Role of TDP-43 and hnRNP Proteins in the Regulation of Different RNA Targets

Thesis

How to cite:

For guidance on citations see FAQs.

© 2019 The Author

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.21954/ou.ro.0000efd5

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Role of TDP-43 and hnRNP proteins in the regulation of different RNA targets

Sara Cappelli

A thesis submitted in fulfilment for the requirements of the Open University (UK) for the degree of Doctor of Philosophy

Life Sciences

International Centre for Genetic Engineering and Biotechnology (ICGEB)
Trieste, Italy

Director of studies: Dr. Emanuele Buratti
External supervisor: Prof. David Elliott

February 2019
TABLE OF CONTENTS

TABLE OF CONTENTS ................................................................. 2
PUBLICATIONS ............................................................................. 7
LIST OF FIGURES ........................................................................ 8
LIST OF TABLES ........................................................................... 11
ABBREVIATIONS ......................................................................... 13
ABSTRACT .................................................................................. 23

1. INTRODUCTION ...................................................................... 25

1.1 RNA-Binding Proteins and post-transcriptional gene regulation in brain ....... 25
  1.1.1 A general overview of RNA-Binding Proteins ........................................... 25
  1.1.2 Structure of RNA-Binding Proteins .......................................................... 27
  1.1.3 Different functions of RNA-binding proteins ............................................. 30
  1.1.4 Implication of RNA-Binding proteins in brain development, maintenance and disease ................................................................................................................. 33

1.2 Transcriptome analysis using RNA sequencing and its applications in the characterization of RNA binding proteins ................................................................. 35
  1.2.1 From hypothesis-driven approaches to high-throughput screening (HTS) approaches .................................................................................................................. 35
  1.2.2 Next-generation sequencing (NGS): RNA sequencing ................................. 38
  1.2.3. RNA sequencing applied to the characterization of RNA binding proteins .... 40

1.3 The Heterogeneous Ribonucleoprotein (hnRNP) family ............................ 41
  1.3.1 General characteristics of the hnRNP family ........................................... 41
  1.3.2 Multiple functions of hnRNPs ................................................................... 43
1.3.3 Role of hnRNPs in neurodegenerative disorders ............................................................. 45

1.4 Trans-activating Response Region (TAR) DNA-Binding protein of 43 kDa (TDP-43) ................................................................................................................................. 46

1.4.1 Brief introduction on TDP-43 .......................................................................................... 46

1.4.2 Characteristics of TDP-43 ............................................................................................. 47

1.4.3 Functions of TDP-43 ..................................................................................................... 52

1.4.4 TDP-43 interacting proteins ......................................................................................... 55

1.4.5 TDP-43 proteinopathies .............................................................................................. 57

1.4.6 Drosophila as a model for understanding TDP-43 proteinopathies ......................... 59

1.5 DAZ Associated Protein 1 (DAZAP1) ............................................................................ 65

1.5.1 Characteristics of DAZAP1 ......................................................................................... 65

1.5.2 Functions of DAZAP1 ............................................................................................... 67

1.5.3 DAZAP1 interacting proteins ..................................................................................... 68

1.5.4 DAZAP1 and pathologies ......................................................................................... 69

1.6 Heterogeneous Nuclear Ribonucleoprotein Q (hnRNP Q) and Heterogeneous Nuclear Ribonucleoprotein R (hnRNP R): two closely related hnRNP proteins .... 70

1.6.1 Characteristics of hnRNP Q and hnRNP R ................................................................. 70

1.6.2 Functions of hnRNP Q and hnRNP R ......................................................................... 74

1.6.3 hnRNP Q and hnRNP R interacting proteins .............................................................. 76

1.6.4 Implication of hnRNP Q and hnRNP R in pathology .................................................. 76

2. AIM OF THE PROJECT ........................................................................................................ 78

3. MATERIALS AND METHODS ........................................................................................... 80

3.1 Cell cultures, gene knockdown and cell differentiation .................................................... 80

3.1.1 Cell cultures ............................................................................................................... 80

3.1.2 Gene knockdown of SH-SY5Y neuroblastoma cell line .............................................. 80
3.1.3 Gene knockdown of Flp-In HEK293 cell line expressing TDP-43 aggregates ........................................... 81
3.1.4 NSC-34 differentiation .......................................................................................................................... 82

