2012), were elevated in these cells by ~51% and ~66%, respectively (Figures 3A, 3B, and S3). *PUM1* mRNA levels showed no change, but the mRNA levels of *ATXN1*, *E2F3*, and the four other known PUM1 targets—*CDKN1B* (Kedde et al., 2010), *SAE1*, *CDK1*, and *AAMP* (Chen et al., 2012)—were all elevated (Figure 3B [right panel]). *FEV* mRNA that does not have a PUM1 binding site in its 3′ UTR was used as a negative control (Figure 3B).

The T1035S variant (family X, adult-onset disease) reduced PUM1 stability to a lesser degree. Western blot analysis of lymphoblastoid cells from subjects 12 and 13 and three age-matched controls showed PUM1 levels to be ~73% (subject 12) and ~74% (subject 13) of WT levels (Figure 3C). The consequent increase in ATXN1 protein levels in subject 12 was ~49% and in subject 13 was ~53%; E2F3 levels rose by ~41% and ~43% in the two subjects, respectively (Figure 3D). Again, *PUM1* mRNA levels were unaffected, but mRNA levels of PUM1 targets were elevated (Figure 3D [right panel]). (Note that lymphoblastoid cells do not express CDK1, so this factor could not be compared.)

Evolutionary Action Algorithm Predicts Loss of Fitness for Each Missense Mutant

We could not obtain consent from the family of subject 10 (R1139W) to collect cells, but we were able to apply the evolutionary action (EA) algorithm, which uses homology and phylogeny to estimate the overall loss of fitness to an organism likely to be caused by a missense variant (Katsonis and Lichtarge, 2014, 2017). The higher the EA score, on a continuous scale from 0 to 100, the greater the perturbation of protein function and loss of organismal fitness. Among the three variants, R1147W produced the largest EA score (81); R1139W had an EA score almost as high (70), and T1035S (family X) had the lowest EA score (32; see Table S1). These EA scores thus correspond very nicely with the different phenotypic severities we see resulting from these mutations. Applying a simple linear equation to the relative difference in EA scores between R1147W and T1035S (see STAR Methods), we can predict the degree to which these mutations reduce PUM1 protein levels. Taking the observed 55% reduction in levels for p.Arg1147Trp as a starting point, the equation predicts that T1035S would reduce PUM1 levels by 22%, which is quite close to the actual average reduction of 25% we found in the cell lines from subjects 12 and 13. The same equation predicts that R1139W would cause a reduction in protein levels of 48% (Table S1).

PUM1 Variants Are Impaired in Their Repression Activity

We next performed *in vitro* transfection assays with WT and mutant *PUM1* variants at different concentrations to evaluate their ability to inhibit transcription of *ATXN1* and *E2F3* in HEK293T cells. As expected, overexpression of WT *PUM1* at different concentrations reduced the protein and mRNA levels of both ATXN1 and E2F3 compared to cells transfected with an empty vector (Figures 4A, 4B, and S4A–S4C). Given that the R1147W mutation markedly reduced PUM1 protein levels, we suspected that strongly overexpressing this mutant might enable it to function, and indeed this was the case: overexpression of *PUM1* p.Arg1147Trp reduced ATXN1 and E2F3 levels. On the other hand, overexpression was insufficient to overcome the defects of R1139W and T1035S within the PUM-HD (RNA-binding) domain (Figures 4A, 4B, and S4A–S4C).

Elevated levels of PUM1 alter the morphology of induced pluripotent stem cell (iPSC)-derived neurons, most notably by reducing dendritic arborization (Rodrigues et al., 2016). To test whether overexpressing *PUM1* mutants recapitulates this abnormal morphology, and thus to gauge how much function they retain, we evaluated dissociated mouse hippocampal

quantification of *PUM1* and its targets in fibroblasts from subject 11 compared to three age-matched fibroblast control cell lines. *FEV*, which does not have PUM1 binding sites in its 3′ UTR, served as a negative control.

(C and D) (C) Representative western blot and (D) (left) quantification of PUM1 and its targets in patient-derived lymphoblastoid cell lines from subjects 12 and 13 (late-onset ataxia), compared to three age-matched lymphoblastoid control cell lines, showing a PUM1 decrease of about 25%. (Right) quantification of mRNA of *PUM1* and its targets in patient-derived lymphoblasts from subjects 12 and 13 compared to three age-matched lymphoblast controls. *FEV* again served as a negative control.

All experiments were performed six times, blinded to genotype (data represent mean ± SEM). Data were normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) protein or mRNA levels as appropriate. p values were calculated by Student's t test; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

See also Figure S3 for human ATXN1 antibody validation.

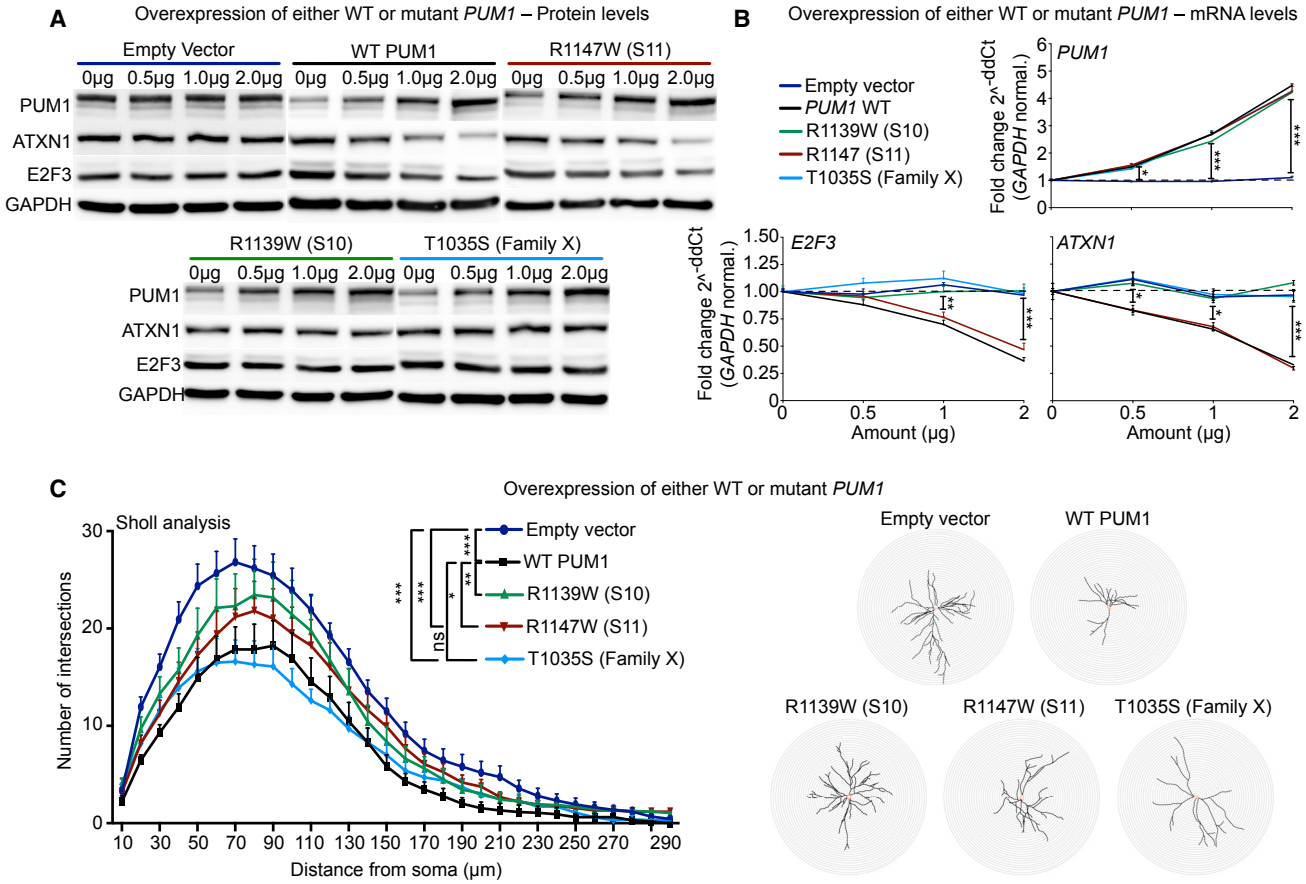


Figure 4. The Missense Mutations Alter Neuronal Morphology and Impair PUM1's Ability to Suppress Its Targets

(A) Protein levels in HEK293T cells transfected with different amounts (from 0 to 2 μg) of either WT or mutant *PUM1*. Cells transfected with *PUM1* bearing a mutation in the PUM-HD (either R1139W or T1035S) cannot suppress levels of ATXN1 and E2F3, two well-known PUM targets. See Figures S4A–S4C for the relative quantification.

(B) mRNA quantification from HEK293T cells transfected as described in (A) showing that PUM1 mRNA repression is lost when the mutation falls inside the PUM-HD.

Empty vector transfection was used as a negative control for (A) and (B). All experiments were performed in triplicate (data represent mean ± SEM); p values were calculated by Student's t test. Data in (A) and (B) were normalized to GAPDH protein or mRNA levels, respectively.

(C) (Left) Sholl analysis of primary mouse hippocampal neurons after overexpression of either WT or *PUM1* mutants quantified the number of intersections at various distances from the soma. (Right) Representative images of dendritic branching as influenced by overexpression of WT or mutant PUM1. Empty vector served as a negative control. Data represent mean ± SEM from 38–44 neurons per transfection. Statistical analysis was performed by two-way ANOVA with Tukey's multiple comparisons test. *p < 0.05, **p < 0.01, ***p < 0.001.

See also Figure S4.

neurons. After confirming that either 3xHA-*PUM1* WT or the mutants were all localized at the cell body of hippocampal neurons, we overexpressed the same amount of 3xHA WT or mutant *PUM1* in hippocampal neurons and measured dendritic complexity (Figure S4D). Sholl analysis showed that, as expected, overexpression of WT *PUM1* in neurons to about twice normal levels impaired dendritic branching (Figure 4C). Overexpression of each mutant suppressed dendritic arborization somewhat compared to empty vector, particularly between 50 and 130 μm from the soma, but only T1035S approached the degree of suppression seen with WT PUM1 (Figure 4C). This is consistent with the milder deficits in family X and suggests that PUM1 may have other functions independent of RNA binding (Darnell and Richter, 2012).

