Supplementary figures for:

Connecting *HTT* intermediate alleles and microRNA dysregulation to enhanced tauopathy in Late-Onset Alzheimer's Disease

Juan Castilla-Silgado[1,2,6,±], Sergio Perez-Oliveira[2,4,±], Paola Pinto-Hernandez[1,2], Manuel Fernandez-Sanjurjo[1,2], Maria Daniela Corte-Torres[2,7], Eduardo Iglesias-Gutierrez[1,2], Manuel Menendez-Gonzalez[2,3,5], Victoria Alvarez[2,4,#,*],Cristina Tomas-Zapico[1,2,#]

[1] Department of Functional Biology (Physiology), University of Oviedo, 33006, Oviedo, Spain.

[2] Health Research Institute of the Principality of Asturias (ISPA), 33011, Oviedo, Spain.

[3] Department of Neurology, Central University Hospital of Asturias (HUCA), 33011, Oviedo, Spain.

[4] Genetics Laboratory, Central University Hospital of Asturias (HUCA), 33011, Oviedo, Spain.

[5] Department of Medicine, University of Oviedo, 33006, Oviedo, Spain.

[6] Parkinson Asturias Association, 33011, Oviedo, Spain.

[7] Biobank of the Principality of Asturias, Central University Hospital of Asturias (HUCA), 33011, Oviedo, Spain.

± These authors contributed equally as first authors.