3.2 DNA plasmids for protein overexpression .............................................................................................. 83
  3.2.1 Generation of competent DH5α Escherichia coli cells ............................................................. 83
  3.2.2 Transformation of competent DH5α cells ................................................................................ 84
  3.2.3 Small scale preparation of plasmid DNA (Miniprep) ................................................................. 85
  3.2.4 Middle scale preparation of plasmidic DNA (Midiprep) .......................................................... 85
  3.2.5 Overexpression of Flag-tagged siRNA resistant TDP-43 and Flag-tagged DAZAP1 in HeLa cells ......................................................................................................................................................... 86

3.3 Protein detection analysis ......................................................................................................................... 88
  3.3.1 Western blot analysis for checking gene knockdown ................................................................. 88
  3.3.2 TDP-43 and DAZAP1 co-immunoprecipitations (CO-IP) analysis ........................................ 90
  3.3.3 Nuclear and cytoplasmic extraction using NER-PER kit ...................................................... 91

3.4 pre-RNA splicing, gene expression and RNA-seq analysis ................................................................ 93
  3.4.1 RNA extraction ................................................................................................................................. 93
  3.4.2 Reverse transcription reaction ...................................................................................................... 94
  3.4.3 pre-mRNA splicing analysis ......................................................................................................... 95
  3.4.4 RNA immunoprecipitation (RNA-IP) analysis ........................................................................ 97
  3.4.5 Gene expression analysis ............................................................................................................... 98
  3.4.6 RNA-seq and analysis of differentially expressed genes (DEGs) ......................................... 102
  3.4.7 Statistical analysis of data ............................................................................................................. 102

3.5 Immunofluorescence (IF) analysis ......................................................................................................... 104
  3.5.1 Immunofluorescence analysis of Flp-In HEK293 cell line expressing TDP-43 aggregates ................................................................. 104
  3.5.2 Immunofluorescence analysis of SH-SY5Y cells differentiated NSC-34 cells ............................ 105
4. RESULTS .......................................................................................................................................... 107

4.1 Effects of DAZAP1, hnRNP Q and hnRNP R suppression on TDP-43 activity in
SH-SY5Y cell line ................................................................................................................................... 107
  4.1.1 DAZAP1, hnRPQ and hnRNPR can regulate and rescue TDP-43 controlled
splicing events without altering the expression of TDP-43 ................................................................. 107
  4.1.2 DAZAP1, hnRNP Q and hnRNP R do not alter TDP-43 nuclear localization. ................................... 115

4.2 Cellular model of TDP-43 pathology .................................................................................................. 116
  4.2.1 Analysis of siRNA-mediated depletion of DAZAP1, hnRNP Q and hnRNP R
using a cellular model of TDP-43 loss-of-function .............................................................................. 116

4.3 Characterization of TDP-43 and DAZAP1 interaction ....................................................................... 118
  4.3.1 DAZAP1 does not bind to TDP-43 but can bind in vivo to TDP-43 controlled
mRNAs ................................................................................................................................................ 118
  4.3.2 TDP-43 and DAZAP1 can alter the expression of neuronal related mRNAs in
SH-SY5Y cells ..................................................................................................................................... 120

4.4 Characterization of hnRNP Q and hnRNP R .................................................................................... 127
  4.4.1 Analysis of protein sequence identity and similarity ................................................................. 127
  4.4.2 Subcellular localization of hnRNP Q and hnRNP R in SH-SY5Y cells .......... 128
  4.4.3 Subcellular localization of hnRNP Q and hnRNP R is not altered by their
reciprocal silencing ............................................................................................................................... 131
  4.4.4 Subcellular localization of hnRNP Q and hnRNP R is not altered by retinoic
acid (RA)-induced neuronal differentiation in NSC-34 cells ............................................................... 133
  4.4.5 Depletion of hnRNP Q and hnRNP R alters neuronal-related genes ............ 135
  4.4.6 hnRNP Q and hnRNP R show different and common features after GO
enrichment and KEGG pathway analysis .............................................................................................. 142
4.5 An overview of the entire transcriptome status after silencing of TDP-43, DAZAP1, hnRNP Q and hnRNP R ................................................................. 146

4.5.1 TDP-43, DAZAP1 and hnRNP Q share common regulated-transcripts associated with ALS pathology ...................................................... 146

