



OPEN ACCESS

EDITED BY

Etienne Herzog,
Centre National de la Recherche Scientifique
(CNRS), France

REVIEWED BY

Lydia Castelli,
The University of Sheffield, United Kingdom
Lara Marrone,
Evotec, Germany

*CORRESPONDENCE

Spyros Petrakis
✉ spetrak@certh.gr

RECEIVED 22 August 2023

ACCEPTED 16 November 2023

PUBLISHED 06 December 2023

CITATION

Gkekas I, Vagiona A-C, Pechlivanis N,
Kastrinaki G, Pliatsika K, Iben S,
Xanthopoulos K, Psomopoulos FE,
Andrade-Navarro MA and Petrakis S (2023)
Intranuclear inclusions of polyQ-expanded
ATXN1 sequester RNA molecules.
Front. Mol. Neurosci. 16:1280546.
doi: 10.3389/fnmol.2023.1280546

COPYRIGHT

© 2023 Gkekas, Vagiona, Pechlivanis,
Kastrinaki, Pliatsika, Iben, Xanthopoulos,
Psomopoulos, Andrade-Navarro and Petrakis.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Intranuclear inclusions of polyQ-expanded ATXN1 sequester RNA molecules

Ioannis Gkekas^{1,2}, Aimilia-Christina Vagiona³,
Nikolaos Pechlivanis¹, Georgia Kastrinaki⁴, Katerina Pliatsika^{1,2},
Sebastian Iben⁵, Konstantinos Xanthopoulos^{1,2},
Fotis E. Psomopoulos¹, Miguel A. Andrade-Navarro³ and
Spyros Petrakis^{1*}

¹Centre for Research and Technology Hellas, Institute of Applied Biosciences, Thessaloniki, Greece,

²Laboratory of Pharmacology, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece, ³Faculty of Biology, Johannes Gutenberg University Mainz, Mainz, Germany, ⁴Aerosol and Particle Technology Laboratory, Centre for Research and Technology Hellas, Chemical Process and Energy Resources Institute, Thessaloniki, Greece, ⁵Department of Dermatology and Allergic Diseases, University of Ulm, Ulm, Germany

Spinocerebellar ataxia type 1 (SCA1) is an autosomal dominant neurodegenerative disease caused by a trinucleotide (CAG) repeat expansion in the ATXN1 gene. It is characterized by the presence of polyglutamine (polyQ) intranuclear inclusion bodies (IIBs) within affected neurons. In order to investigate the impact of polyQ IIBs in SCA1 pathogenesis, we generated a novel protein aggregation model by inducible overexpression of the mutant ATXN1(Q82) isoform in human neuroblastoma SH-SY5Y cells. Moreover, we developed a simple and reproducible protocol for the efficient isolation of insoluble IIBs. Biophysical characterization showed that polyQ IIBs are enriched in RNA molecules which were further identified by next-generation sequencing. Finally, a protein interaction network analysis indicated that sequestration of essential RNA transcripts within ATXN1(Q82) IIBs may affect the ribosome resulting in error-prone protein synthesis and global proteome instability. These findings provide novel insights into the molecular pathogenesis of SCA1, highlighting the role of polyQ IIBs and their impact on critical cellular processes.

KEYWORDS

spinocerebellar ataxia type 1, ataxin-1, intranuclear inclusion bodies, RNA sequestration, RNA-seq, protein-protein interaction network, ribosome

Introduction

Spinocerebellar ataxia type 1 (SCA1) is a rare neurodegenerative disease belonging to the group of polyglutamine (polyQ) diseases. It is caused by trinucleotide (CAG) repeat expansions in the *ATXN1* gene resulting in the production of an abnormal polyQ tract in the ataxin-1 (ATXN1) protein (Chung et al., 1993). The length of the polyQ tract directly correlates with the age of onset and the severity of the disease (Orr et al., 1993). In healthy individuals, it typically ranges from 4 to 39 glutamines whereas in SCA1 patients it is significantly expanded, ranging from 40 to 83 glutamines. SCA1 is characterized by cerebellar atrophy, as well as degeneration of the brainstem and the spinal cord (Guerrini et al., 2004; Döhlinger et al., 2008; Pedrosa and Barsottini, 2013). Disease pathology involves loss of

Purkinje cells and reduction in the number of granule cells or other neuronal populations but may also extend beyond the cerebellum affecting other brain regions (Coffin et al., 2023).

ATXN1 is involved in transcriptional regulation and RNA splicing through its interaction with transcription factors and RNA-binding proteins (Mizutani et al., 2005; Tsuda et al., 2005; Zhang et al., 2020; Coffin et al., 2023). However, expansion of the polyQ tract in the mutant protein results in the formation of intranuclear inclusion bodies (IIBs), a striking feature of SCA1 (Cummings et al., 1999). ATXN1 IIBs sequester a variety of proteins, including proteasome and ribosome subunits, ubiquitin-related enzymes, kinases and heat-shock proteins, potentially impairing critical cellular functions and contributing to SCA1 pathology (Laidou et al., 2020). ATXN1 is also characterized by RNA-binding activity and may interact with multiple RNA targets. This feature is partially affected by the length of the polyQ tract (Yue et al., 2001; Chen et al., 2022); however, the extent to which this mechanism contributes to the pathogenesis of SCA1 remains to be elucidated.

Numerous findings support the notion that pathogenic protein inclusions contain RNA. RNA sequences have been found in inclusions derived from the hippocampus of Alzheimer's disease (AD) patients and their number was significantly enriched compared to preparations from healthy individuals (Shmookler Reis et al., 2021). Furthermore, tau inclusions, detected in AD and Pick's disease exhibit positive staining for RNA and are enriched in small nuclear and nucleolar RNAs (snRNAs and snoRNAs) (Ginsberg et al., 1997, 1998; Lester et al., 2021). Although significant research efforts have been made so far, the exact molecular composition and structure of protein inclusions associated with neurodegeneration remain poorly understood. The heterogeneity and complexity of these inclusions, along with the lack of efficient protocols for their isolation, present substantial challenges in characterizing their precise molecular organization.

Here, we generated human neuroblastoma SH-SY5Y cells with inducible overexpression of the SCA1-causing *ATXN1(Q82)* isoform and developed a protocol for the efficient isolation of polyQ-expanded ATXN1 IIBs. Fourier-transform infrared spectroscopy (FTIR) indicated that these IIBs are enriched in RNA, a finding that was further validated by RNA-seq. Protein interaction network (PIN) analysis aiming to identify perturbed cellular processes in SCA1 cells, indicated that the sequestration of critical RNA transcripts in ATXN1(Q82) IIBs may affect the ribosome resulting in error-prone protein synthesis. The identification of RNA molecules within polyQ-expanded ATXN1 IIBs could provide valuable insights for the molecular mechanisms underlying SCA1 disease pathology.

Materials and methods

Generation of Tet-On YFP-ATXN1(Q82) and Venus SH-SY5Y cells

Human neuroblastoma SH-SY5Y cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS, Biowest, France) and penicillin-streptomycin (1X, Biowest, France) at 37°C in a humidified atmosphere containing 5% CO₂. Tet-On YFP-ATXN1(Q82) SH-SY5Y cells were generated as previously described (Laidou et al., 2020). In brief, cells were

seeded in a 6-well plate (2×10^5 cells per well) and transfected with a mixture of pT2-CMV/TetO₂-YFP-ATXN1(Q82), pT2-TetR-neo^R (Laidou et al., 2020) and pCMV(CAT)T7-SB100 (Mátés et al., 2009) transposon plasmids at a 7:2:1 ratio using Xfect reagent (Clontech, USA). Venus SH-SY5Y cells were generated by transfection with pT2-GAGGS-Venus-neo^R transposon (Dori et al., 2017) and pCMV(CAT)T7-SB100 plasmids at a 5:1 ratio. Transfected cells were selected at day 4 post-transfection using 200 µg/mL G418 (InvivoGen, USA). For transgene induction, doxycycline (Dox, 2 µg/mL, Sigma-Aldrich, USA) was added to the culture medium and replenished every 48 h. Genetically modified cells were differentiated into neuron-like cells by culturing for 7 days in medium containing 10 µM retinoic acid (RA).

Fluorescence microscopy

Tet-On YFP-ATXN1(Q82) SH-SY5Y cells were seeded in a 24-well plate and cultured in the presence of Dox. Cells were fixed with 4% formaldehyde PBS for 10 min and permeabilized for 10 min with 0.1% Triton-X 100 (Sigma-Aldrich, USA) PBS. Nuclei were stained with DAPI for 5 min at room temperature. Fluorescent cells were observed in a ZOE Fluorescent cell imager equipped with three fluorescence channels and an integrated digital camera (Bio-Rad, USA).

Flow cytometry

Cells expressing the *YFP-ATXN1(Q82)* transgene were measured on a CYTEK NL-CLC flow cytometer (CYTEC, USA). For each sample, 30,000 events were recorded with a rate of 5,000 events per second. Data was acquired and analyzed with the SpectroFlo software.

