

## Materials and Methods

### *Electrophysiological studies*

The protocol included nerve conduction studies, F-waves, EMG, autonomic nervous system test and SSEPs. Motor nerve conduction studies included bilateral deep peroneal and posterior tibial nerves and sensory nerve studies included bilateral sural, superficial peroneal and median (digit II) and ulnar (digit V) sensory nerves. F waves were studied for median, tibial and deep peroneal nerves. EMG study was also performed from anterior tibial muscle. Sympathetic skin response (SSR) and the R-R interval were determined to assess sympathetic and parasympathetic autonomic nervous system functions, respectively. Bilateral tibial nerve SSEPs were recorded on popliteal fossa, lower spine and scalp (Cz'/Fz according to 10-20 International System). Reference control values were obtained from Kimura and collaborators.<sup>1</sup>

### *Genetic subtyping of known SCA genes*

Genomic DNAs of IV:2 and V:1 probands were tested for pathogenic DNA repeat expansions within the SCA1, SCA2, SCA3, SCA6, SCA7, SCA8, SCA10, SCA12, SCA17, SCA36, SCA37 and DRPLA associated genes. Mutations in *CACNA1A* were also ruled out by Sanger sequencing in IV:14 affected patient referring recurrent vertigo episodes.

### *Genome-wide linkage analysis*

In the age-dependent penetrance model, liability classes of penetrance were generated assuming a normal distribution of the age at onset in affected individuals, estimating a mean age at onset of  $\mu = 42.56$  years and a standard deviation of  $\sigma = 18.01$ . A normal density function was generated according to the normalized variable corresponding values ( $Z$ ) from our data of age at onset ( $X$ ) from the  $N = 9$  affected individuals of the pedigree ( $Z = (X - \mu) / \sigma$ ). The penetrance value for each individual of the pedigree was calculated as the probability to be affected given a heterozygous (+/m) or homozygous (m/m) genotype (m = mutant allele; + = wild-type allele) at the respective actual age, according to the normal distribution function from the normalized data. Using the tables of probability of the normal distribution function provided,<sup>2</sup> we assigned the probability to be affected for each of the five healthy and at-risk individuals of the pedigree (absence of cerebellar atrophy on MRI, nystagmus or hyperreflexia since these are the initial signs of the disease revealed in all affected patients) to five liability classes (LC) corresponding to their actual age as follows: LC1 = 0.95 (IV:3, 72 years); LC2 = 0.82 (IV:10, 59 years); LC3 = 0.80 (IV:12, 58 years); LC4 = 0.62 (IV:4, 48 years); LC5 = 0.55 (V:2, 45 years).

### *Whole-exome sequencing (WES)*

Paired-end, 100 bp long reads were generated on a HiSeq2000 platform (Illumina, Inc.). Reads alignments were performed using GEM<sup>3</sup> to Hs37d5 genome, allowing for up to four

mismatches when necessary, followed by local-realignment with GATK 3.1-1.<sup>4</sup> Variant calling was performed using Samtools 0.1.19.<sup>5</sup> Identified variants were annotated using SNPEff,<sup>6</sup> Annovar<sup>7</sup> and dbNSFP<sup>8</sup> where appropriate. Variants with a minor allele frequency (MAF)  $\geq 1\%$  in genetic databases including the Exome Aggregation Consortium,<sup>9</sup> 1,000-Genomes Project, dbSNP<sup>10</sup> and gnomAD<sup>11</sup> were excluded for further analysis. The genetic heterogeneity model<sup>12</sup> was applied in order to identify potential candidate genes. Candidate variants were assessed for their computationally predicted pathogenicity by SIFT,<sup>13</sup> PolyPhen-2,<sup>14</sup> Condel,<sup>15</sup> FATHMM,<sup>16</sup> MutationAssessor,<sup>17</sup> MutationTaster,<sup>18</sup> Provean,<sup>19</sup> and CADD,<sup>20</sup> and prioritized in accordance with the clinical characteristics shared by affected individuals. Segregation of candidate variant was confirmed by Sanger sequencing.

#### *Fibroblasts and neuroblastoma cell lines cultures*

Three mm-skin punch biopsies from patients IV:14, V:1 and two unrelated age and gender matched controls were collected in Hank's Balanced Salt Solution (HBBS; (Gibco) supplemented with 1 mM sodium pyruvate (Thermo Fisher Scientific), 10 mM HEPES (Thermo Fisher Scientific) and penicillin-streptomycin (Thermo Fisher Scientific). Fibroblasts and SH-SY5Y neuroblastoma cells were grown in Easyflask T75 with filter (Thermo Fisher Scientific) with Dulbecco's modified Eagle medium (Thermo Fisher Scientific) supplemented with 10% of Fetal bovine serum (Thermo Fisher Scientific), and 1% of penicillin-streptomycin-glutamine (Thermo Fisher Scientific).

#### *SDS-PAGE and immunoblotting*

Proteins were extracted from human fibroblasts samples by homogenization in RIPA lysis buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% Na Deoxycholate, 0.1% SDS, 1% Triton, 50 mM NaF, 5 mM NaVO<sub>3</sub>, 2 mM EDTA and protein inhibitor cocktail (Roche). Protein concentrations were determined using the DC-BioRad protein assay (BioRad). Fifty micrograms of protein sample were mixed with 4X protein loading buffer (125 mM Tris-HCl pH 6.8, 50% glycerol, 4% SDS and 0.2% (w/v) Orange G; LICOR) with 10% DTT, separated by electrophoresis at constant 20 mA, transferred onto PVDF membranes with a trans-blot turbo transfer system (BioRad) at 25 V (1 A) for 30 min and blocked with casein for 60 min.

Mitochondrial fractions were obtained as previously reported with minor modifications.<sup>21</sup> Briefly, fibroblast cells and whole zebrafish larvae were homogenized in hypotonic lysis buffer containing 320 mM sucrose, 10 mM Hepes, 10 mM KCL, 1 mM EGTA, 2 mM EDTA, 5 mM MgCl, 25 mM NaF, 2.5 mM NaVO<sub>3</sub> and protein inhibitor cocktail (Roche). Homogenized samples were incubated on ice for 15 min and centrifuged at low speed (500 g) for 10 min at 4 °C to separate the cytoplasm containing mitochondria, and other organelles in the supernatant from the pellet containing membranes and nuclear fraction contained in the nuclei. Supernatants were centrifuged at 3,000 g at 4 °C to obtain a mitochondria enriched fraction (pellet). Triton was added to a final 1% concentration to solubilize cytoplasmic and mitochondrial fractions. Protein concentration was determined (using the DC-BioRad protein assay) as stated above and 10 µg of protein from each sample were prepared for Western Blot analysis as described in the main text.

*Human SAMD9L expression*

Using RNeasy Mini Kit (Qiagen). RNA integrity was measured using RNA Screen Tape assay (Agilent Technologies). cDNAs were synthesised from 1 µg total RNA with PrimeScript RT reagent Kit (Takara Bio Inc.). To quantify the expression levels, cDNAs were mixed with SYBR® Green PCR master mix (Thermo Fisher Scientific) and specific primers (Supplementary Table 1) for a partial product which includes partial exons 4 and 5 to avoid amplification of DNA contamination. GAPDH was amplified as internal control. Cycles and analysis were performed on the LightCycler 480 (Roche) and relative cDNAs fold changes were normalised to GAPDH cDNA and calculated using the  $2^{-\Delta\Delta CT}$  method.<sup>22</sup>

*Generation of SAMD9L DNA plasmids for SAMD9L zebrafish injections*

The SAMD9L cDNA coding region was amplified by PCR using PfuI polymerase (Promega) from patient's white blood cells cDNA as template. The resulting fragment was then in-frame subcloned into the pCS2<sup>+</sup> expression vector downstream of an SP6 promoter containing a 3'polyA signal.<sup>23</sup> The resulting pCS2-SAMD9LWT construct included 5'-CCGCCACC-3' Kozak sequence previous to cDNA SAMD9L constitutional ATG. To introduce the c.1877C>T (p.Ser626Leu) point mutation, in vitro mutagenesis was performed using PCR and specific primers (Supplementary Table 1). DNA sequences of pCS2-SAMD9LWT and pCS2-SAMD9L.S626L were confirmed by Sanger DNA sequencing. Plasmid DNAs were amplified by transformation of Escherichia coli HST08 competent cells, digested with NotI restriction enzyme, synthesized using mMessage mMachine SP6 Transcription Kit (Thermo Fisher Scientific) following manufacturer's instructions, purified, and resuspended in RNase-free water.

*Cell transfection for plasmid expression validation*

Both pCS2-SAMD9LWT and pCS2-SAMD9L.S626L DNA constructs were transfected using lipofectamine into SH-SY5Y cells according to manufacturer's instructions (Thermo Fisher Scientific) to test SAMD9L expression. Transfected cells were incubated with the vector and lipofectamine mixture for 4 h in OPTIMEM (Thermo Fisher Scientific) and growth for 48 h in Neurobasal medium (Thermo Fisher Scientific) with B27 supplement and penicillin-streptomycin either on 6-well plate or flask.

*Zebrafish ataxia model*

Adult wild-type heterozygous zebrafish (Danio rerio; AB background) were maintained in the animal facility at 27 °C on a light cycle of 12 h light / 12 h dark. Zebrafish embryos obtained in a cross between wild-type AB individuals were collected in E3 1x medium Petri dishes. Fertilized embryos were microinjected into one-cell stage embryos with 1 nl of wild-type or mutant human SAMD9L mRNA at 25, 50 and 100 ng/µl. Embryos were grown at 28 °C and 2 h post injections non-fertilized eggs were discarded. 24- and 96-hours post-fertilization (hpf) survival and teratogenesis were scored.

*hSAMD9L cDNA levels in microinjected zebrafish larvae*

qRT-PCR was carried out using five larvae of 5 dpf for each microinjected group to quantify the total of human SAMD9L mRNA. RNA was extracted using RNeasy Mini Kit (Qiagen) following manufacturer instructions. A 1 µg portion of total RNA was reverse transcribed (PrimeScript RT Reagent kit, Takara) and PCR was carried out for human SAMD9L mRNA and zebrafish *tbp* targets using the oligonucleotide primers specified in Supplementary Table 1. Three replicates for each group were performed. The Shapiro-Wilk test was used to assess for normal distribution followed by the parametric t-test.

*Mitochondrial DNA damage quantification*

Sequencing of the mtDNA genome from fibroblasts samples from IV:14 and V:1 M-SCA affected patients and two age-matched controls, were performed using the Illumina MiSeq platform (Illumina, Inc.). Enrichment of the entire mtDNA was performed by nine long-range PCRs (Supplementary Table 2) using LA Taq DNA polymerase with GC Buffer I (Takara Bio Inc.) and 100 ng of total genomic DNA. Library preparation was performed by the Illumina Nextera XT kit (Illumina, Inc.) according to the manufacturer's instructions. Four indexed DNA libraries were equimolarly pooled and sequenced in a single lane of 1 MiSeq flow-cell using paired 150 nt reads. Reads were mapped to the revised Cambridge Reference Sequence (rCRS; GenBank ID NC\_012920.1) using the MiSeq Reporter (Illumina, Inc.), which uses a Burrows-Wheeler Aligner (BWA)<sup>24</sup> and the Genome Analysis Tool Kit (GATK)<sup>4</sup> for variant calling of single nucleotide polymorphisms (SNPs) and short indels. Heteroplasmy levels were calculated by the read depth of variations vs. reference nucleotides for all nucleotide positions. Quantification of mtDNA copy number and depletion were performed by qPCR using primers for two short D-LOOP and MT-CO3 mtDNA regions and a nuclear GAPDH gene region (Supplementary Table 2) in combination with TB Green Premix Ex Taq II (Takara Bio Inc.). Mitochondria DNA lesions were quantified using an adapted long-run qPCR for DNA-damage quantification (LORD-Q) method previously described, which calculates DNA-damage rates per 10 kb using amplification efficiency from the standard curve based on Ct values.<sup>25</sup> In the adapted version, TB Green Premix Ex Taq II (Takara Bio Inc.) reagent was used for the amplification of the long 3,723 bp mtDNA instead of the combination of an independent fluorescent dye with a Hot Start Taq.

*ATP determination*

The ATP content of mitochondrial fractions from two patient's fibroblasts (IV:14 and V:1) and two age-matched controls were determined using four replicates for each sample and the ATPlite Luminescence ATP Detection Assay System (Perkin Elmer) according to manufacturer's instructions. The luminescent intensity was measured in quadruplicate wells with the microplate reader Varioskan® Flash (Thermo Fisher Scientific).

*Immunohistochemistry and immunofluorescence studies*

For immunohistochemistry studies, human cerebellar sections from a 31-year-old unaffected male were processed following standard protocols. For tissue immunocytochemistry, sections were deparaffinised with Clearene (Leica Biosystems), hydrated in 100% and 96% ethanol, and ddH<sub>2</sub>O and then immersed in 3% hydrogen peroxide in methanol for 30 min to block endogenous peroxide activity. Next, slides were rinsed and microwaved in Lab Vision™ citrate buffer for heat-induced epitope retrieval (pH 6.0, 10 min) (Thermo Fisher Scientific), cooled down at room temperature (RT), and rinsed in phosphate buffered saline (PBS) with 0.1% Tween. Non-specific binding was blocked for 1 h with SuperBlock Blocking Buffer (Thermo Fisher Scientific) and incubated with anti-SAMD9L protein primary antibody (Reference: 25173-1-AP; Protein Tech) and anti-ATP5B (ab14730; abcam). They were then rinsed in 0.1% Tween in phosphate-buffered saline (PBS) before incubation with the secondary Goat anti-Rabbit IgG (H+L) Superclonal antibody (Thermo Fisher Scientific). Slides were covered with a DAB Enhancer (Dako) and counterstained with haematoxylin/eosin before dehydration, clearing and mounting in DPX (Merck).

For immunofluorescence, cultured cells on coverslips were quantified, washed and permeabilised after which they were incubated overnight at 4 °C with primary antibodies. Following a washing step, coverslips were incubated with secondary antibodies (3% BSA in 1x PBS) for 45 min at RT. Finally, a 10 min incubation with Hoechst was performed for nuclear staining. For staining of mitochondria with 100 nm MitoTracker™ Red CMXRos (Thermo Fisher Scientific) was diluted in DMEM (Dulbecco's Modified Eagle Medium; Thermo Fisher Scientific) prior to addition to cells for 45 min at 37 °C in 5% CO<sub>2</sub> atmosphere. After the incubation period, cells were washed three times in cell culture media prior to fixation in 4% paraformaldehyde/PBS at room temperature for 20 min and permeabilized and blocked in 0.2% Triton X-100/PBS for 1 h before primary antibody incubation when applicable. Visualisation was performed using a Carl Zeiss Axio Scope 2 or an Axio Observer Z1 microscope coupled with a LSM710 ZEN confocal module (Zeiss) and processed using ImageJ<sup>26</sup> and analysed using JACOP and Coloc2 plugins. The mitochondrial network was analysed with Image J Plugin MiNA software.<sup>27</sup>

*Immunohistochemistry*

For tissue immunocytochemistry, sections were deparaffinised with Clearene (Leica Biosystems), hydrated in 100% and 96% ethanol, and ddH<sub>2</sub>O and then immersed in 3% hydrogen peroxide in methanol for 30 min to block endogenous peroxide activity. Next, slides were rinsed and microwaved in Lab Vision™ citrate buffer for heat-induced epitope retrieval (pH 6.0, 10 min) (Thermo Fisher Scientific), cooled down at room temperature (RT), and rinsed in phosphate buffered saline (PBS) with 0.1% Tween. Non-specific binding was blocked for 1 h with SuperBlock Blocking Buffer (Thermo Fisher Scientific).

*Whole-mount larvae fluorescent immunostaining*

Larvae were fixed with 4% PFA at 4 °C overnight. For whole mount immunostaining fixed larvae were equilibrated in 1 × PBS-Tw (PBS containing 0.1% Tween 20) for 10 min three times, and then were incubated in 150 mM Tris-HCl at pH 9.0 for 5 min, followed by

heating at 70 °C for 15 min. Larvae were directly washed after the heating treatment in 1 × PBS-Tw for 2 × 10 min and then washed in dH<sub>2</sub>O for 2 × 5 min. Subsequently to enhance tissue permeabilization, for whole mount immunostaining all larvae were treated with acetone for 20 min at – 20 °C. After six sequential washes in 1 × PBS-Tw (5 min each), larvae were then incubated in the blocking buffer (SuperBlock Blocking Buffer; Thermo Fisher Scientific) for 3 h at 4 °C. After blocking, the larvae were incubated with the respective primary antibodies in the Lab Vision™ OP Quanto incubation buffer (Thermo Fisher Scientific) at 4 °C for three days in agitation. To remove residual primary antibody, larvae were sequentially washed with PBS-Tw for 3 × 1 h. Larvae were incubated with secondary antibodies and Hoechst 33342 (1:10,000; Thermo Fisher Scientific) for two and half days, and then washed with PBS-Tw for 3 × 1 h and dH<sub>2</sub>O for 1 h. Larvae were embedded in 1% low-melting agarose on a 35 mm imaging dish (Ibidi), were visualized using a Carl Zeiss Axio Scope 2 or an Axio Observer Z1 microscope coupled with a LSM710 ZEN confocal module (Zeiss), processed with ImageJ,<sup>26</sup> and analysed using JACOP and Coloc2 plugins.