Pum1 Mutant Mice Recapitulate Human *PUM1*-Associated Features

Patients with a ~50% reduction in PUM1 protein levels (subjects 1–11) developed a phenotype very similar to that of *Pum1* knockout mice (Gennarino et al., 2015b). These mice not only have hyperactivity (noted in a few of the PADDAS patients) and progressive cerebellar signs involving gross and fine motor incoordination, but they also show marked growth defects and spontaneous seizures at 16 weeks of age; those that survive gestation (they are born at a lower frequency than expected for Mendelian ratios) are smaller than normal at birth and fall farther behind on the growth curve over the course of their lives (Gennarino et al., 2015b). As noted in Table S1, many of the PADDAS subjects—particularly the two missense mutation patients, for whom we

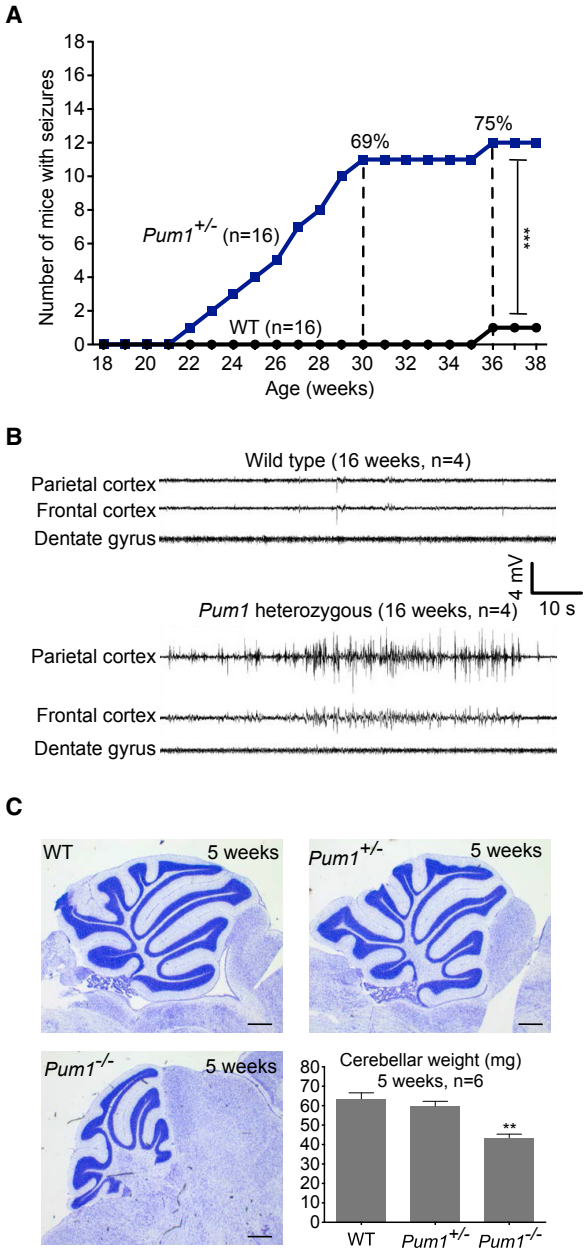


Figure 5. *Pum1* Mutant Mice Manifest Spontaneous Seizures, Abnormal EEG, and Cerebellar Hypoplasia

(A) *Pum1*$^{+/-}$ mice (n = 12) began to manifest spontaneous seizures at 22 weeks of age, and by 30 weeks, the majority had seizures (11 out of 16). See also Movie S5. Data are shown here as cumulative events.

(B) Representative EEG traces. The *Pum1*$^{+/-}$ mice (n = 4) predominantly showed prolonged hyperexcitability discharges in the neocortex. Generalized epileptiform spikes typically lasted over 10 s, indicating a state of subclinical seizure. Neither hyperexcitability discharges nor electrographic seizures were observed in any of the recorded brain regions in WT littermates (n = 4). The scale bars are for both WT and *Pum1*$^{+/-}$ mice.

(C) Nissl staining at 5 weeks of age shows that *Pum1*$^{-/-}$ mice (n = 6) had cerebellar hypoplasia (as confirmed by cerebellar weight, right bottom panel) in comparison to WT and *Pum1*$^{+/-}$ mice. Scale bars, 200 μm. *p < 0.05, **p < 0.01, ***p < 0.001.

have the most reliable clinical data—are small for their age. The two male subjects have cryptorchidism, which is consistent with the testicular hypoplasia reported in *Pum1* knockout mice (Chen et al., 2012). These patients are not null for *PUM1*, however, so we re-examined the phenotype of our *Pum1* haploinsufficient mice, which we had reported as having progressive ataxia, hyperactivity, and Purkinje cell degeneration (Gennarino et al., 2015b).

We found that 11 out of 16 (69%) heterozygous *Pum1*$^{+/-}$ mice exhibited spontaneous seizures by the end of the 30th week of age (Figure 5A and Movie S5) and showed abnormal EEG activity with generalized epileptiform spikes, indicating a state of subclinical seizures, by the age of 16 weeks (Figure 5B). We performed cresyl violet staining (Nissl staining) on cerebella from WT, *Pum1*$^{+/-}$, and *Pum1*$^{-/-}$ mice at 5 weeks of age and found that *Pum1*$^{-/-}$ mice have smaller-than-normal cerebella (Figure 5C). These two *Pum1* mutant lines are therefore a reliable model of *PUM1* haploinsufficiency in humans.

DISCUSSION

The identification of these patients supports our initial hypothesis that *PUM1* would prove to be involved in human neurological disease. The difference PUM1 protein levels make in these individuals is striking: subject 11, with ~45% of normal PUM1 levels, developed severe disease in infancy, whereas members of family X, who retained ~75% of normal PUM1 levels, developed a very mild ataxia in adulthood. Both missense mutations, however, failed to properly suppress PUM1's six known targets, whose protein and mRNA levels were significantly elevated.

Even as we strive to identify the full complement of PUM1 targets, it will be useful to understand the particular roles these six proteins play in the developing brain. ATXN1 and E2F3 may prove particularly challenging, as both are involved in transcription: ATXN1 is a transcriptional suppressor, whereas E2F3 is an activator, particularly of genes with functions in cellular proliferation and differentiation (Chong et al., 2009). Each influences the expression of multiple targets, many of which have yet to be identified. The other PUM1 targets are involved in protein metabolism (*SAE1*, sumoylation), angiogenesis (*AAMP*), or cell-cycle regulation (*CDKN1B* and *CDK1*). In this context, it is worth noting that PUM1 has a close homolog, PUM2, which has been shown to affect neurological development and function in the rat (Darnell and Richter, 2012; Vessey et al., 2006, 2010) and, most recently, in the mouse (Zhang et al., 2017).

The mutations resulting in infantile disease were not identified in the ExAC database (release version: August 8, 2016) (Lek et al., 2016), but the mutation associated with adult-onset disease was found in one exome from ExAC. This could reflect the milder, later onset of the disorder associated with this variant: the individual in ExAC provided his or her clinical information only once, and it remains to be seen if s/he will develop ataxia in later life or if s/he represents a case of incomplete penetrance.

The incomplete penetrance in family X, along with the small number of family members that we were able to directly evaluate, leads us to be cautious in making claims about the pathogenicity of the T1035S variant. T1035S did lose its ability to repress transcription of PUM1 targets in the cellular studies, and the symptoms in the adult-onset PRCA are consistent across the

family, though the age of onset varied over a ~20-year period. Such wide variance indicates the presence of factors modulating onset or penetrance. Subject 17 and the single individual with this variant in ExAC could also be resilient individuals in the face of this milder mutation (Tarailo-Graovac et al., 2017). Further studies will be necessary to understand what other molecular or environmental factors confer risk for (or protection against) PRCA. It would be very interesting to determine whether variations in polyglutamine tract length in ATXN1 or other proteins are risk factors for PRCA, as for example has been shown for ATXN1, ATXN2, and amyotrophic lateral sclerosis (Conforti et al., 2012). Ultimately, we must await the identification of other mild variants in unrelated individuals with adult-onset ataxia to confirm the contribution of PUM1 to a late-onset ataxia.

Genes encoding RBPs are increasingly appreciated as potential causes of disease (Cookson, 2017). Given that RBPs regulate hundreds or thousands of mRNAs, it is conceivable that alterations or mutations in RBPs will be at the root of many complex neurological diseases (Gerstberger et al., 2014; Lukong et al., 2008). The identification of factors that regulate the levels of disease-driving proteins, particularly those whose excess or insufficiency are detrimental to neurons (such as APP and SNCA), would be a reasonable place to start. Besides deepening our understanding of pathogenesis, there is a strong likelihood that such studies would also reveal new candidate disease genes, as with the discovery of ATXN1's regulation by PUM1. Understanding RBPs and their associated pathways could also lead to molecular means of modulating protein levels in neurological disorders in which protein accumulation is central to pathogenesis.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human studies
  - Exome sequencing and Sanger confirmation
  - In vivo animal studies
  - Fibroblast generation and culturing
  - Lymphoblastoid cell culture
  - Cell culture, transfection and mutagenesis
- METHOD DETAILS
  - In silico variant impact prediction
  - Protein quantification and western blot analysis
  - RNA extraction and quantitative real-time PCR
  - Purification and generation of human ATXN1 antibody #534
  - Primary hippocampal neuron culture, transfection and immunostaining
  - Surgery and EEG recordings
  - Cresyl Violet (Nissl) staining
  - Primers
  - Experimental design
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, three tables, and five movies and can be found with this article online at https://doi.org/10.1016/j.cell.2018.02.006.