# These authors contributed equally as senior authors.

***Corresponding author:**

Victoria Álvarez, PhD

Genetics Laboratory, Central University Hospital of Asturias,

33011 Oviedo, Asturias, Spain

e-mail address: victoria.alvarez@sespa.es

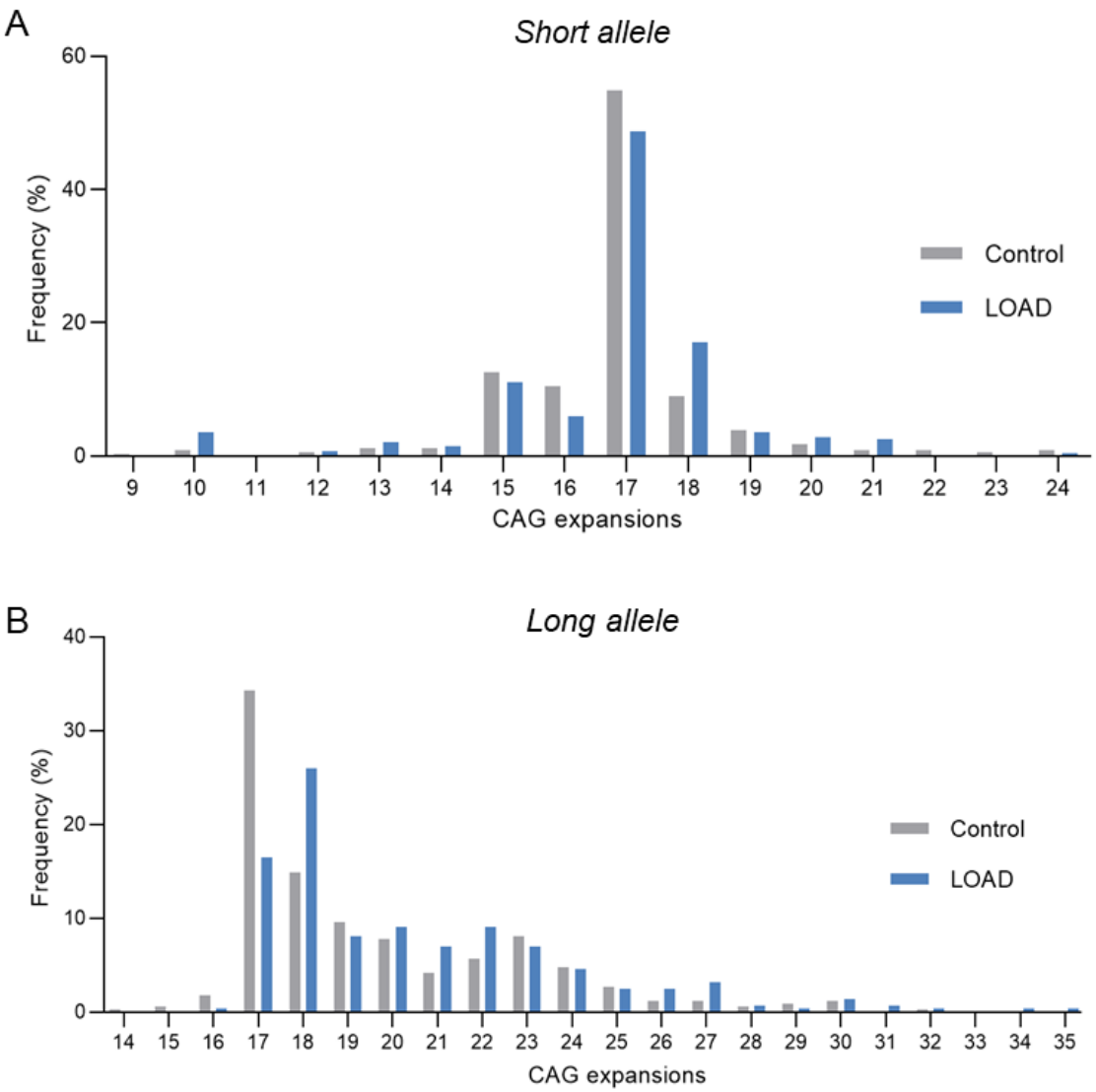

Figure S1. Frequency of CAG expansions in the *HTT* gene in control and LOAD groups. **A**. Within the minor allele of the *HTT* gene, a high frequency of 17 repeats was observed in both groups, with no expansion found in the intermediate range. **B**. The distribution of CAG expansions in the major allele revealed that the most frequent number was 17 in the control group and 18 in LOAD patients. Both groups showed a homogeneous distribution within the nonpathological range of 17-24 CAGs, although there are cases of expansions above 27 CAGs. Controls: *N* = 325; LOAD: *N* = 323.

Abbreviations: LOAD, late-onset Alzheimer's disease; *HTT*, huntingtin gene.

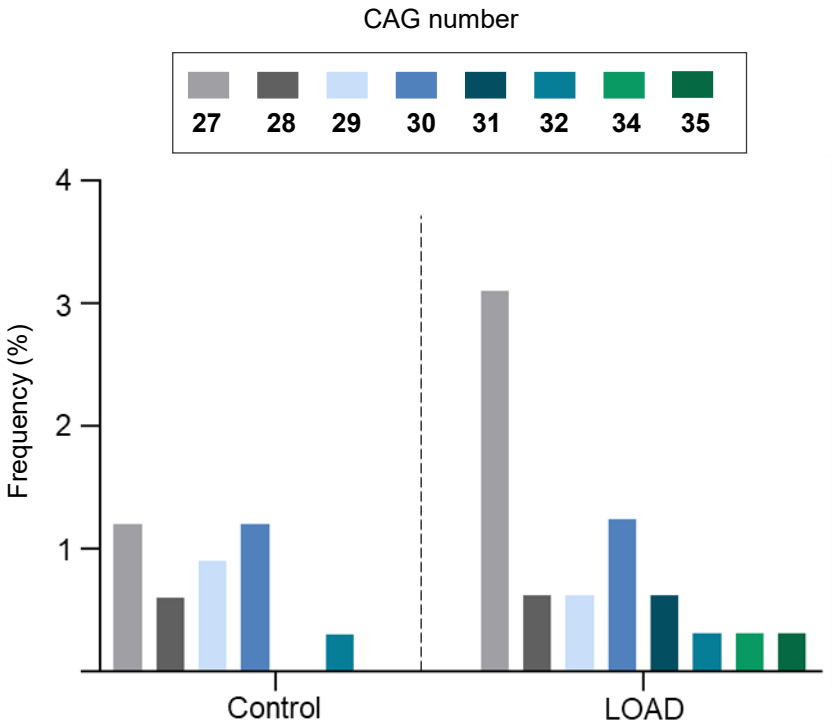

Figure S2. Intermediate CAG expansions distribution in the *HTT* gene in control subjects and LOAD patients. Control subjects presented a homogeneous distribution with a very low frequency, with 30 being the highest number of repeats observed. In contrast, LOAD patients showed repetitions throughout the intermediate range (27-35), with 27 being the most frequent number.

Abbreviations: LOAD, late-onset Alzheimer's disease. Controls: *N* = 14; LOAD: *N* = 23.