## Results

### *Genomic studies*

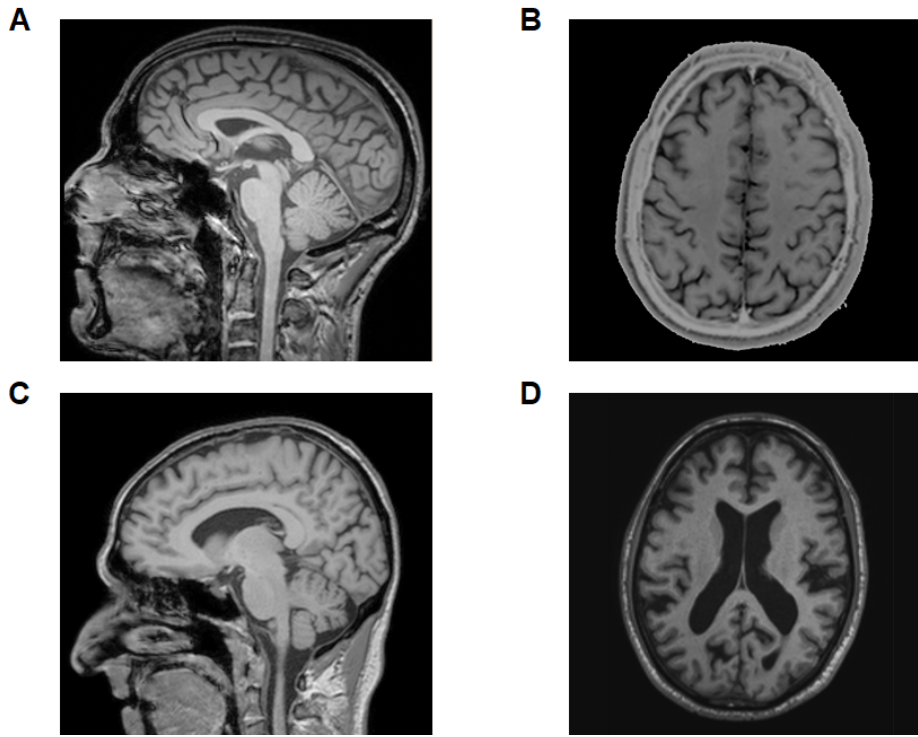
Two point and multipoint genetic linkage analyses were performed with MLINK and LINKMAP modules (version 5.10), from the FASTLINK (version 4.1P) of the LINKAGE program.<sup>28,29</sup> We assumed a prevalence of 1.5:1,000,000 (1/10 of the global ataxia prevalence) for this SCA phenotype, resulting in a genetic frequency of 7.5:10,000,000. Penetrance values for each individual of the pedigree were calculated as the probability of being affected, given a heterozygote or homozygote genotype at the individual's current age considering that the disease could initiate either with cerebellar atrophy on MRI, nystagmus or hyperreflexia as the initial signs. In the 2-point linkage analysis the proper SAMD9L gene putative pathogenic variant c.1877C>T (p.Ser626Leu) was used which was absent in 80 individuals from the Menorca population (160 chromosomes). Four identified shared candidate variants in the *SLC9B1*, *RPI*, *FRG1B* genes located in regions with positive, but non-significant, linkage values ( $\geq 2$  and  $\leq 3$ ) for the initial genome-wide linkage analysis were excluded by segregation analysis. For multipoint linkage analysis, we used the framework map from Marshfield.<sup>30</sup> The Haldane and Kosambi map functions were used to transform the genetic map distances into recombination fractions.<sup>31</sup> As depicted in the Marshfield map framework, the SAMD9L gene is located between two SNPs: 0.001 cM from rs2374628 and 0.011 cM from rs42501. We obtained the multipoint LOD scores at each locus taking the genetic marker framework map as fixed (Marshfield) and moving the disease locus successively at each point of the map. We estimated the marker allele frequencies from our general Spanish population controls as follows (number of allele, frequency): rs17393952 (1, 0.82; 2, 0.18), rs4455763 (1, 0.70; 2, 0.30); rs10263800 (1, 0.88; 2, 0.12). A multi-point LOD score analysis with the three SNPs (rs17393952, rs4455763 and rs10263800) and the candidate variant in the 7q21 candidate region, resulted in a maximum significant multipoint LOD score  $Z_{\max}$  of 3.46 ( $P < 3.28 \times 10^{-5}$ ). Furthermore, linkage analysis using an age-dependent penetrance model with five multiple liability classes also resulted in significant two-point ( $Z_{\max} = 3.10$ ;  $\theta = 0.00$ ;

$P < 7.89 \times 10^{-5}$ ) and multi-point ( $Z_{\max} = 3.12$ ;  $\theta = 0.00$ ;  $P < 7.51 \times 10^{-5}$ ) LOD scores using the same markers.

*Mitochondrial and lysosome dysregulation in patients' fibroblasts*

Mitotracker Red CMXRos fluorescent images obtained from affected ( $n = 12$ ) and control fibroblasts ( $n = 12$ ) were analysed using the ImageJ macro tool Mitochondrial Network Analysis (MiNA) toolset.<sup>27</sup> Branch length means, the sum of all branch lengths divided by the number of independent skeletons, network branches mean and number of donuts were compared between affected and control fibroblasts. No significant differences between affected and control fibroblasts regarding branch length means ( $F_{(1,20)} = 0.048$ ,  $P = 0.830$ ), summed branch lengths ( $F_{(1,20)} = 0.039$ ,  $P = 0.845$ ), network branches ( $F_{(1,20)} = 1.331$ ,  $P = 0.262$ ) or number of donuts ( $F_{(1,20)} = 1.354$ ,  $P = 0.258$ ).

**Supplementary Figure 1. Age and gender matched control MRI images.** Sagittal and transverse T1-weighted magnetic resonance imaging scan from a control (**A** and **B**) age- and gender-matched with patient V:1 and from a control (**C** and **D**) age- and gender-matched with patient IV:2.





**Supplementary Figure 2. Volumetry assessment demonstrates decreased cerebellar volume in affected patients.** CERES volumetry assessment report for patient IV:8 indicates decreased cerebellar volume relative to the total intracranial volume compared to age-and gender matched control ranges. Values in red box were used for statistical analysis.

## CERES Volumetry Report. version 1.0 release 20-06-2016

Patient ID	Sex	Age	Report Date	
job91939	Female	62	16-Jul-2018	
Image Information				
Orientation		radiological		
Scale factor		0.82		
SNR		36.85		
Total intracranial volume (cm <sup>3</sup> )		1381.44		
Volumes	Total (cm <sup>3</sup> /%)	Right (cm <sup>3</sup> /%)	Left (cm <sup>3</sup> /%)	Asym.(%)
Cerebellum	91.73 (6.6400) [7.8646, 10.6512]	45.26 (3.2765) [3.9011, 5.3197]	46.46 (3.3635) [3.9545, 5.3405]	-2.6182 [-4.2786, 2.6357]
Lobule I-II	0.05 (0.0037) [0.0034, 0.0144]	0.02 (0.0012) [0.0015, 0.0069]	0.04 (0.0026) [0.0016, 0.0078]	-73.0159 [-49.1292, 25.9580]
Lobule III	1.62 (0.1171) [0.0593, 0.1420]	0.89 (0.0648) [0.0280, 0.0716]	0.72 (0.0523) [0.0294, 0.0724]	21.2982 [-26.7391, 21.7160]
Lobule IV	2.47 (0.1787) [0.2337, 0.4186]	1.18 (0.0856) [0.1149, 0.2159]	1.29 (0.0931) [0.1094, 0.2122]	-8.4413 [-22.7758, 28.5291]
Lobule V	4.94 (0.3578) [0.2337, 0.4186]	2.67 (0.1936) [0.1149, 0.2159]	2.27 (0.1643) [0.1094, 0.2122]	16.3597 [-22.7758, 28.5291]
Lobule VI	11.54 (0.8351) [0.9758, 1.6099]	5.25 (0.3797) [0.4738, 0.8027]	6.29 (0.4554) [0.4851, 0.8241]	-18.1301 [-18.6962, 13.5853]
Lobule Crus I	15.22 (1.1017) [1.4151, 2.3857]	7.28 (0.5271) [0.6928, 1.1905]	7.94 (0.5747) [0.7078, 1.2096]	-8.6441 [-14.2757, 10.5193]
Lobule Crus II	11.15 (0.8071) [0.8482, 1.5048]	6.25 (0.4527) [0.4107, 0.7498]	4.90 (0.3544) [0.4199, 0.7726]	24.3637 [-21.2144, 15.5625]
Lobule VIIIB	6.41 (0.4639) [0.5015, 0.8592]	3.24 (0.2343) [0.2407, 0.4308]	3.17 (0.2296) [0.2471, 0.4422]	2.0476 [-23.4386, 18.2994]
Lobule VIIIA	9.51 (0.6882) [0.6283, 1.0487]	4.58 (0.3312) [0.3047, 0.5431]	4.93 (0.3570) [0.3035, 0.5257]	-7.5058 [-19.6200, 23.6346]
Lobule VIIIB	5.28 (0.3819) [0.4338, 0.7282]	2.37 (0.1714) [0.2134, 0.3752]	2.91 (0.2105) [0.2051, 0.3683]	-20.5224 [-21.1875, 26.2162]
Lobule IX	6.68 (0.4837) [0.3518, 0.7145]	3.44 (0.2493) [0.1680, 0.3540]	3.24 (0.2344) [0.1815, 0.3629]	6.1380 [-16.1633, 7.2567]
Lobule X	0.93 (0.0673) [0.3518, 0.7145]	0.50 (0.0362) [0.1680, 0.3540]	0.43 (0.0311) [0.1815, 0.3629]	15.0044 [-16.1633, 7.2567]

\*All the volumes are presented in absolute value (measured in cm<sup>3</sup>) and in relative value (measured in relation to the ICV).

\*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

\*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

\*Result images located in the MNI space (neurological orientation).

**Supplementary Figure 3. Grey matter assessment revealed decreased cerebellar grey matter relative to the total brain grey matter.** CERES grey matter assessment report for patient IV:8 indicates decreased cerebellar grey matter relative to the total brain grey matter compared to age-and gender matched control ranges. Values in red box were used for statistical analysis (Supplementary Table 1).

Grey matter vol.	Total ( $\text{cm}^3/\%$ )	Right ( $\text{cm}^3/\%$ )	Left ( $\text{cm}^3/\%$ )	Asym.(%)
<i>Cerebellum</i>	62.05 (4.4915) [5.6588, 7.9533]	31.35 (2.2691) [2.8216, 3.9756]	30.70 (2.2224) [2.8268, 3.9880]	2.0833 [-4.7413, 4.1877]
<i>Lobule I-II</i>	0.02 (0.0017) [0.0019, 0.0087]	0.01 (0.0008) [0.0007, 0.0041]	0.01 (0.0009) [0.0009, 0.0050]	-25.2248 [-98.5493, 39.7842]
<i>Lobule III</i>	1.04 (0.0749) [0.0428, 0.1059]	0.60 (0.0432) [0.0205, 0.0540]	0.44 (0.0318) [0.0207, 0.0536]	37.0976 [-37.5830, 37.4351]
<i>Lobule IV</i>	2.04 (0.1478) [0.1999, 0.3641]	0.98 (0.0711) [0.0981, 0.1894]	1.06 (0.0767) [0.0931, 0.1835]	-9.3069 [-32.6071, 43.3704]
<i>Lobule V</i>	3.88 (0.2808) [0.1999, 0.3641]	2.10 (0.1518) [0.0981, 0.1894]	1.78 (0.1290) [0.0931, 0.1835]	19.8516 [-32.6071, 43.3704]
<i>Lobule VI</i>	9.58 (0.6933) [0.8539, 1.4318]	4.37 (0.3166) [0.4172, 0.7182]	5.20 (0.3767) [0.4210, 0.7292]	-21.1535 [-25.1724, 21.6277]
<i>Lobule Crus I</i>	11.62 (0.8413) [1.0801, 1.9232]	5.89 (0.4267) [0.5321, 0.9620]	5.73 (0.4146) [0.5327, 0.9766]	3.5105 [-22.8176, 19.8920]
<i>Lobule Crus II</i>	9.13 (0.6611) [0.6735, 1.2395]	5.18 (0.3749) [0.3287, 0.6174]	3.95 (0.2862) [0.3297, 0.6372]	32.7382 [-30.6785, 24.2944]
<i>Lobule VIIIB</i>	5.73 (0.4147) [0.4132, 0.7306]	2.86 (0.2070) [0.1960, 0.3624]	2.87 (0.2076) [0.2057, 0.3798]	-0.3491 [-36.9189, 23.6035]
<i>Lobule VIIIA</i>	8.49 (0.6147) [0.5285, 0.9021]	4.13 (0.2989) [0.2591, 0.4668]	4.36 (0.3158) [0.2534, 0.4512]	-6.7125 [-25.9555, 33.9477]
<i>Lobule VIIIB</i>	4.14 (0.2999) [0.3644, 0.6328]	1.99 (0.1440) [0.1793, 0.3268]	2.15 (0.1559) [0.1714, 0.3198]	-9.7034 [-31.2122, 39.3605]
<i>Lobule IX</i>	4.69 (0.3396) [0.2779, 0.5662]	2.45 (0.1772) [0.1332, 0.2787]	2.24 (0.1624) [0.1420, 0.2903]	10.6166 [-26.6666, 12.3970]
<i>Lobule X</i>	0.89 (0.0644) [0.2779, 0.5662]	0.48 (0.0349) [0.1332, 0.2787]	0.41 (0.0295) [0.1420, 0.2903]	20.2450 [-26.6666, 12.3970]

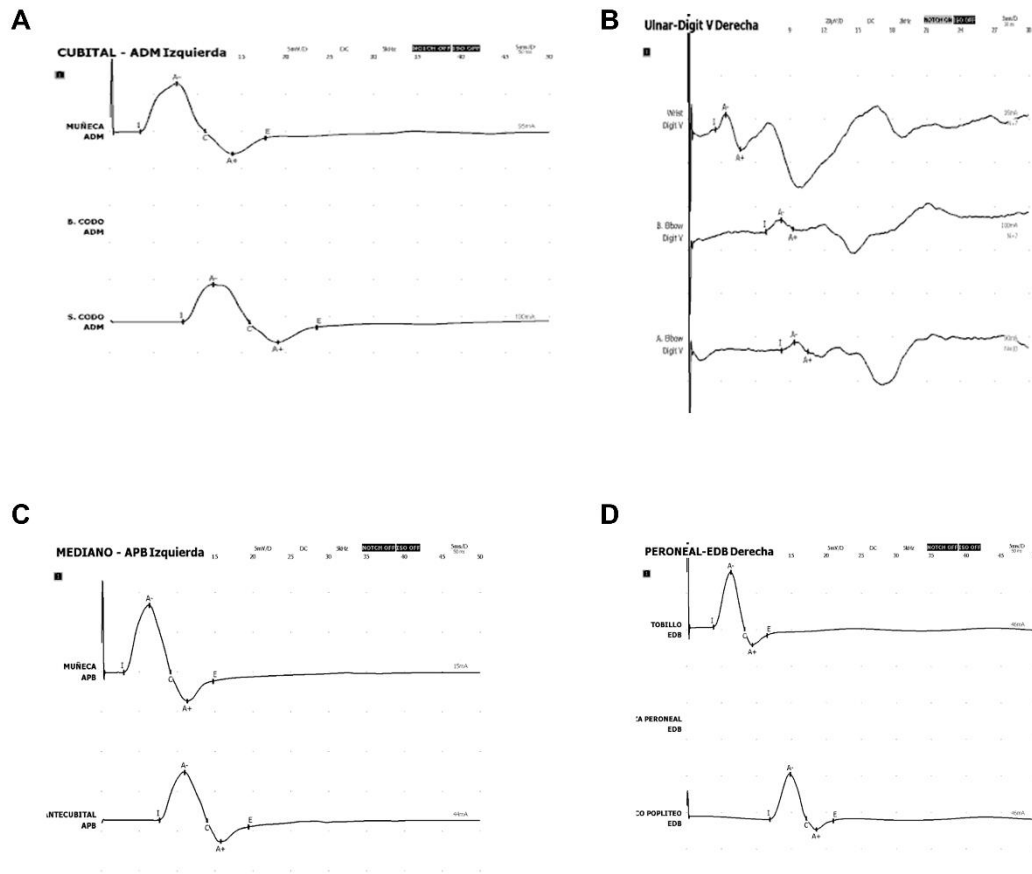
\*All the volumes are presented in absolute value (measured in  $\text{cm}^3$ ) and in relative value (measured in relation to the ICV).

\*The Asymmetry Index is calculated as the difference between right and left volumes divided by their mean (in percent).

\*Cortical thickness is given in absolute value (mm) and also normalized in relation to the cube root of the intracranial volume (adimensional).