ACKNOWLEDGMENTS

We thank the patients and their families for participating in our study and the clinical teams—particularly our co-author, J.A.P., who devoted herself to this work even during the final months of her fatal illness. We thank J.Y. Jones for re-analyzing MRI images; M. Buckley, G. Mullan, W. Lo, G. Elakis, and C. Walsh from NSW Health Pathology Genetics Laboratory for their technical assistance; staff of the *In Vivo* Neurophysiology Core at the Jan and Dan Duncan Neurological Research Institute (NRI); E.H. Crutcher for growing human cells; the Protein and Monoclonal Antibody Production Core at Baylor College of Medicine (NCI-CA125123) for generating ATXN1 monoclonal antibody; C. Alcott, S. Yamamoto, M.F. Wangler, E.H. Crutcher and J.P. Orengo for reading the paper; and V. Brandt for essential input on the manuscript. The work was supported by the Howard Hughes Medical Institute (HHMI to H.Y.Z.), National Institute of Neurological Disorders and Stroke (NINDS) (2R37NS027699 to H.Y.Z. and NS091381 to J.L.H.), National Institute of General Medical Sciences (GM066099 and GM079656 to O.L.), the Robbins Foundation (J.L.H.), Joan and Stanford Alexander Family (C.P.S.), CPRIT RP160283 (A.K.), the National Ataxia Foundation/Young Investigator Research Grant 2017 (to V.A.G.); the Neuroconnectivity Core at the NRI and Baylor College of Medicine (IDDRC U54HD083092), Eunice Kennedy Shriver National Institute of Child Health & Human Development, Care4Rare Canada Consortium funded by Genome Canada, the Canadian Institute of Health Research, the Ontario Genomics Institute, Ontario Research Fund, Genome Quebec, Children's Hospital of Eastern Ontario Foundation, NHMRC post-graduate scholarship (GNT11149630 to E.E.P.), NHMRC CJ Martin Overseas Post-Doctoral Fellowship (GNT0512123 to T.R.), and Toupin Foundation at the University of Alberta (O.S.). We thank the Exome Aggregation Database (ExAC) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. The content is the sole responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development, the National Institutes of Health, or any other funding entity. The Department of Molecular and Human Genetics at Baylor College of Medicine receives income from clinical genetic testing performed at Baylor Genetics Laboratories.

AUTHOR CONTRIBUTIONS

V.A.G. conceived the study, performed and supervised molecular experiments, analyzed the data, and wrote the manuscript. H.Y.Z. analyzed data and supervised the project. J.L.H., C.P.S., U.M., J.A.R., J.A.P., H.A.D., E.E.P., E.P.K., R.K.S., R.M., A.M.B., B.B., K.-R.D., L.B.H., M.T.C., P.S., T.K., A.B., T.R., M.C., M.E.D., Y.Z., K.Y., P.M., J. Tao, L.M.M., O.C., O.S., H.Y.Z., and K.M.B. were involved in recruiting subjects, collecting clinical information, and exome sequencing. J. Tang and S.H. recorded and analyzed mouse EEG data. L.W., C.-A.C., L.S., A.M., and C.J.A. contributed to the molecular studies. M.J.C. oversaw the bioinformatics pipeline. A.K. and O.L. performed evolutionary action analysis of missense mutations. All authors read and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 17, 2017
Revised: August 23, 2017
Accepted: February 1, 2018
Published: February 22, 2018

REFERENCES

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A.; 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al.; International HapMap 3 Consortium (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58.

Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F., and Hamosh, A. (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res. *43*, D789–D798.

Anheim, M., López-Sánchez, U., Giovannini, D., Richard, A.C., Touhami, J., N'Guyen, L., Rudolf, G., Thibault-Stoll, A., Frebourg, T., Hannequin, D., et al. (2016). XPR1 mutations are a rare cause of primary familial brain calcification. J. Neurol. *263*, 1559–1564.

Banerjee, S., Oldridge, D., Poptsova, M., Hussain, W.M., Chakravarty, D., and Demichelis, F. (2011). A computational framework discovers new copy number variants with functional importance. PLoS ONE *6*, e17539.

Barch, M.J.; Association of Cytogenetic Technologists (1991). The ACT Cytogenetics Laboratory Manual, Second Edition (New York: Raven Press).

Beaulieu, C.L., Majewski, J., Schwartzentruber, J., Samuels, M.E., Fernandez, B.A., Bernier, F.P., Brudno, M., Knoppers, B., Marcadier, J., Dyment, D., et al.; FORGE Canada Consortium (2014). FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. Am. J. Hum. Genet. *94*, 809–817.

Blake, J.A., Eppig, J.T., Kadin, J.A., Richardson, J.E., Smith, C.L., and Bult, C.J.; the Mouse Genome Database Group (2017). Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. Nucleic Acids Res. *45* (D1), D723–D729.

Burright, E.N., Clark, H.B., Servadio, A., Matilla, T., Feddersen, R.M., Yunis, W.S., Duvick, L.A., Zoghbi, H.Y., and Orr, H.T. (1995). SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. Cell *82*, 937–948.

Campbell, C.D., Sampas, N., Tsalenko, A., Sudmant, P.H., Kidd, J.M., Malig, M., Vu, T.H., Vives, L., Tsang, P., Bruhn, L., and Eichler, E.E. (2011). Population-genetic properties of differentiated human copy-number polymorphisms. Am. J. Hum. Genet. *88*, 317–332.

Chartier-Harlin, M.C., Kachergus, J., Roumier, C., Mouroux, V., Douay, X., Lincoln, S., Levecque, C., Larvor, L., Andrieux, J., Hulihan, M., et al. (2004). Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. Lancet *364*, 1167–1169.

Chen, D., Zheng, W., Lin, A., Uyhazi, K., Zhao, H., and Lin, H. (2012). Pumilio 1 suppresses multiple activators of p53 to safeguard spermatogenesis. Curr. Biol. *22*, 420–425.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the functional effect of amino acid substitutions and indels. PLoS ONE *7*, e46688.

Chong, J.L., Tsai, S.Y., Sharma, N., Opavsky, R., Price, R., Wu, L., Fernandez, S.A., and Leone, G. (2009). E2f3a and E2f3b contribute to the control of cell proliferation and mouse development. Mol. Cell. Biol. *29*, 414–424.

Conforti, F.L., Spataro, R., Sproviero, W., Mazzei, R., Cavalcanti, F., Condino, F., Simone, I.L., Logroscino, G., Patitucci, A., Magariello, A., et al. (2012). Ataxin-1 and ataxin-2 intermediate-length PolyQ expansions in amyotrophic lateral sclerosis. Neurology *79*, 2315–2320.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al.; Wellcome Trust Case Control Consortium (2010). Origins and functional impact of copy number variation in the human genome. Nature *464*, 704–712.

Cookson, M.R. (2017). RNA-binding proteins implicated in neurodegenerative diseases. Wiley Interdiscip. Rev. RNA *8*, 1.

Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. Nat. Genet. *43*, 838–846.

Darnell, J.C., and Richter, J.D. (2012). Cytoplasmic RNA-binding proteins and the control of complex brain function. Cold Spring Harb. Perspect. Biol. *4*, a012344.

Eppig, J.T., Smith, C.L., Blake, J.A., Ringwald, M., Kadin, J.A., Richardson, J.E., and Bult, C.J. (2017). Mouse Genome Informatics (MGI): resources for mining mouse genetic, genomic, and biological data in support of primary and translational research. Methods Mol. Biol. *1488*, 47–73.

Fernandez-Funez, P., Nino-Rosales, M.L., de Gouyon, B., She, W.C., Luchak, J.M., Martinez, P., Turiegano, E., Benito, J., Capovilla, M., Skinner, P.J., et al. (2000). Identification of genes that modify ataxin-1-induced neurodegeneration. Nature *408*, 101–106.

Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. Am. J. Hum. Genet. *84*, 524–533.

Fukao, A., and Fujiwara, T. (2017). The coupled and uncoupled mechanisms by which trans-acting factors regulate mRNA stability and translation. J. Biochem. *161*, 309–314.

Gennarino, V.A., Alcott, C.E., Chen, C.A., Chaudhury, A., Gillentine, M.A., Rosenfeld, J.A., Parikh, S., Wheless, J.W., Roeder, E.R., Horovitz, D.D., et al. (2015a). NUDT21-spanning CNVs lead to neuropsychiatric disease and altered MeCP2 abundance via alternative polyadenylation. eLife *4*, e10782.

Gennarino, V.A., Singh, R.K., White, J.J., De Maio, A., Han, K., Kim, J.Y., Jafar-Nejad, P., di Ronza, A., Kang, H., Sayegh, L.S., et al. (2015b). Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. Cell *160*, 1087–1098.

Gerstberger, S., Hafner, M., Ascano, M., and Tuschl, T. (2014). Evolutionary conservation and expression of human RNA-binding proteins and their role in human genetic disease. Adv. Exp. Med. Biol. *825*, 1–55.

Hao, S., Tang, B., Wu, Z., Ure, K., Sun, Y., Tao, H., Gao, Y., Patel, A.J., Curry, D.J., Samaco, R.C., et al. (2015). Forniceal deep brain stimulation rescues hippocampal memory in Rett syndrome mice. Nature *526*, 430–434.

International Schizophrenia Consortium (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature *455*, 237–241.

Itsara, A., Cooper, G.M., Baker, C., Girirajan, S., Li, J., Absher, D., Krauss, R.M., Myers, R.M., Ridker, P.M., Chasman, D.I., et al. (2009). Population analysis of large copy number variants and hotspots of human genetic disease. Am. J. Hum. Genet. *84*, 148–161.

Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., et al. (2008). Genotype, haplotype and copy-number variation in worldwide human populations. Nature *451*, 998–1003.

Kaminsky, E.B., Kaul, V., Paschall, J., Church, D.M., Bunke, B., Kunig, D., Moreno-De-Luca, D., Moreno-De-Luca, A., Mulle, J.G., Warren, S.T., et al. (2011). An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genet. Med. *13*, 777–784.

Katsonis, P., and Lichtarge, O. (2014). A formal perturbation equation between genotype and phenotype determines the Evolutionary Action of protein-coding variations on fitness. Genome Res. *24*, 2050–2058.

Katsonis, P., and Lichtarge, O. (2017). Objective assessment of the evolutionary action equation for the fitness effect of missense mutations across CAGI-blinded contests. Hum. Mutat. *38*, 1072–1084.

