Supplemental Text:
Type II Alexander disease caused by splicing errors and aberrant overexpression of an uncharacterized GFAP isoform.

Guy Helman,[1,2,*] Asako Takanohashi,[3,*] Tracy L. Hagemann,[4] Ming D. Perng,[5] Marzena Walkiewicz,[1] Sarah Woidill,[3] Sunetra Sase,[3] Zachary Cross,[3] Yangzhu Du,[6] Ling Zhao,[6] Amy Waldman,[3] Bret C. Haake,[7] Ali Fatemi,[8] Michael Brenner,[9] Omar Sherbini,[3] Albee Messing,[4,10,**] Adeline Vanderver,[3,11,#,**] Cas Simons[1,2,#,**]

1. Murdoch Children's Research Institute, The Royal Children's Hospital, Parkville, Melbourne, Australia
2. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia
3. Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA, USA
4. Waisman Center, University of Wisconsin-Madison, Madison, WI, USA
5. Institute of Molecular Medicine, College of Life Sciences, National Tsing Hua University, Hsinchu, Taiwan
6. Human Immunology Core, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
7. Regions Hospital, Saint Paul, MN, USA
8. Moser Center for Leukodystrophies, Kennedy Krieger Institute, Johns Hopkins University, Baltimore, MD, USA
9. Department of Neurobiology, University of Alabama at Birmingham, Birmingham, AL, USA
10. Department of Comparative Biosciences, University of Wisconsin-Madison, Madison, WI, USA
11. Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

To whom correspondence should be addressed
* shared first authorship
** shared senior authorship

Correspondence to:
Adeline Vanderver: vandervera@email.chop.edu
Cas Simons: cas.simons@mcri.edu.au

Supplemental Methods:

Patient Selection

Informed consent for all participating individuals was obtained under IRB approval by the Institutional Review Board at the Children's Hospital of Philadelphia (IRB#14-011236). Protein expression studies were performed under institutional review board approvals from the University of Alabama at Birmingham, Albert Einstein College of Medicine, and the University of Wisconsin-Madison. Available patient medical records and neuroimaging were reviewed (Figure S1 and Table S1).

Plasma GFAP Screening

GFAP protein levels were tested from plasma samples isolated from patients, unaffected family members, and controls using an ultrasensitive Single Molecule Array (Simoa, also called digital ELISA, Quanterix™, #102336) GFAP immunoassay as previously described (Abdelhak, Huss, Kassubek, Tumani, & Otto, 2018). A one-way ANOVA with multiple comparisons was used for multigroup comparisons and for two group comparisons an unpaired t-test was performed and plotted using Prism software (Graphpad, Inc.).

Clinical Sequencing and Segregation Analysis

Single gene sequencing was performed in separate CLIA-approved laboratories for each family. Segregation analysis for Family 1 was performed by Sanger sequencing of available blood DNA samples on an Applied Biosystems 3730XL system (Thermo Fisher Scientific).

Mini-gene splicing assay

Mini-gene splicing assays were employed to assess the splicing impact of four *GFAP* variants predicted to impact splicing. We cloned the majority of genomic sequence of *GFAP* (Exons 1-9, Amino Acids 57-433) and introduced variants using three fragment HiFi DNA Assembly. The mutant-bearing fragments were isolated using the Monarch DNA Gel Extraction Kit (New England BioLabs, #T1020L) and introduced into a linearized EYFP-C1 vector using the NEBuilder HiFi DNA Assembly Kit (New England BioLabs, #E5520S) (see Table S2 for primers), resulting in an insert size of approximately 8.1 kb and a total vector size of 12kb.

The assembled product was transformed using NEB 5-alpha Competent *E. coli* cells and plated on an LB agar plate containing kanamycin. Plasmids for each mutant and wild-type construct

were Sanger-sequenced to confirm the presence of each variant in mutated constructs and the absence of variants in the wild-type construct (see Table S3 for primers).