Filter retardation assay

Cell extracts were mixed with an equal volume of a 4% SDS (AppliChem, Denmark), 100 mM DTT solution. Samples were heated at 95°C for 10 min, diluted with 100 µL 0.2% SDS and filtered through a 0.2 µm cellulose acetate membrane (Whatman, Merck). SDS-resistant inclusions retained on the membranes were detected using the anti-ATXN1 SA4645 antibody (1:1,000) (Petrakis et al., 2012).

Isolation of polyQ protein inclusions

PolyQ protein inclusions were purified from technical triplicates of Tet-On YFP-ATXN1(Q82) SH-SY5Y at day 5 post-induction. Cells (4×10^6) were harvested by trypsin-EDTA (Biowest, France), collected by centrifugation at 1,000 g for 5 min and washed twice with PBS. The resulting cell pellet was resuspended in hypotonic buffer containing 0.5% Triton-X 100 PBS and incubated on ice for 20 min. Intact nuclei were collected after centrifugation at 12,000 g for 10 min at 4°C. The resulting supernatant containing the cytoplasmic fraction was collected and stored at -20°C. The nuclear pellet was resuspended in

RIPA buffer supplemented with protease/phosphatase inhibitors (Thermo Fisher Scientific, Denmark) and benzonase (Calbiochem-Novagen, USA) and incubated on ice for 40 min. PolyQ inclusions were purified from the nuclear fraction using an Amicon Ultra-0.5 concentrator with a 100 kDa molecular weight cut-off (Merck) and concentrated to 20 μ L.

Immunoblotting

Samples from cells lysed in RIPA buffer containing protease/phosphatase inhibitors and benzonase, and cytoplasmic or nuclear fractions were resolved in a 10% SDS-PAGE and transferred to polyvinylidene difluoride (PVDF) membranes (Thermo Fisher Scientific, Denmark). Protein bands were detected with antibodies against ataxin-1 (SA4645), GFP (2956, Cell Signaling Technologies, USA), lamin A/C (4777, Cell Signaling Technologies, USA) and GAPDH (5171, Cell Signaling Technologies, USA) and visualized using either NBT/BCIP (AppliChem, Denmark) or Luminata Western HRP Chemiluminescence Substrate (Merck).

FTIR spectroscopy

YFP-ATXN1(Q82) protein inclusions were seeded on MirrIR low-e-glass slides (Kevley Technologies, USA) and dried in ambient conditions. Then glass slides were measured by ATR-FTIR method at a Jasco spectrometer (Jasco FTIR-6700, Japan). Spectra were collected in the range of 4,000–400 cm^{-1} by 180 scans with 4 cm^{-1} resolution and analyzed using Spectra Manager 2.15.12 software (Jasco Corporation, Japan).

RNA-sequencing

RNA was extracted from purified polyQ IIBs or control samples using TRIzol (Thermo Fisher Scientific, Denmark). The obtained RNA was quantified by NanoDrop (Thermo Fisher Scientific, Denmark) and its integrity was checked by Bioanalyzer 2100 (Agilent, USA). Sequencing libraries were prepared using the NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs, USA), according to manufacturer's instructions. Library preparations were QC checked by electrophoresis (Bioanalyzer 2100, Agilent, USA) and quantified by Qubit Fluorometer (Thermo Fisher Scientific, Denmark). Paired-end sequencing was performed in a NextSeq 500 Illumina platform using the NextSeq 500 System Mid-Output Kit (Illumina). RNA-seq data were deposited in the European Nucleotide Archive (ENA) under accession number PRJEB64635.

Data analysis and visualization

Quality control was performed based on the *FastQC* software with raw reads being trimmed and filtered using the *Trim Galore!* software suite. Trimmed reads were aligned against the *hg38* assembly of the human genome using the *HISAT2* protocol with an overall mapping rate ranging from 56.74 to 62.65% (Kim et al., 2019). Unmapped reads were later aligned against 45S,

18S, 5.8S, 5S rRNA, and tRNA resulting in a high mapping rate for 45S above 90%. The *Samtools* suite was used for the processing of the BAM files (Danecek et al., 2021). The final gene quantification (*hg38*) was performed using the *featureCounts* program (Liao et al., 2014). Subsequent data and statistical analysis were carried out in R v4.2.2.

Differential expression analysis was performed using the *DESeq2* package (Love et al., 2014). Genes with a *p*-value lower than 0.05 were considered significant. Enriched genes ($\log_{2}FC > 0$) were used for downstream analyses and visualized in a volcano plot. Principal component analysis (PCA) was performed using the *prcomp* function on the significantly enriched genes. Samples were projected on the resulting space of the first two principal components (PC1, PC2). Functional analysis was performed using the *pathfinder* package against the gene ontology resource (Cellular Component and Biological Process terms) (Ulgen et al., 2019). Specific enriched terms were visualized using the *ggplot2* library.

RT-qPCR

RT-qPCR was performed using Luna Universal One-Step RT-qPCR Kit (New England Biolabs, USA) in an AriaMx Real-Time PCR (qPCR) instrument (Agilent, USA). Primers were designed using NCBI Primer-BLAST or selected from the literature (Zhu et al., 2023) and their sequences are shown in [Supplementary Table 1](#). The correct size of amplified RT-qPCR products was verified by electrophoresis in a 2% agarose gel. Data were analyzed by the $2^{-\Delta\Delta C_t}$ method using *GAPDH* as a housekeeping gene.

Motif discovery

The transcript sequences (including 3',5' UTRs and introns) of genes were extracted from the Ensembl BioMart database, using the human genome assembly, GRCh38.p13 (Cunningham et al., 2022). Long sequences were split into 9,497 smaller overlapping fragments of 100nt to the adjacent regions. Transcripts were analyzed with the STREME tool of the MEME suite to discover patterns (ungapped motifs) that were enriched with respect to a control set generated by shuffling the input dataset (Bailey, 2021). The motif width was set between 3 and 30 nucleotides length. STREME applies a statistical test at *p*-value threshold < 0.05 to determine the enrichment of motifs in the input transcripts compared to the control set. The common motif sequences were visualized as logos using enoLOGOS. Identified motifs were scanned for putative binding sites in the CIS-BP RNA Database (Ray et al., 2013). Search was performed via the RNA Scan tool using the species parameter "Homo sapiens" and the motif model was set to the standard scoring system option which is position weight matrices (PWMs) – log-odds (Stormo, 1990). The log-odds threshold was set to its standard value of 6.

Construction of the hPIN and embedding into the hyperbolic space

The human protein interaction network (hPIN) is a subset of release 2.3 of the Human Integrated Protein–Protein Interaction

reference (HIPPIE). The raw version of this network is available in the Download section of the HIPPIE database (Schaefer et al., 2012; Alanis-Lobato et al., 2017). After discarding self-interactions and extracting the network's largest connected component (LCC), a hPIN was generated, consisting of 186,196 high-confidence interactions (score ≥ 0.71) between 15,587 proteins. Next, the hPIN was embedded into the two-dimensional hyperbolic plane using the R package "NetHypGeom" which implements the LaBNE + HM algorithm (Alanis-Lobato et al., 2016b). This algorithm combines manifold learning and maximum likelihood estimation to uncover the hidden geometry of complex networks (Papadopoulos et al., 2015; Alanis-Lobato et al., 2016a,b). The network was embedded into the H^2 to infer the hyperbolic coordinates of each protein, with parameters $\gamma = 2.97$, $T = 0.83$ and $w = 2\pi$.

SCA1 network in the hyperbolic space and clustering in the similarity dimension

The list of protein interactors was obtained from the HIPPIEv2.3 database and their position in the hPIN was determined (Schaefer et al., 2012; Alanis-Lobato et al., 2017). Groups of proteins were created in the network along the angular similarity dimension. To determine the start and the end of each group, proteins were sorted increasingly by their inferred angular coordinate (θ) and the difference between θ_i and θ_{i+1} was computed. The gap size (g) separating protein clusters produced sectors with a minimum of five components ($g = 0.0346$). Gene ontology (GO) (Ashburner et al., 2000) and KEGG pathway enrichment analysis (Kanehisa and Goto, 2000) for the proteins in each sector of the network was performed and enriched terms with p -value ≥ 0.05 were considered significant.

Error-prone protein synthesis

Tet-On YFP-ATXN1(Q82) or Venus SH-SY5Y cells were cultured in white 96-well plates (4×10^4 per well) and transfected with either wild-type or mutant (K529N) pCI-neo plasmids (500 ng per well) using Xfect reagent. The plasmids encode a dual reporter system, with the Renilla luciferase acting as the control reporter and the Firefly luciferase serving as the experimental reporter, both residing on a single pCI-neo plasmid. Luminescence produced by the Firefly and Renilla luciferases was quantified using a Dual-Glo assay kit (Promega, USA). Error-prone protein synthesis was determined based on the Firefly/Renilla ratio, as previously described (Alupei et al., 2018).

BisANS assay

Cells were harvested by trypsinization and pelleted by centrifugation at 1,000 g for 5 min. The resulting pellet was resuspended in TNE buffer (50 mM Tris HCl, 100 mM NaCl, 1 mM EDTA) and cells were lysed by sonication (3×30 seconds at 25% amplitude). Cell extracts were centrifugated at 14,000 g for 20 min at 4°C. Protein concentration in the supernatant was quantified using NanoDrop (Thermo Fisher Scientific, Denmark). Next, 100 μ g of protein was denatured by incubation in 2M urea for 2 h

at room temperature. Then, BisANS dye (Cayman Chemical) was added (30 μ M final concentration) and the mixture was incubated for 20 min to facilitate dye binding. Fluorescence was quantified with an excitation wavelength at 375 nm and emission data were recorded at 500 nm. Proteome instability was calculated by dividing the fluorescence signal after addition of BisANS to the background fluorescence of cells.