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J.A., Elkon, R., and Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3′ UTR controls miR-221 and miR-222 accessibility. Nat. Cell Biol. *12*, 1014–1020.

Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K.K., Holmans, P., Craddock, N., Owen, M.J., and O'Donovan, M.C.; International Schizophrenia Consortium; Wellcome Trust Case Control Consortium (2009). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. Hum. Mol. Genet. *18*, 1497–1503.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics *23*, 1289–1291.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. *4*, 1073–1081.

La Cognata, V., Morello, G., D'Agata, V., and Cavallaro, S. (2017). Copy number variability in Parkinson's disease: assembling the puzzle through a systems biology approach. Hum. Genet. *136*, 13–37.

Lee, S., Kopp, F., Chang, T.C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. Cell *164*, 69–80.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell *14*, 385–393.

Legati, A., Giovannini, D., Nicolas, G., López-Sánchez, U., Quintáns, B., Oliveira, J.R., Sears, R.L., Ramos, E.M., Spiteri, E., Sobrido, M.J., et al. (2015). Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. Nat. Genet. *47*, 579–581.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

Lin, W.H., Giachello, C.N., and Baines, R.A. (2017). Seizure control through genetic and pharmacological manipulation of Pumilio in Drosophila: a key component of neuronal homeostasis. Dis. Model. Mech. *10*, 141–150.

Lukong, K.E., Chang, K.W., Khandjian, E.W., and Richard, S. (2008). RNA-binding proteins in human genetic disease. Trends Genet. *24*, 416–425.

Mata, J., Marguerat, S., and Bähler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. Trends Biochem. Sci. *30*, 506–514.

Miles, W.O., Tschöp, K., Herr, A., Ji, J.Y., and Dyson, N.J. (2012). Pumilio facilitates miRNA regulation of the E2F3 oncogene. Genes Dev. *26*, 356–368.

Moura, D.A., and Oliveira, J.R. (2015). XPR1: a gene linked to primary familial brain calcification might help explain a spectrum of neuropsychiatric disorders. J. Mol. Neurosci. *57*, 519–521.

Palmer, E.E., Hayner, J., Sachdev, R., Cardamone, M., Kandula, T., Morris, P., Dias, K.R., Tao, J., Miller, D., Zhu, Y., et al. (2015). Asparagine Synthetase Deficiency causes reduced proliferation of cells under conditions of limited asparagine. Mol. Genet. Metab. *116*, 178–186.

Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.

Petrovski, S., Gussow, A.B., Wang, Q., Halvorsen, M., Han, Y., Weir, W.H., Allen, A.S., and Goldstein, D.B. (2015). The Intolerance of regulatory sequence to genetic variation predicts gene dosage sensitivity. PLoS Genet. *11*, e1005492.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. *29*, e45.

Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., García-García, J., Sanz, F., and Furlong, L.I. (2017). DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Res. *45* (D1), D833–D839.

Pinto, D., Marshall, C., Feuk, L., and Scherer, S.W. (2007). Copy-number variation in control population cohorts. Hum. Mol. Genet. *16*, R168–R173, *Spec No. 2*.

Rackham, O.J., Shihab, H.A., Johnson, M.R., and Petretto, E. (2015). EvoTol: a protein-sequence based evolutionary intolerance framework for disease-gene prioritization. Nucleic Acids Res. *43*, e33.

Ramocki, M.B., Peters, S.U., Tavyev, Y.J., Zhang, F., Carvalho, C.M., Schaaf, C.P., Richman, R., Fang, P., Glaze, D.G., Lupski, J.R., and Zoghbi, H.Y. (2009). Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. Ann. Neurol. *66*, 771–782.

Rodrigues, D.C., Kim, D.S., Yang, G., Zaslavsky, K., Ha, K.C., Mok, R.S., Ross, P.J., Zhao, M., Piekna, A., Wei, W., et al. (2016). MECP2 Is post-transcriptionally regulated during human neurodevelopment by combinatorial action of RNA-binding proteins and miRNAs. Cell Rep. *17*, 720–734.

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Le Meur, N., Laquerrière, A., Vital, A., Dumanchin, C., Feuillette, S., Brice, A., Vercelletto, M., et al. (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat. Genet. *38*, 24–26.

Schmahmann, J.D., Gardner, R., MacMore, J., and Vangel, M.G. (2009). Development of a brief ataxia rating scale (BARS) based on a modified form of the ICARS. Mov. Disord. *24*, 1820–1828.

Schweers, B.A., Walters, K.J., and Stern, M. (2002). The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability. Genetics *161*, 1177–1185.

Shaikh, T.H., Gai, X., Perin, J.C., Glessner, J.T., Xie, H., Murphy, K., O'Hara, R., Casalunovo, T., Conlin, L.K., D'Arcy, M., et al. (2009). High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. Genome Res. *19*, 1682–1690.

Simon-Sanchez, J., Scholz, S., Fung, H.C., Matarin, M., Hernandez, D., Gibbs, J.R., Britton, A., de Vrieze, F.W., Peckham, E., Gwinn-Hardy, K., et al. (2007). Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. Hum. Mol. Genet. *16*, 1–14.

Sleegers, K., Brouwers, N., Gijselinck, I., Theuns, J., Goossens, D., Wauters, J., Del-Favero, J., Cruts, M., van Duijn, C.M., and Van Broeckhoven, C. (2006). APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. Brain *129*, 2977–2983.

Sobreira, N., Schiettecatte, F., Boehm, C., Valle, D., and Hamosh, A. (2015a). New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. Hum. Mutat. *36*, 425–431.

Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015b). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.

Spassov, D.S., and Jurecic, R. (2003). The PUF family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function? IUBMB Life *55*, 359–366.

Tanaka, A.J., Cho, M.T., Millan, F., Juusola, J., Retterer, K., Joshi, C., Niyazov, D., Garnica, A., Gratz, E., Deardorff, M., et al. (2015). Mutations in SPATA5 are associated with microcephaly, intellectual disability, seizures, and hearing loss. Am. J. Hum. Genet. *97*, 457–464.

Tarailo-Graovac, M., Zhu, J.Y.A., Matthews, A., van Karnebeek, C.D.M., and Wasserman, W.W. (2017). Assessment of the ExAC data set for the presence of individuals with pathogenic genotypes implicated in severe Mendelian pediatric disorders. Genet. Med. *19*, 1300–1308.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3–new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. *3*, RESEARCH0034.

Vessey, J.P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M.A., and Macchi, P. (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. J. Neurosci. *26*, 6496–6508.

Vessey, J.P., Schoderboeck, L., Gingl, E., Luzi, E., Riefler, J., Di Leva, F., Karra, D., Thomas, S., Kiebler, M.A., and Macchi, P. (2010). Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. Proc. Natl. Acad. Sci. USA *107*, 3222–3227.

Vogler, C., Gschwind, L., Röthlisberger, B., Huber, A., Filges, I., Miny, P., Auschra, B., Stetak, A., Demougin, P., Vukojevic, V., et al. (2010). Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. PLoS ONE *5*, e15246.

Watase, K., Weeber, E.J., Xu, B., Antalffy, B., Yuva-Paylor, L., Hashimoto, K., Kano, M., Atkinson, R., Sun, Y., Armstrong, D.L., et al. (2002). A long CAG repeat in the mouse Sca1 locus replicates SCA1 features and reveals the impact of protein solubility on selective neurodegeneration. Neuron *34*, 905–919.

Wilson, B.T., Omer, M., Hellens, S.W., Zwolinski, S.A., Yates, L.M., and Lynch, S.A. (2015). Microdeletion 1p35.2: a recognizable facial phenotype with developmental delay. Am. J. Med. Genet. A. *167A*, 1916–1920.

Xiao, L., Lanz, R.B., Frolov, A., Castro, P.D., Zhang, Z., Dong, B., Xue, W., Jung, S.Y., Lydon, J.P., Edwards, D.P., et al. (2016). The germ cell gene TDRD1 as an ERG target gene and a novel prostate cancer biomarker. Prostate *76*, 1271–1284.

Xie, W.R., Jen, H.I., Seymour, M.L., Yeh, S.Y., Pereira, F.A., Groves, A.K., Klisch, T.J., and Zoghbi, H.Y. (2017). An Atoh1-S193A phospho-mutant allele causes hearing deficits and motor impairment. J. Neurosci. *37*, 8583–8594.

Zamore, P.D., Williamson, J.R., and Lehmann, R. (1997). The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. RNA *3*, 1421–1433.

Zhang, M., Chen, D., Xia, J., Han, W., Cui, X., Neuenkirchen, N., Hermes, G., Sestan, N., and Lin, H. (2017). Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins. Genes Dev. *31*, 1354–1369.