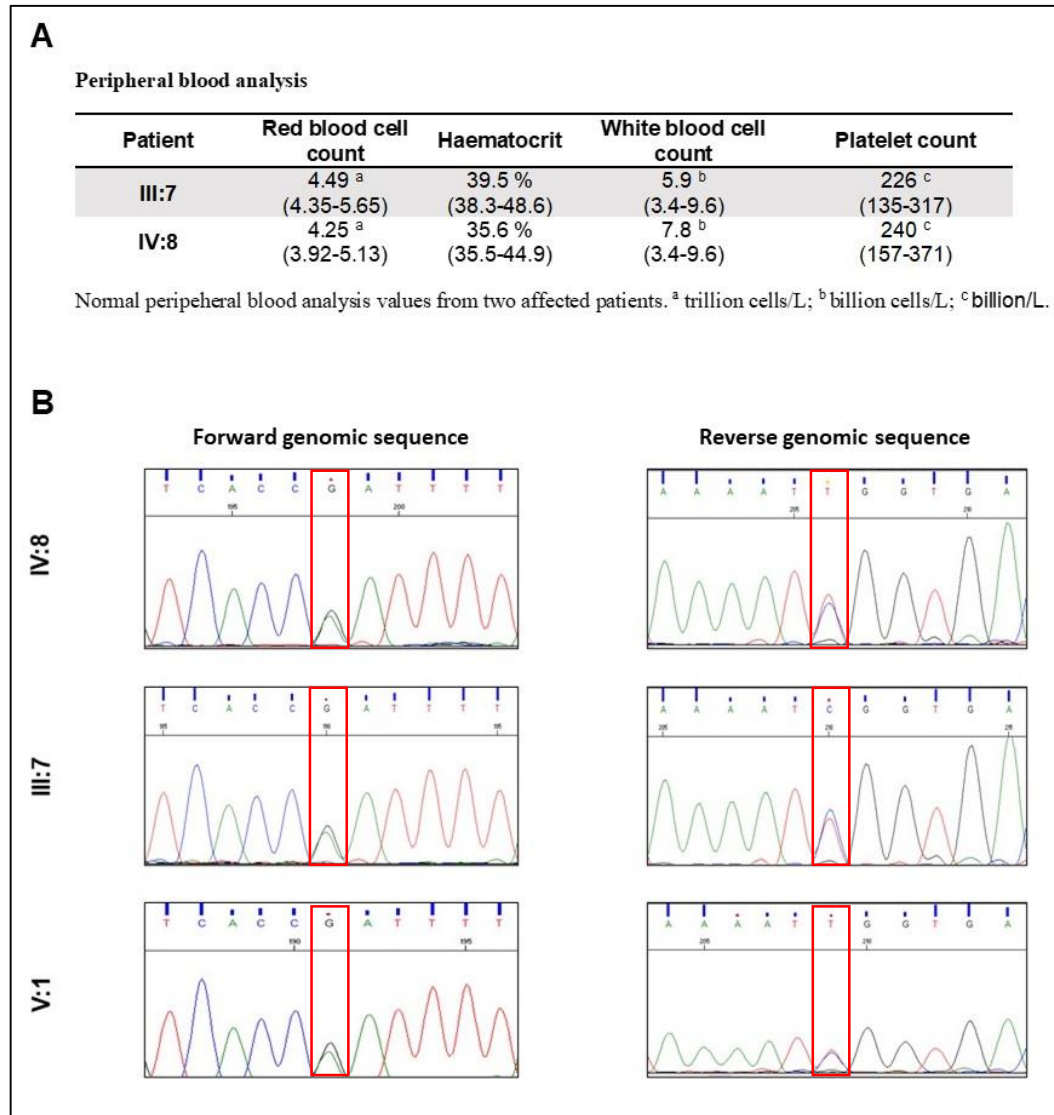
\*Result images located in the MNI space (neurological orientation).

**Supplementary Figure 4. Reduction of sensory nerve amplitudes in upper limbs at later stages of disease progression.** (A and B) Patient IV:2 showed reduction of sensory nerve amplitudes in upper limbs at later stages of disease progression. (C and D) Control records using reference values from Kimura and collaborators.<sup>1</sup>



**Supplementary Figure 5. Haematologic evaluation and heterozygosity of *SAMD9L* c.1877C>T (p.Ser626Leu) Sanger sequencing.**

(A) Complete blood counts and red cell indices were normal without hematologic manifestations in two patients. (B) DNA isolated from peripheral blood shows *SAMD9L* heterozygous c.1877C>T (p.Ser626Leu) mutation in affected patients IV:8, III:7 and V:1, discarding decrease of the mutant allele in haematopoietic lineage.



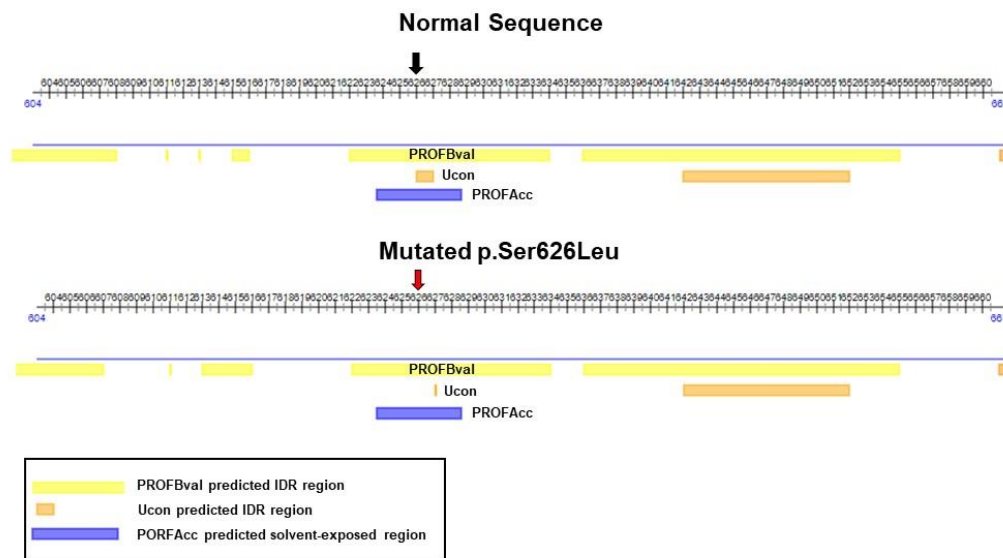
## Supplementary Figure 6. Putative effects of the p.Ser626Leu amino acid change in SAMD9L structural predicted region.

(A) PROFBval and Ucon predictors from PredictProtein algorithm identified intrinsic disorder region (IDR) surrounding SAMD9L Ser626. P.Ser626Leu change reduce the IDR predicted by Ucon. PROFAcc identified solvent-exposed region around SAMD9L Ser626.

(B) The p.Ser626Leu mutation in SAMD9L would disrupt IDR predicted by mobdib-lite at InterPro protein sequence predictors.

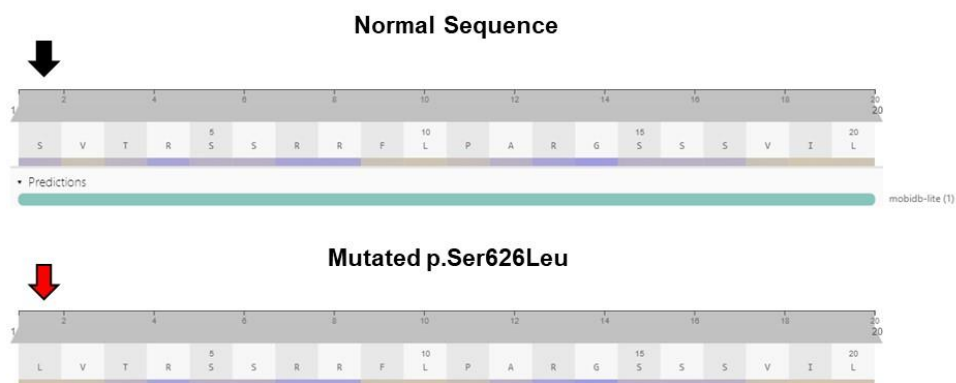
**A**

### Protein Disorder and Flexibility (PredictProtein)



**B**

### InterPro (MobiDB-lite) analysis of Ser626 surrounding region





**Supplementary Figure 7. Phosphorylation prediction of SAMD9L Ser626.**

NetPhos 3.1 and NetworKIN predicted SAMD9L Ser626 to be potentially phosphorylated by PKC.

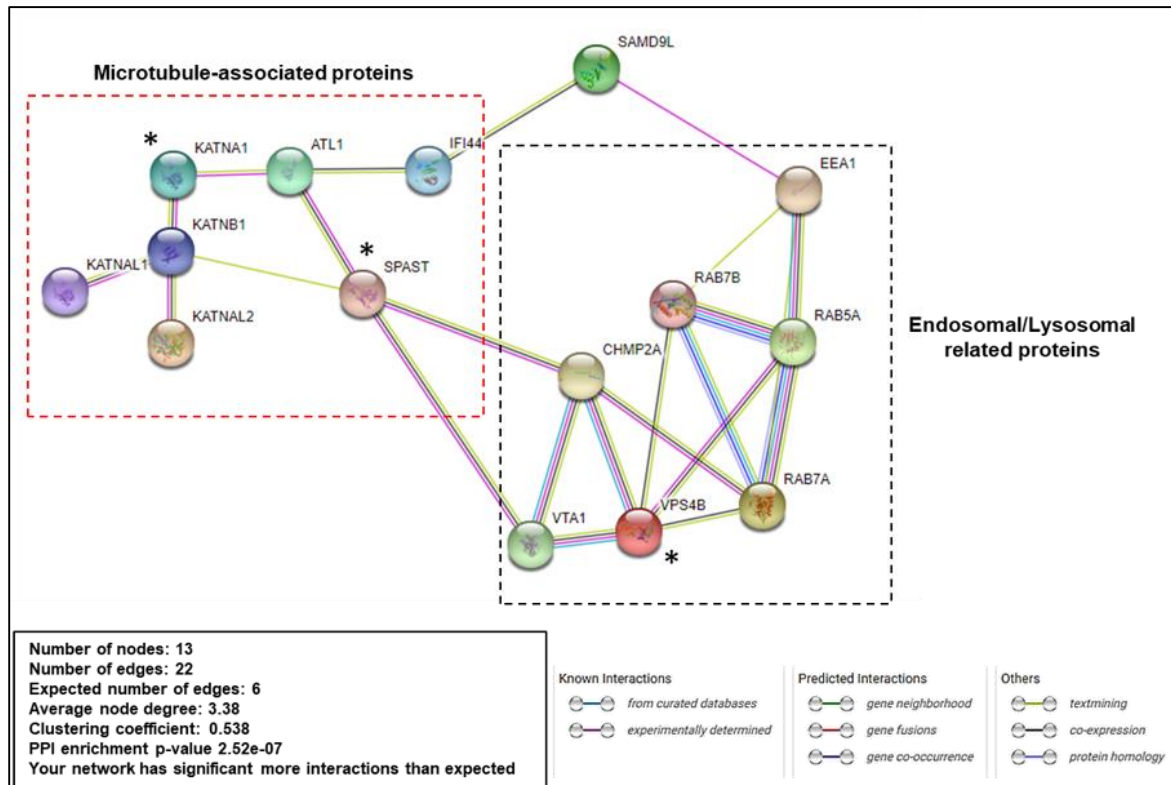
**NetPhos 3.1**

--						
# Sequence	626	S	LKLKSVTRS	0.981	unsp	YES
# Sequence	626	S	LKLKSVTRS	0.712	PKC	YES
# Sequence	626	S	LKLKSVTRS	0.460	CaM-II	.
# Sequence	626	S	LKLKSVTRS	0.459	cdc2	.
# Sequence	626	S	LKLKSVTRS	0.443	GSK3	.
# Sequence	626	S	LKLKSVTRS	0.378	p38MAPK	.
# Sequence	626	S	LKLKSVTRS	0.372	CKI	.
# Sequence	626	S	LKLKSVTRS	0.369	RSK	.
# Sequence	626	S	LKLKSVTRS	0.342	DNAPK	.
# Sequence	626	S	LKLKSVTRS	0.296	ATM	.
# Sequence	626	S	LKLKSVTRS	0.249	CKII	.
# Sequence	626	S	LKLKSVTRS	0.246	PKG	.
# Sequence	626	S	LKLKSVTRS	0.230	cdk5	.
# Sequence	626	S	LKLKSVTRS	0.164	PKA	.
# Sequence	626	S	LKLKSVTRS	0.098	PKB	.

**NetworKIN**

sp Q8IVG5 SAM9L_HUMAN (SAMD9L)						
S626	Kinase	PKCbeta	2.49	ILKLKSVTRSS		
		PKCalpha	2.40	ILKLKSVTRSS		

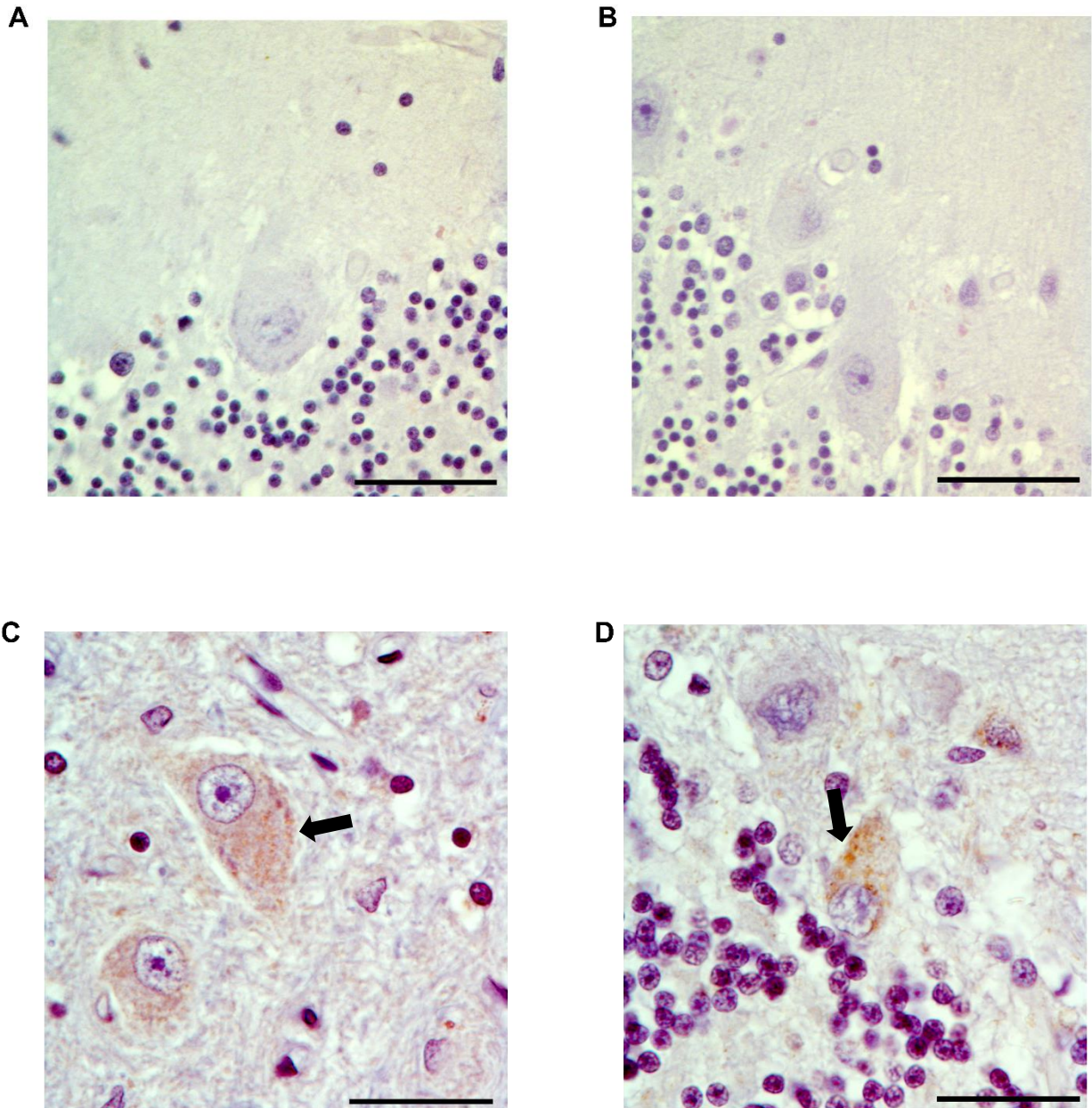
**Supplementary Figure 8. STRING protein-protein (PPI) interaction network of SAMD9L with structurally homologous predicted proteins identified by the HHpred protein sequence profile predictor algorithm.**



STRING was analysed by implementing the KATNA1, SPAST, VPS4B proteins identified by the HHpred protein sequence profile predictor algorithm (Fig. 3B) and with the previously reported Samd9l-interacting proteins in mice Eea1 and Rab5. The SAMD9L and SPAST network interaction was identified via IFI44I, also known as microtubule-associated protein 44, and ATL1 proteins. SPAST and ATL1 proteins have been previously associated with mitochondrial function.<sup>32,33</sup> The PPI interaction enrichment analysis revealed significantly more interactions in the indicated proteins than expected ( $P < 2.52 \times 10^{-7}$ ). Co-expression and interaction analyses link both IFI44I and EEA1 with SAMD9L. (\*) Predicted proteins with structural homology with SAMD9L. Background: whole human genome.