Statistical analysis

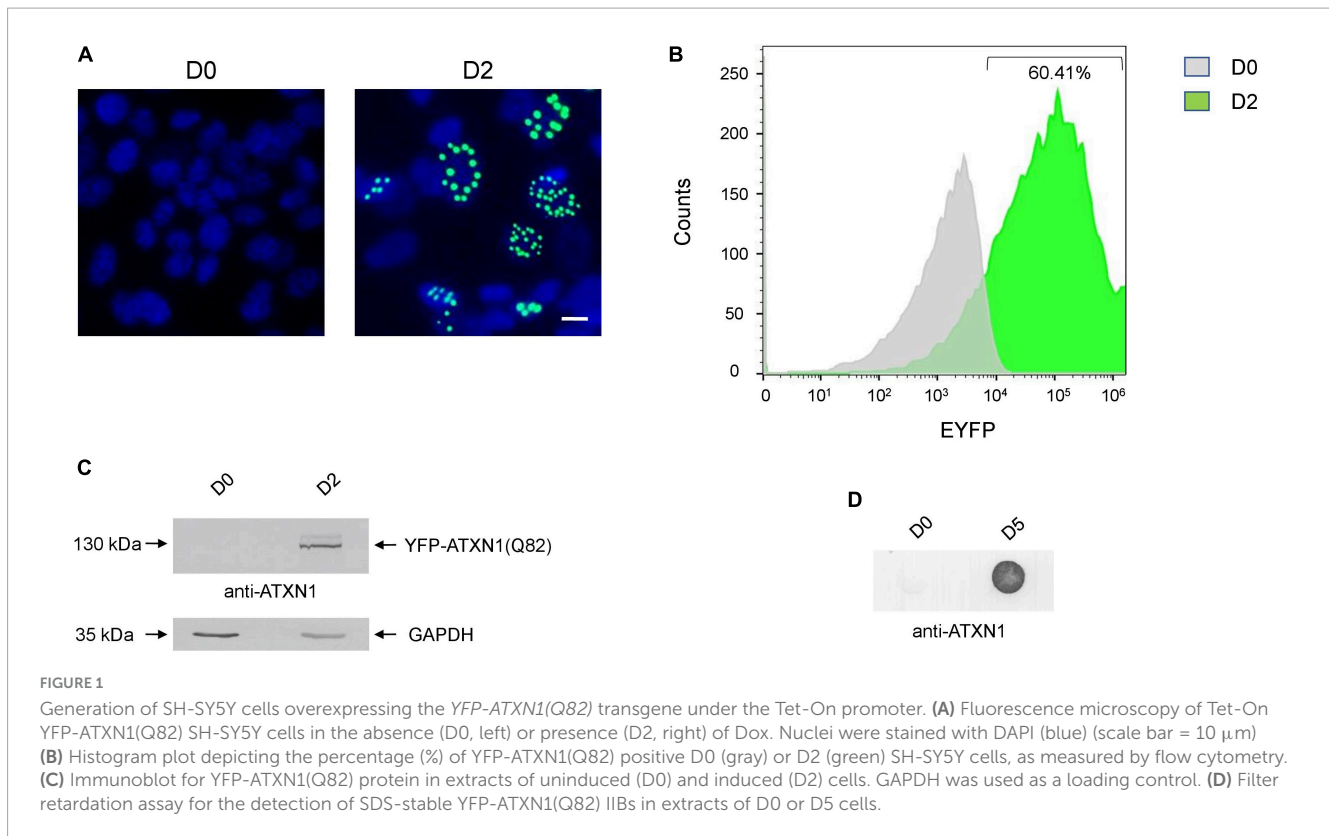
Statistical analysis was performed using the GraphPad Prism software v9 (San Diego, USA). All experiments were performed in triplicates and results are shown as mean \pm SD calculated by a t -test.

Results

Generation of Tet-On YFP-ATXN1(Q82) SH-SY5Y cells

SH-SY5Y cells expressing YFP-ATXN1(Q82) under the control of the Tet-On promoter were generated using the sleeping beauty transposon technology. After G418 selection, genetically modified cells were characterized. Induction of transgene expression for 48 h resulted in the formation of IIBs of YFP-ATXN1(Q82) protein, whereas no fluorescence was observed in the absence of Dox (Figure 1A). Flow cytometry indicated that at least 60.41% of Tet-On YFP-ATXN1(Q82) SH-SY5Y cells were strongly YFP-positive 48 h post-induction (Figure 1B) while this percentage increased to 78% by D5 of induction (Supplementary Figures 1A, B). At this time point, two distinct populations were observed which may correspond to cells expressing low- or high-copy numbers of the transgene (Mátés et al., 2009). Production of YFP-ATXN1(Q82) was verified by SDS-PAGE and immunoblotting analysis in cell extracts of uninduced and induced cells. As expected, the recombinant protein was detected at 130 kDa only in extracts of induced cells whereas no such protein was detected in uninduced cells (Figure 1C). Immunoblotting using anti-GFP antibody indicated that YFP-ATXN1(Q82) is not cleaved but is rather constantly produced as a fusion protein (Supplementary Figure 1C).

SCA1 is characterized by the formation of insoluble protein inclusions, mainly consisting of mutant ATXN1 (Zoghbi and Orr, 2009). The solubility of polyQ-expanded ATXN1 IIBs in Tet-On YFP-ATXN1(Q82) SH-SY5Y cells was assessed using a filter retardation assay. Insoluble inclusions of YFP-ATXN1(Q82) were readily immunodetected in extracts of Tet-On YFP-ATXN1(Q82) cells producing the mutant protein for 5 days; in contrast, no such inclusions were observed in uninduced cells (Figure 1D). Interestingly, fluorescent polyQ IIBs were also detected in neuron-like Tet-On YFP-ATXN1(Q82) SH-SY5Y cells differentiated in the presence of RA (Supplementary Figure 2A). These findings demonstrate that Tet-On YFP-ATXN1(Q82) SH-SY5Y cells accumulate insoluble IIBs of polyQ-expanded ATXN1, in consistency with other cellular models of SCA1 (Laidou et al., 2020).



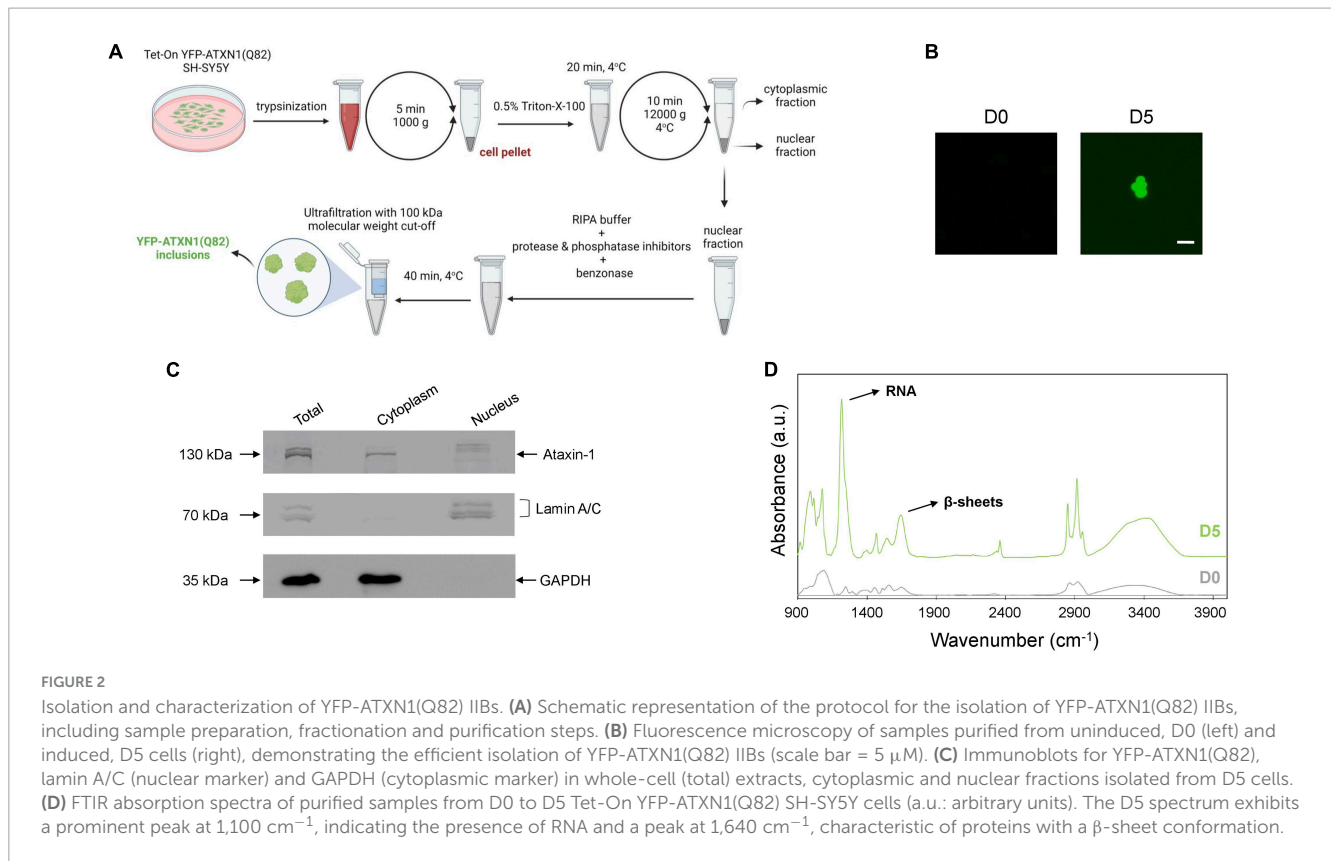
Isolation and characterization of insoluble IIBs of polyQ-expanded ATXN1