Zogopoulos, G., Ha, K.C., Naqib, F., Moore, S., Kim, H., Montpetit, A., Robidoux, F., Laflamme, P., Cotterchio, M., Greenwood, C., et al. (2007). Germline DNA copy number variation frequencies in a large North American population. Hum. Genet. *122*, 345–353.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal anti-PUM1	Bethyl Laboratories	Cat# A300-201A; RRID: AB_2253218
Rabbit polyclonal anti-E2F3	Abcam	Cat# ab50917; RRID: AB_869541
Human monoclonal anti-ATXN1	This paper	N/A
Mouse monoclonal anti-GAPDH	Millipore	Cat# CB1001-500UG; RRID: AB_2107426
Chicken polyclonal anti-GFP	Abcam	Cat# ab13970; RRID: AB_300798
Mouse monoclonal anti-HA.11	Biolegend	Cat# 901514; RRID: AB_2565336
Biological Samples		
Subject 11 fibroblast cells	Sydney Children's Hospital https://www.schn.health.nsw.gov.au/	N/A
Subjects 12 and 13 lymphoblastoid cells	Care4Rare Canada http://care4rare.ca	https://www.phenomecentral.org/
Critical Commercial Assays		
QuikChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent Technologies	NC9893312
miRNeasy kit	QIAGEN	Cat# 217004; RRID: AB_776329
Quantitect Reverse Transcription kit	QIAGEN	205311
Agilent Clinical Research Exome kit	Agilent Technologies	5190-9492
Nextera rapid capture expanded exome kit	Illumina	FC-140
PowerUP SYBR Green Master Mix	Thermo Fisher Scientific	A25743
Experimental Models: Cell Lines		
Human embryonic kidney immortalized 293 cells (HEK293T)	Zoghbi Lab	N/A
Human medulloblastoma cells (DAOY)	Zoghbi Lab	N/A
Experimental Models: Organisms/Strains		
B6/129 mixed background *Pum1* mouse models	Chen et al., 2012 This paper	Lin Lab
Oligonucleotides		
Primer sequences for qPCR, see Figures 3 and 4	This paper	N/A
Human *GAPDH*: Forward, 5′-CGACCACTTTGTCAAGCTCA-3′;	This paper	N/A
Human *GAPDH*: Reverse, 5′-TTACTCCTTGGAGGCCATGT-3′	This paper	N/A
Human *PUM1*: Forward, 5′- GCCCCAGTCTTTGCAATTTA-3′;	This paper	N/A
Human *PUM1*: Reverse. 5′- AATCACTCGGCAGCCATAAG-3′	This paper	N/A
Human *ATXN1*: Forward, 5′-CCAGGTCAGCGTTGAAGTTT-3′;	This paper	N/A
Human *ATXN1*: Reverse, 5′- CAAAGAGCTGGCTGGTTCTC-3′	This paper	N/A
Human *E2F3*: Forward, 5′-AAAGCCCCTCCAGAAACAAG-3′;	This paper	N/A
Human *E2F3*: Reverse, 5′-AATGGGCCCTTGGGTACTT-3′	This paper	N/A
Sequences for mutagenesis, see Figures 4 and S4	This paper	N/A
Subject 10: 5′-CATGACGATCTTCCACTGGCCTGGCTCC-3′	This paper	N/A
Subject 11: 5′-GCGATGTGGGGCCAGATCTTATGCATGACGA-3′	This paper	N/A
Subject 12/13/14: 5′-GTACAAGCTGCTCTGAGTGCTGGT GAAGCTC-3′	This paper	N/A
Recombinant DNA		
*PUM1*cDNA expression vector	OmickLink Expression Vector, GeneCopoeia	pEZ-M02, EX-E2337-M02-10
Empty expression vector	OmickLink Expression Vector, GeneCopoeia	EX-NEG-M02

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Primer3	Whitehead Institute	http://bioinfo.ut.ee/primer3-0.4.0/ Untergasser et al., 2012
GraphPrism 6	GraphPad	https://www.graphpad.com/scientific-software/prism/
Evolutionary Action (EA)	Olivier Lichtarge	Katsonis and Lichtarge, 2014, 2017
GeneMatcher	Center for Mendelian Genomics	https://genematcher.org/
Xome Analyzer	GeneDx	https://www.genedx.com/
HaplotypeCaller	GATK	https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_haplotypecaller_HaplotypeCaller.php
SAMtools	Github social coding	http://samtools.sourceforge.net/
BCFtools	Github social coding	https://samtools.github.io/bcftools/
QuikChange software	Stratagene	https://www.genomics.agilent.com/primerDesignProgram.jsp
Mouse Genome Informatics (MGI)	The Jackson Laboratory	http://www.informatics.jax.org/
DisGeNET (v5.0)	Piñero et al., 2017	http://www.disgenet.org/web/DisGeNET/menu;jsessionid=qqvf9r16hk99w6v7mzc4ikth
Exome Aggregation Consortium (ExAC)	Broad Institute	http://exac.broadinstitute.org/
BRAVO	University of Michigan	https://bravo.sph.umich.edu/freeze5/hg38/
SIFT	Kumar et al., 2009	http://sift.jcvi.org/
PROVEAN	Choi et al., 2012	http://provean.jcvi.org/index.php
pcGERP, GERP++	Petrovski et al., 2015	http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005492
RVIS	Petrovski et al., 2013	http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1003709
EvoTol	Rackham et al., 2015	https://academic.oup.com/nar/article/43/5/e33/2453165
OMIM, Online Mendelian Inheritance in MAN	Amberger et al., 2015	https://www.ncbi.nlm.nih.gov/omim
Integrative Genomic Viewer (IGV)	Thorvaldsdóttir et al., 2013	http://software.broadinstitute.org/software/igv/
NHLBI Exome Sequencing Project Exome Variant Server (EVS)	University of Washington	http://evs.gs.washington.edu/EVS/
PolyPhen2	Adzhubei et al., 2010	http://genetics.bwh.harvard.edu/pph2/
Combined Annotation Dependent Depletion (CADD)	University of Washington	http://cadd.gs.washington.edu/
NCBI database (dbSNP)	National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov/snp/
The International Standard Cytogenomic Array (ISCA)	Kaminsky et al., 2011	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000205.v6.p2
DECIPHER	Wellcome Sanger Institute	https://decipher.sanger.ac.uk/
Other		
siRNAs, see Figures 4 and S4	N/A	N/A
ON-TARGET plus Duplex, human ATXN1	Thermo Fisher Scientific (Dharmacon)	J-004510-06-0005
ON-TARGET plus Duplex, non-targeting siRNA#3	Thermo Fisher Scientific (Dharmacon)	D-001810-03-02-50
Ham's F10 transport medium	Thermo Fisher Scientific	11550043
AmnioMAX	Thermo Fisher Scientific	11269016
RPMI 1640 medium	Thermo Fisher Scientific (Invitrogen)	11875093
Fetal Bovine Serum	Atlanta Biologicals	S11195
Fetal Bovine Serum Heat Inactivated	Atlanta Biologicals	S11195H
DMEM	Thermo Fisher Scientific (Invitrogen)	11966-025

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Huda Y Zoghbi (hzoghbi@bcm.edu). All reagents, cell lines and mouse models used in this manuscript are available upon request without restrictions.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human studies

Subjects 1, 2, 4, 5 and 6 were identified by using The International Standard Cytogenomic Array (ISCA, as of October 2015) (Kaminsky et al., 2011) public database with the following accession numbers: nssv582298 (Subject 1), nssv576394 (Subject 2), nssv1603425 (Subject 4), nssv582880 (Subject 5) and nssv577199 (Subject 6). Table 1 lists whatever information on the sex and age of these patients that was available in the database. Of approximately 52,000 patients referred to the Baylor Genetics (BG) Laboratories for clinical array comparative genomic hybridization (aCGH) analysis between April 2007 and February 2013, there were two probands (Subjects 7 and 8) with deletions affecting *PUM1*. The search was limited to copy-number variants (CNVs) <20 megabases. When possible, we reviewed medical records and neuropsychological testing following informed consent in accordance with a protocol approved by the Institutional Review Board (IRB) for Human Subject Research at Baylor College of Medicine (H-25466). The CNVs had been detected by clinical aCGH on the following platforms: Oligo V6.5, Oligo V8.1.1, and CMA-HR + SNP V9.1.1 at Baylor College of Medicine. Deletion coordinates and clinical details were obtained from Wilson BT., et al., 2015 (Wilson et al., 2015) for subject 9. Subject 3 was discovered initially through Decipher (258365) (Firth et al., 2009) and subsequently seen by co-author U.M. The 41,345 healthy controls used here to calculate the statistical likelihood that *PUM1* deletion is pathogenic were drawn from multiple studies (Banerjee et al., 2011; Campbell et al., 2011; Conrad et al., 2010; Cooper et al., 2011; Abecasis et al., 2012; Altshuler et al., 2010; International Schizophrenia Consortium, 2008; Itsara et al., 2009; Jakobsson et al., 2008; Kirov et al., 2009; Pinto et al., 2007; Shaikh et al., 2009; Simon-Sanchez et al., 2007; Vogler et al., 2010; Zogopoulos et al., 2007).

Candidate pathogenic mutations of *PUM1* (Subjects 10-18) were identified by using the online tool GeneMatcher (Sobreira et al., 2015a; Sobreira et al., 2015b). Subject 10: The proband, with her parents, had clinically indicated trio exome sequencing performed at GeneDx (Gaithersburg, MD). Subject 11: The proband was enrolled, with her parents, in a trio exome sequencing project performed at the Kinghorn Centre for Clinical Genomics (KCCG) as part of a cohort of patients with epileptic encephalopathy of unknown cause, coordinated by the Departments of Clinical Genetics and Pediatric Neurology at Sydney Children's Hospital, Sydney, Australia (Palmer et al., 2015). Informed consent for exome sequencing and skin biopsy was obtained, and the research was approved by the ethics committee from The Sydney Children's Hospital Network and the Prince of Wales Hospital Campus, Sydney, Australia (HREC ref no 13/094 and LNR/13/SCHN/112) and the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals (approval H-34578). Electroencephalograms and brain MRI were performed as diagnostic procedures at the Prince of Wales Hospital Campus, Sydney, Australia.

Family X belongs to a cohort of patients with autosomal dominant spinocerebellar ataxia of unidentified cause enrolled for study by the Care4Rare Canada Consortium. Whole exome sequencing was performed on Subjects 12, 13, and their unaffected mother (Subject 16) at McGill University and Genome Quebec Innovation Centre (Montreal, Canada). Sanger sequencing was performed on Subjects 12-18 from this family. The study, including generation of lymphoblastoid cell lines from blood samples, was approved by Research Ethics Board at the Children's Hospital of Eastern Ontario. MRIs on Subjects 12 and 13 were performed as diagnostic procedures at their local hospitals (The Ottawa Hospital and University of Alberta Hospital, respectively). See Tables 1, S1 and S3 for more details.

The consent form for all individuals specifically allows for sharing of medical information and physical exam findings; the sharing of cell lines from Subjects 11, 12, 13, and the controls was approved under the Baylor College of Medicine IRB H-34578.