Wild-type and mutant plasmids were transfected into the HEK293T cell line (Lipofectamine™ 2000 reagent (Thermo Fisher Scientific, #11668019)) and cultured for 48 hours. RNA was isolated from the transfected cells using the RNeasy Plus Mini Kit (Qiagen) and cDNA was synthesized using a High-Capacity cDNA Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. RT-PCR was performed using *GFAP* exonic primers (see Table S2 for primers). The resulting bands were purified using the Monarch DNA Gel Extraction Kit and cloned using the TOPO TA Cloning kit (Thermo Fisher Scientific, #K4500-01 before Sanger sequencing (see Table S3 for primers).

RT-PCR from Patient Brain

RNA was isolated from the brain tissue sample of a deceased member of Family 1 carrying the c.1289G>A variant. RNA was isolated using the RNeasy Lipid Tissue Mini Kit (Qiagen # 74804) and treated with DNase I. cDNA was then generated using the SuperScript III First-Strand Synthesis System (Invitrogen #18080-051). RT-PCR was performed using Q5 Hot Start High-Fidelity DNA polymerase (New England BioLabs #M0493) and gene-specific primers (Table S4) for 35 cycles. The resulting PCR products were run on E-Gel Size Select II Agarose Gel, 2% (Life Technologies # G661012), isolated and purified using Genomic DNA Clean & ConcentratorTM-10 (Zymo Research, #D4010). The final PCR products were purified with ExoSAP-IT (Life Technologies #78201.1mL). Purified PCR products were analyzed by Sanger sequencing (see Table S3 for primers).

Human brain samples

Clinical and genetic details of human brain samples used for protein studies are presented in Table S5. AxD human tissues were obtained from the NIH Neurobiobank with approval from the institutional review board at the University of Wisconsin-Madison.

Preparation of the intermediate filament-enriched fractions

Brain tissues (0.05-0.1 mg) were thawed on ice and dounce-homogenized in low salt buffer (20 mM Tris-HCl, pH 7.4, 5 mM EDTA, 1% (v/v), Triton X-100, and 150 mM KCl). All buffers were supplemented with a cocktail of protease inhibitors, containing 10 μM ALLN, 2 μg/ml leupeptin, 5 μg/ml aprotinin, and 2 mM PMSF. Homogenates were centrifuged at 14,000 rpm for 15

minutes at 4°C, and the resulting pellets were extracted in high salt buffer (20 mM Tris-HCl, pH 7.4, 5 mM EDTA, 0.5% (v/v) Triton X-100, and 1.5 M KCl). After centrifugation at 19,000 g for 15 min at 4°C, the pellets were resuspended in urea buffer (6 M urea, 10 mM Tris-HCl, pH 7.4, and 1 mM EDTA) and further extracted at 4°C overnight. The extracts were centrifuged at 19,000 g at 4°C for 5 minutes, and the urea-soluble fractions were taken as the IF-enriched fractions. After concentration determination by BCA protein assay, protein samples were resolved by 12% (w/v) SDS-PAGE using a Tris/glycine electrophoresis system prior to analysis by immunoblotting.

Immunoblotting

Immunoblotting was performed using the wet electrophoretic transfer system (Biorad, Hercules, CA) according to the manufacturer's instructions. Following electrophoretic transfer, the nitrocellulose membranes (Pall Corporation, Ann Arbor, MI) were washed three times with Tris-buffered saline (TBS: 150 mM NaCl, 20 mM Tris-HCl, pH 7.4), followed by blocking in TTBS ((0.1% (v/v) Tween-20 in TBS) containing 3% (w/v) bovine serum albumin (BSA)) for 1 h at room temperature. After blocking, the membranes were incubated at 4°C overnight with mouse anti-GFAP-α (clone 52; BD Biosciences) and rabbit anti-GFAP-ε antibodies (Roelofs et al., 2005). The specificity of these antibodies was tested by immunoblotting using recombinant human GFAP-α and GFAP-ε proteins purified as described previously (Perng et al., 2008). After being washed with TTBS, membranes were incubated with horseradish peroxidase–conjugated goat anti-mouse or anti-rabbit secondary antibodies (Jackson ImmunoResearch Laboratories) for 2 h at room temperature. All antibodies were diluted in TTBS containing 1% (w/v) BSA. Antibody labeling was detected by enhanced chemiluminescence substrate (Western Lightning Plus, PerkinElmer; Waltham, MA) with use of a luminescent image analyzer (LAS 4000; GE Healthcare).