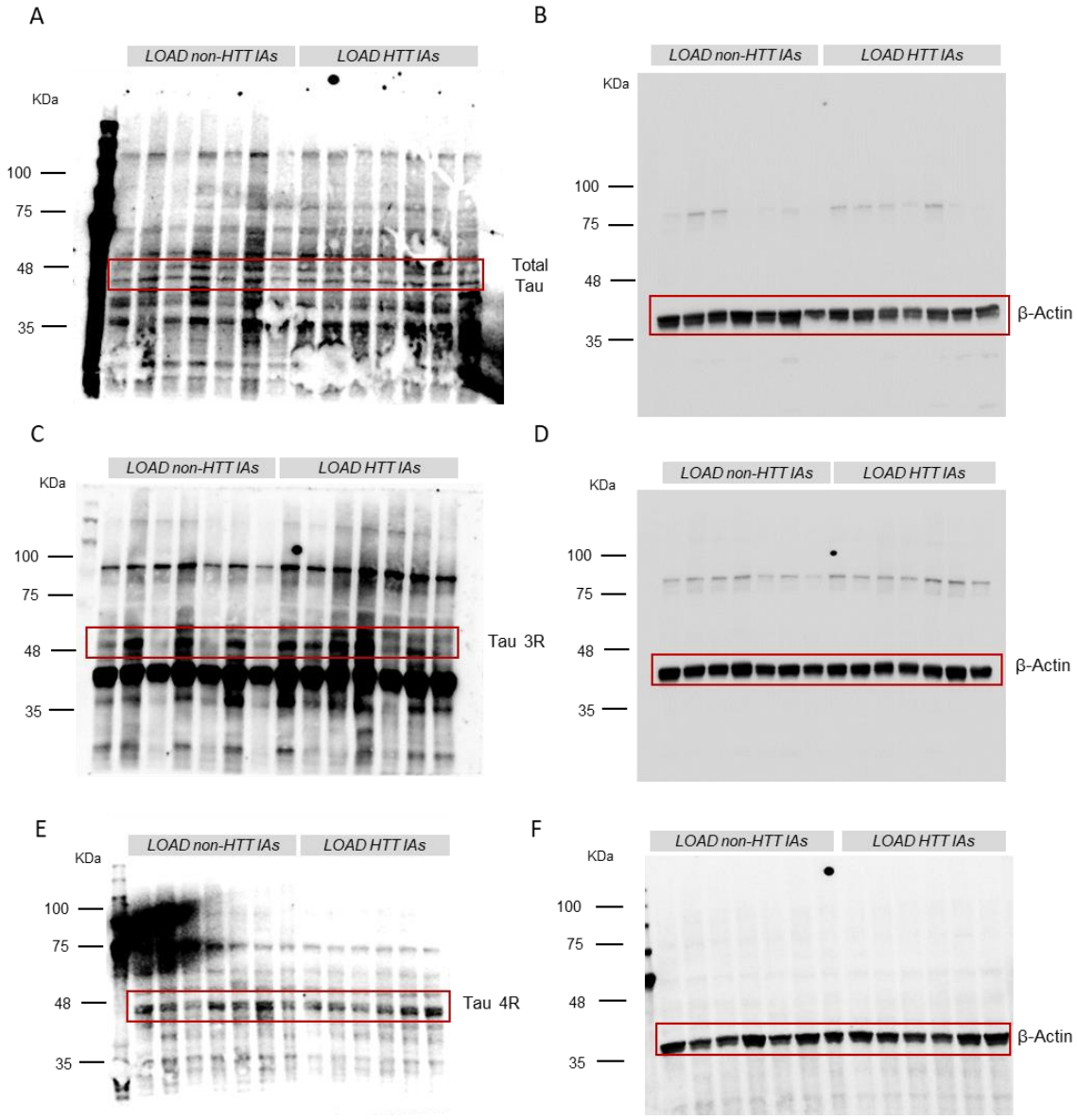

Figure S3. Uncropped western blot image in figure 3. A. Total Tau protein expression**. B.** β-actin protein expression on the same membrane used for Total tau detection. **C.** Tau 3R protein expression. **D**. β-actin protein expression on the same membrane used for Tau 3R detection. **E.** Tau 4R protein expression. **F.** β-actin protein expression on the same membrane used for tau 4R detection. Red boxes correspond to the cropped area to the main figure. β-actin was detected on the same membranes as their respective proteins of interest, following a stripping protocol to remove the primary antibodies. LOAD non-*HTT IA* carriers, *N* = 7; LOAD *HTT IAs, N* = 7