Exome sequencing and Sanger confirmation

Subject 10: Genomic DNA was extracted from whole blood from the affected individual and her parents. The Agilent Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA, USA) was used to target the exonic regions and flanking splice junctions of the genome. These targeted regions were sequenced simultaneously by massively parallel (NextGen) sequencing on an Illumina HiSeq 2000 sequencing system with 100bp paired-end reads (Tanaka et al., 2015). Bi-directional sequence was assembled, aligned to reference gene sequences based in human genome build GRCh37/UCSC hg19, and analyzed for sequence variants using Xome Analyzer from GeneDx (https://www.genedx.com/test-catalog/medical-specialty/xomedx/). GeneDx also performed capillary sequencing to confirm all potentially pathogenic variants identified in the proband and parents' samples.

Subject 11: High-quality DNA was obtained by extraction from peripheral blood in EDTA. Next Generation Sequencing was performed using a Nextera rapid capture expanded exome kit (Illumina, San Diego, CA), with libraries analyzed on an Illumina HiSeq2500. Bidirectional Sanger sequencing was performed on DNA from the proband and both parents to verify and segregate candidate variants. Reads were aligned to Human Genome Reference Sequence GRCh37/UCSC hg19 using BWA MEM (http://bio-bwa.sourceforge.net/), and single nucleotide and short insertion/deletion variants were identified using HaplotypeCaller from GATK (https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_haplotypecaller_HaplotypeCaller.php). Data filtering and variant prioritization were performed using aligned BAM files and the GEMINI (v0.5.1b)

platform (https://gemini.readthedocs.io/en/latest/), incorporating data from the NCBI database (dbSNP, version 138), the 1000 Genomes Project (release date 21/05/2011), and the NHLBI Exome Sequencing Project Exome Variant Server (EVS) database (ESP6500SI-V2; http://evs.gs.washington.edu/EVS/), and the Combined Annotation Dependent Depletion database (CADD, http://cadd.gs.washington.edu/). Variants were filtered out if they were in dbSNP, had a frequency of > 1% in both 1000 Genomes and EVS, were present in an in-house control, or were predicted to have a low impact on protein function. Any candidate variants from prioritization were further assessed for pathogenicity using *in silico* prediction tools—SIFT, PolyPhen2, PROVEAN (Choi et al., 2012), and CADD—and were manually checked on the Binary Alignment Map files through Integrative Genomic Viewer (IGV) (Thorvaldsdóttir et al., 2013; Palmer et al., 2015). Informed consent for exome sequencing was obtained and the research was approved by the ethics committee from The Sydney Children's Hospital Network and the Prince of Wales Hospital Campus, Sydney, Australia (HREC ref no 13/094 and LNR/13/SCHN/112).

Subjects 12, 13, and 16: Exome sequencing was performed on blood-derived DNA. Exome target enrichment was performed using the Sure Select All Exon V5 capture kit, followed by sequencing on the Illumina HiSeq 2000. Reads were quality trimmed and subsequently aligned to hg19 using BWA. Duplicate reads were marked and excluded using Picard (http://picard.sourceforge.net). Average coverage of the consensus coding sequence was calculated using Genome Analysis ToolKit (GATK, https://software.broadinstitute.org/gatk/). Mean coverage was approximately 130x for Subjects 12 and 13 after duplicate read removal and 89x for Subject 16. Variant calling and annotation was performed with GATK, Sequence Alignment/Map Tools (SAMtools, http://samtools.sourceforge.net/) and BCFtools (https://samtools.github.io/bcftools/) (Beaulieu et al., 2014). The data were filtered to identify missense, nonsense and splice variants. These were screened against variants in ExAC (http://exac.broadinstitute.org/), the NHLBI Exome Sequencing Project Exome Variant Server (EVS, http://evs.gs.washington.edu/EVS/), the NCBI database (dbSNP, https://www.ncbi.nlm.nih.gov/snp/), and in-house variant databases (Care4Rare, CHEO Research Institute, Ottawa, Ontario, Canada). In addition, only variants shared by Subjects 12 and 13 were considered; variants inherited from the unaffected mother (Subject 16) were filtered out. Sanger sequencing of the *PUM1* variant was performed on available DNA from Subjects 12-18.

In vivo animal studies

All animal procedures were approved by the Institutional Animal Care and Use Committee for Baylor College of Medicine and Affiliates. Mice were maintained on a 12-hr light, 12-hr dark cycle with regular mouse chow and water *ad libitum*. *Pum1* mutant mice used here were previously generated in a B6/129 mixed background by the University of Connecticut Gene Targeting and Transgenic Facility (Chen et al., 2012). For the experiments in Figure 5 we used 16 mice per genotype (8 males and 8 females) from 18 to 38 weeks of age for panel A; 4 mice per genotype for panel B (2 males and 2 females); and 6 mice per genotype for panel C (3 males and 3 females).

Fibroblast generation and culturing

Primary fibroblasts were isolated from a skin biopsies taken from Subject 11 (9-year old female) and three age-matched controls using standard methodology (Barch and Association of Cytogenetic Technologists, 1991), and placed in a transport medium (Ham's F10, Thermo Fisher Scientific). The skin tissue specimen was later removed from the transport medium using a sterile technique (in a Class II biohazard cabinet) and transferred to a sterile Petri dish where it was cut into small pieces (< 0.5 mm) using sterile scalpel blades. These pieces were later transferred to a lower surface of a 25 cm^2 culture flask (6-8 pieces per flask) which had been pre-moistened with 1-2 mL of AmnioMAX Complete Medium (Thermo Fisher Scientific) supplemented with 1% penicillin/streptomycin. Cell cultures were maintained at 37°C in a humidified incubator supplemented with 5% CO_2. When cell growth was observed around the edges of the tissue, usually 3 to 5 days later, 2 to 3 mL were added. Once growth was established and the tissue was anchored to the flask, another 8 mL of AmnioMAX Complete Medium was added. Thereafter, the medium was renewed every 3 to 4 days until ready for sub-culturing.

Lymphoblastoid cell culture

Venous blood of Subjects 12 (59-year old female), 13 (58-year old female), and three age-matched controls was drawn into ACD solution A tubes. Buffy coat was prepared, lymphocytes were pelleted and transformed with Epstein–Barr virus and cyclosporin A following standard procedures. Cell lines were grown in RPMI 1640 medium (Invitrogen, Carlsbad, CA) supplemented with 10% heat-inactivated fetal bovine serum (Atlanta Biological, Flowery Branch, GA) and 1% penicillin/streptomycin. Cell cultures were maintained at 37°C in a humidified incubator supplemented with 5% CO_2. Medium was renewed every 2 to 3 days to maintain the cell density between 1×10^5 and 2×10^6 cells/ml. Subjects 12 and 13 consented to participate in the Care4Rare Canada research protocol as described above.

Cell culture, transfection and mutagenesis

Human embryonic kidney immortalized 293 cells (HEK293T) and human medulloblastoma cells (DAOY) were grown in DMEM (Invitrogen), supplemented with 10% of heat-inactivated fetal bovine serum (FBS) and 1% penicillin/streptomycin. All cells were incubated at 37°C in a humidified chamber supplemented with 5% CO_2. Transfection of HEK293T cells was performed using jetPRIME

Transfection Reagent (Polyplus-transfection, New York, NY) and DAOY cells by DharmaFECT Transfection Reagent (Thermo Fisher Scientific, Springfield Township, NJ) according to the manufacturers' protocols. HEK293T cells were seeded in 12-well plates before transfection for 24 h and then transfected with zero, 0.5 μg, 1.0 μg, or 2.0 μg of either full or mutant cDNA of *PUM1* (3567 nt) cloned into a mammalian expression vector OmickLink™ Expression Vector pEZ-M02, EX-E2337-M02-10 (GeneCopoeia, Rockville, MD). Empty vector OmickLink™ Expression Vector EX-NEG-M02 (GeneCopoeia) was used here as a negative control. After 48 h, cells were collected and processed for RNA or protein extraction. DAOY cells were seeded in 6-well plates before transfection for 24 h and then transfected with 20 nM of either ON-TARGET plus Duplex human Atxn1- siRNA or ON-TARGET plus non-targeting siRNA#3 respectively (Thermo Fisher Scientific). Mutagenesis reactions were performed using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies). Primers for single mutagenesis analysis were automatically designed by QuikChange software (Stratagene, San Diego, CA). See Key Resources Table for primer sequences.

METHOD DETAILS

In silico variant impact prediction

The three missense mutations were assessed for their likely effect on protein function using SIFT, PolyPhen2, and Evolutionary Action (EA) (Katsonis and Lichtarge, 2014, 2017). EA scores were also used to predict the percent reduction of protein levels for pArg1139Trp, according to the equation $Var1_{EA}/Var2_{EA} = Var1_{\%\ reduction}/Var2_{\%\ reduction}$.

Protein quantification and western blot analysis

Lymphoblastoid or fibroblast cell suspension and HEK293T cell cultures were collected at 6 × 10^6 confluence and processed for protein extraction. Cell pellets were lysed with modified RIPA buffer (25 mM Tris-HCL, pH 7.6, 150 mM NaCl, 1.0% NP-40, 1.0% sodium deoxycholate, 0.1% SDS, and complete protease inhibitor cocktail [Roche, Basel, Switzerland]) by pipetting them up and down with a p1000 tip and then placed for 15 min on ice followed by rotisserie shaker for 15 min at 4 °C. Proteins were quantified by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and resolved by high resolution NuPAGE 4%–12% Bis-Tris Gel (Thermo Fisher Scientific) according to the manufacturer's instruction. Antibodies: Goat α-PUM1 (1:5000, A300-201A [Bethyl Laboratories, Montgomery, TX]); rabbit α-E2F3 (1:2000, ab50917 [Abcam, Cambridge, United Kingdom]); human α-ATXN1 (1:1000, Zoghbi, #534, unpublished; for more detail, see Figure S3 and purification section below); mouse α-GAPDH (1:10,000, CB1001-500UG [Millipore, Billerica, MA]).