In silico prediction of alteration in mRNA splicing

The SpliceAI algorithm was used to identify ClinVar variants predicted to alter GFAP mRNA splicing (Jaganathan et al., 2019). Pre-computed file of SpliceAI scores for all possible single nucleotide substitutions with 50nt of Gencode exons was downloaded from https://basespace.illumina.com/s/5u6ThOblecrh on April 18 2019. Vcfanno (Pedersen, Layer, & Quinlan, 2016) was used to annotate all GFAP variants classified as "Pathogenic" or Likely Pathogenic" in the ClinVar database (downloaded 18th June 2019) with annotations from SpliceAI. Variants were filtered to identify those with a SpliceAI score greater than 0.75 for any

given splicing event (Splice scores range from 0-1, where “confidently predicted cryptic splice variants” are those with a score > 0.5 (Jaganathan et al., 2019)).

Supplemental Results:

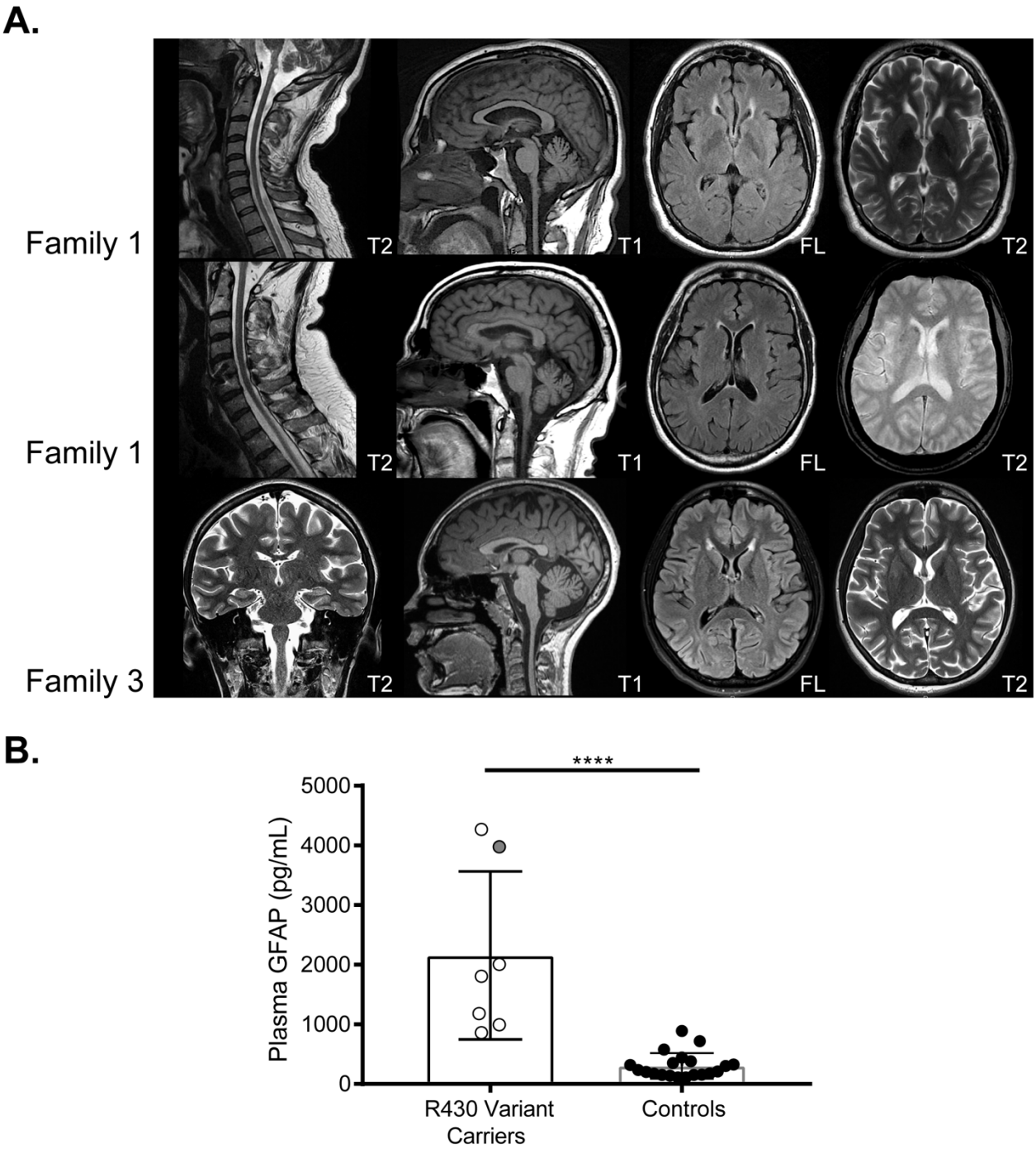

Figure S1 - MRI and plasma GFAP levels of select affected individuals. (A) MRI of selected affected individuals. Imaging of two separate affected individuals from the same nuclear family of Family 1 (Row 1 aged 48 years and Row 2 aged 62 years) shows medullary atrophy. In all affected individuals shown there are periventricular hyperintensities in the region of the frontal horns of the lateral ventricles on FLAIR and T_2 sequences. Imaging from Individual 1 in Family 3

(Row 3) was performed at 13 years of age. (B) Plasma GFAP levels were determined by digital ELISA. As a group, six individuals heterozygous for the c.1289G>A variant (annotated in white with black outline) and a single individual with the c.1290C>A variant (annotated in grey with black outline) demonstrated significantly elevated blood plasma GFAP levels relative to unaffected controls (p=0.0001).