5. DISCUSSION ................................................................................................. 156

5.1 TDP-43 activity is modulated by the interaction with other hnRNP proteins ..... 156
5.2 hnRNP Q and hnRNP R have evolved to perform different functions within cells .............................................................................................................. 166
5.3 TDP-43, DAZAP1 and hnRNP Q potentially regulate same transcripts important for ALS pathogenesis ................................................................................. 175

6. CONCLUSION AND FUTURE DIRECTIONS ............................................. 180

7. BIBLIOGRAPHY ............................................................................................... 181
PUBLICATIONS


LIST OF FIGURES

Figure 1. Classification of RNA-binding proteins ........................................................... 26
Figure 2. Common RNA-binding domains ................................................................. 29
Figure 3. Function of RBPs in RNA metabolism ......................................................... 31
Figure 4. High-throughput screening (HTS) approaches .............................................. 37
Figure 5. Structure of the canonical hnRNP family members ........................................ 42
Figure 6. Localization of TARDBP gene on human chromosome 1 ............................. 48
Figure 7. Schematic representation of TDP-43 protein domains and their associated functions ............................................................... 49
Figure 8. Schematic representation of TDP-43 functions .............................................. 52
Figure 9. Schematic representation of the TDP-43 gain-of-function model .................... 58
Figure 10. Schematic representation of TDP-43 loss-of-function model ......................... 59
Figure 11. Comparison of human TDP-43 and Drosophila TBPH proteins .................... 61
Figure 12. Gal4-UAS system ....................................................................................... 64
Figure 13. Localization of DAZAP1 gene on human chromosome 19 ............................ 65
Figure 14. Schematic representation of DAZAP1 domains ............................................ 66
Figure 15. Localization of SYNCRIP gene on human chromosome 6 .......................... 70
Figure 16. Schematic representation of protein domains and major isoforms of hnRNP Q ................................................................. 71
Figure 17. Localization of HNRNPR gene on human chromosome 1 ........................... 72
Figure 18. Schematic representation of protein domains and major isoforms of hnRNP R ................................................................. 73
Figure 19. plasmid DNA for overexpression ............................................................... 87
Figure 20. PCR conditions used in the pre-mRNA splicing assay ................................. 97
Figure 21. Real-Time PCR conditions for gene expression analysis ........................... 102
Figure 22. Western Blot analysis of siRNA treatment on SH-SY5Y cells. ..........110
Figure 23. Effects of different siRNA treatment on TDP-43-regulated events. ...... 110
Figure 24. Rescue of TDP-43 controlled events. ........................................112
Figure 25. Schematic diagram of TDP-43, DAZAP1, hnRNQ and hnRNP activity on RNA splicing of *POPLIDP3, STAG2, TNIK* and *MADD* gene. ..............113
Figure 26. Effects of DAZAP1 and hnRNP Q/R depletion in the absence (A-C) or in the presence (D-F) of TDP-43 co-silencing on gene expression events controlled by TDP-43. ................................................................................114
Figure 27. Localization of endogenous TDP-43 in SH-SY5Y cells. .....................117
Figure 28. Effects of hnRNP silencing in Flp-In HEK293 cells expressing FLAG-TDP43-12XQ/N aggregates ........................................................................117
Figure 29. Probing the DAZAP1 - TDP-43 interaction.......................................119
Figure 30. Validation of the TDP-43 or DAZAP1 silencing and comparison between RNA-seq and qRT-PCR results.................................................................123
Figure 31. Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for TDP-43................................................................................124
Figure 32. Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for DAZAP1............................................................... 125
Figure 33. GO functional enrichment analysis of the genes significantly co-regulated by TDP-43 and DAZAP1 ................................................................. 126
Figure 34. Structure of *Drosophila melanogaster* CG17838 (isoform F)/Syp variant. ........................................................................................................... 128
Figure 35. Cellular localization of endogenous hnRNP Q and hnRNP R in SH-SY5Y cells. ....................................................................................................... 129
Figure 36. Nuclear and cytoplasmic fractions of endogenous human hnRNP Q and human hnRNP R. ................................................................. 131

Figure 37. mRNA levels and cellular localization of endogenous hnRNP Q and hnRNP R after siRNA treatment in SH-SY5Y cells. ........................................... 132

Figure 38. Endogenous localization of hnRNP Q and hnRNP R is not affected by differentiation with 1 µM of RA in NSC-34 cells. ............................................. 134

Figure 39. Validation of hnRNP Q silencing and comparison between RNA-seq and RT-qPCR results. ................................................................. 137