RNA extraction and quantitative real-time PCR

HEK293T were seeded in 12-well and transfected with 1000 ng of either WT or mutant cDNA of PUM1. Empty vector and non-transfection were used as negative controls (see cell cultures, transfection and mutagenesis for more details). Human patient or control fibroblast cell lines were seeded in 6-well plates and harvested at 1.3 X 10^6 cells, while lymphoblastoid cell suspension cultures were collected at 6 X 10^6 confluence before proceeding for RNA extraction. For all cells, total RNA was obtained using the miRNeasy kit (QIAGEN, Venlo, Netherlands) according to the manufacturer's instructions. RNA was quantified using NanoDrop 1000 (Thermo Fisher Scientific). Quality of RNA was assessed by gel electrophoresis. cDNA was synthesized using Quantitect Reverse Transcription kit (QIAGEN) starting from 1 μg of DNase-treated RNA. Quantitative RT-polymerase chain reaction (qRT-PCR) experiments were performed using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA) with PowerUP SYBR Green Master Mix (Thermo Fisher Scientific). Real-time PCR results were analyzed using the comparative Ct method normalized against the housekeeping gene GAPDH (Vandesompele et al., 2002). To ensure the efficacy of the genomic DNA elimination, we ran negative control samples in the qRT-PCR that did not have reverse transcriptase (–RT) in the cDNA synthesis reaction.

Purification and generation of human ATXN1 antibody #534

Three gBlocks were designed to encode human Ataxin-1 with codons optimized for expression in *E. coli* K-12 (IDT technologies, Coralville, IA). Gibson cloning was performed using the pET28a vector and the three gBlocks to generate a protein expression vector with a 6X-His tag and a TEV cleavage site. This construct was transformed in BL21AI One Shot *E. coli* and Ataxin-1 [30Q] expression was induced with 0.5 mM IPTG and 0.2% L-arabinose for 4 hours at 37°C. The bacteria were lysed using sonication and 1% Triton X-100 in lysis buffer (6 M urea 20 mM BME 0.5 M NaCl 30 mM Imidazole 50 mM $NaPO_4$ pH 7.4). The supernatant was then loaded onto a pre-packed 5 mL high-performance Ni Sepharose column (GE Healthcare, Little Chalfont, United Kingdom). Elution was performed using a step gradient of 330 mM imidazole in lysis buffer. Fractions with > 80% purity were concentrated and used for mouse immunization. Mouse immunizations, spleen removal, PEG-mediated cell fusion with mouse myelomas, propagation of hybridomas and ELISA screening of hybridomas were performed by the BCM Protein and Monoclonal Antibody Production Core (Xiao et al., 2016). Secondary screening of cloned cell lines using immunoblotting and immunoprecipitation were also performed to confirm specificity of antibodies produced. Cells were cultured in IMDM (Corning, Corning NY) + 15% fetal bovine serum + penicillin and streptomycin until cell density reached 1x10^6 cells per ml. Supernatant was collected and filtered with 0.2 μm filter then loaded onto a GE protein G HP 5 mL column and eluted with 125 mM Glycine pH 2.3. See Figure S3 for a representative western blot validation.

Primary hippocampal neuron culture, transfection and immunostaining

Hippocampal neurons were prepared each of two WT postnatal day 0-1 FVB/N mice. Cells from each mouse were plated independently on poly-D-lysine/mouse laminin-coated coverslips (BD Biosciences, Franklin Lakes, NJ) in Neurobasal medium supplemented with GlutaMAX (Invitrogen), B-27 and 1% FBS. At days *in vitro* (DIV) 9, neurons were transfected with 300ng pEGFP-C1 plus either 700ng empty vector or vectors expressing 3xHA-*PUM1* variants. At DIV 14, neurons were fixed with 4% formaldehyde/4% sucrose in PBS, and permeabilized/blocked by 2% goat serum and 0.1% Triton X-100 in PBS (blocking buffer). Samples were then stained with anti-GFP antibody (Abcam, ab13970) and anti-HA. 11 antibody (901514 [Biolegend, San Diego, CA]) at 4°C overnight, then secondary Alexa-conjugated antibodies at room temperature for 1hr. Z stack images were acquired by LSM710 (Zeiss, Oberkochen, Germany) confocal microscope under the same parameter settings. Neurons were traced and Sholl analysis was performed using Neurolucida 360 (MBF Bioscience, Williston, VT) in a blinded manner. This experiment was performed in triplicate for each mouse. 3xHA-*PUM1* WT and variants were all localized at the cell body of hippocampal neurons. The overexpression level for each variant was quantified based on the fluorescence signal intensity of the HA tag. Specifically, cells expressing 3xHA-*PUM1* were identified and analyzed using Imaris (http://www.bitplane.com/download); the mean intensity of HA at the cell body was used to quantify *PUM1* overexpression for each variant.

Surgery and EEG recordings

The methods were modified from previous publication (Hao et al., 2015). Adult mice at 12 weeks of age were anesthetized with 1%–2% isofluorane. Under aseptic conditions, each animal was surgically implanted with cortical EEG recording electrodes (Teflon-coated silver wire, 127 μm diameter) in the subdural space of the left parietal cortex and left frontal cortex, respectively, with the reference electrode positioned in the occipital region of the skull. The third recording electrode (Teflon-coated tungsten wire, 50 μm diameter) was aimed at the dentate gyrus (P2.0R1.8H1.8) with the reference electrode at the corpus callosum. All electrode wires were attached to a miniature connector (Harwin Connector, Nicomatic, Warminster, PA) and secured on the skull by dental cement. After 2 weeks of post-surgical recovery, simultaneous EEG activity (filtered between 0.1 Hz and 1.0 kHz, sampled at 2 kHz), and behavior were recorded in freely-moving mice for a total of 28 hours over 7 days.

Cresyl Violet (Nissl) staining

Serial paraffin sagittal sections (40 μm) from WT, *Pum1*$^{+/-}$ and *Pum1*$^{-/-}$ mice at 5 weeks of age were cut on a cryostat, then collected and processed free-floating in PBS. Free-floating tissue sections were washed with PBD and Nissl-stained in a solution containing 0.1% thionin in acetic acid, pH 5.5. Four mice (2 males and 2 females) were used for each genotype, for a total of 12 mice.

Primers

In order to unambiguously distinguish spliced cDNA from genomic DNA contamination, specific exon primers were designed to amplify across introns of the genes tested. The primers for all target genes tested were designed with Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012). See Key Resources Table for primer sequences.

Experimental design

For protein and RNA quantification from patient-derived cell lines, we used values from six independent experiments with three biological replicates for each experiment. At every stage of the study, for each experiment, the experimenter was blinded to the identity of control and patient-derived cell lines. For example, for the data in Figure 3, Experimenter #1 made a list of samples and controls to be tested, Experimenter #2 randomized this list and re-labeled the tubes; this was the only person who knew the key to identify the samples. These samples were then distributed to Experimenter #3 to culture the cells, then to Experimenter #1 to perform western blots and qRT-PCRs, and lastly Experimenters #1 and #4 analyzed the data. Only then was the key applied to identify the samples.

For mouse experiments, we used equal numbers of male and female animals. The experimenters as noted above were randomized and blinded for performing the experiments, doing statistical analysis, and interpreting the results.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical significance was analyzed using GraphPad Prism Software. All data are presented as mean ± SEM. *$p < 0.05$; **$p < 0.01$; ***$p < 0.0001$ in all figures.

The range of expression levels in qPCR was determined from six independent experiments with three biological replicates by calculating the standard deviation of the ΔCt (Pfaffl, 2001). We considered genes to be down- or upregulated if they showed a change in their expression with a p value less than < 0.01.

The range of expression levels in western blots was determined from six independent experiments with three biological replicates. P values were calculated by Student's t test or analysis of variance with Tukey's post hoc analysis.

For the Sholl analysis, data represent mean ± SEM from 38-44 neurons in each of three transfections. Statistical analysis was performed by two-way ANOVA with Tukey's multiple comparisons test.

The number of animals used (n), and the specific statistical tests used are indicated for each experiment in the figure legends. Sample size was based on previous experience using the same mice (Gennarino et al., 2015b). Mice were randomly assigned to vehicle or treatment groups using Excel software to generate a table of random numbers, and the experimenter was always blinded to the treatment.

DATA AND SOFTWARE AVAILABILITY

No software was generated for this project. All software used in this study is publicly available and links are provided as appropriate in different sections of the methods.