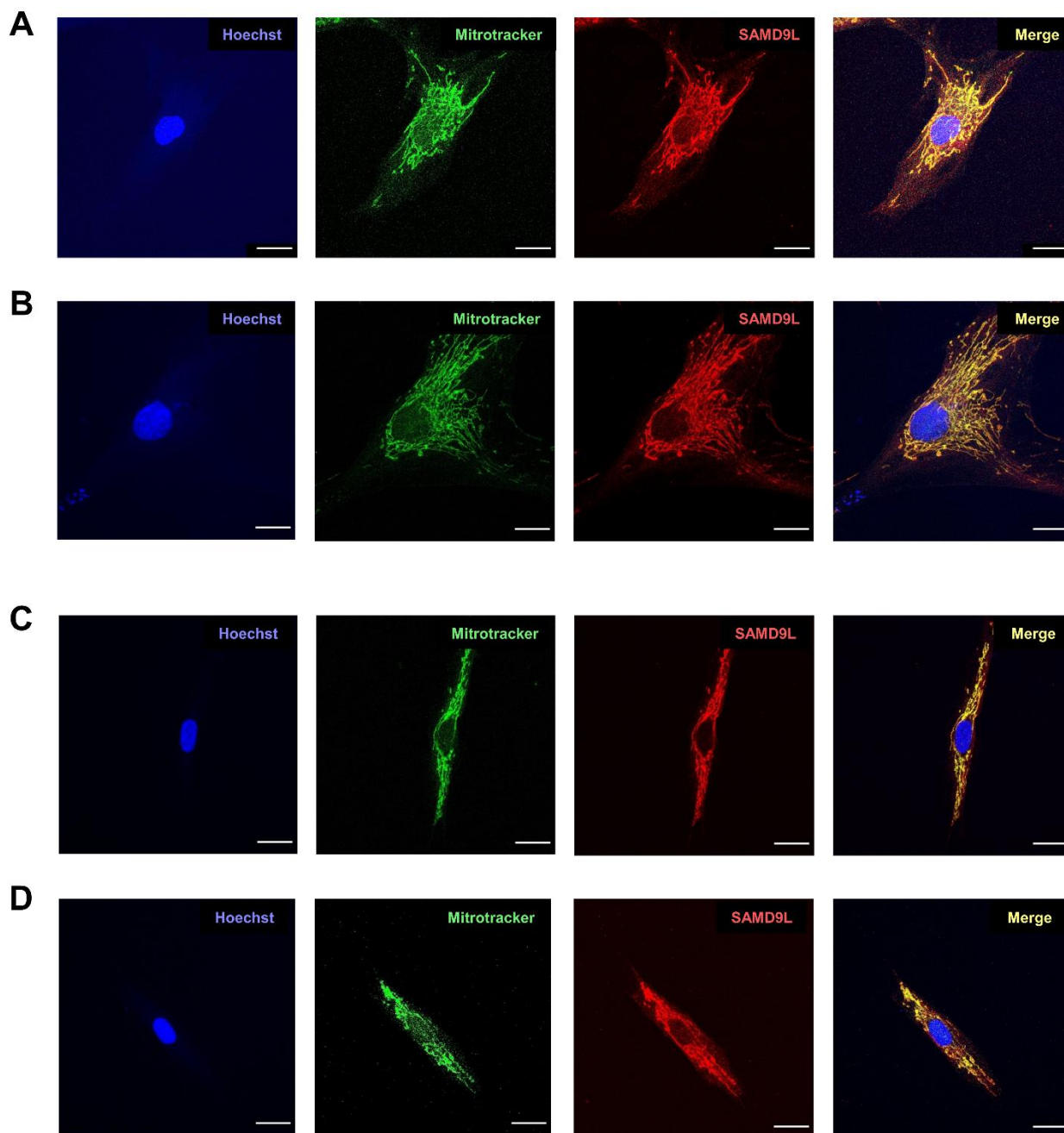


**Supplementary Figure 9. Specificity of antibodies and mitochondrial staining of ATP5B in human cerebellar Purkinje cells and multipolar neurons of the cerebellar dentate nucleus. (A)** Primary anti-SAMD9L antibody specificity without the secondary goat anti-Rabbit IgG (H+L) superclonal antibody. **(B)** Secondary anti-Rabbit IgG (H+L) superclonal antibody specificity without primary anti-SAMD9L antibody. ATP5B staining in human multipolar neurons of the cerebellar dentate nucleus **(C)** and cerebellar Purkinje cells **(D)** showing mitochondria. Magnification bars: 50  $\mu$ m.

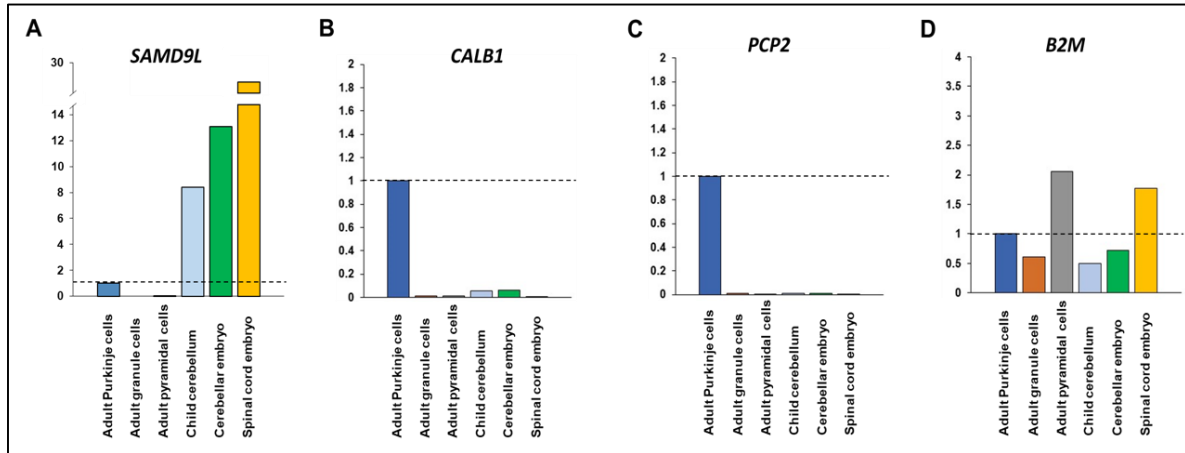




**Supplementary Figure 10. Mitochondrial localisation of SAMD9L in fibroblasts.** Immunofluorescence staining of fibroblasts from human control (A and B) and an M-SCA affected patient (C and D) showed co-localisation of SAMD9L to mitoTracker Red CMXRos demonstrating mitochondrial co-localisation. Magnification bars: 20  $\mu$ m.

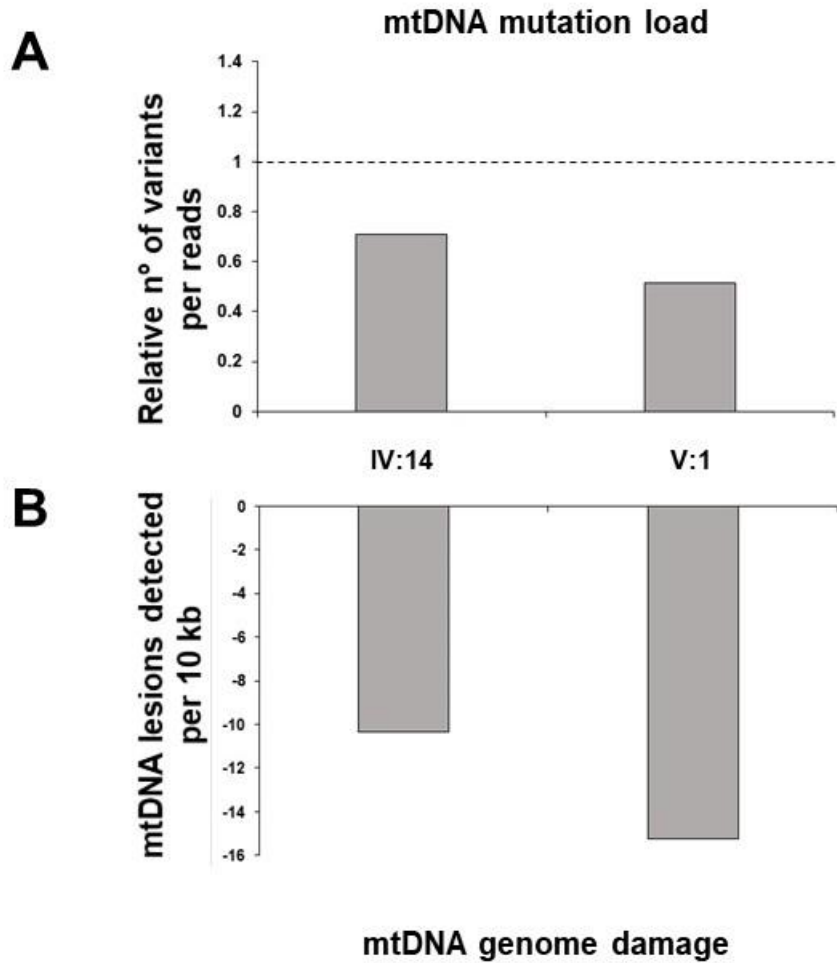


**Supplementary Figure 11. Relative expression levels of *SAMD9L* mRNA in human cerebellar Purkinje, granule and pyramidal cells, cerebellum, and spinal cord tissues.**



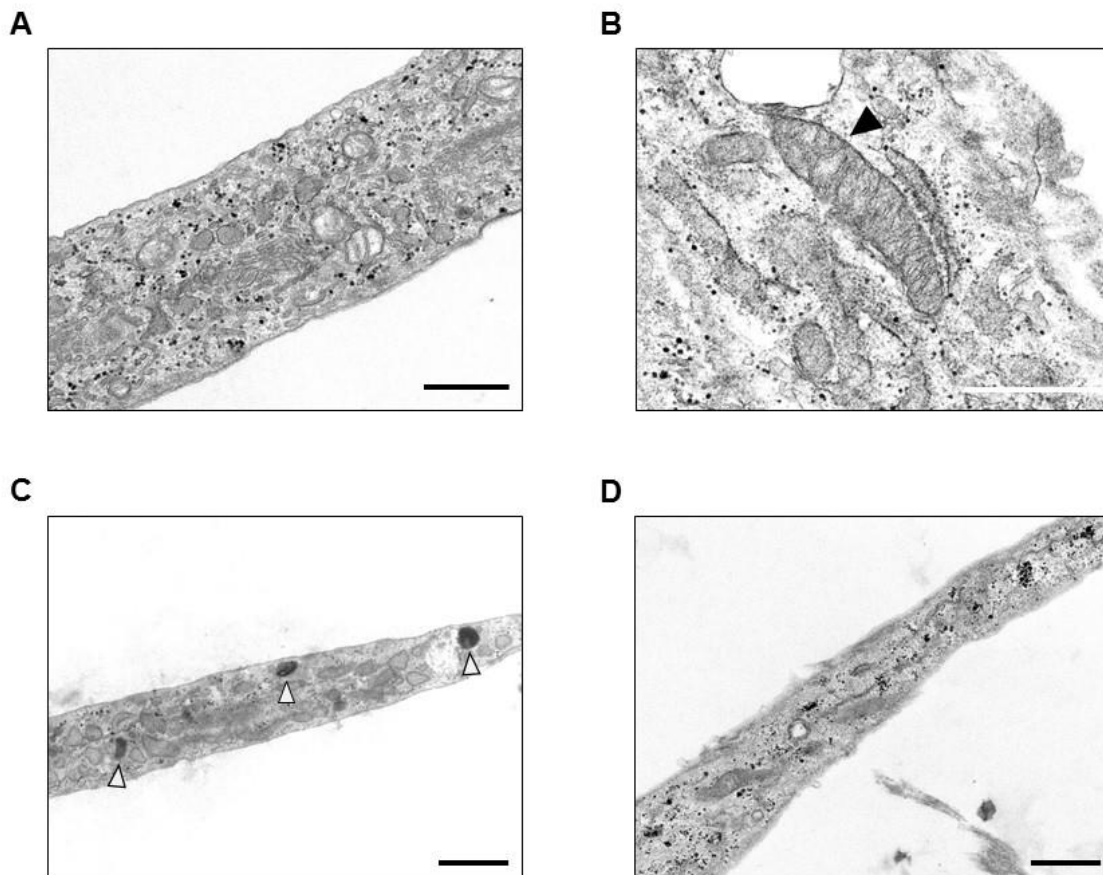
ENCODE *SAMD9L*, *CALB1*, *PCP2* and *B2M* mRNA sequences reads were compared on BAM files obtained from adult cerebellar Purkinje cells (dotted line), adult granular and pyramidal cells, 6-years-old child cerebellum, and cerebellar and spinal cord human embryos. *GAPDH* sequences reads were used to normalise intrasample differential expression. (A) *SAMD9L* was found only relatively expressed in adult male Purkinje cells and almost no expression was identified in cerebellar granular and pyramidal cells. mRNA tissue expression shows higher levels of *SAMD9L* in child cerebellum and cerebellar and spinal cord embryos in consonance with high mitochondrial activity and tissues involved in the pathogenesis of this new ataxia subtype. (B and C) Expression of *CALB1* and *PCP2*, two specific Purkinje cell markers, were mainly identified in adult Purkinje cells. (D) Expression of *B2M* gene, predominantly expressed in motor neurons, was accordingly higher in adult pyramidal cells and spinal cord embryo.

**Supplementary Figure 12. No alteration in mtDNA mutation load and damage identified in patient's fibroblasts.**



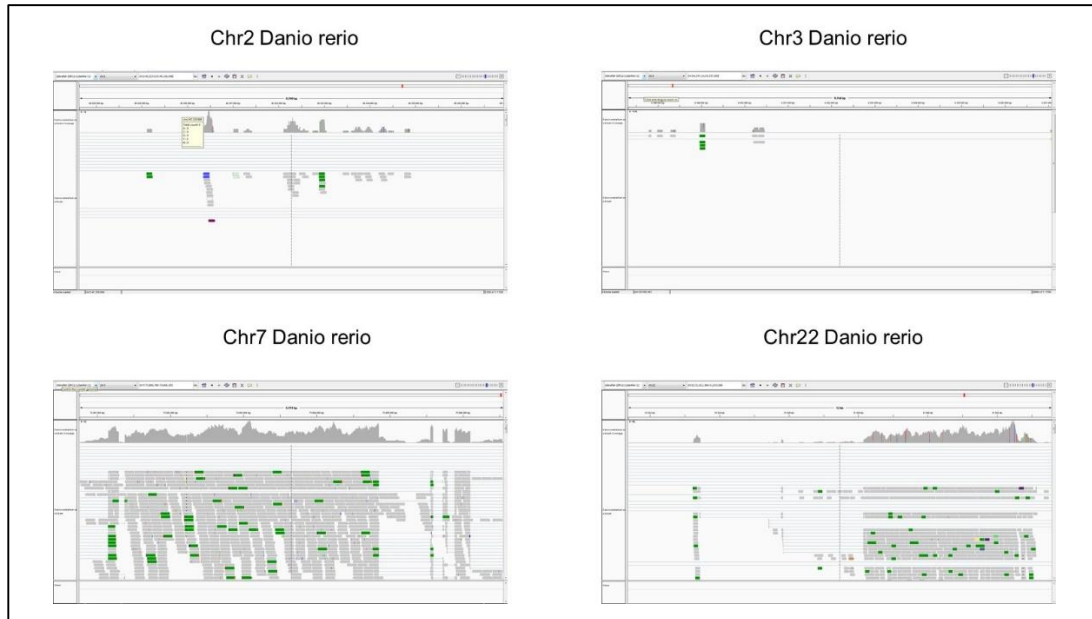
NGS analysis showed no alteration in variant frequency in affected patients' fibroblasts (variation frequency  $2.1 \times 10^{-2}$ / reads) compared to four age-matched controls (variation frequency =  $3.1 \times 10^{-2}$ / reads) (**A**) Values shown for patients IV:14 and V:1 are relative to the obtained control value set to 1 (dotted line). None of the identified variants were predicted to be pathogenic. (**B**) LORD-Q method denotes similar mitochondrial genome damage representing number of detected lesions per 10 kb in affected fibroblasts compared to controls, discarding a possible oxidative stress effect.