The precise cellular impact of ATXN1 inclusions remains poorly understood due to the difficulties associated with their isolation. Here, we developed a protocol for the isolation of insoluble IIBs from Tet-On *YFP-ATXN1(Q82)* SH-SY5Y cells. Insoluble *YFP-ATXN1(Q82)* IIBs were isolated from genetically modified cells at day 5 post-induction. The protocol consisted of discrete steps for: (1) the isolation of intact nuclei containing IIBs by centrifugation, (2) lysis of the pelleted nuclei and prolonged incubation with benzonase to remove residual nucleic acids, and (3) purification of insoluble IIBs using concentrators with high molecular-weight cut-off (Figure 2A). As shown in Figure 2B, this process reproducibly yielded large globular fluorescent inclusions of *YFP-ATXN1(Q82)*, validating the effectiveness of the experimental protocol. Similarly, insoluble polyQ IIBs were isolated from neuron-like Tet-On *YFP-ATXN1(Q82)* SH-SY5Y cells at D5 post-induction (Supplementary Figure 2B). In contrast, application of the same protocol in preparations from control (D0) samples did not result in the isolation of similar structures.

Next, we evaluated the efficiency of the purification protocol. Equal amounts of cytoplasmic and nuclear extracts (Figure 2A) were analyzed by SDS-PAGE while whole-cell extracts served as control. In order to assess the purity of the fractions, marker proteins known to specifically localize in the nucleus (lamin A/C) or the cytoplasm (GAPDH) were detected by immunoblots. As expected, both marker proteins were detected in whole-cell extracts (total fraction), whereas lamin A/C was detected only in the nuclear fraction and GAPDH only in the cytoplasmic fraction. These

results validate the efficient isolation of intact nuclei and the lack of contamination with cytoplasmic components. ATXN1 localizes in both the cytoplasm and the nucleus; therefore, recombinant *YFP-ATXN1(Q82)* was detected in both extracts. Interestingly, multiple bands were observed in whole-cell extracts from day 5 cells containing insoluble IIBs. The cytoplasmic fraction contained only a single band at approximately 130 kDa, corresponding to soluble *YFP-ATXN1(Q82)*, whereas distinct bands corresponding to higher molecular weight, partially insoluble forms of *YFP-ATXN1(Q82)* protein were observed in the nuclear fraction (Figure 2C).

Given the well-known propensity of the pathological ATXN1 to adopt a beta-sheet conformation, we employed FTIR to obtain structural insights on *YFP-ATXN1(Q82)* IIBs. First, absorption spectra (400–4,000 cm^{-1}) were collected from purified IIBs and control samples prepared from uninduced cells, using the previously established isolation protocol. This range of the FTIR spectrum includes the fundamental vibrations of the molecular bonds and provides information on the chemical composition and structure of the sample. Compared to control samples, the spectrum of *YFP-ATXN1(Q82)* IIBs indicated the presence of beta-sheets, as shown by a peak at 1,640 cm^{-1} corresponding to the amide I band. This observation is in agreement with a previous study showing that this peak characterizes the presence of beta-sheets in insoluble IIBs of *YFP-ATXN1(Q82)* (Laidou et al., 2020). Remarkably, a pronounced peak at approximately 1,100 cm^{-1} , corresponding to ribonucleic acid (RNA), was also detected (Figure 2D), suggesting that these inclusions may contain RNA molecules. This peak was not detected in control samples prepared using the same IIB isolation protocol. Furthermore, the use of the endonucleolytic enzyme benzonase, which specifically



degrades free DNA and RNA, excludes the possibility that this peak represents an artifact of the purification protocol or contamination with residual RNA. It rather reflects the presence of RNA, specifically associated with polyQ IIBs, that is protected from degradation by benzonase.

In order to validate this observation, we generated control overexpression SH-SY5Y cells stably producing Venus fluorescent protein, a variant of YFP (Nagai et al., 2002). Venus SH-SY5Y cells were uniformly fluorescent and produced the recombinant protein, as indicated by flow cytometry and immunoblotting (Supplementary Figure 3). These cells were further utilized for the isolation of RNA, according to the protocol described in Figure 2A. Extracted RNA was analyzed in a Bioanalyzer system providing a comprehensive view of RNA integrity and size distribution which indicated that no RNA can be isolated from Venus SH-SY5Y cells. In contrast, RNA with a similar pattern was isolated from polyQ IIBs of normal and neuron-like Tet-On YFP-ATXN1(Q82) SH-SY5Y cells (Supplementary Figure 2C).

Identification of RNA molecules bound on IIBs of polyQ-expanded ATXN1

In order to determine the identity of the RNA molecules bound on insoluble polyQ IIBs, RNA was isolated from YFP-ATXN1(Q82) IIBs (Q82), followed by RNA sequencing. To ensure the reliability of the results and to rule out the possibility of detecting non-specific products or artifacts, RNA was also isolated and sequenced from control samples (CTRL) which do not contain any inclusions (Figure 2B). Principal component analysis (PCA)

indicated distinct clustering of Q82 from CTRL samples, suggesting that the transcriptomes of these two groups were significantly different (Figure 3A). Computational analysis indicated 96 genes that were specifically enriched in polyQ IIBs compared to control samples (p -value < 0.05) (Supplementary Table 2). The top enriched genes, including *COL3A1*, *PXDN*, *FLNA*, *FN1*, and *MRC2* are shown in Table 1 and Figure 3B; their position in the volcano plot reflects their significance and magnitude of differential enrichment in polyQ IIBs. Furthermore, the presence of selected RNA transcripts (*COL3A1*, *PXDN*, *FLNA*, *FN1*, *MRC2*, and *FAT1*) in YFP-ATXN1(Q82) IIBs was validated by RT-qPCR (Figure 3C), whereas no such transcripts were detected in control preparations.

Gene ontology (GO) analysis of the 96 identified genes revealed significant enrichment for various biological processes. These include positive regulation of cell migration and miRNA transcription and negative regulation of gene expression. Additionally, the analysis identified several enriched cellular components, including collagen-containing extracellular matrix, actin cytoskeleton, basement membrane and chromatin (Figure 3D).

We then sought to identify common motif sequences in enriched RNA transcripts using the STREME tool, which enables the identification of common ungapped motifs ranging from 3–30 nucleotide (nt) length in transcript sequences. To this end, we utilized the transcript sequences of the top-44 highly significant genes (adj. p -value < 0.05). Out of 65 computationally identified motifs, three of them ranging from 13 to 25 nucleotides were significantly identified (based on their E-value) in the majority of enriched transcripts. In detail, the first motif was present in 38 out of the 44 transcripts while the second and the third motifs