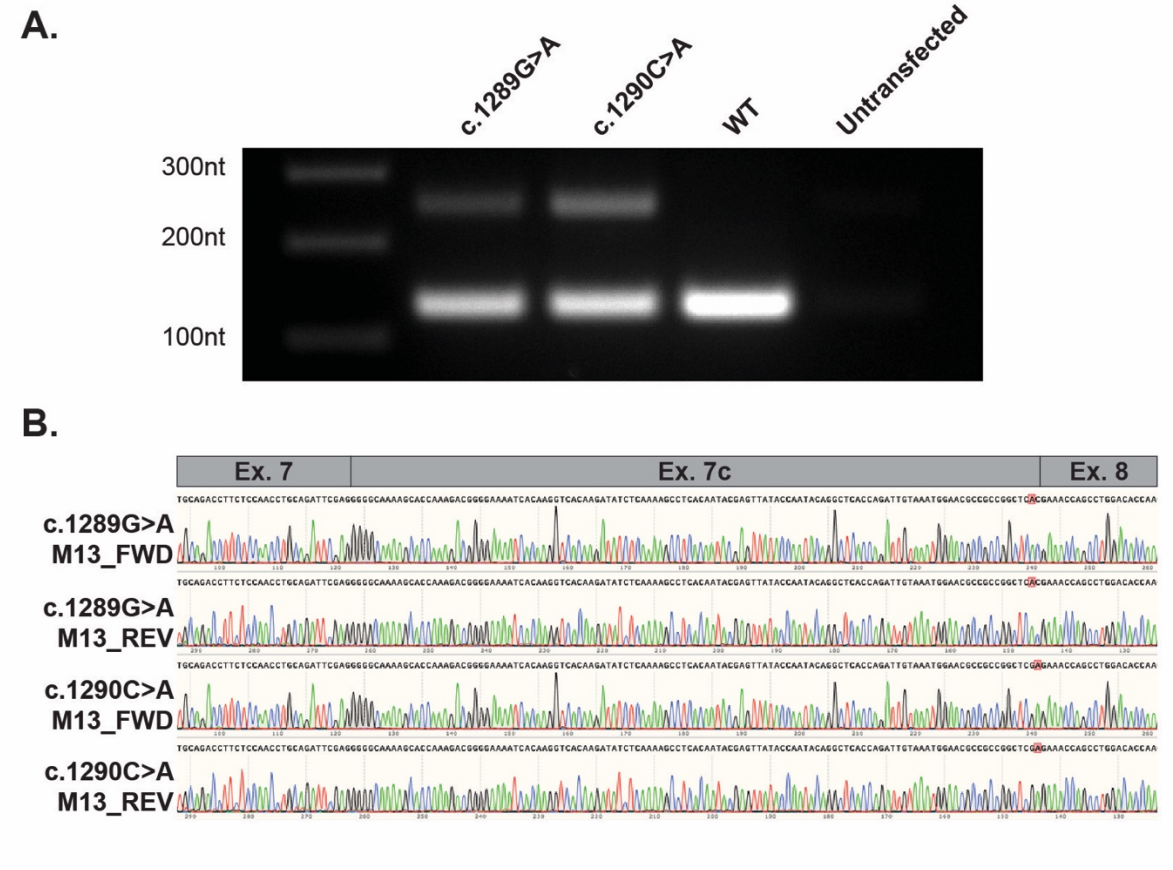

Figure S3 – **Aberrant *GFAP* splicing due to the c.1289G>A and c.1290C>A variants results in production of the GFAP-λ isoform**. (A) RT-PCR from RNA extracted from HEK cells transfected with GFAP mini-gene plasmids. An amplicon consistent with the expected size for GFAP-*α* (130 nt) was generated from the c.1289G>A (Lane 1), c.1290C>A (Lane 2) and WT plasmids (Lane 3). A larger band consistent with the expected size for GFAP-λ (252 nt) is only visible from the mutant plasmids c.1289G>A and c.1290C>A. (B) Sanger sequencing of the cloned 252 nt amplicon from each mutant plasmid demonstrates inclusion of exon 7c between exon 7 and exon 8, representative of the GFAP-λ transcript. The position of each variant is highlighted in red.

Figure S4 - Aberrant *GFAP* splicing of mini-gene constructs bearing either the c.731C>T or the c.1171+5G>A variant. (A) Amplicons shown in Figure 2E were gel extracted and Sanger sequenced. The c.731C>T variant results in a 51 nt truncation of Exon 4. The intervening sequence from Intron 4 is shown truncated at the dashed