**Supplementary Figure 13. Absence of dilated endoplasmic reticulum and diffused mitochondrial crests in control fibroblasts TEM images (A-B and D), as compared to affected (C) fibroblasts.**



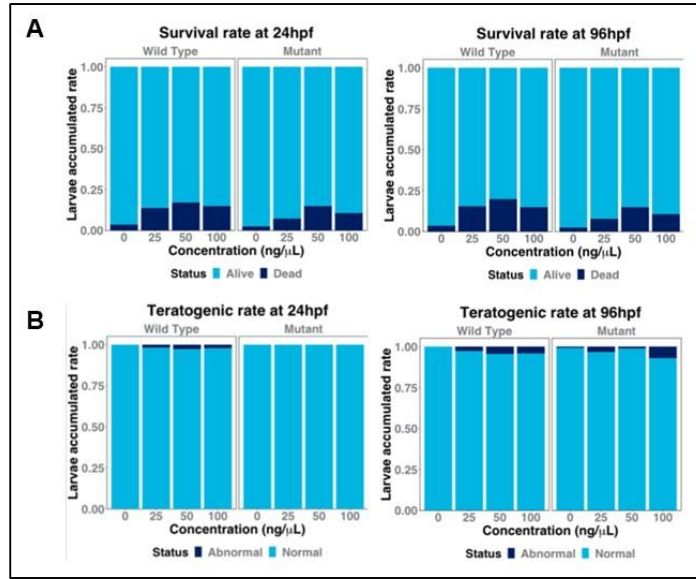
Dilated endoplasmic reticulum was absent in control fibroblasts (A and D) compared to patients' (Fig. 6E and F) and C. Mitochondrial crests appear markedly in the mitochondrial matrix from control fibroblasts (B) as compared to affected (Fig. 6C). Number of lysosomes were significantly increased in patients' fibroblasts (C) compared to controls (A, B and D). White arrowheads point to lysosomes. Black arrowhead points to mitochondria with defined crests.

**Supplementary Figure 14. Characterisation of human *SAMD9L* ortholog gene in zebrafish by analysis of RNA-seq zebrafish cerebellar expression in previously reported candidate regions from databases.**



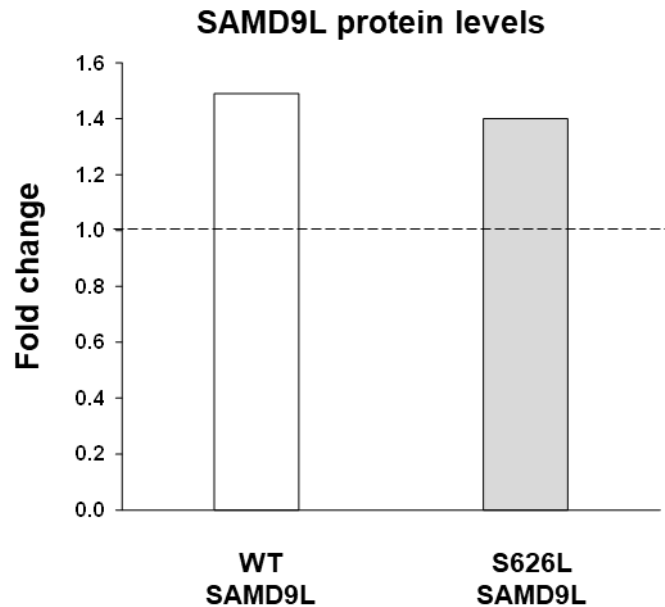
Candidate regions were selected regarding previously generated transgenic zebrafishes reported to be candidate for human *SAMD9L* orthologs (<https://zmp.buschlab.org/>). Zebrafish cerebellum RNA-seq data from SRA database (Accession: SRX4184229) were aligned using HISAT2 and candidate regions analysed.

**Supplementary Figure 15. *SAMD9L* mRNA concentrations tested do not affect accumulated survival or teratogenic rate in zebrafish larvae.**



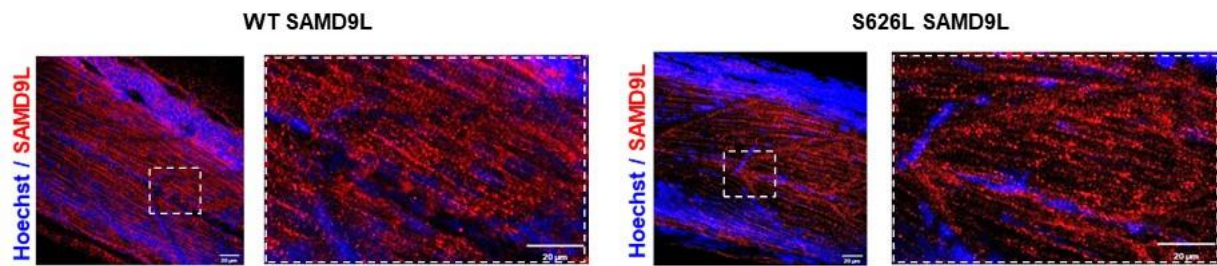
Accumulated survival (A) and teratogenic rate (B) at 24 and 96 hpf were not significantly impaired with any mRNA concentration tested.

**Supplementary Figure 16. Similar levels of both wild-type and mutant human *SAMD9L* protein in zebrafish larvae.**



Overexpression of *SAMD9L* protein levels were confirmed in both WT *SAMD9L* and S626L zebrafish larvae (N = 10), without any significant differences. Mitochondrial VDAC was used as internal loading control and *SAMD9L* relative levels were normalised to lynGFP group endogenous expression.

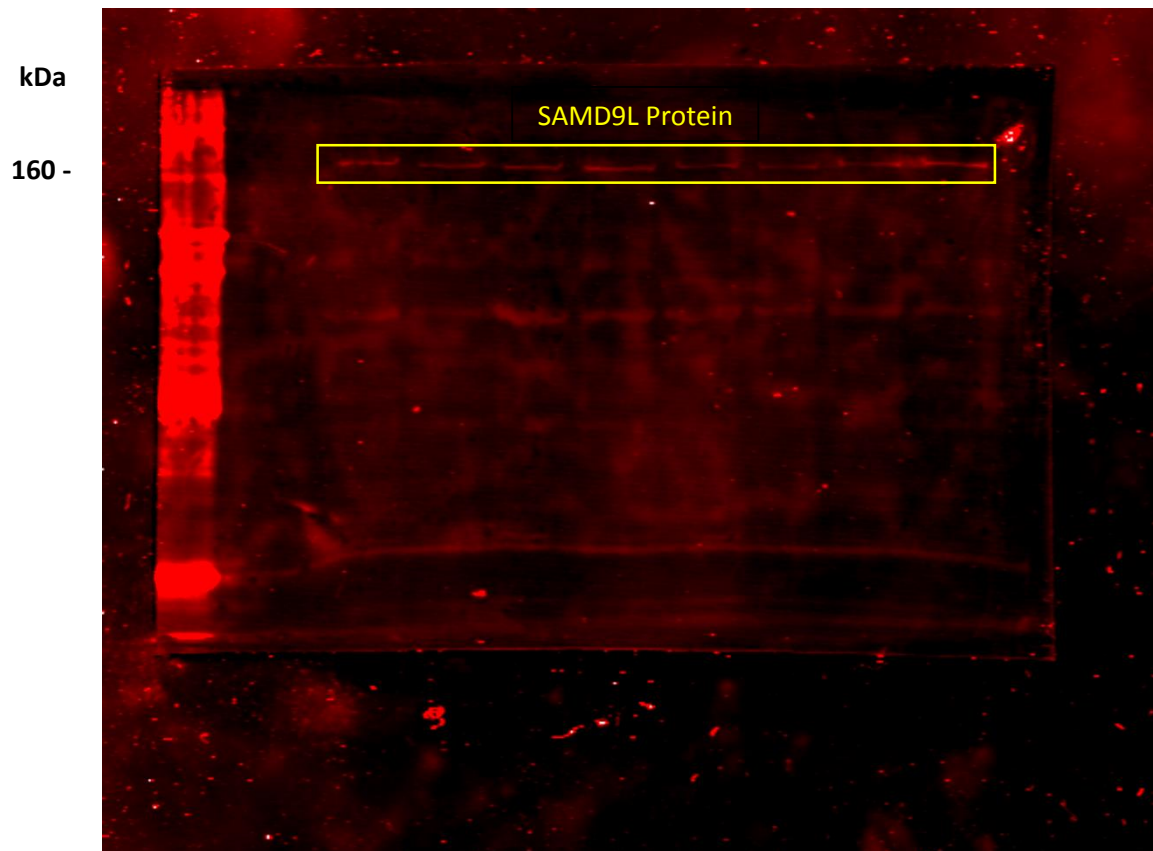
**Supplementary Figure 17. Overexpressed SAMD9L mitochondrial staining in the zebrafish spinal cord and peripheral nerves.**



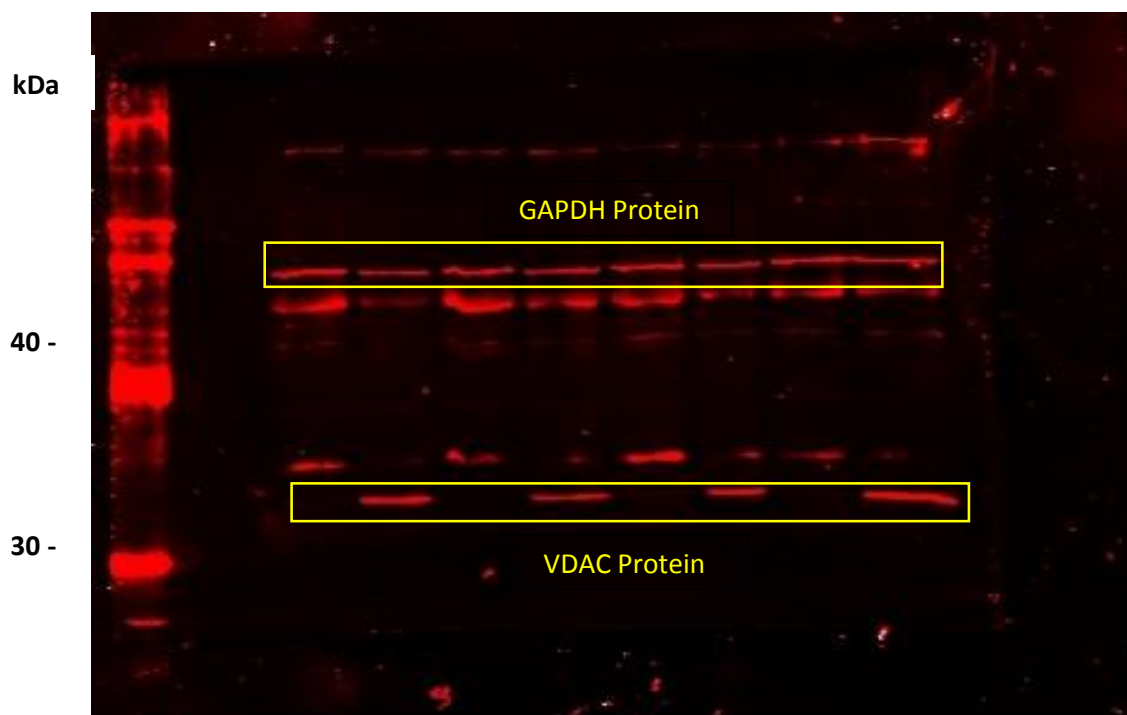
SAMD9L staining in spinal cord and peripheral nerves in zebrafish embryos suggest mitochondrial staining. This is corroborated by SAMD9L co-localisation with mitochondrial ATP5B protein shown in Fig. 7F.



**Supplementary Figure 18. Uncropped Figure 5A SAMD9L protein**

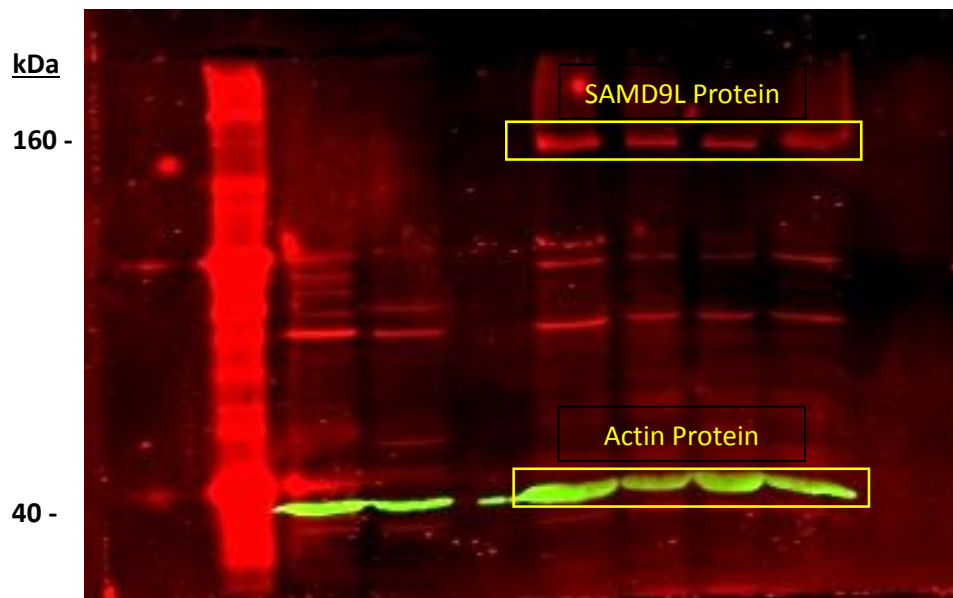


**Supplementary Figure 19. Uncropped Figure 5A GAPDH protein** (same membrane previously blotted with specific antibodies for additional three proteins including SAMD9L)

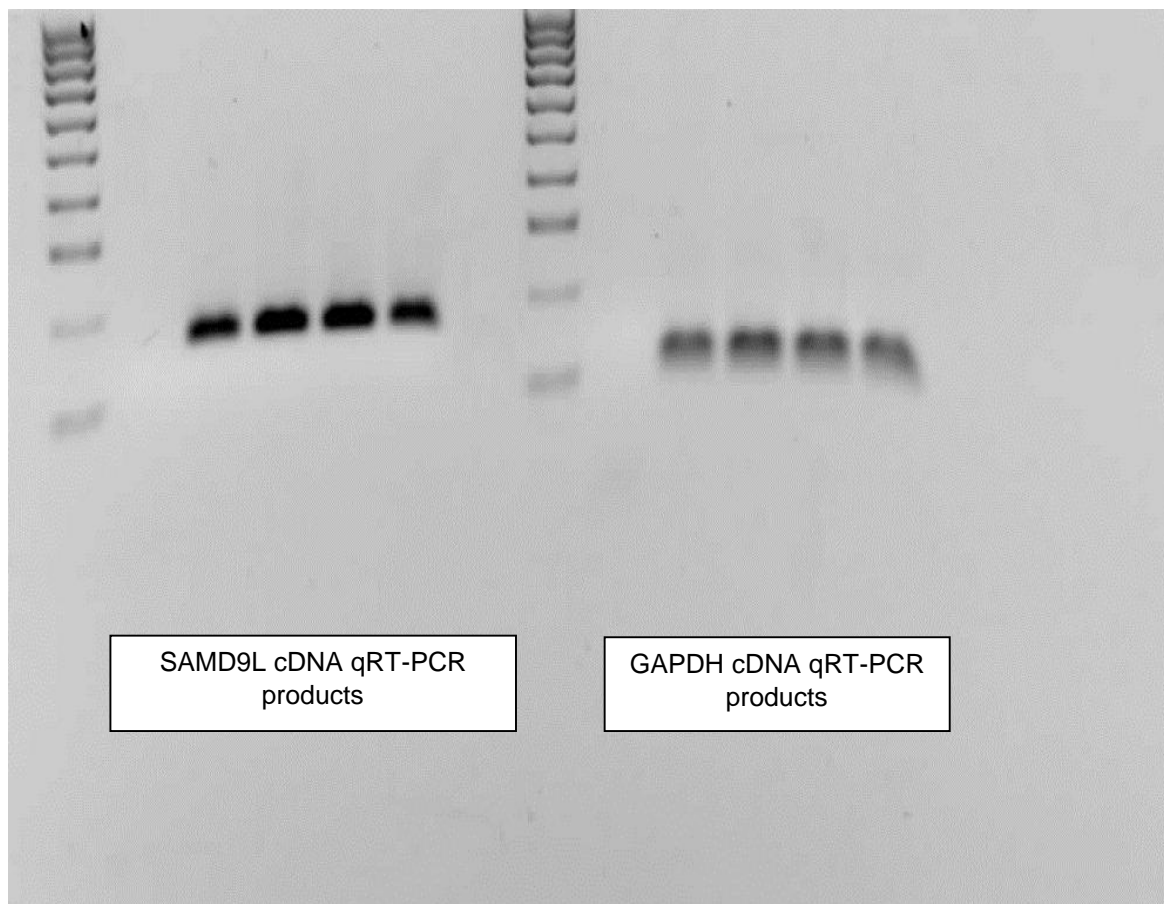




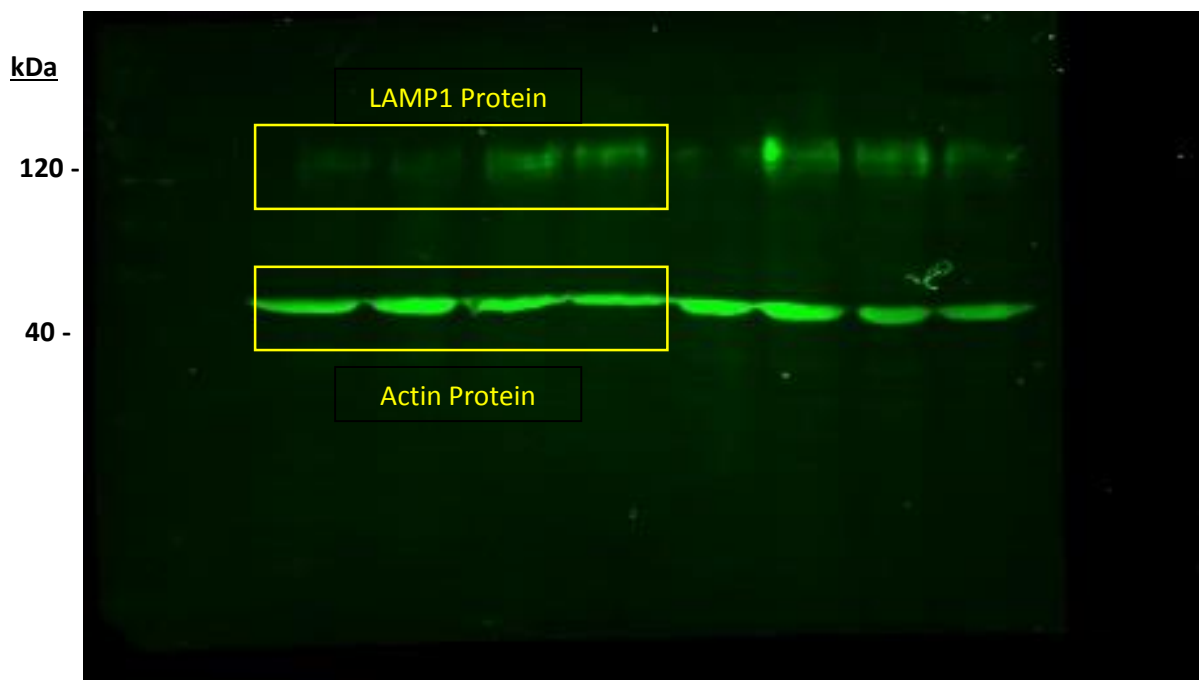
**Supplementary Figure 20. Uncropped Figure 5B SAMD9L and Actin proteins** (same membrane previously blotted with specific antibodies for additional proteins)