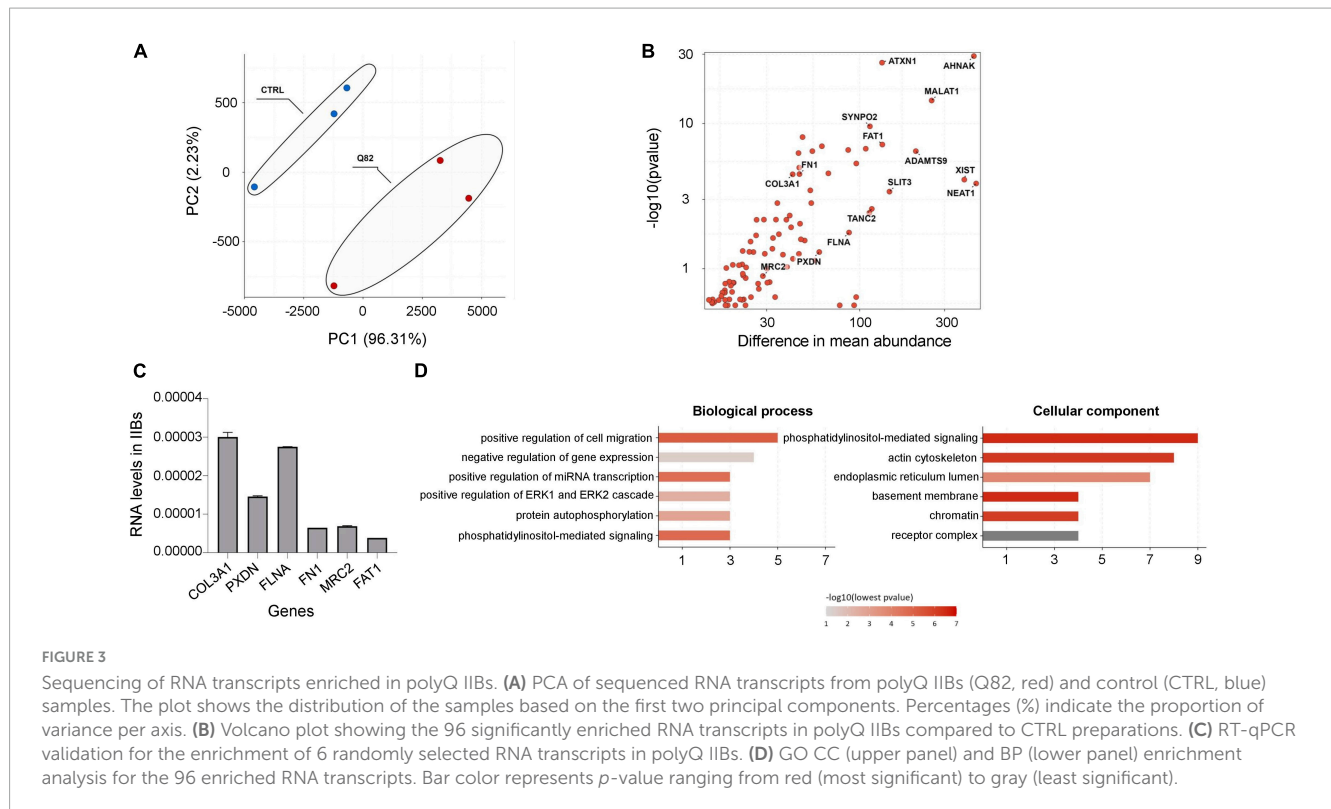


TABLE 1 Top enriched genes in YFP-ATXN1(Q82) IIBs.

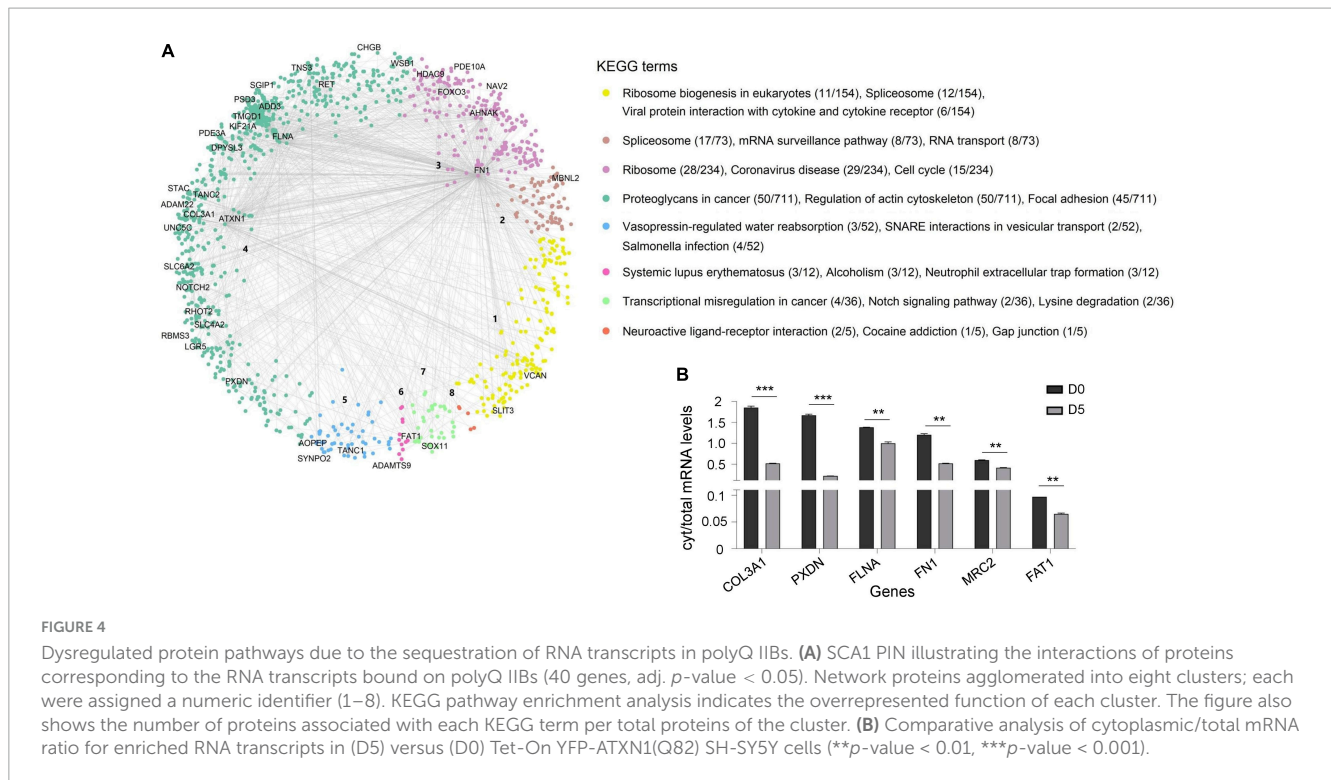
Gene symbol	Difference in mean abundance	p -value
NEAT1	451.36	< 0.0001
AHNAK	438.16	< 0.0001
XIST	387.15	< 0.0001
MALAT1	253.84	< 0.0001
ADAMTS9	206.82	< 0.0001
SLIT3	147.02	< 0.0001
FAT1	134.08	< 0.0001
ATXN1	133.59	< 0.0001
TANC2	117.03	< 0.0001
SYNP02	114.16	< 0.0001
FLNA	87.16	< 0.001
PXDN	59.29	< 0.01
FN1	45.88	< 0.0001
COL3A1	42.02	< 0.0001
MRC2	28.58	< 0.05

Selection criteria were $\log_2FC > 0$ and p -value < 0.05.

were detected in 39 out of the 44 transcripts. Sequence logos of these motifs, along with their E-value and the list of genes in which they were identified are shown in [Supplementary Figure 4](#) and [Supplementary Table 3](#). Interestingly, these motifs were also identified as potential binding sites for proteins belonging to the PUF, RRM, and KH families ([Supplementary Table 4](#)).

RNA sequestration on polyQ IIBs dysregulates protein complexes

RNA transcripts are typically synthesized in the nucleus and then exported to the cytoplasm for translation into proteins. Therefore, we hypothesized that sequestration of RNA on polyQ IIBs might disturb synthesis and stoichiometry of proteins, resulting in dysregulation of relevant protein complexes. To determine dysregulated processes in Tet-On YFP-ATXN1(Q82) SH-SY5Y cells, we generated a hPIN using as input interactions of human proteins corresponding to the RNA transcripts bound on polyQ IIBs. The resulting network was embedded into the two-dimensional hyperbolic plane H^2 . Then, a list of proteins interacting with the products of the 40/44 most significant (adj. p -value < 0.05) protein coding genes was obtained; four RNA genes that did not encode a protein product were excluded from downstream analysis. The resulting SCA1 PIN consisted of 1,277 proteins, which were connected by 1,482 interactions ([Figure 4A](#)). The similarity component of the PS model (angular coordinates of nodes in H^2) abstracts the characteristics that make a node similar to neighboring proteins participating in related biological process ([Härtner et al., 2018](#); [Vagiona et al., 2022](#)). Functional modularity of the generated SCA1 PIN, was detected by the presence of big gaps between consecutive inferred angles. Therefore, protein components of the network agglomerated into 8 clusters. The overrepresented function of each cluster was determined by KEGG pathway enrichment analysis, highlighting the functional heterogeneity of the sectors and the similarity-based protein agglomeration. Indicatively, proteins of cluster 1 were found to participate in ribosome biogenesis while spliceosome, mRNA surveillance and RNA transport were the



majorly enriched pathways of cluster 2. Proteins participating in cluster 3 and 4 were highly associated with the ribosome and focal adhesion (Figure 4A), in agreement with previous studies reporting perturbation of the protein synthesis machinery and extracellular matrix remodeling in SCA1 (Lee et al., 2011b; Vagiona et al., 2020).