Figure 40. Validation of hnRNP R silencing and comparison between RNA-seq and RT-qPCR results. ................................................................. 138

Figure 41. Volcano plot of hnRNP Q and hnRNP R. .................................................. 140

Figure 42. Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for hnRNP Q and hnRNP R. .................................................. 141

Figure 43. GO enrichment analysis of hnRNP Q and hnRNP R DEGs. ..................... 144

Figure 44. KEGG pathway analysis of hnRNP Q and hnRNP R ......................... 145

Figure 45. Venn diagramm of RNA-seq data. ..................................................... 146

Figure 46. Pathways of genes co-regulated only by TDP-43, DAZAP1 and hnRNP Q. ................................................................................................. 155
LIST OF TABLES

Table 1. siRNA sequence from 5'-prime to 3'-prime are listed. .........................................................81
Table 2. Composition of TSS 1X .............................................................................................................84
Table 3. Composition of stock 4X sample buffer. ..................................................................................89
Table 4. Composition of 10% SDS-acrylamide gel. .............................................................................89
Table 5. Specifics of AbI and AbII used for checking gene knockdown. .............................................90
Table 6. Composition of IP buffer. .........................................................................................................91
Table 7. List of AbI and AbII used for CO-IP experiment ..................................................................91
Table 8. Reagent volumes for different packed cell volumes..............................................................92
Table 9. List of AbI and AbII antibodies used for nuclear and cytoplasmic fractionation. 93
Table 10. Volume of reagent using for RNA extraction according to cell dish......................94
Table 11. Preparation of RNA sample mixture.....................................................................................95
Table 12. Mixture for reverse transcription reaction. .................................................................95
Table 13. PCR reaction mixture...........................................................................................................96
Table 14. List of genes and relative primers (forward and reverse sequences) used for pre-mRNA splicing analysis. .................................................................................................96
Table 15. Composition of HEGN buffer...............................................................................................98
Table 16. Composition of iQ mix solution.............................................................................................99
Table 17. List of target genes and relative primers (forward and reverse sequences) used for gene expression analysis. .................................................................100
Table 18. List of housekeeping genes and relative primers (forward and reverse sequences) used for normalization of gene expression analysis. .................................101
Table 19. List of AbI and AbII for IF analysis of Flp-In HEK293 expressing TDP-43

12XQ/N..............................................................................................................................................104
Table 20. List of AbI and AbII for IF analysis of SH-SY5Y cells and differentiated NSC-34 cells. ........................................................................................................................................... 106

Table 21. Percentage of identity and similarity of human hnRNP Q and hnRNP R with respect to *Drosophila* CG17838 (isoform F)/Syp variant. ....................................................... 128

Table 22. Co-regulated transcripts between TDP-43, DAZAP1 and hnRNP Q. ............ 147