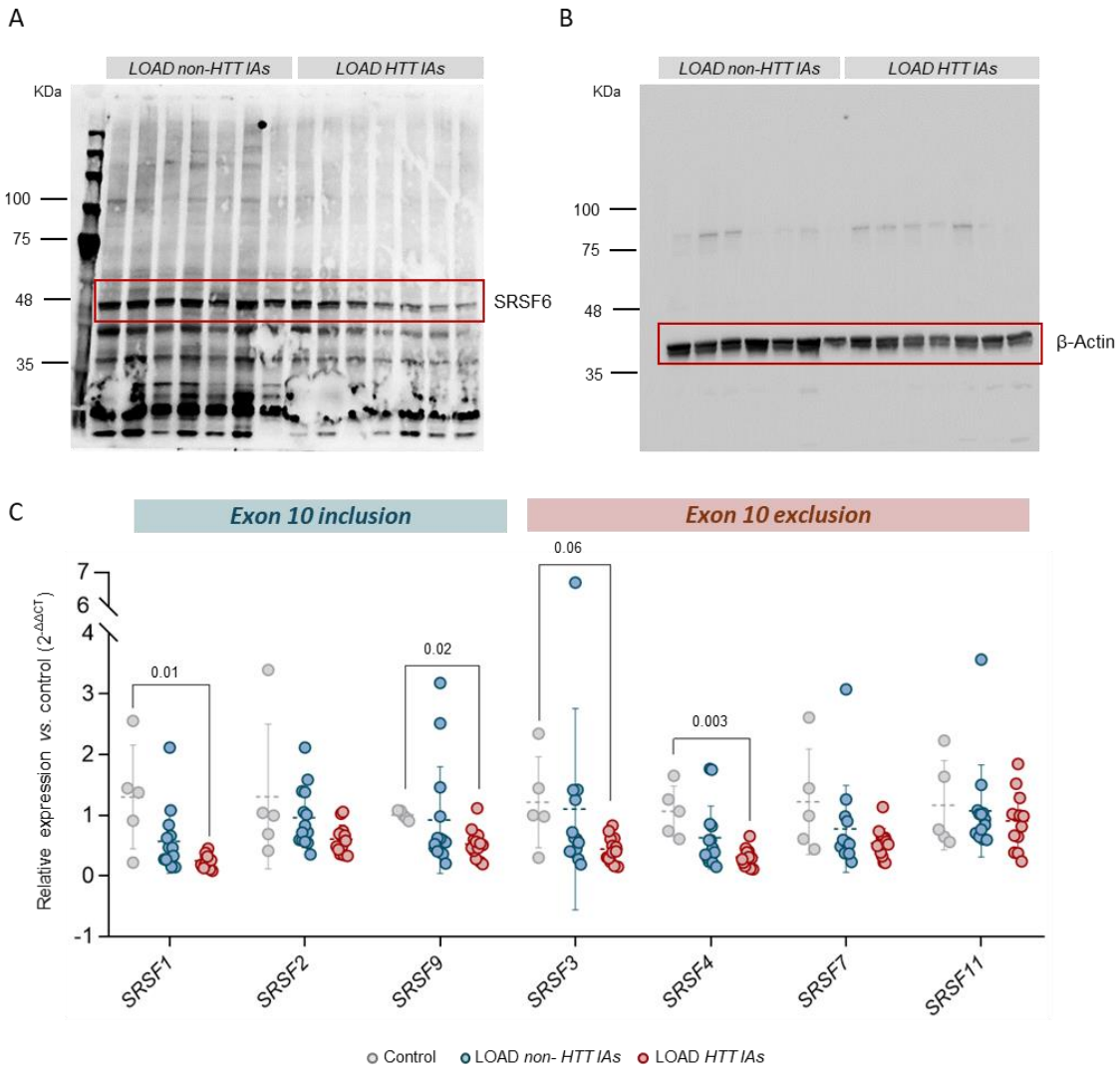

Figure S4. Uncropped western blot image in figure 5 and SRSF splicing factor family analysis. A. SRSF6 protein expression**. B.** β-actin protein expression on the same membrane used for SRSF6 detection. Red boxes correspond to the cropped area to the main figure. β-actin was detected on the same membranes as their respective proteins of interest, following a stripping protocol to remove the primary antibodies. LOAD non-*HTT IA* carriers, *N* = 7; LOAD *HTT IAs, N* = 7 C. Analysis of the relative mRNA expression levels of several SRSF family members was assayed by RT-qPCR in control subjects (*N* = 4), and LOAD patients (non-*HTT IAs* carriers, *N* = 14; *HTT IAs, N* = 13). GAPDH mRNA was used as normalizer. Relative mRNA expression was performed using the comparative Ct method ($2^{-\Delta\Delta Ct}$), considering control group as a reference. The results are shown as the median ± SD, with each point representing an individual subject. Statistical analysis was performed using Kruskal-Wallis's test followed by Dunn's multiple comparison test. Significative *p-values* are indicated in the graphs.