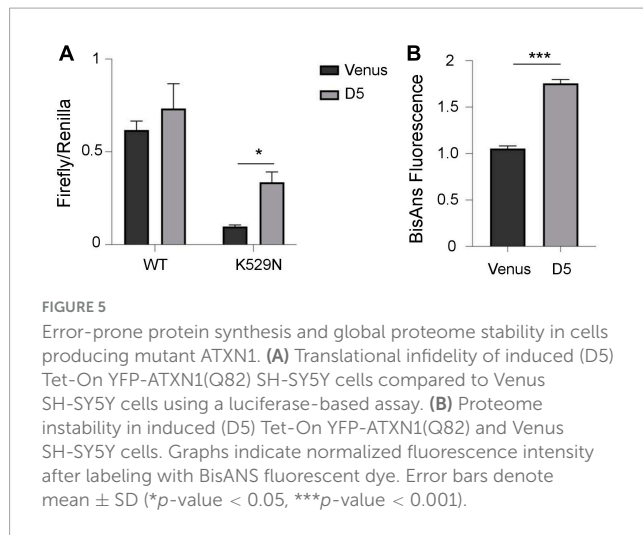
We then sought to validate the partial cytoplasmic depletion of RNA transcripts corresponding to critical nodes of the dysregulated SCA1 PIN. Six genes were selected that either represented top-linked nodes (*FN1*, *FLNA*, and *FAT1* in clusters 3, 4 and 7, respectively) or were significantly enriched in ATXN1(Q82) IIBs (*COL3A1*, *PXDN*, and *MRC2*) (Supplementary Table 5). A comparative analysis of cytoplasmic/total mRNA levels was performed in Tet-On YFP-ATXN1(Q82) SH-SY5Y cells containing IIBs (D5) compared to control (D0) cells. As shown in Figure 4B, the cytoplasmic/total mRNA ratio of all selected genes was significantly decreased in D5 versus D0 cells, which did not contain inclusions. No such reduction was observed for the housekeeping gene *GAPDH* (Supplementary Figure 5). Additionally, total expression of *FN1*, *FLNA*, *FAT1*, *COL3A1*, *PXDN*, and *MRC2* genes was not down-regulated in D5 vs. D0 cells (Supplementary Figure 6), indicating the specificity of the previous findings. These results support our hypothesis that sequestration of RNA transcripts within polyQ-expanded ATXN1 IIBs affects their cytoplasmic availability and transcription, contributing to the functional dysregulation of specific protein complexes.

PolyQ IIBs affect ribosomal activity

Multiple lines of evidence indicate that aggregation of mutant polyQ proteins disrupts the protein synthesis machinery (Laidou et al., 2020; Eshraghi et al., 2021). This notion is consistent

with network analysis highlighting the dysregulation of the ribosome and its biogenesis in our cell model (clusters 1 and 3 of the hPIN shown in Figure 4A). To address this prediction, we assessed the elongation activity of RNA polymerase I, related to the pre-rRNA abundance of 5.8S/ITS2 and 28S/ETS, and the processing of the primary pre-rRNA transcript by measuring ITS1 levels. RT-qPCR analysis indicated significantly lower levels of all markers in Tet-On YFP-ATXN1(Q82) SH-SY5Y at D5 vs. Venus cells (Supplementary Figure 7); these disturbances might be indicative of problems with ribosome biogenesis and assembly, as described in the literature (Zhu et al., 2023).

Ribosomes decode the genetic information of mRNA and synthesize proteins by accurately matching each RNA codon to the corresponding amino acid. Errors due to ribosomal dysfunction may cause incorporation of incorrect amino acids into proteins, resulting in error-prone protein synthesis (Kapur and Ackerman, 2018). Therefore, we assessed translation accuracy in cells containing polyQ-expanded ATXN1 IIBs using a luciferase activity assay (Alupej et al., 2018). This assay detects the error rate of translation at the ribosome through a reporter system with a defined mutation in the active center of firefly luciferase (K529N). Correct translation due to the incorporation of asparagine results in minimum luciferase activity. However, incorrect, error prone translation re-activates luciferase activity by the random incorporation of the activating amino acid lysine. Activity of exogenous wild-type luciferase was similar in both induced (D5) and uninduced (D0) cells, as expected. However, the activity of mutant luciferase harboring the K529N point-mutation was partially restored only in D5 Tet-On YFP-ATXN1(Q82) SH-SY5Y (Supplementary Figure 8). Next, we compared mutant luciferase activity in D5 and Venus cells. Again, the activity of mutant luciferase was significantly higher in D5 cells, suggesting



that the restoration of its function is not generically affected by overexpression of a random transgene in SH-SY5Y cells (Figure 5A).

Finally, we quantified global proteome stability in Tet-On YFP-ATXN1(Q82) SH-SY5Y by measuring the amount of exposed hydrophobic side chains, as previously described (Alupej et al., 2018). A significant elevation of misfolded proteins was observed in D5 vs. Venus cells (Figure 5B), suggesting that the proteome of these cells might be destabilized. Collectively, these results indicate the occurrence of error-prone protein synthesis in cells producing mutant ATXN1, directly affecting the quality of protein synthesis. In conclusion, here we show that sequestration of RNA transcripts within insoluble IIBs of polyQ-expanded ATXN1 affects various cellular processes and potentially, the assembly and function of the ribosome, the protein synthesis machinery.

Discussion

A cell model allowing the isolation of ATXN1(Q82) IIBs

Cell models provide valuable insights on the pathogenesis of neurodegenerative disorders, including SCA1. Induced pluripotent stem cells (iPSCs) derived from patient material enable the generation of disease-relevant cell types, facilitating detailed investigations and experimentation (Buijsen et al., 2018; Jin and Nan, 2023). However, reprogramming of somatic cells usually results in developmental reset and partial rejuvenation which may disrupt the aggregation of polyQ-expanded proteins in disease-specific neurons.

In an alternative approach for disease modeling, we overexpressed polyQ-expanded ATXN1 in SH-SY5Y cells, which belong to the neuronal lineage, and demonstrate a relevant gene expression pattern. Tet-On YFP-ATXN1(Q82) SH-SY5Y cells allow the controlled expression of the transgene and most importantly, accumulate insoluble IIBs, which characterize SCA1. The ability to monitor the formation of IIBs in real-time further enhances the utility of this model for investigating SCA1 pathogenesis. Of note, our cell-model has certain limitations. Compared to patient-derived cellular models, it lacks the potential to differentiate

into disease-relevant neuronal subtypes while overexpression of the polyQ-expanded ATXN1 isoform may augment pathogenic phenotypes. On the other hand, our cell model can be used for the reproducible isolation of polyQ IIBs, also from neuron-like cells, enabling the study of their structure and composition. Existing protocols for the purification of pathogenic inclusions from human disease material are inefficient and lack reproducibility.

Sequestration of RNA transcripts in polyQ IIBs affects important cellular processes

Compelling evidence substantiate the sequestration of RNA molecules within pathological aggregates associated with neurodegeneration. Indeed, protein aggregates characterizing Alzheimer's and Pick's disease were previously shown to contain RNA molecules (Ginsberg et al., 1997, 1998; Shmookler Reis et al., 2021). Furthermore, RNA binding drives the conformational change of tau and the oligomerization of TDP-43, associated with Alzheimer's disease and FTD/ALS, respectively (McMillan et al., 2023; Pérez-Berlanga et al., 2023). Despite the well-established role of ATXN1 as an RNA-binding protein, the interplay between RNA and polyQ IIBs remains largely unknown. Here, we documented the sequestration of RNA transcripts within ATXN1 polyQ IIBs, highlighting the potential involvement of RNA in the pathogenesis of SCA1.

The presence of RNA molecules within polyQ IIBs raises intriguing questions regarding their composition and function. Motif analysis indicated shared sequences among enriched RNA transcripts suggesting their specific sequestration into polyQ IIBs. These motifs may be recognized by ATXN1 and be partially responsible for the global dysregulation of gene expression in SCA1 models. Interestingly, the top identified motif is strikingly similar to the binding site (UGUAAUC) of proteins belonging to the PUF family. It has been previously shown that PUM1 is neuroprotective in cell models of SCA1 (Petrakis et al., 2012) while its haploinsufficiency causes SCA1-like neurodegeneration (Petrakis et al., 2012; Gennarino et al., 2015). PUM1 is also an RNA-binding protein and a known interactor of ATXN1 (Petrakis et al., 2012). Whether mutant ATXN1 and PUM1 proteins compete for binding to the same RNA molecules, along with the potential role of such a competition for SCA1 pathogenesis remains to be investigated.

Among the top statistically significant transcripts, we identified NEAT1, AHNAK, and SYNPO2 as potential contributors to SCA1 neurodegeneration. NEAT1 regulates the nuclear architecture and the formation of paraspeckles, nuclear bodies that are involved in the regulation of gene expression. Although the precise role of NEAT1 in neurodegeneration remains elusive, its abnormal expression has been associated with various conditions, including amyotrophic lateral sclerosis (ALS), Alzheimer's, Parkinson's, and Huntington's disease (Li and Wang, 2023). AHNAK encodes a nucleoprotein participating in intracellular calcium homeostasis, a process that is disrupted in polyQ diseases (Giacomello et al., 2013). SYNPO2 is involved in the maintenance of the neuronal cytoskeleton, contributing to the stability of synaptic connections. Dysregulation of cytoskeletal dynamics and impaired synaptic

function are common features in polyQ diseases (Lee et al., 2011a; Lim et al., 2017).

Gene ontology (GO) analysis of the 96 transcripts identified in polyQ IIBs provides novel insights for dysregulated cellular processes. Interestingly, computational analysis indicated that these genes participate in positive regulation of miRNA transcription, a regulatory process governing the production of microRNAs, which modulate gene expression; their dysregulation is a common feature of neurodegenerative diseases, including HD and SCA17 (Roshan et al., 2017; Catanesi et al., 2020; Tung et al., 2021). In terms of cellular components, this analysis highlighted the enrichment of genes associated with chromatin suggesting their potential involvement in chromatin remodeling and gene regulation, which may contribute to the dysregulation of gene expression in polyQ diseases.