line. (B) The amplicon 360nt amplicon (WT) and the 713nt amplicon (c.1171+5G>A) shown in Figure 2D were gel extracted and Sanger sequenced. The WT product demonstrates canonical splicing between exons 7 and exon 8 while the c.1171+5G>A weakens the Exon 7 canonical donor splice site to generate the *GFAP-κ* isoform.

Table S1: Summary of clinical features

	Family 1	Family 2	Family 3		Melchionda et al.		Karp et al.
# of affected family members	> 20* in 5 generations	1	2		2		1
Variant Transcript	c.1289G>A p.Arg430His NM_001363846.1	c.1289G>A p.Arg430His NM_001363846.1	c.1290C>A p.Arg430Arg NM_001363846.1		c.1289G>A p.Arg430His NM_001363846.1		c.1289G>A p.Arg430His NM_001363846.1
Affected individuals with segregation confirmed	6/6	1/1	2/2		N/A		N/A
Age at Onset	42 Y (range: 20 – 66 Y)	45 Y	8 Y	UNK	55 Y	46 Y	32 Y
Age at Last Assessment or Report	-	59 Y^	14 Y	42 Y	68 Y	60 Y	58 Y
Cognitive involvement	N	N/A	N	N	Y	N	N
Psychiatric Symptoms	N	N	N	N	Y	N	N
Upper motor signs (e.g. spasticity)	8/16	Y	N	N/A	Y	Y	Y
Hypotonia	N/A	N/A	N	N/A	N	Y	N
Movement Disorder: tremor, dystonia or chorea	N/A	N/A	N	N/A	Y	N	N