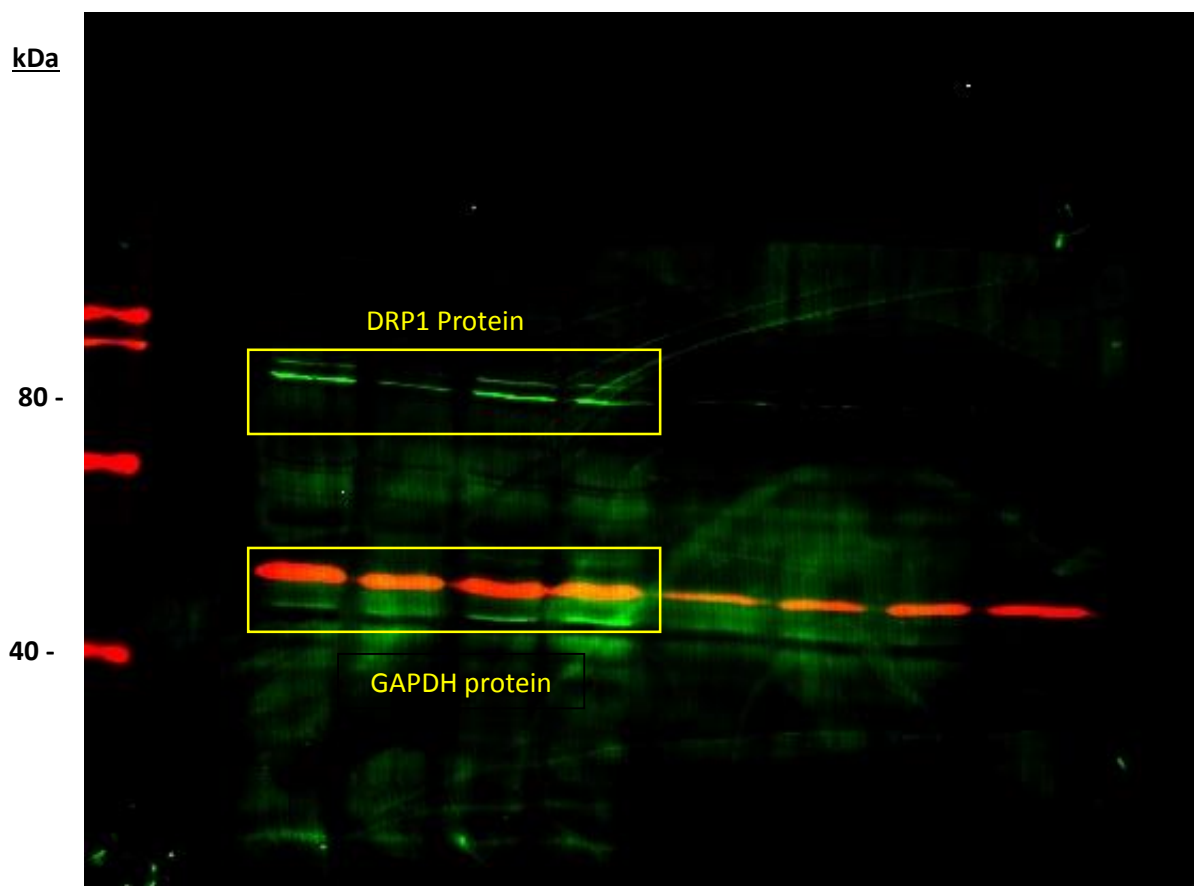
**Supplementary Figure 21. Uncropped Figure 5C qRT-PCR products SAMD9L and GAPDH**



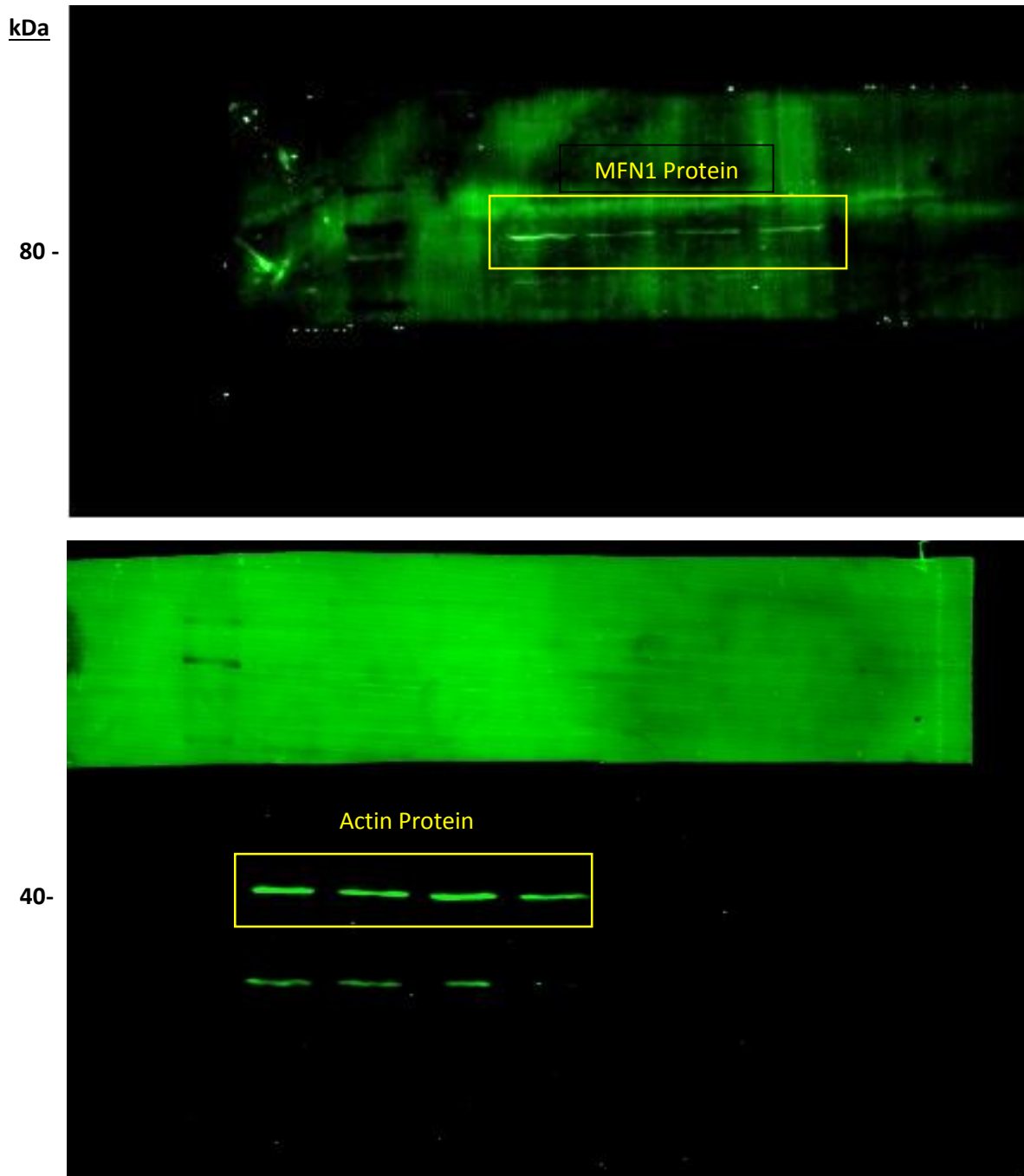
**Supplementary Figure 22. Uncropped Figure 6A LAMP1 and Actin proteins**



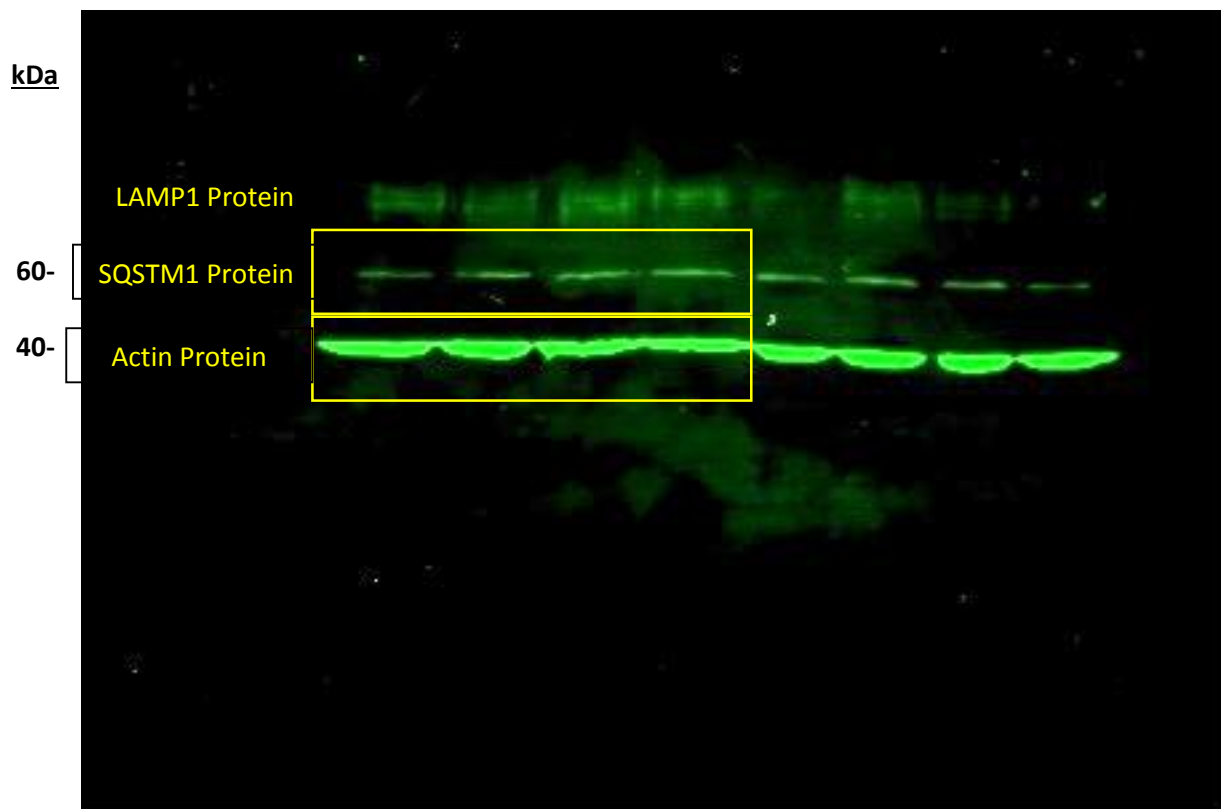
**Supplementary Figure 23. Uncropped Figure 6A DRP1 and GAPDH proteins**



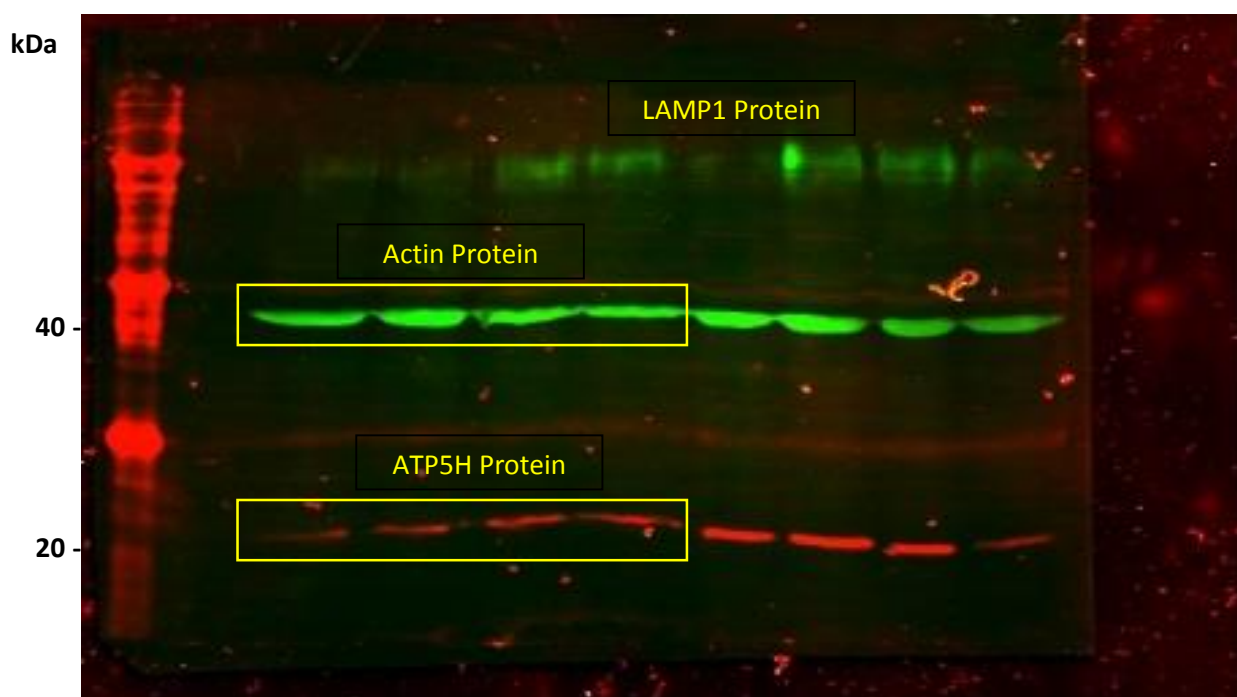
**Supplementary Figure 24. Uncropped Figure 6A MFN1 and Actin proteins**



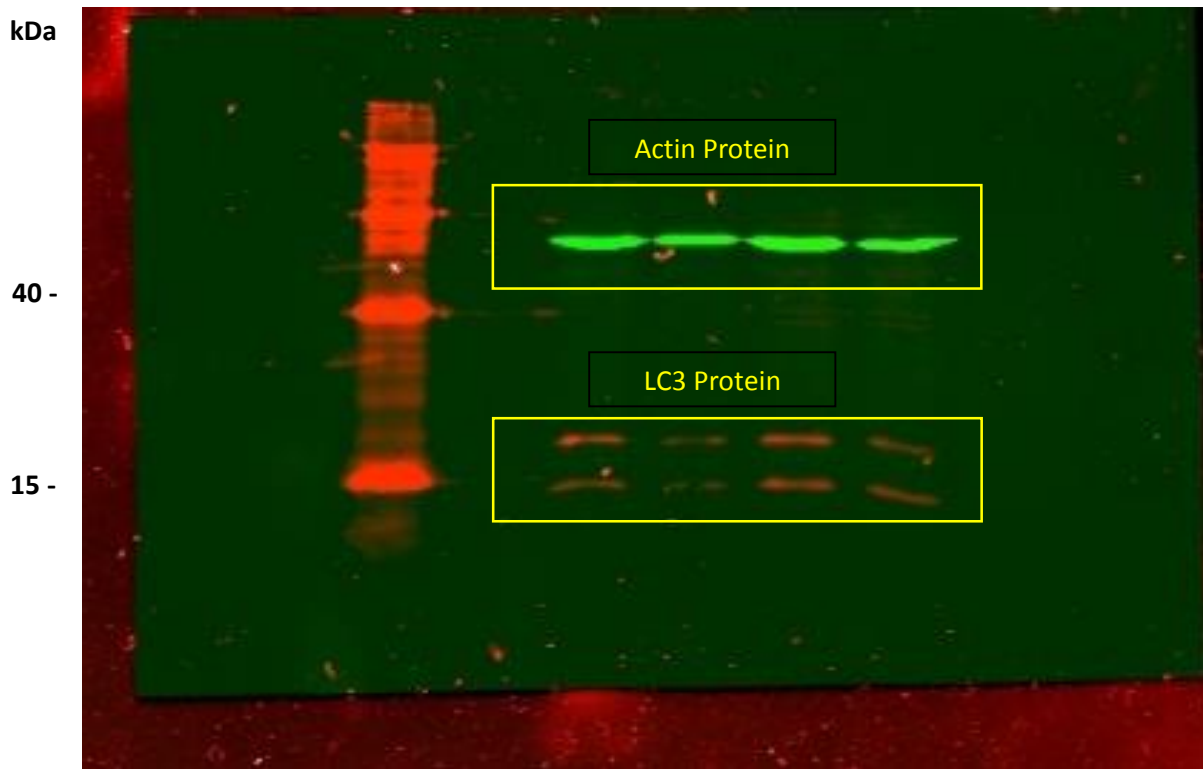
**Supplementary Figure 25. Uncropped Figure 6A LAMP1, SQSTM1 and Actin proteins**