We further hypothesized that RNA sequestration in polyQ IIBs may affect the function of critical protein complexes for SCA1 progression. Consequently, we focused on the ribosome which is frequently dysregulated in neurodegenerative diseases. We showed that sequestration of RNA transcripts in insoluble polyQ IIBs induces error-prone protein synthesis and enhances proteome instability, potentially affecting global protein synthesis, as previously documented for various polyQ diseases (Hübener et al., 2011). Our observation is in agreement with other studies reporting that mutant huntingtin stalls ribosomes and represses protein synthesis (Eshraghi et al., 2021) or indicating down-regulation of ribosomal proteins in SCA1 (Laidou et al., 2020).

ATXN1 is a widely expressed protein which normally, translocates between the cytoplasm and the nucleus where it binds RNA molecules. However, polyQ-expanded ATXN1 is not capable of nuclear export and its nuclear retention, along with its bound RNA molecules, may be partially responsible for its cytotoxicity (Irwin et al., 2005). Such pathogenic interplays may be widely common among neurodegenerative diseases. For example, the aggregation of TDP-43, an RNA-binding protein which is involved in ALS is modulated by its bound RNA molecules (Louka et al., 2020), highlighting its significance for disease progression. Concerning SCA1, neuronal vulnerability is due to complex mechanisms in different brain regions which are characterized by differential gene expression patterns (Coffin et al., 2023). The sequestration of RNA transcripts on polyQ IIBs, as shown here, may contribute to region-specific defects. In brief, different RNA transcripts may be sequestered on polyQ IIBs and trapped to the nucleus in the various brain regions. The impact of such events in SCA1 pathogenesis remains to be investigated.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in this article/[Supplementary material](#).

Author contributions

IG: Conceptualization, Methodology, Validation, Visualization, Writing – original draft. A-CV: Conceptualization, Formal analysis,

Visualization, Writing – original draft. NP: Conceptualization, Formal analysis, Visualization, Writing – review and editing. GK: Conceptualization, Methodology, Visualization, Writing – original draft. KP: Methodology, Formal analysis, Visualization. SI: Methodology, Writing – review and editing. KX: Conceptualization, Writing – review and editing. FP: Conceptualization, Formal analysis, Writing – review and editing. MA-N: Conceptualization, Formal analysis, Writing – review and editing. SP: Conceptualization, Methodology, Formal analysis, Investigation, Supervision, Writing – original draft.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by: (1) the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Innovation (GSRI) under the 1st Call for research projects to support Postdoctoral Researchers (Grant Agreement No. 122), (2) the HFRI under the 3rd Call for HFRI PhD Fellowships (Fellowship Number: 6654), and (3) the STREAMLINE project funded by Horizon Europe programme Widening Participation and Spreading Excellence (Grant Agreement No. 101060201).

Acknowledgments

We wish to thank Dr. Zsuzsanna Izsvák and Dr. Zoltán Ivics for kindly providing components of the SB100X transposon system, the Gorbunova group of the University of Rochester for providing the wild-type and mutant (K529N) pCl-neo plasmids and Dr. Marina Gerousi for technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnmol.2023.1280546/full#supplementary-material>

References

- Alanis-Lobato, G., Andrade-Navarro, M. A., and Schaefer, M. H. (2017). HIPPIE v2.0: Enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res.* 45, D408–D414. doi: 10.1093/nar/gkw985
- Alanis-Lobato, G., Mier, P., and Andrade-Navarro, M. A. (2016b). Manifold learning and maximum likelihood estimation for hyperbolic network embedding. *Appl. Netw. Sci.* 1:10. doi: 10.1007/s41109-016-0013-0
- Alanis-Lobato, G., Mier, P., and Andrade-Navarro, M. A. (2016a). Efficient embedding of complex networks to hyperbolic space via their Laplacian. *Sci. Rep.* 6:30108. doi: 10.1038/srep30108
- Alupe, M. C., Maity, P., Esser, P. R., Krikki, I., Tuorto, F., Parlato, R., et al. (2018). Loss of proteostasis is a pathomechanism in cockayne syndrome. *Cell. Rep.* 23, 1612–1619. doi: 10.1016/j.celrep.2018.04.041
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Bailey, T. L. (2021). STREME: Accurate and versatile sequence motif discovery. *Bioinformatics* 37, 2834–2840. doi: 10.1093/bioinformatics/btab203
- Buijssen, R. A. M., Gardiner, S. L., Bouma, M. J., van der Graaf, L. M., Boogaard, M. W., Pepers, B. A., et al. (2018). Generation of 3 spinocerebellar ataxia type 1 (SCA1) patient-derived induced pluripotent stem cell lines LUMCi002-A, B, and C and 2 unaffected sibling control induced pluripotent stem cell lines LUMCi003-A and B. *Stem Cell Res.* 29, 125–128. doi: 10.1016/j.scr.2018.03.018
- Catanesi, M., D'angelo, M., Tupone, M. G., Benedetti, E., Giordano, A., Castelli, V., et al. (2020). Micromas dysregulation and mitochondrial dysfunction in neurodegenerative diseases. *Int. J. Mol. Sci.* 21:5986. doi: 10.3390/ijms21175986
- Chen, J. M., Chen, S. K., Jin, P. P., and Sun, S. C. (2022). Identification of the ataxin-1 interaction network and its impact on spinocerebellar ataxia type 1. *Hum. Genomics* 16:29. doi: 10.1186/s40246-022-00404-0
- Chung, M. Y., Ranum, L. P. W., Duvick, L. A., Servadio, A., Zoghbi, H. Y., and Orr, H. T. (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat. Genet.* 5, 254–258. doi: 10.1038/ng1193-254
- Coffin, S. L., Durham, M. A., Nitschke, L., Xhako, E., Brown, A. M., Revelli, J. P., et al. (2023). Disruption of the ATXN1-CIC complex reveals the role of additional nuclear ATXN1 interactors in spinocerebellar ataxia type 1. *Neuron* 111, 481.e8–492.e8. doi: 10.1016/j.neuron.2022.11.016
- Cummings, C. J., Orr, H. T., and Zoghbi, H. Y. (1999). Progress in pathogenesis studies of spinocerebellar ataxia type 1. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 1079–1081. doi: 10.1098/rstb.1999.0462
- Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., et al. (2022). Ensembl 2022. *Nucleic Acids Res.* 50, D988–D995. doi: 10.1093/nar/gkab1049
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. doi: 10.1093/gigascience/giab008
- Döhlinger, S., Hauser, T. K., Borkert, J., Luft, A. R., and Schulz, J. B. (2008). Magnetic resonance imaging in spinocerebellar ataxias. *Cerebellum* 7, 204–214. doi: 10.1007/s12311-008-0025-0
- Dori, I., Petrakis, S., Giannakopoulou, A., Bekiari, C., Grivas, I., Siska, E. K., et al. (2017). Seven days post-injury fate and effects of genetically labeled adipose-derived mesenchymal cells on a rat traumatic brain injury experimental model. *Histol. Histopathol.* 32, 1041–1055. doi: 10.14670/HH-11-864
- Eshraghi, M., Karunadharma, P. P., Blin, J., Shahani, N., Ricci, E. P., Michel, A., et al. (2021). Mutant Huntingtin stalls ribosomes and represses protein synthesis in a cellular model of Huntington disease. *Nat. Commun.* 12:1461. doi: 10.1038/s41467-021-21637-y
- Gennarino, V. A., Singh, R. K., White, J. J., De Maio, A., Han, K., Kim, J.-Y., et al. (2015). Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels. *Cell* 160, 1087–1098. doi: 10.1016/j.cell.2015.02.012
- Giacomello, M., Oliveros, J. C., Naranjo, J. R., and Carafoli, E. (2013). Neuronal Ca²⁺ dyshomeostasis in Huntington disease. *Prion* 7, 76–84. doi: 10.4161/pri.23581
- Ginsberg, S. D., Crino, P. B., Lee, V. M. Y., Eberwine, J. H., and Trojanowski, J. Q. (1997). Sequestration of RNA in Alzheimer's disease neurofibrillary tangles and senile plaques. *Ann. Neurol.* 41, 200–209. doi: 10.1002/ana.410410211
- Ginsberg, S. D., Galvin, J. E., Chiu, T. S., Lee, V. M. Y., Masliah, E., and Trojanowski, J. Q. (1998). RNA sequestration to pathological lesions of neurodegenerative diseases. *Acta Neuropathol.* 96, 487–494. doi: 10.1007/s004010050923
- Guerrini, L., Lolli, F., Ginestrone, A., Belli, G., Della Nave, R., Tessa, C., et al. (2004). Brainstem neurodegeneration correlates with clinical dysfunction in SCA1 but not in SCA2. A quantitative volumetric, diffusion and proton spectroscopy MR study. *Brain* 127, 1785–1795. doi: 10.1093/brain/awh201
- Härtner, F., Andrade-Navarro, M. A., and Alanis-Lobato, G. (2018). Geometric characterisation of disease modules. *Appl. Netw. Sci.* 3:10. doi: 10.1007/s41109-018-0066-3
- Hübener, J., Vauti, F., Funke, C., Wolburg, H., Ye, Y., Schmidt, T., et al. (2011). N-terminal ataxin-3 causes neurological symptoms with inclusions, endoplasmic reticulum stress and ribosomal dislocation. *Brain* 134, 1925–1942. doi: 10.1093/brain/awr118
- Irwin, S., Vandelft, M., Pinchev, D., Howell, J. L., Graczyk, J., Orr, H. T., et al. (2005). RNA association and nucleocytoplasmic shuttling by ataxin-1. *J. Cell. Sci.* 118, 233–242. doi: 10.1242/jcs.01611
- Jin, Y., and Nan, G. (2023). Generation of induced pluripotent stem cell(iPSC)line CJUHi001-A derived peripheral blood mononuclear cells of spinocerebellar ataxia type 1(SCA1) the CAG repeat mutation in ATXN1 gene. *Stem Cell Res.* 66:102975. doi: 10.1016/j.scr.2022.102975
- Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kapur, M., and Ackerman, S. L. (2018). mRNA translation gone awry: Translation fidelity and neurological disease. *Trends Genet.* 34, 218–231. doi: 10.1016/j.tig.2017.12.007
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Laidou, S., Alanis-Lobato, G., Pribyl, J., Raskó, T., Tichy, B., Mikulasek, K., et al. (2020). Nuclear inclusions of pathogenic ataxin-1 induce oxidative stress and perturb the protein synthesis machinery. *Redox Biol.* 32:101458. doi: 10.1016/j.redox.2020.101458
- Lee, Y., Fryer, J. D., Kang, H., Crespo-Barreto, J., Bowman, A. B., Gao, Y., et al. (2011b). ATXN1 protein family and CIC regulate extracellular matrix remodeling and lung alveolarization. *Dev. Cell.* 21, 746–757. doi: 10.1016/j.devcel.2011.08.017
- Lee, S. B., Bagley, J. A., Lee, H. Y., Jan, L. Y., and Jan, Y. N. (2011a). Pathogenic polyglutamine proteins cause dendrite defects associated with specific actin cytoskeletal alterations in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16795–16800. doi: 10.1073/pnas.1113573108
- Lester, E., Ooi, F. K., Bakkar, N., Ayers, J., Woerman, A. L., Wheeler, J., et al. (2021). Tau aggregates are RNA-protein assemblies that mislocalize multiple nuclear speckle components. *Neuron* 109, 1675.e9–1691.e9. doi: 10.1016/j.neuron.2021.03.026
- Li, K., and Wang, Z. (2023). lncRNA NEAT1: Key player in neurodegenerative diseases. *Ageing Res. Rev.* 86:101878. doi: 10.1016/j.arr.2023.101878
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656
- Lim, R. G., Salazar, L. L., Wilton, D. K., King, A. R., Stocksdale, J. T., Sharifabad, D., et al. (2017). Developmental alterations in Huntington's disease neural cells and pharmacological rescue in cells and mice. *Nat. Neurosci.* 20, 648–660. doi: 10.1038/nn.4532
- Louka, A., Zacco, E., Temussi, P. A., Tartaglia, G. G., and Pastore, A. (2020). RNA as the stone guest of protein aggregation. *Nucleic Acids Res.* 48, 11880–11889. doi: 10.1093/nar/gkaa822
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Mátés, L., Chuah, M. K. L., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., et al. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* 41, 753–761. doi: 10.1038/ng.343
- McMillan, P. J., Benbow, S. J., Uhrich, R., Saxton, A., Baum, M., Strovast, T., et al. (2023). Tau-RNA complexes inhibit microtubule polymerization and drive disease-relevant conformation change. *Brain* 146, 3206–3220. doi: 10.1093/brain/awad032
- Mizutani, A., Wang, L., Rajan, H., Vig, P. J. S., Alaynick, W. A., Thaler, J. P., et al. (2005). Boat, an AXH domain protein, suppresses the cytotoxicity of mutant ataxin-1. *EMBO J.* 24, 3339–3351. doi: 10.1038/sj.emboj.7600785
- Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 20, 87–90. doi: 10.1038/nbt0102-87
- Orr, H. T., Chung, M., Banfi, S., Kwiatkowski, T. J., Servadio, A., Beaudet, A. L., et al. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* 4, 221–226. doi: 10.1038/ng0793-221
- Papadopoulos, F., Aldecoa, R., and Krioukov, D. (2015). Network geometry inference using common neighbors. *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.* 92:022807. doi: 10.1103/PhysRevE.92.022807
- Pedroso, J. L., and Barsottini, O. G. P. (2013). Spinal cord atrophy in spinocerebellar ataxia type 1. *Arq. Neuropsiquiatr.* 71:977. doi: 10.1590/0004-282X20130187