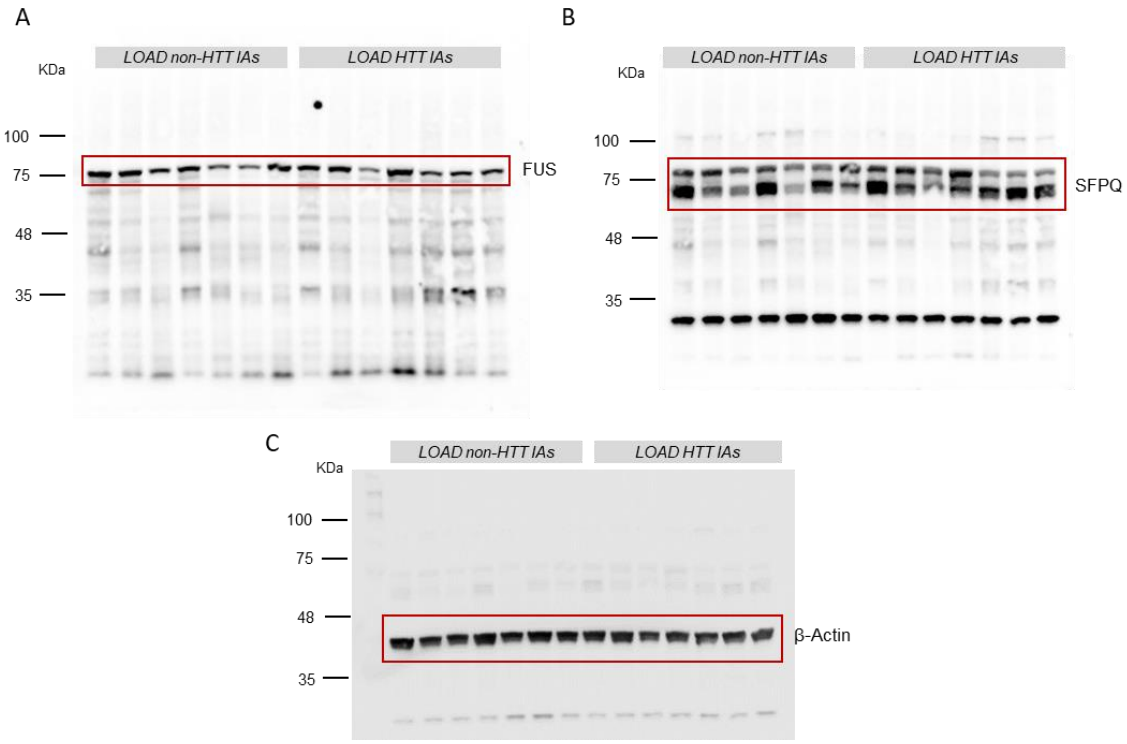

Figure S5. Uncropped western blot image in figure 6. A. FUS protein expression**. B.** SFPQ protein expression. **C.** β-actin protein expression. Red boxes correspond to the cropped area to the main figure. Both FUS and SFPQ protein were detected on the same membrane after stripping protocol to remove the primary antibodies. β-actin was detected on the same membranes as their respective proteins of interest, following a stripping protocol to remove the primary antibodies. LOAD non-*HTT IA* carriers, *N* = 7; LOAD *HTT IAs, N* = 7