Table 23. Genes commonly regulated by depletion of TDP-43, DAZAP1 and hnRNP Q in SH-SY5Y cells which showed an involvement in brain disorders and/or ALS pathology. ......................................................................................................................................... 176
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>12XQ/N</td>
<td>TDP-43 Q/N rich amino acid sequence 331–369 repeated 12 times</td>
</tr>
<tr>
<td>ABAT</td>
<td>4-aminobutyrate Aminotransferase</td>
</tr>
<tr>
<td>ABCC9</td>
<td>ATP Binding Cassette Subfamily C Member 9</td>
</tr>
<tr>
<td>AcD</td>
<td>Acidic Domain</td>
</tr>
<tr>
<td>ACHE</td>
<td>Acetylcholinesterase</td>
</tr>
<tr>
<td>ACP5</td>
<td>ACP5, Acid phosphatase 5, Tartrate Resistant</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ADCY1</td>
<td>Adenylate Cyclase 1</td>
</tr>
<tr>
<td>ADD3</td>
<td>Adducin 3</td>
</tr>
<tr>
<td>AGO2</td>
<td>Argonaute 2</td>
</tr>
<tr>
<td>AGP</td>
<td>α1-Acid Glycoprotein</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute Lymphoblastic Leukemia</td>
</tr>
<tr>
<td>ALS</td>
<td>Amyotrophic Lateral Sclerosis</td>
</tr>
<tr>
<td>AP-1</td>
<td>Active Activator Protein-1</td>
</tr>
<tr>
<td>apoB</td>
<td>Apolipoprotein B</td>
</tr>
<tr>
<td>ApoII</td>
<td>Apolipoprotein AII</td>
</tr>
<tr>
<td>ARE</td>
<td>AU-rich elements</td>
</tr>
<tr>
<td>ARHGAP36</td>
<td>ARHGAP36, Rho GTPase activating Protein 36</td>
</tr>
<tr>
<td>ATG7</td>
<td>Autophagy Related 7</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia Teleangectasia Mutated</td>
</tr>
<tr>
<td>BDNF</td>
<td>Brain Derived Neurotrophic Factor</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast Cancer 1</td>
</tr>
<tr>
<td>BRD8</td>
<td>Bromodomain Containing 8</td>
</tr>
<tr>
<td>Brunol</td>
<td>Bruno-like</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>c-Fos</td>
<td>Cellular Oncogene Fos</td>
</tr>
<tr>
<td>C25</td>
<td>C-terminal region of 25 amino acids involved in DAZAP1’s nuclear localization</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>Voltage-dependent L-type Calcium Channel Subunit α1C</td>
</tr>
<tr>
<td>CAMs</td>
<td>Cell Adhesion Molecules</td>
</tr>
<tr>
<td>CAPN5</td>
<td>Calpain 5</td>
</tr>
<tr>
<td>CARTPT</td>
<td>CARTPT, CART Prepropeptide</td>
</tr>
<tr>
<td>CASP3</td>
<td>Caspase 3</td>
</tr>
<tr>
<td>CDK5R1</td>
<td>Cyclin Dependent Kinase 5 Regulatory Subunit 1</td>
</tr>
<tr>
<td>CDK6</td>
<td>Cyclin-Dependent Kinase 6</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CELF</td>
<td>CUGBP, Elav-like family</td>
</tr>
<tr>
<td>CELF5</td>
<td>CELF5, CUGBP, Elav-like Family Member 5</td>
</tr>
<tr>
<td>CFTR</td>
<td>Cystic Fibrosis Transmembrane Conductance Regulator</td>
</tr>
<tr>
<td>CLIP-seq</td>
<td>Cross-Linking Immunoprecipitation Sequencing</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>Co-IP</td>
<td>Co-Immunoprecipitation</td>
</tr>
<tr>
<td>CPSF</td>
<td>Polyadenylation Specificity Factor</td>
</tr>
<tr>
<td>Cry</td>
<td>Cytochrome 1</td>
</tr>
<tr>
<td>CT</td>
<td>Pyrimidine-Rich Elements</td>
</tr>
<tr>
<td>CT55</td>
<td>CT55, Cancer/testis Antigen 5</td>
</tr>
<tr>
<td>CTFs</td>
<td>C-terminal fragments</td>
</tr>
<tr>
<td>CTSS</td>
<td>Cathepsin S</td>
</tr>
<tr>
<td>CUG-BP</td>
<td>CUG binding protein</td>
</tr>
<tr>
<td>DAB1</td>