- Pérez-Berlanga, M., Wiersma, V. I., Zbinden, A., De Vos, L., Wagner, U., Foglieni, C., et al. (2023). Loss of TDP-43 oligomerization or RNA binding elicits distinct aggregation patterns. *EMBO J.* 42:e111719. doi: 10.15252/embj.2022111719
- Petrakis, S., Raskó, T., Russ, J., Friedrich, R. P., Stroedicke, M., Riechers, S. P., et al. (2012). Identification of human proteins that modify misfolding and proteotoxicity of pathogenic Ataxin-1. *PLoS Genet.* 8:e1002897. doi: 10.1371/journal.pgen.1002897
- Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177. doi: 10.1038/nature12311
- Roshan, R., Choudhary, A., Bhambri, A., Bakshi, B., Ghosh, T., and Pillai, B. (2017). microRNA dysregulation in polyglutamine toxicity of TATA-box binding protein is mediated through STAT1 in mouse neuronal cells. *J. Neuroinflamm.* 14:155. doi: 10.1186/s12974-017-0925-3
- Schaefer, M. H., Fontaine, J. F., Vinayagam, A., Porras, P., Wanker, E. E., and Andrade-Navarro, M. A. (2012). Hippie: Integrating protein interaction networks with experiment based quality scores. *PLoS One* 7:e31826. doi: 10.1371/journal.pone.0031826
- Shmookler Reis, R. J., Atluri, R., Balasubramaniam, M., Johnson, J., Ganne, A., and Ayyadevara, S. (2021). “Protein aggregates” contain RNA and DNA, entrapped by misfolded proteins but largely rescued by slowing translational elongation. *Aging Cell* 20:e13326. doi: 10.1111/acel.13326
- Stormo, G. D. (1990). Consensus patterns in DNA. *Methods Enzymol.* 183, 211–221. doi: 10.1016/0076-6879(90)83015-2
- Tsuda, H., Jafar-Nejad, H., Patel, A. J., Sun, Y., Chen, H.-K., Rose, M. F., et al. (2005). The AXH domain of Ataxin-1 mediates neurodegeneration through its interaction with Gfi-1/Senseless proteins. *Cell* 122, 633–644. doi: 10.1016/j.cell.2005.06.012
- Tung, C. W., Huang, P. Y., Chan, S. C., Cheng, P. H., and Yang, S. H. (2021). The regulatory roles of microRNAs toward pathogenesis and treatments in Huntington’s disease. *J. Biomed. Sci.* 28:59. doi: 10.1186/s12929-021-00755-1
- Ulgen, E., Ozisik, O., and Sezerman, O. U. (2019). pathfindR: An R package for comprehensive identification of enriched pathways in omics data through active subnetworks. *Front. Genet.* 10:858. doi: 10.3389/fgene.2019.00858
- Vagiona, A. C., Andrade-Navarro, M. A., Psomopoulos, F., and Petrakis, S. (2020). Dynamics of a protein interaction network associated to the aggregation of polyQ-expanded ataxin-1. *Genes (Basel)* 11:1129. doi: 10.3390/genes11101129
- Vagiona, A. C., Mier, P., Petrakis, S., and Andrade-Navarro, M. A. (2022). Analysis of Huntington’s disease modifiers using the hyperbolic mapping of the protein interaction network. *Int. J. Mol. Sci.* 23:5853. doi: 10.3390/ijms23105853
- Yue, S., Serra, H. G., Zoghbi, H. Y., and Orr, H. T. (2001). The spinocerebellar ataxia type 1 protein, ataxin-1, has RNA-binding activity that is inversely affected by the length of its polyglutamine tract. *Hum. Mol. Genet.* 10, 25–30. doi: 10.1093/hmg/10.1.25
- Zhang, S., Williamson, N. A., Duvick, L., Lee, A., Orr, H. T., Korlin-Downs, A., et al. (2020). The ataxin-1 interactome reveals direct connection with multiple disrupted nuclear transport pathways. *Nat. Commun.* 11:3343. doi: 10.1038/s41467-020-17145-0
- Zhu, G., Khalid, F., Zhang, D., Cao, Z., Maity, P., Kestler, H. A., et al. (2023). Ribosomal dysfunction is a common pathomechanism in different forms of trichothiodystrophy. *Cells* 12:1877. doi: 10.3390/cells12141877
- Zoghbi, H. Y., and Orr, H. T. (2009). Pathogenic mechanisms of a polyglutamine-mediated neurodegenerative disease, spinocerebellar ataxia type 1. *J. Biol. Chem.* 284, 7425–7429. doi: 10.1074/jbc.R800041200