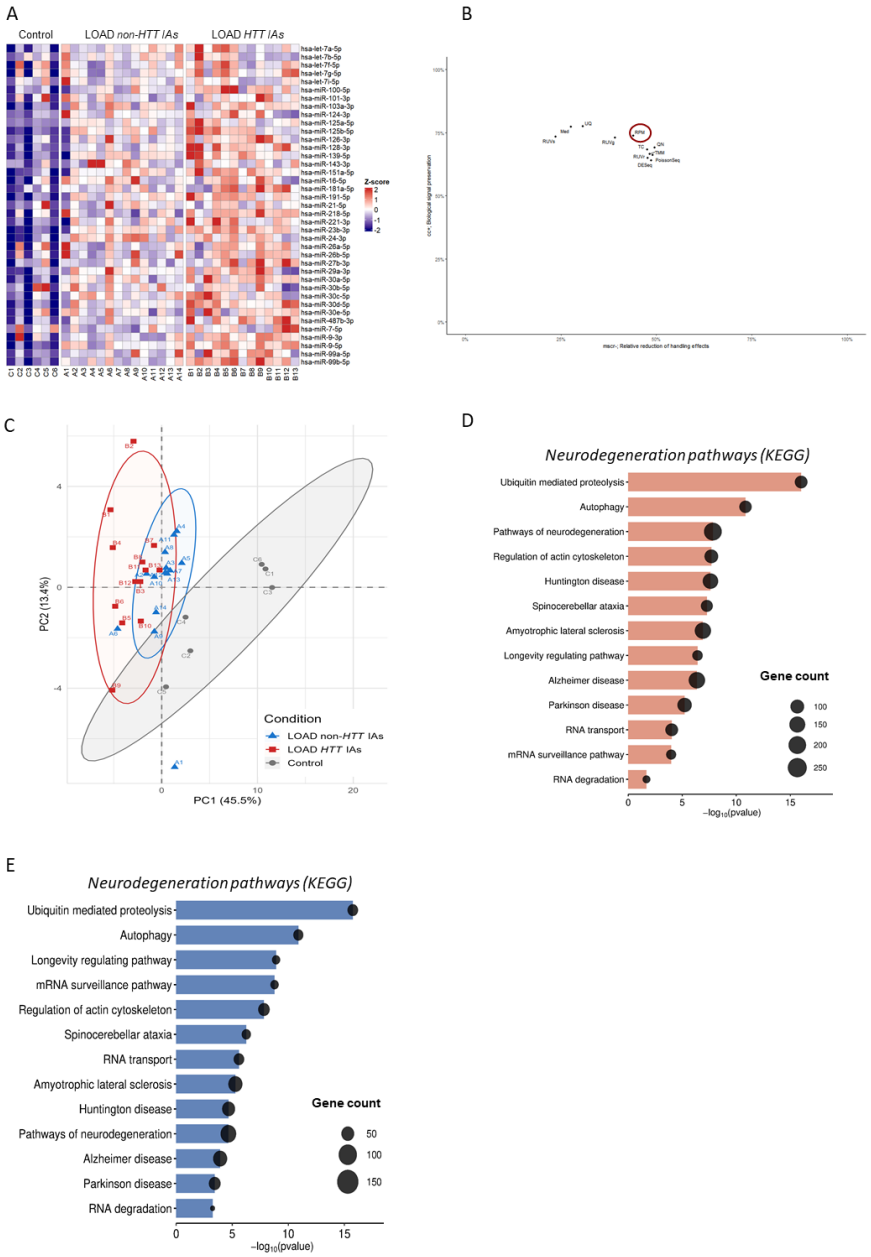

Figure S6. Differential expression and KEGG pathway enrichment analysis of miRNAs. **A** Heatmap shows normalized expression levels (Z-scores) of screened miRNAs, across control and LOAD groups, including both *HTT IAs* carriers and non-carriers. **B.** Performance comparison of different normalization strategies using the DANA framework. Each point represents a distinct normalization method, plotted according to its relative reduction of handling effects (x-axis) and its ability to preserve biological signal (y-axis). The RPM-based normalization used in this study is highlighted, demonstrating a favorable balance between batch effect mitigation and biological relevance. **C.** Principal component analysis (PCA) of expression profiles from screened miRNAs after small RNA-sequencing shows a clear separation between control subjects and LOAD patients, regardless of *HTT IAs* status. (Controls: *N = 6*; LOAD: *non-HTT IAs, N = 14; HTT IAs, N = 13).* **D.** KEGG pathway enrichment analysis of neurodegeneration-related pathways targeted by miRNAs differentially expressed among the three groups of study. **E.** KEGG pathway enrichment analysis of neurodegeneration-related pathways targeted by miRNAs differentially expressed between LOAD groups. Bar length represents -$\log_{10}$(*p-value*). Pathways positions higher on the plot exhibit increased statistical significance. Circle size indicates the number of predicted target genes per pathway.