<td>Disabled Homologue 1</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>DAZ</td>
<td>Deleted in Azoospermia</td>
</tr>
<tr>
<td>DAZAP1</td>
<td>DAZ Associated Protein 1</td>
</tr>
<tr>
<td>DAZL</td>
<td>DAZ-like</td>
</tr>
<tr>
<td>DDX20</td>
<td>DEAD-Box Helicase 20</td>
</tr>
<tr>
<td>DDX58</td>
<td>DExD/H-Box Helicase 58</td>
</tr>
<tr>
<td>DEG(s)</td>
<td>Differential Expressed Gene(s)</td>
</tr>
<tr>
<td>DGE</td>
<td>Differential Gene Expression</td>
</tr>
<tr>
<td>DM1</td>
<td>Myotonic Dystrophy of Type 1</td>
</tr>
<tr>
<td>DNA</td>
<td>DeoxyriboNucleic Acid</td>
</tr>
<tr>
<td>DSB</td>
<td>Double Strand Breaks</td>
</tr>
<tr>
<td>dsRBD</td>
<td>Double-stranded RNA Binding Domain</td>
</tr>
<tr>
<td>DUOXA1</td>
<td>DUOXA1, Dual Oxidase Maturation Factor 1</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular Matrix</td>
</tr>
<tr>
<td>EFEMP1</td>
<td>EFEMP1, EGF containing Fibulin-like Extracellular Matrix Protein 1</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal Growth Factor</td>
</tr>
<tr>
<td>ELAV</td>
<td>Embryonic Lethal Abnormal Division</td>
</tr>
<tr>
<td>ELAVL3</td>
<td>ELAV like Neuron-specific RNA Binding Protein 3</td>
</tr>
<tr>
<td>Elk-1</td>
<td>ETS Domain-containing Protein Elk-1</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electromobility Shift Assay</td>
</tr>
<tr>
<td>ERG</td>
<td>Early-Response Genes</td>
</tr>
<tr>
<td>ERK2</td>
<td>Extracellular Signal Regulated Kinase 2</td>
</tr>
<tr>
<td>ESCs</td>
<td>Self-renewal of Embryonic Stem Cells</td>
</tr>
<tr>
<td>ETF1</td>
<td>Eukaryotic Translation Termination Factor 1</td>
</tr>
<tr>
<td>ETR-3</td>
<td>Elav-type RNA Binding Protein 3</td>
</tr>
<tr>
<td>FGF2</td>
<td>Fibroblastic Growth Factor 2</td>
</tr>
</tbody>
</table>
FOSB, FBJ Murine Osteosarcoma Viral Oncogene Homolog B
Frontotemporal Lobar Degeneration
Fused in Sarcoma/translocated in Liposarcoma
Fragile X-associated Tremor/Ataxia Syndrome
Gamma-aminobutyric Acid
Glyceraldehyde-3-phosphate Dehydrogenase
Gene Ontology
Gain-of-Function
Histone Deacetylase 6
Human Embryonic Kidney 293 Cell Line
Human Epithelial Cell Line
Human Immunodeficiency Type 1
HMOX1, Heme Oxygenase (Decycling) 1
Heterogeneous Nuclear RNA (pre-mRNA)
Heterogeneous Ribonucleoprotein A/B
Heterogeneous Ribonucleoprotein A1
Heterogeneous Ribonucleoprotein A2
Heterogeneous Ribonucleoprotein C
Heterogeneous Ribonucleoprotein D
Heterogeneous Ribonucleoprotein H
Heterogeneous Ribonucleoprotein I
Heterogeneous Ribonucleoprotein K
Heterogeneous Ribonucleoprotein L
Heterogeneous Ribonucleoprotein M
Heterogeneous Ribonucleoprotein Q
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnRNP R</td>
<td>Heterogeneous Ribonucleoprotein R</td>
</tr>
<tr>
<td>hnRNP U</td>
<td>Heterogeneous Ribonucleoprotein U</td>
</tr>
<tr>
<td>hnRNP(s)</td>
<td>Heterogeneous Ribonucleoprotein(s)</td>
</tr>
<tr>
<td>HPRT1</td>
<td>Hypoxanthine Phosphoribosyltransferase 1</td>
</tr>
<tr>
<td>HS3ST2</td>
<td>Heparan Sulfate-Glucosamine 3-Sulfotransferase 2</td>
</tr>
<tr>
<td>hTDP-43</td>
<td>human TDP-43</td>
</tr>
<tr>
<td>HTS</td>
<td>High-Throughput Screening</td>
</tr>
<tr>
<td>IBMPFD</td>
<td>Inclusion Body Myopathy associated with Paget’s Disease of Bone and Frontotemporal Dementia</td>
</tr>
<tr>
<td>ICAM1</td>
<td>Intercellular Adhesion Molecule 1</td>
</tr>
<tr>
<td>ICAM5</td>