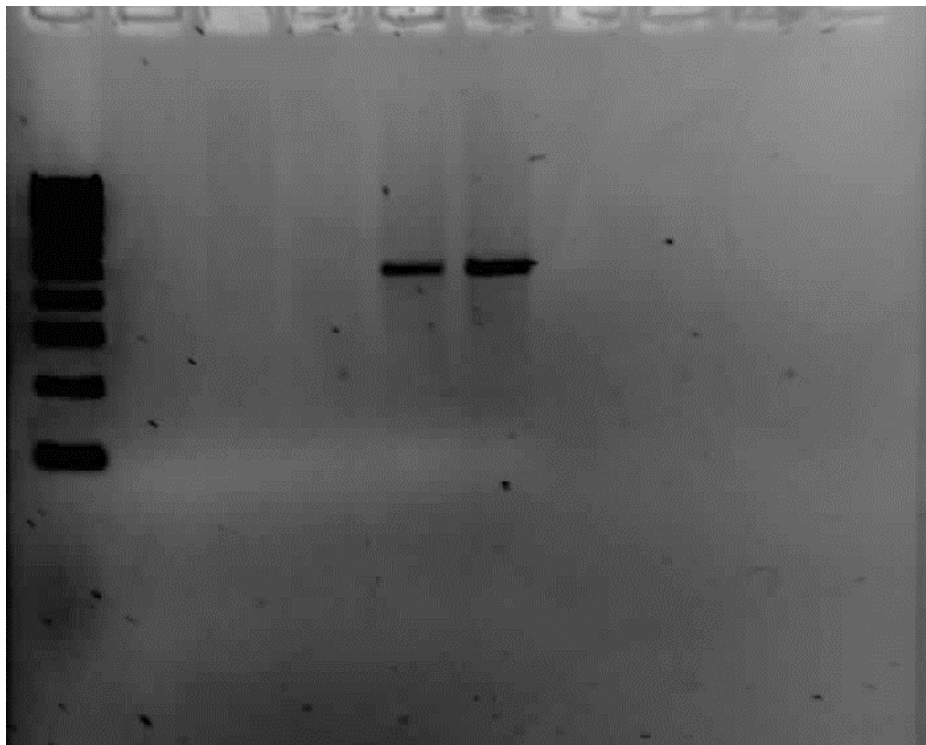
**Supplementary Figure 26. Figure 6A ATP5H and Actin proteins (membrane previously probed with anti-LAMP1 Ab)**



**Supplementary Figure 27. Figure 6A LC3 and Actin proteins**

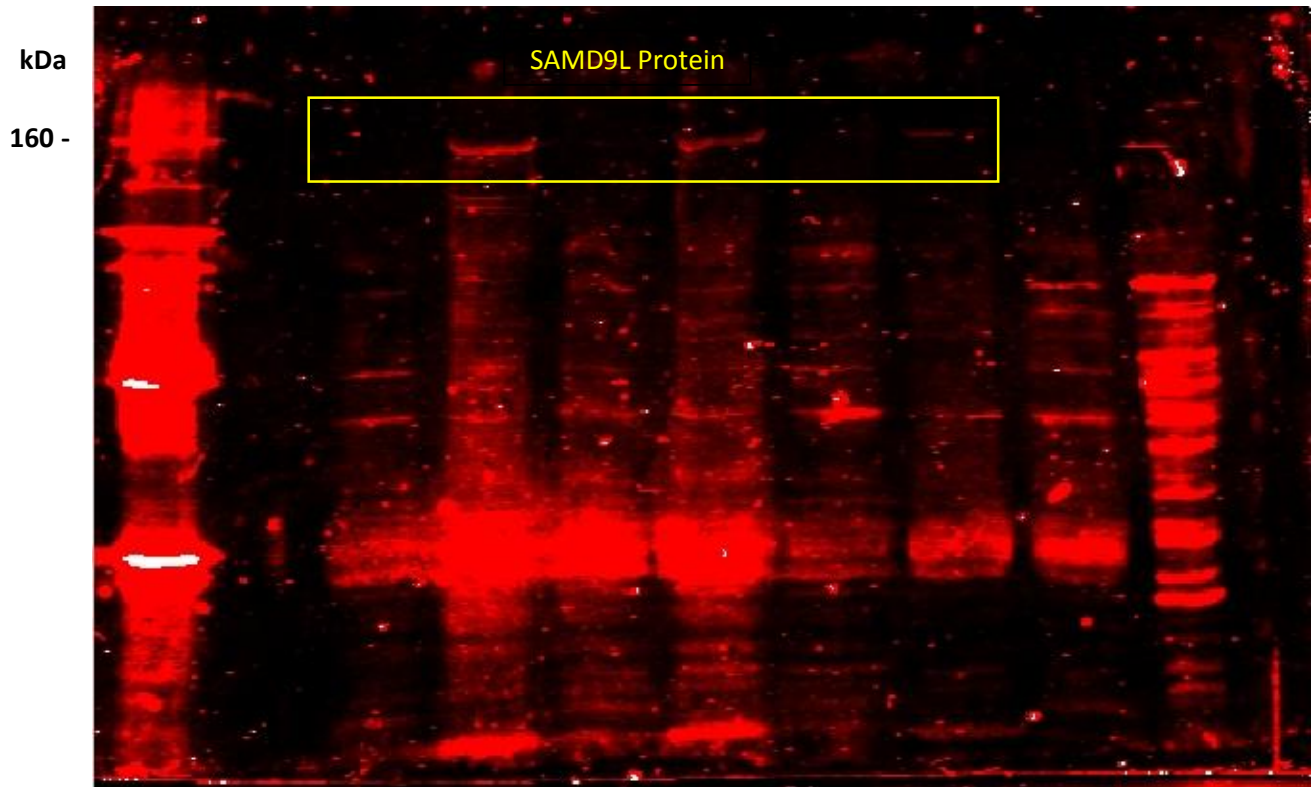


**Supplementary Figure 28. Figure 7C SAMD9L cDNA products from zebrafish embryos**

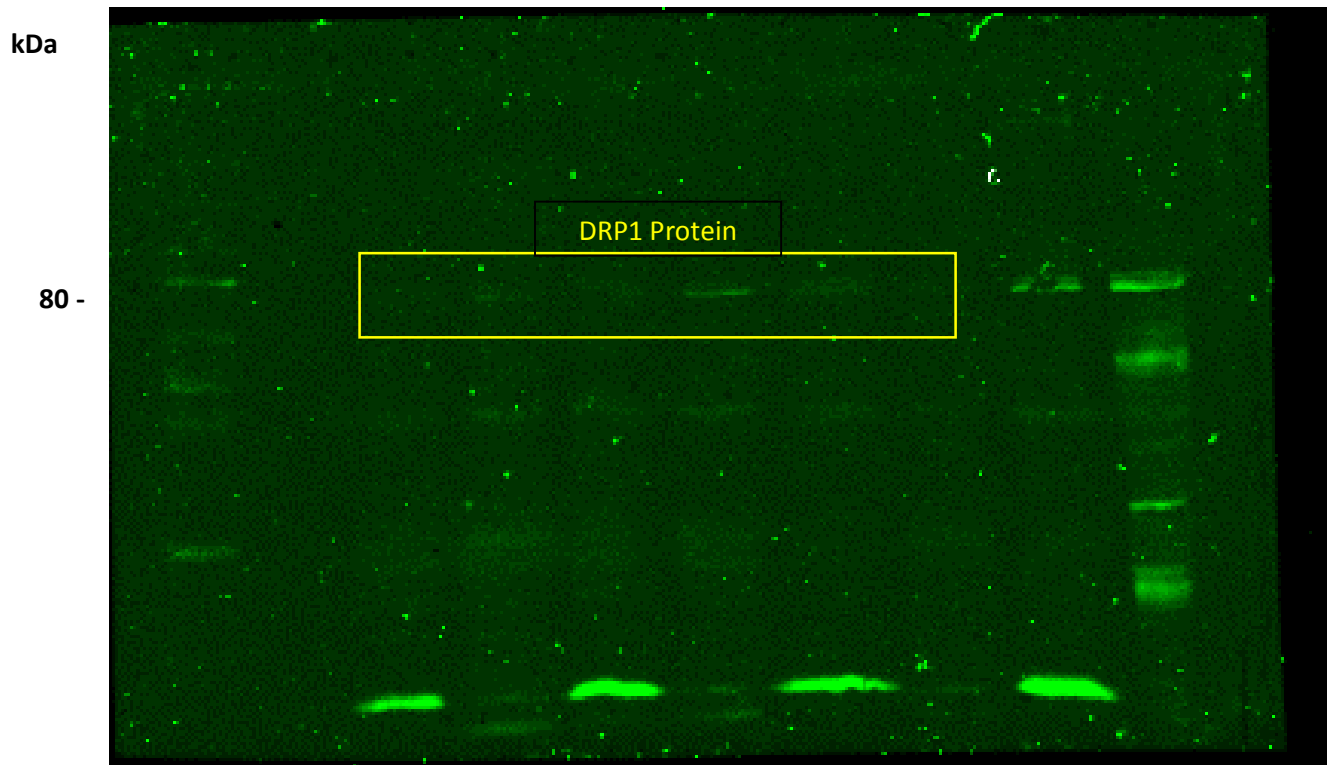




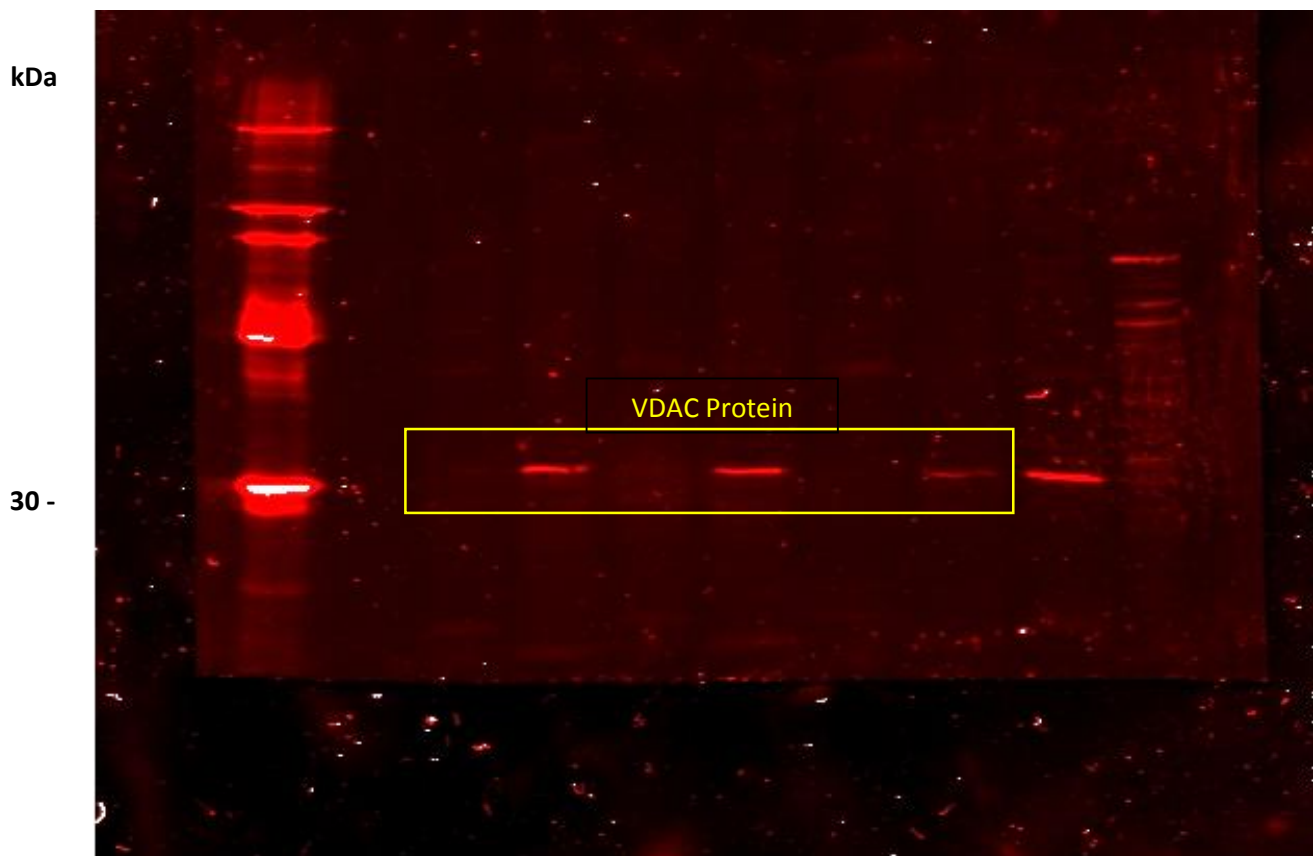
**Supplementary Figure 29. Figure 7D SAMD9L on zebrafish embryos (very low amounts of loaded protein from cellular fractions to distinguish overexpressing human recombinant SAMD9L with endogenous fish SAMD9L)**



**Supplementary Figure 30. Figure 7D DRP1 from zebrafish embryos (previously blotted with anti-SAMD9L Ab)**



**Supplementary Figure 31. Figure 7D VDAC from zebrafish embryos**



**Supplementary Table 1. Oligonucleotide primer sequences and PCR conditions used in *SAMD9L* studies.**

Oligonucleotide Primer ID	Sequence (5'-3')
<b><i>hSAMD9L</i> c.1877C&gt;T (p.Ser626Leu) Sanger Sequencing</b>	
hSAMD9L.S626L_FW	GAAAAATAGAAGTTCCACCAGGA
hSAMD9L.S626L_RV	TTTCTATTACTCTCTTCAGTGGAAAGC
<b><i>hSAMD9L</i> transcript Real Time qRT-PCR</b>	
hSAMD9L_cDNA_Exon4_FW	GGCACGGAAAGTGAGTGAGT
hSAMD9L_cDNA_Exon5_RV	AGAATTTGCCCGTATTGCTCATT
<b><i>hGAPDH</i> Real Time qRT-PCR</b>	
GAPDH_FW	CTTGTCATCAATGGAAATCCCA
GAPDH_RV	CAGTGGACTCCACGACGTACTC
<b><i>hSAMD9L</i> CDS cloning</b>	
hSAMD9L.cloning.ATG.FW	ATGAGTAAACAAGTATCTCTACCTGAAA
hSAMD9L.cloning.TAA.RV	TTAAATTACTTCTATATCATATGCCAGAG
<b><i>DrTbp</i> Real Time qRT-PCR</b>	
Dr_tbp_FW_cDNA	ACACCGCAGCCTGTGCAGAA
Dr_tbp_RV_cDNA	TGGCCTGAACCTCCCACCAT

One-hundred ng of DNA or cDNA were included in a 20 µl reaction. PCR conditions were as follows: initial denaturation at 95 °C for 5 min with a Taq DNA Polymerase, recombinant (ThermoFisher Scientific) followed by 35 cycles (95 °C for 30 sec, 58 °C for 40 sec, 72 °C for 50 sec), and a final denaturation at 72 °C for 10 min. Real Time qRT-PCR were performed using TB Green Premix Ex Taq II (Takara Bio) and PCR conditions were as follows: 95 °C for 2 min, and 40 cycles (95 °C for 5 sec, 60 °C for 20 sec, 72 °C for 20 sec). The  $2^{-\Delta\Delta C_t}$  method<sup>34</sup> was used to analyse the relative cDNA levels data obtained by Real Time qRT-PCR.