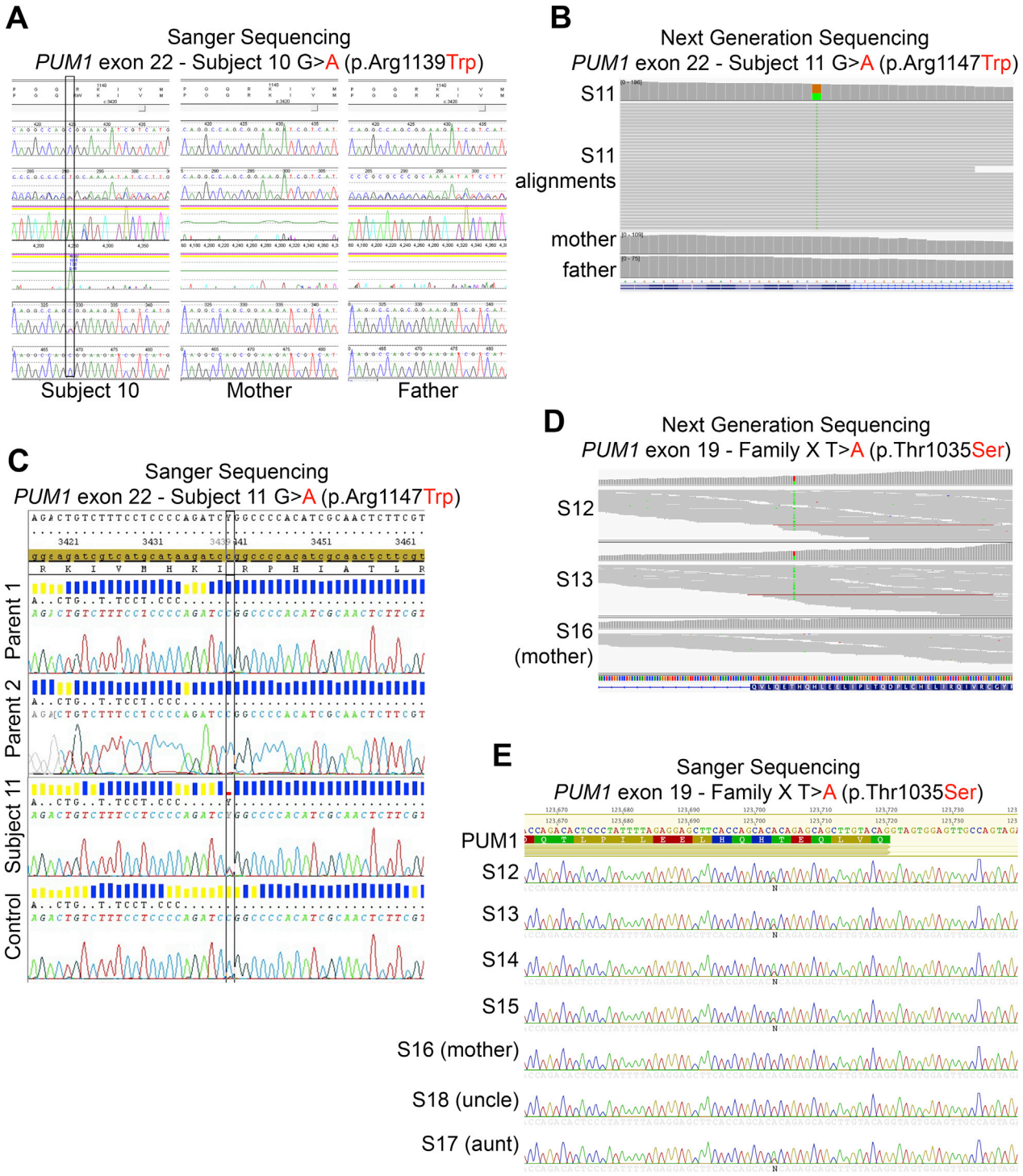


Figure S1. Sanger Sequencing Next Generation Sequencing Pileup for All the Identified Patients with a Single-Nucleotide Variant in *PUM1*, Related to Figure 1

(A) Subject 10 shows the variant G > A (C > T on the coding strand) in *PUM1*. Both parents were homozygous for the wild-type *PUM1* allele, confirming the variant was *de novo*.

(B) Next-generation sequencing (NGS) of Subject 11's family. Subject 11 is heterozygous for PUM1 variant G > A (C > T on the coding strand); both parents were homozygous for the wild-type PUM1 allele, confirming the variant was *de novo*.

(C) Sanger sequencing of Subject 11 confirmed the variant G > A in *PUM1*. Both parents and a healthy control were used to confirm the variant was *de novo*.

(D and E) Representative snapshots of NGS pileup and Sanger sequence of Subjects 12-18 (Family X) showing the variant T > A in *PUM1*.

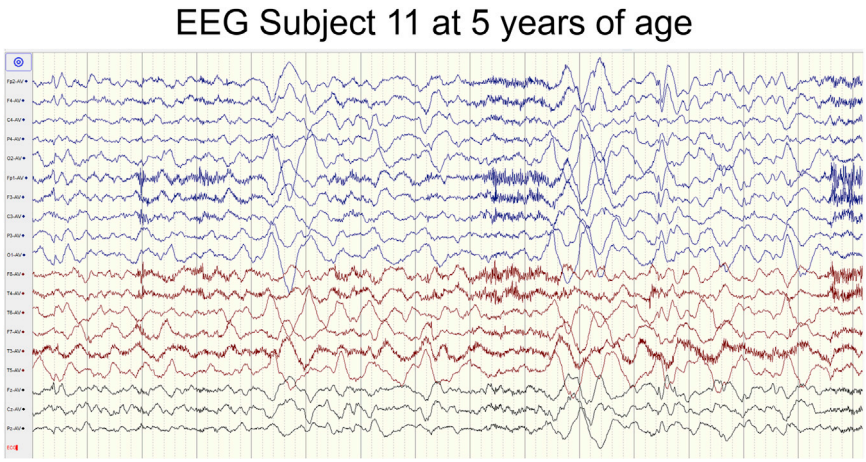


Figure S2. Electroencephalogram of Subject 11, Related to Figure 2

EEG performed at 5 years of age shows poorly organized background slowing and occasional multifocal epileptiform activity.

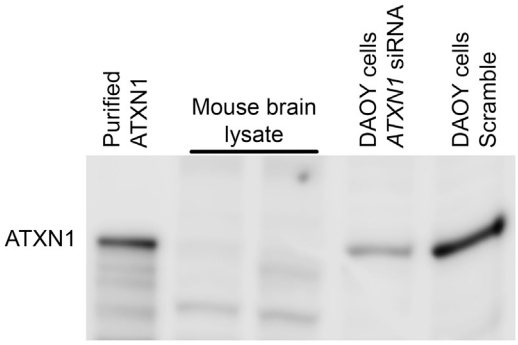


Figure S3. Validation of Human ATXN1 Monoclonal Antibody, Related to Figure 3

The purified monoclonal antibody #534 was confirmed to react with purified human ATXN1 (lane 1) and human ATXN1 expressed in DAOY cells (lane 4 and 5). Furthermore, a reduction in signal is observed when human ATXN1 is knocked down in DAOY cells upon treatment with a small interfering RNA against *ATXN1* (si*ATXN1*). There was no signal observed from mouse brain lysate (lanes 2 and 3), indicating that the purified antibody specifically recognized human ATXN1.

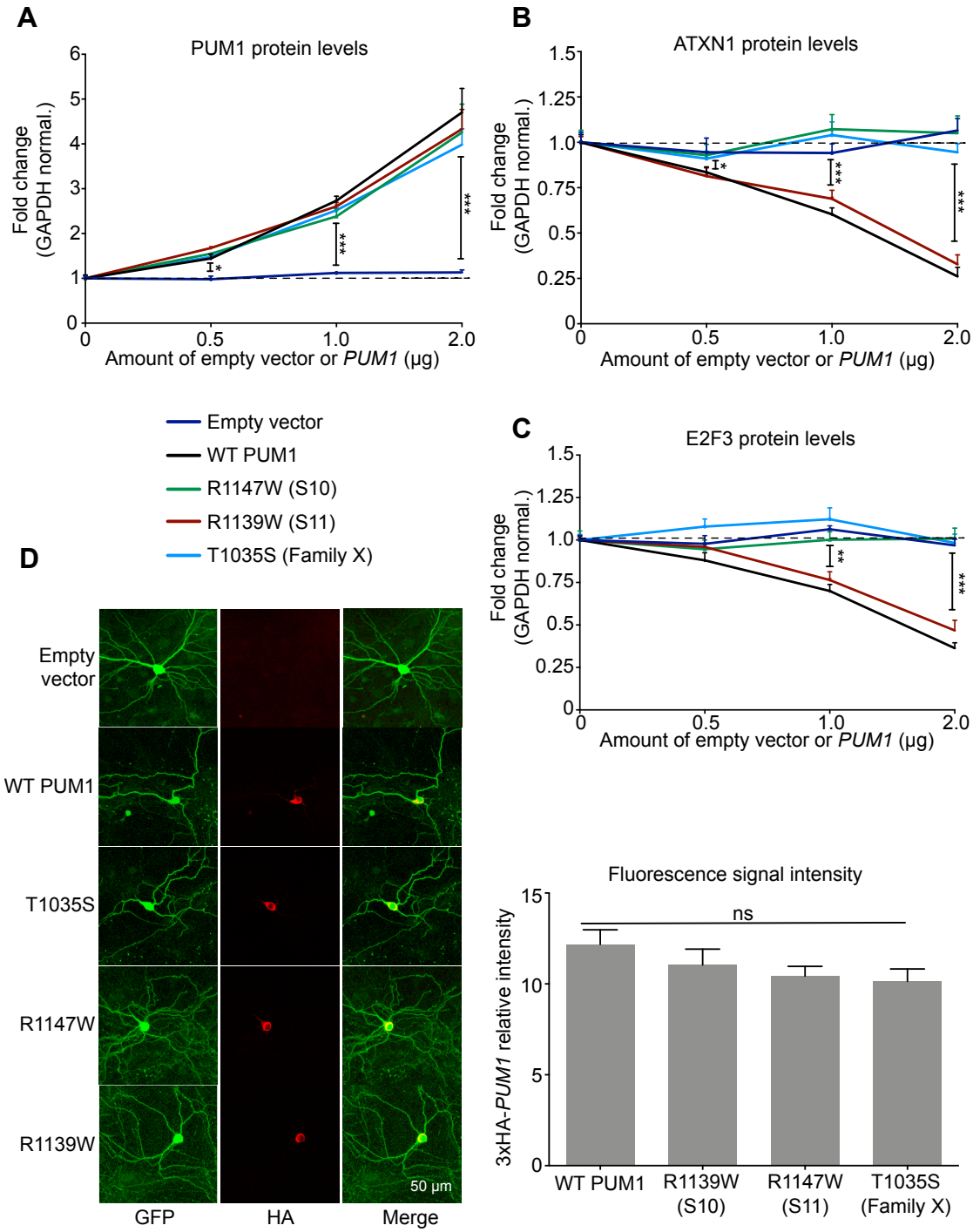


Figure S4. Western Blot and 3xHA-*PUM1* Quantification in HEK293T and Hippocampal Neurons upon Overexpression of Either WT or Mutant *PUM1*, Related to Figure 4

(A–C) Quantification of (A) PUM1, (B) ATXN1, and (C) E2F3 protein levels, see Figure 4A for corresponding western blots. All the experiments were performed in triplicate (data represent mean ± SEM); *p* values were calculated by Student's t test; *p < 0.05, **p < 0.01, ***p < 0.0001. Data were normalized to GAPDH protein levels.

(D) *Left panel:* Representative pictures of 3xHA-*PUM1* localization in hippocampal neurons. *Right panel:* Quantification of *PUM1* overexpression in hippocampal neurons shows that 3xHA-PUM1 WT and mutants were all transfected at the same level. Data represent mean ± SEM from 38-44 neurons for 3 independent animals per transfection. Statistical analysis was performed by two-way anova with Turkey's multiple comparisons test. ns, not statistically significant.