<td>ICAM5, Intercellular Adhesion Molecule 5, Telencephalin</td>
</tr>
<tr>
<td>IF</td>
<td>Immunofluorescence</td>
</tr>
<tr>
<td>IGF2</td>
<td>Insulin-like Growth Factor 2</td>
</tr>
<tr>
<td>IGF2BP1</td>
<td>Insulin-like Growth Factor-II mRNA-binding Proteins 1</td>
</tr>
<tr>
<td>iPSC</td>
<td>Induced Pluripotent Stem Cells</td>
</tr>
<tr>
<td>ISE</td>
<td>Intronic Splicing Enhancer</td>
</tr>
<tr>
<td>JAG1</td>
<td>JAG1, Jagged 1</td>
</tr>
<tr>
<td>K/R</td>
<td>Lysine and Arginine</td>
</tr>
<tr>
<td>KCNAB1</td>
<td>KCNAB1, Potassium Voltage-gated Channel, Shaker-related Subfamily, Beta Member 1</td>
</tr>
<tr>
<td>KH</td>
<td>K-homology</td>
</tr>
<tr>
<td>KLF4</td>
<td>KLF4, Kruppel-like Factor 4 (gut)</td>
</tr>
<tr>
<td>KLHL4</td>
<td>KLHL4, Kelch-like Family Member 4</td>
</tr>
<tr>
<td>LOF</td>
<td>Loss-of-Function</td>
</tr>
<tr>
<td>MADD</td>
<td>MAP Kinase-activating Death Domain Protein</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>MEF2D</td>
<td>Myocyte Enhancer Factor 2D</td>
</tr>
<tr>
<td>MEK/Erk</td>
<td>Extracellular Signal-regulated Protein Kinase</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>N42</td>
<td>42 amino acid residues located at the N-terminal</td>
</tr>
<tr>
<td>ncRNA</td>
<td>non-coding RNA</td>
</tr>
<tr>
<td>NE</td>
<td>Neurotransmitter Norepinephrine</td>
</tr>
<tr>
<td>NEFL</td>
<td>Low Molecular Weight Neurofilament</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear Export Signal</td>
</tr>
<tr>
<td>NGS</td>
<td>Next-Generation Sequencing</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear Localization Signal</td>
</tr>
<tr>
<td>NMJ</td>
<td>Neuromuscular Junction</td>
</tr>
<tr>
<td>NOVA2</td>
<td>Neuro-oncological Ventral Antigen 2</td>
</tr>
<tr>
<td>NRG3</td>
<td>NRG3, Neuregulin 3</td>
</tr>
<tr>
<td>NSC-34</td>
<td>Mouse Motor Neuron-like NSC-34 Cell Line</td>
</tr>
<tr>
<td>P-bodies</td>
<td>Processing Bodies</td>
</tr>
<tr>
<td>PABP</td>
<td>Polyadenylate Binding Protein</td>
</tr>
<tr>
<td>PABPN1</td>
<td>Nuclear Poly(A) Binding Protein</td>
</tr>
<tr>
<td>PCR</td>
<td>Polimerase Chain Reaction</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s Disease</td>
</tr>
<tr>
<td>PDZ</td>
<td>Postsynaptic density protein, Discs-large, Zona occludens</td>
</tr>
<tr>
<td>PENK</td>
<td>PENK, Proenkephalin</td>
</tr>
<tr>
<td>PIWI</td>
<td>P-element Induced Wimpy Testis</td>
</tr>
<tr>
<td>PND</td>
<td>Paraneoplastic Neurologic Syndrome</td>
</tr>
<tr>
<td>PNS</td>
<td>Paraneoplastic Neurologic Syndrome</td>
</tr>
<tr>
<td>POLDIP3</td>
<td>Polymerase Delta-interacting Protein 3</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------</td>
</tr>
<tr>
<td>POLDIP3/SKAR</td>
<td>Polymerase Delta Interacting Protein/S6 Kinase 1 Aly/REF-like Target</td>
</tr>
<tr>
<td>POLR2A</td>
<td>RNA Polymerase II Subunit A</td>
</tr>
<tr>
<td>polyA</td>
<td>Polyadeylation</td>
</tr>
<tr>
<td>POMA</td>
<td>Opsoclonus Myoclonus Ataxia</td>
</tr>
<tr>
<td>PPP1R15A</td>
<td>Protein Phosphatase 1 Regulatory Subunit 15A</td>
</tr>
<tr>
<td>pre-mRNA</td>
<td>precursor mRNA</td>
</tr>
<tr>
<td>pre-rRNA</td>
<td>pre-ribosomal RNA</td>
</tr>
<tr>
<td>Pro-Rich</td>
<td>Proline rich region</td>
</tr>
<tr>
<td>Prrp</td>
<td>Proline-rich RNA Binding Protein</td>
</tr>
<tr>
<td>PTBP1</td>
<td>Polypyrimidine Tract-binding Protein 1</td>
</tr>
<tr>