**Supplementary Table 2. Oligonucleotide primer sequences used and PCR conditions for mtDNA amplification.**

Oligonucleotide Primer ID	Sequence (5'-3')
<b>mtDNA PCR amplification</b>	
hMT_15058_Frag1_FW	CGGATCATTTCTCTACTCAGAAA
hMT_208_Frag1_RV	ACACACTTTAGTAAGTATGTTCCGCC
hMT_16516_Frag2_FW	GGGTCATAAAGCCTAAATAGC
hMT_1751_Frag2_RV	TCGCCTATACTTTATTTGGGTA
hMT_1404_Frag3_FW	ACTTAAGGGTCGAAGGTGGATT
hMT_3947_Frag3_RV	TCGATGTTGAAGCCTGAGACTA
hMT_3734_Frag4_FW	AAGTCACCCTAGCCATCATTCTAC
hMT_6739_Frag4_RV	GATATCATAGCTCAGACCATACCTATG
hMT_6364_Frag5_FW	GTGTCTCCTCTATCTTAGGGGCCAT
hMT_9785_Frag5_RV	GTAGATGCCGTCGGAAATGGTG
hMT_8655_Frag6_FW	GTACGTAGTCTAGGCCATATGTGT
hMT_10740_Frag6_RV	CACCACCCAACAATGACT
hMT_10361_Frag7_FW	TCTGGCCTATGAGTGACTACAAAAAGG
hMT_12373_Frag7_RV	TCAGGGTTAGGGTGGTTATAGTAGTGTGC
hMT_11977_Frag8_FW	CTCCCTCTACATATTTACCACAACACA
hMT_13830_Frag8_RV	AAGTCCTAGGAAAGTGACAGCGA
hMT_13469_Frag9_FW	TAGCATTAGCAGGAATACCTTTCC
hMT_15347_Frag9_RV	GCAAGAATAGGAGGTGGAGTGTT
<b>mtDNA qRT-PCR</b>	
hMT_DLOOP_FW	CAGGCGAACATACTTACTAAAGTGTGTTA
hMT_DLOOP_RV	GTGACTGTATAAAAGTGCATACCGC
hMT_MT-CO3_FW	ATCAGGAGTATCAATCACCTGAGCTCA
hMT_MT-CO3_RV	GTAGATGCCGTCGGAAATGGTG
hGAPDH_DNA_FW	ACGTAGCTCAGGCCTCAAGACCT
hGAPDH_DNA_RV	TTCTCTCCGCCCGTCTTCAC
<b>LORD-Q method for mitochondrial genome damage</b>	
AS2.F_mtDNA	GGCCACAGCACTTAAACACA
AS2.R_mtDNA	TGGTTAGGCTGGTGTAGGG
CL5.F_mtDNA	ATCGTAGCCTTCTCCACTTC

For mtDNA amplification one-hundred ng of DNA were included in a 20 µl reaction. PCR conditions were as follows: initial denaturation at 94 °C for 3 min with a LA Taq DNA Polymerase with GC Buffer I (Takara) followed by 35 cycles (94 °C for 30 sec, 58 °C for 30 sec, 72 °C for 4 min), and a final denaturation at 72 °C for 10 min. mtDNA Real Time qRT-PCR were performed using TB Green Premix Ex Taq II (Takara Bio) and PCR conditions were as follows: 95 °C for 2 min, and 40 cycles (95 °C for 5 sec, 60 °C for 20

sec, 72 °C for 20 sec). The  $2^{-\Delta\Delta C_t}$  method<sup>34</sup> was used to analyse the relative cDNA levels data obtained by Real Time qRT-PCR. Adapted LORD-Q (long-run real-time PCR-based DNA-damage quantification) method amplified a long 3,723 bp mtDNA fragment using only one reagent, TB Green Premix Ex Taq II (Takara Bio) instead of an independent fluorescent dye in combination with a Hot Start Taq. Data analysis was based on the amplification efficiencies of long (EL) and short (ES) probes, number of base pairs of the long fragment (a), number of quantification cycles (Cp) and the number of reference samples (n).<sup>25</sup>

**Supplementary Table 3. MRI assessment in five patients revealed significant decrease in cerebellar volume and cerebellar grey matter.**

Patient ID	Age at MRI	Relative cerebellar volume (%)	Age-matched controls median cerebellar volume (%)	Age-matched controls range cerebellar volumes (%)	Relative grey matter (%)	Age-matched controls median relative grey matter (%)	Age-matched controls range relative cerebellar volumes (%)
IV:8	62	6.64	9.26	7.86 - 10.65	4.49	6.80	5.65 - 7.95
IV:6	63	7.05	9.33	7.93 - 10.72	4.64	6.84	5.70 - 7.99
IV:14	47	7.07	9.26	7.93 - 10.59	4.78	6.91	5.82 - 8.00
IV:2	65	5.76	8.84	7.51 - 10.17	3.59	6.59	5.50 - 7.68
V:1	37	6.26	9.43	8.10 - 10.75	4.12	7.07	5.99 - 8.16
Mean		6.56	9.22		4.32	6.84	
STD		0.55	0.22		0.47	0.17	
Oneway ANOVA		$F_{(1,8)} = 99.506$ ; $P < 0.0001$			$F_{(1,8)} = 122.136$ ; $P < 0.0001$		

**Supplementary Table 4. SAMD9L domains and motifs identified by Hhpred for protein homology detection and structure prediction.**

Hit	Name	Probability	E-value	Position in the query sequence	Target Length	Conclusion
4PZO_A	Polyhomeotic-like protein 3; SAM domain {Homo sapiens}	97.23	0.000047	6-85	82	SAM
1Z6T_B	Apoptotic protease activating factor 1; Apaf-1, caspase activation, ADP, nucleotide; {Homo sapiens};	97.2	0.0014	716-1018	591	Apaf-1 ADP
2DKZ_A	hypothetical protein LOC64762; CELL-FREE PROTEIN SYNTHESIS, PROTEIN REGULATION; NMR {Homo sapiens}	96.97	0.00015	2-79	84	SAM_PNT-domain
5UJ7_E	Origin recognition complex subunit 1; Replication, DNA-binding, AAA+ ATPase, DNA; {Homo sapiens};	96.77	0.0025	716-872	284	DNA-binding, AAA+ ATPase,
5UJ7_B	Origin recognition complex subunit 1; Replication, DNA-binding, AAA+ ATPase, DNA; {Homo sapiens};	96.55	0.0052	711-873	391	AAA+ ATP-ase
3UK6_B	RuvB-like 2 (E.C.3.6.4.12); Hexameric AAA+ ATP-ase, DNA unwinding; {Homo sapiens};	96.54	0.0048	711-873	368	AAA+ ATP-ase
2XSZ_B	RUVB-LIKE 1 (E.C.3.6.4.12), RUVB-LIKE 2; HYDROLASE, AAA+ PROTEINS, HELICASE, CHROMATIN; OA {HOMO SAPIENS};	96.52	0.042	708-946	367	HYDROLASE
6BLB_A	Adenosylhomocysteinase {Pseudomonas aeruginosa};	96.51	0.0031	711-873	355	HYDROLASE
5VHO_A	26S proteasome non-ATPase regulatory subunit; p28, 26S proteasome, regulatory particle; {Homo sapiens};	96.43	0.0062	715-910	267	HYDROLASE
6P07_B	<b>Spastin (E.C.5.6.1.1), polyglutamate peptide; AAA+ ATPase, Homohehexamer, Microtubule Severing; {Drosophila melanogaster};</b>	96.38	0.019	689-904	494	AAA+ ATPase
5UJ7_C	Origin recognition complex subunit 1; Replication, DNA-binding, AAA+ ATPase, DNA; {Homo sapiens};	96.35	0.0096	708-873	436	DNA-binding, AAA+ ATPase
1XWI_A	<b>SKD1 protein; VPS4B, SKD1, AAA ATPase, PROTEIN; {Homo sapiens}</b>	96.27	0.013	692-902	322	AAA ATPase
6B5C_A	<b>Katanin p60 ATPase-containing subunit A-like; Katanin, AAA ATPase, Microtubule severing; {Homo sapiens};</b>	96.23	0.0083	715-873	307	AAA ATPase
2X8A_A	NUCLEAR VALOSIN-CONTAINING PROTEIN-LIKE, NUCLEAR PROTEIN; {HOMO SAPIENS}	96.22	0.018	690-873	274	HYDROLASE
5J1S_A	<b>Torsin-1A (E.C.3.6.4.-), Torsin-1A-interacting protein 2; AAA+ ATPase, Torsin, endoplasmic reticulum; {Homo sapiens};</b>	96.2	0.019	711-873	284	AAA+ ATPase
6MSB_E	26S proteasome; Proteasome, HYDROLASE; {Homo sapiens};	96.19	0.03	713-909	403	
5VHN_F	26S proteasome non-ATPase regulatory subunit; p28, 26S proteasome, regulatory particle; {Homo sapiens};	96.19	0.025	716-910	267	HYDROLASE
5UJM_A	Origin recognition complex subunit 1; ORC, replication, ATPase; {Homo sapiens}	96.17	0.022	713-873	522	ATPase
6G2Z_A	Transitional endoplasmic reticulum ATPase (E.C.3.6.4.6); ATPase, p97, protein degradation, hydrolase; HET: ADP, MPD, EJW; 1.923A {Homo sapiens};	96.03	0.024	708-873	306	ATPase

hSAMD9L NCBI Reference Sequence: NP\_001290425.1; HHpred – default parameters with HHblits, against NCBI\_Conserved\_Domains(CD)\_v3.16, PDB\_mmCIF70\_4\_Feb, Pfam-A\_v32.0 and SMART\_V6.0 databases. Sequence similarity results obtained with E-Value < 0.05. In bold proteins found functionally associated with SAMD9L.

**Supplementary Table 5. Microtubule associated motifs significantly present in microtubule associated proteins identified in hSAMD9L.**

Motif	Start	End	p-value * for significantly motif presence in microtubule interaction sites/domains
<b>S...LK<b>S</b></b>	619	626	0.009
Q...KR	1261	1266	0.023
R..P...SS	633	641	0.035
SKPS	93	96	0.036
P.SK L	1395	1399	0.042
T..P..P	508	514	0.047
T..P..P	951	957	0.047
H..R....S	686	694	0.048
PK....K	729	735	0.048
KR....K	711	717	0.048
KR....K	1464	1470	0.048
K..SS...S	1450	1458	0.048
K....KPE	152	159	0.048
SSL....R	1453	1460	0.049
E...E.RK	520	527	0.05

hSAMD9L NCBI Reference Sequence: NP\_001290425.1; MAPanalyzer was set to moderate threshold with 90% of sensitivity obtaining a significant combined score of 0.014 which classified hSAMD9L as a microtubule associated protein. In bold, the motif where c.1877C>T (p.Ser626Leu) locates, highlighted in red.

**Supplementary Table 6. Intrinsic disorder region (IDR) predicted around WT SAMD9L Ser626 and mutated p.Ser626Leu SAMD9L.**

Predictor algorithm	Ser626 SAMD9L Disorder region (DR)	Leu626 SAMD9L Disorder region (DR)
PredictProtein (PROFBval)	622-634	622-634
PredictProtein (Ucon)	626-627	627
DisEMBL (Hot loops)	625-650	626-650
InterPro (MobiDB-lite)	626-645	-

PredictProtein (PROFBval and Ucon), DisEMBL (Hot loops) and InterPro (MobiDB-lite) identified an intrinsic disorder (IDR) around the SAMD9L amino acid Ser626. p.Ser626Leu change lead to a prediction change for Ucon and Hot loops and abolish the identified prediction for InterPro (MobiDB-lite). hSAMD9L NCBI Reference Sequence: NP\_001290425.1 was used except for InterPro (MobiDB-lite) where 626-645 region was analysed.

**Supplementary Table 7. Related diseases and protein function of genes associated with microtubules and the endosomal/lysosomal pathway identified in the SAMD9L PPI network by STRING.**

Gene	Disease	Inheritance	Relevant clinical signs	Protein function
<b>ATL1</b>	Autosomal Dominant and Neuropathy, Hereditary Sensory, Type ID	AD	Hyperreflexia, Sensorimotor axonal neuropathy	Dynamain superfamily of large GTPases involved in cellular trafficking patterns
	Spastic Paraplegia 3	AD	Hyperreflexia, Decreased vibratory sense in lower limbs	
<b>SPAST</b>	Spastic Paraplegia 4	AD	Hyperreflexia, Decreased vibratory sense in the lower limbs	Activity in axonal transport and the maintenance of excitatory synapses of motor neurons in the brain cortex and the spinal cord
<b>RAB7A</b>	Charcot-Marie-Tooth disease, type 2B	AD	Hyporeflexia, marked distal sensory impairment	Key regulators of endosomal membrane traffic. functions in retrograde endosomal transport
<b>KATNB1</b>	Lissencephaly 6, with microcephaly	AR	Hyperreflexia	Required for the transport into neuronal processes by microtubule-dependent motor proteins.
<b>IFI44</b>	-	-	-	Aggregates to form microtubular structures
<b>KATNA1</b>	-	-	-	Required for the transport into neuronal processes by microtubule-dependent motor proteins.
<b>KATNAL1</b>	-	-	-	Katnal1 mutant mice revealed numerous morphological abnormalities and defects in neuronal migration and morphology.
<b>KATNAL2</b>	-	-	-	Severs microtubules in vitro in an ATP-dependent manner.
<b>EEA1</b>	-	-	-	Participates in endosomal trafficking
<b>RAB7B</b>	-	-	-	Mediate transport from endosomes to the TGN and/or Golgi, a step that is also needed to accomplish efficient delivery of lysosomal enzymes to the endocytic route
<b>RAB5A</b>	-	-	-	Involved in axonal and dendritic endocytosis, and participates in axonal endosome vesicles.
<b>VPS4B</b>	-	-	-	VPS4B belongs to the AAA (ATPases associated with diverse cellular activities) protein family and is involved in lysosomal/endosomal membrane trafficking
<b>VTA1</b>	-	-	-	Involved in trafficking of the multivesicular body, an endosomal compartment involved in sorting membrane proteins for degradation in lysosomes
<b>CHMP2A</b>	-	-	-	Probable core component of the endosomal sorting required for transport complex III (ESCRT-III) which is involved in multivesicular bodies (MVBs) formation and sorting of endosomal cargo proteins into MVBs.

Mutations in *ATL1*, *SPAST*, *RAB7A* and *KATNB1* have been previously associated with diseases presenting with hyperreflexia, sensory axonal neuropathy or sensory impairment among other clinical signs, overlapping with phenotypic features identified in this new spinocerebellar ataxia subtype. Proteins identified are mainly related to microtubules, cellular transport or to endosomal/lysosomal trafficking. Uniprot and GeneCards databases were used for the protein function description.



## References

1. Kimura J. *Electrodiagnosis in Diseases of Nerve and Muscle*. Oxford University Press; 2013.
2. Feller W, Teichmann T. An Introduction to Probability Theory and Its Applications, Vol. 2. *Phys Today*. 1967;20(5):76-76.
3. Marco-Sola S, Sammeth M, Guigó R, Ribeca P. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods*. 2012;9(12):1185-1188.
4. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
5. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
6. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012;6(2):80-92.
7. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
8. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat*. 2016;37(3):235-241.
9. Karczewski KJ, Weisburd B, Thomas B, et al. The ExAC browser: Displaying reference data information from over 60 000 exomes. *Nucleic Acids Res*. 2017;45(D1):D840-D845.
10. Sherry ST, Ward MH, Kholodov M, et al. DbSNP: The NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29(1):308-311.
11. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
12. Ng SB, Bigham AW, Buckingham KJ, et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*. 2010;42(9):790-793.
13. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc*. 2016;11(1):1-9.
14. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013;Chapter 7(1):Unit7.20.
15. González-Pérez A, López-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet*. 2011;88(4):440-449.
16. Shihab HA, Gough J, Cooper DN, et al. Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. *Hum Mutat*. 2013;34(1):57-65.

17. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* 2011;39(17):e118-e118.
18. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014;11(4):361-362.
19. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. de Brevern AG, ed. *PLoS One.* 2012;7(10):e46688.
20. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 2019;47(D1):D886-D894.
21. Sánchez I, Balagué E, Matilla-Dueñas A. Ataxin-1 regulates the cerebellar bioenergetics proteome through the GSK3 $\beta$ -mTOR pathway which is altered in Spinocerebellar ataxia type 1 (SCA1). *Hum Mol Genet.* 2016;25(18):4021-4040.
22. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.
23. Lennon G, Auffray C, Polymeropoulos M, Soares MB. The I.M.A.G.E consortium: An integrated molecular analysis of genomes and their expression. *Genomics.* 1996;33(1):151-152.
24. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
25. Lehle S, Hildebrand DG, Merz B, et al. LORD-Q: a long-run real-time PCR-based DNA-damage quantification method for nuclear and mitochondrial genome analysis. *Nucleic Acids Res.* 2014;42(6):e41.
26. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev.* 2015;82(7-8):518-529.
27. Valente A, Stuart J. A Generalized Image Workflow for Quantitative Analysis of Mitochondrial Networks. *Free Radic Biol Med.* 2017;112.
28. Cottingham RW, Idury RM, Schäffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet.* 1993;53(1):252-263.
29. Schäffer AA, Gupta SK, Shriram K, Cottingham RW. Avoiding recomputation in linkage analysis. *Hum Hered.* 1994;44(4):225-237.
30. Broman KW, Weber JL. Estimation of pairwise relationships in the presence of genotyping errors [6]. *Am J Hum Genet.* 1998;63(5):1563-1564.
31. Baltimore JO. Analysis of Human Genetic Linkage. *Ann Hum Genet.* 1986;50(1):101-102.
32. Liu X, Guo X, Niu L, et al. Atlastin-1 regulates morphology and function of endoplasmic reticulum in dendrites. *Nat Commun.* 2019;10(1).
33. McDermott CJ, Grierson AJ, Wood JD, et al. Hereditary spastic paraparesis:

- disrupted intracellular transport associated with spastin mutation. *Ann Neurol.* 2003;54(6):748-759.
34. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods.* 2001;25(4):402-408.