Ataxia / gait dysfunction	Y (12/16)	Y	N	Y	Y	Y	Y
Peripheral neuropathy	N/A	N/A	N	N/A	N	Y	N
Autonomic Dysfunction	N/A	N/A	N	N/A	N	N	N
Seizures	N/A	N/A	Y	N/A	N	N	N
Bulbar dysfunction manifesting as swallowing difficulties, choking, or dysarthria	Y (7/16)	Y, dysarthria	Y, laryngospasm and dysarthria	N/A	Y, dysarthria, dysphagia, palatal myoclonus	Y, dysphagia, hypophonia	Y, Dysarthria (mild), tongue fasciculations
Other Parasthesia Nystagmus Sleep Apnea Failure to thrive Gastric dysmotility Headaches	 Y (4/16) Y (2/16) N/A N/A N/A Y (1/16)	N/A	 N N N Y Y Y	 N N Y N N N	N/A	Nocturnal lower limb myoclonus	 N Y Y N N N
Family History	Y	Y 3 affected siblings	Y	Y	N/A	N/A	Sibling with monophasic multiple sclerosis; otherwise unremarkable

Investigations							
MRI							
Cerebral T_2 Lesions	Y	N/A	Y	Y	N/A	N/A	Y
Periventricular hyperintensities	Y	N/A	Y	Y	N/A	N/A	Y
Brainstem lesion	Y, 6/6	N	Y	Y	N/A	N/A	Y
Medulla / Spinal Cord atrophy	Y, 2/6	N	Y	Y	N/A	N/A	Y
EMG	N/A	N/A	N/A	N/A	NML	Y	NML
Nerve Conduction Studies	N/A	N/A	N/A	N/A	N/A	Motor axonal neuropathy	N/A
Secondary diagnoses	N/A	N	N/A	N/A	N/A	*HDAC6* c.2566C>T p.(Pro856Ser)	N/A
Pathology studies	Rosenthal fibers	Rosenthal fibers (present in sibling autopsy)	-	-	N/A	N/A	N/A

All variants are described relative to their position using the RefSeq Transcript: NM_001363846.1* Limited clinical information is available for 16 family members while no clinical information is available for other reportedly affected family members; ^ Deceased; N/A – Not Available. Pinto e Viaro et al. (2018) reports one additional patient with the c.1289G>A variant found in *GFAP-ε / GFAP-λ.* However, insufficient clinical data were provided for inclusion in this table. This individual had positive family history for AxD and MRI lesions affecting the midbrain and atrophy of the medulla. He had a potential secondary diagnosis of hereditary hemorrhagic telangiectasia with a variant of uncertain significance in *ACVRL1*.

Table S2: Mini-gene assembly primers

Primer	Primer Sequence (5' to 3')
EYFP-C1:GFAP_ex1_fwd	gtcagatccgctagcgctaccgatgACTCAATGCTGGCTTCAAGG
EYFP-C1:GFAP_ex9_rev	gtttcaggttcaggggaggtgtgGAGGGGAGCAGCTGGGGTG
GFAP_1289G-A_fwd	GCCGGCTCaCGGTTAGCTGCCTGCCTCTC
GFAP_1289G-A_rev	GCAGCTAACCGtGAGCCGGCGGCGTTCC
GFAP_1290C-A_fwd	GCCGGCTCGaGGTTAGCTGCCTGCCTCTC
GFAP_1290C-A_rev	GCAGCTAACCtCGAGCCGGCGGCGTTCC
GFAP_WT_fwd	GCCGGCTCGCGGTTAGCTGCCTGCCTCTC
GFAP_WT_rev	GCAGCTAACCGCGAGCCGGCGGCGTTCC
GFAP_1171+5G-A_fwd	CAGATTCGAGGTCAaTACAGCAGGGCCTCG
GFAP_1171+5G-A _rev	GCCCTGCTGTATtGACCTCGAATCTGC
GFAP_731G-A_fwd	CAGTATGAGGtAATGGCGTCCAGCAAC
GFAP_731G-A_rev	CTGGACGCCATTaCCTCATACTGCGTGC

Table S3: Sanger sequencing primers

Primer	Primer Sequence (5' to 3')
GFAP_ex7_fwd	GGGCAAAAGCACCAAAG
GFAP_ex6_fwd	GAGATCGCCACCTACAGGAAGC
GFAP_ex3_fwd	GATTGAGTCGCTGGAGGAGG

Table S4: RT-PCR of Family 1 Autopsy Tissue

Primer	Primer Sequence (5' to 3')
FWD1	CCAAGCACGAAGCCAACGACTAC
REV1	CTCTCCATCCCGCATCTCC
FWD2	GATTGAGTCGCTGGAGGAGG
REV2	CCGTCTTTGGTGCTTTTGCC
FWD3	GCACGAAGCCAACGACTA
REV3	CTGTAGGTGGCGATCTCGATGTC
FWD4	ACGGGGAAAATCACAAGGTCA
REV4	TTCTCTCCTTCCTCCTCATTCT
REV5	CTATCCTGCTTCTGCTCGGG

Table S5. Clinical and genetic details of AxD control and Family 2 patient samples

Case	ID Number	Disease	Age at death	GFAP mutation	Sex
C1	#13 in Li et al. (2005)	Adult-onset/type II	38 y	Glu210Lys	F
C2	#34 in Li et al. (2005)	Juvenile-onset /type II	20 y	Leu359Val	M
C3	#4 in Li et al. (2005)	Infantile-onset/type I	10.5 y	Leu76Val	F
F2-1	#43 in Li et al. (2005)	Adult-onset/type II	59 y	Arg430His/Arg 430Cys	M

References:

Abdelhak, A., Huss, A., Kassubek, J., Tumani, H., & Otto, M. (2018). Serum GFAP as a biomarker for disease severity in multiple sclerosis. *Scientific Reports, 8*(1), 14798. doi:10.1038/s41598-018-33158-8

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., . . . Farh, K. K.-H. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell, 176*(3), 535-548.e524. doi:https://doi.org/10.1016/j.cell.2018.12.015

Li, R., Johnson, A. B., Salomons, G., Goldman, J. E., Naidu, S., Quinlan, R., . . . Brenner, M. (2005). Glial fibrillary acidic protein mutations in infantile, juvenile, and adult forms of Alexander disease. *Ann Neurol, 57*(3), 310-326. doi:10.1002/ana.20406

Pedersen, B. S., Layer, R. M., & Quinlan, A. R. (2016). Vcfanno: fast, flexible annotation of genetic variants. *Genome Biol, 17*(1), 118. doi:10.1186/s13059-016-0973-5

Perng, M.-D., Wen, S.-F., Gibbon, T., Middeldorp, J., Sluijs, J., Hol, E. M., . . . Pollard, T. D. (2008). Glial Fibrillary Acidic Protein Filaments Can Tolerate the Incorporation of Assembly-compromised GFAP-δ, but with Consequences for Filament Organization and αB-Crystallin Association. *Molecular Biology of the Cell, 19*(10), 4521-4533. doi:10.1091/mbc.e08-03-0284

Roelofs, R. F., Fischer, D. F., Houtman, S. H., Sluijs, J. A., Van Haren, W., Van Leeuwen, F. W., & Hol, E. M. (2005). Adult human subventricular, subgranular, and subpial zones contain astrocytes with a specialized intermediate filament cytoskeleton. *Glia, 52*(4), 289-300. doi:10.1002/glia.20243