<td>PTBP2</td>
<td>Polypyrimidine Tract-binding Protein 2</td>
</tr>
<tr>
<td>PUFAs</td>
<td>Polyunsaturated Fatty Acids</td>
</tr>
<tr>
<td>Q/N</td>
<td>Glutamine/Asparagine</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative PCR</td>
</tr>
<tr>
<td>R/G</td>
<td>Arginine and Glycine</td>
</tr>
<tr>
<td>RA</td>
<td>Retinoic Acid</td>
</tr>
<tr>
<td>RAB26</td>
<td>RAB26, Member RAS Oncogene Family</td>
</tr>
<tr>
<td>RAB31</td>
<td>RAB31, Member RAS Oncogene Family</td>
</tr>
<tr>
<td>RBD(s)</td>
<td>RNA-binding domain(s)</td>
</tr>
<tr>
<td>RBFOX1</td>
<td>RNA-binding Protein FOX1 Homologue 1</td>
</tr>
<tr>
<td>RBP(s)</td>
<td>RNA-binding protein(s)</td>
</tr>
<tr>
<td>RELN</td>
<td>Reelin</td>
</tr>
<tr>
<td>RGG</td>
<td>Arginine-Glycine-Glycine box</td>
</tr>
<tr>
<td>RIP-seq</td>
<td>RNA Immunoprecipitation Sequencing</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced Silencing Complex</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RNA-IP</td>
<td>RNA Immunoprecipitation</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA Sequencing</td>
</tr>
<tr>
<td>RNA(s)</td>
<td>Ribonucleic Acid(s)</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA Interference</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>RNP1</td>
<td>Ribonucleoprotein Motif-1</td>
</tr>
<tr>
<td>RNP2</td>
<td>Ribonucleoprotein Motif-1</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>RPL13A</td>
<td>Ribosomal Protein L13a</td>
</tr>
<tr>
<td>RRM</td>
<td>RNA-Recognition Motif</td>
</tr>
<tr>
<td>RXRG</td>
<td>Retinoid X Receptor Gamma</td>
</tr>
<tr>
<td>S/R</td>
<td>Serine and Arginine</td>
</tr>
<tr>
<td>SDCBP2</td>
<td>SDCBP2, Syndecan Binding Protein (Syntenin) 2</td>
</tr>
<tr>
<td>SDHA</td>
<td>Succinate Dehydrogenase Complex Flavoprotein Subunit A</td>
</tr>
<tr>
<td>SG(s)</td>
<td>Stress Granule(s)</td>
</tr>
<tr>
<td>SH-SY5Y</td>
<td>Human Neuroblastoma Cell Line</td>
</tr>
<tr>
<td>SHANK3</td>
<td>SH3 and Multiple Ankyrin Repeat Domains Protein 3</td>
</tr>
<tr>
<td>SLIRP</td>
<td>Steroid Receptor RNA Activator Stem-loop Interacting RNA Binding Protein</td>
</tr>
<tr>
<td>SMA</td>
<td>Spinal Muscular Atrophy</td>
</tr>
<tr>
<td>SMN</td>
<td>Survival of Motor Neurons</td>
</tr>
<tr>
<td>SNAP25</td>
<td>Synaptosomal-associated Protein of 25 kDa</td>
</tr>
<tr>
<td>SNARE</td>
<td>Soluble N-ethylmaleimide Sensitive Factor Attachment Protein Receptor</td>
</tr>
<tr>
<td>snoRNA</td>
<td>small nucleolar RNA</td>
</tr>
<tr>
<td><strong>SNPs</strong></td>
<td>Single-Nucleotide Polymorphisms</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>snRNA</strong></td>
<td>small nuclear RNA</td>
</tr>
<tr>
<td><strong>SORT1</strong></td>
<td>Sortilin 1</td>
</tr>
<tr>
<td><strong>SR protein</strong></td>
<td>Serine Arginine Rich Splicing Factors</td>
</tr>
<tr>
<td><strong>SRRM4</strong></td>
<td>Ser/Arg Repetitive Matrix Protein 4</td>
</tr>
<tr>
<td><strong>SRSFs</strong></td>
<td>Serine and Arginine Rich Splicing Factors</td>
</tr>
<tr>
<td><strong>ssDNA</strong></td>
<td>single strand DNA</td>
</tr>
<tr>
<td><strong>STAG2</strong></td>
<td>Stromal Antigen 2</td>
</tr>
<tr>
<td><strong>STX3</strong></td>
<td>Syntaxin 3</td>
</tr>
<tr>
<td><strong>Syp</strong></td>
<td>Syncrip/hnRNP Q</td>
</tr>
<tr>
<td><strong>SYT14</strong></td>
<td>Synaptotagmin 14</td>
</tr>
<tr>
<td><strong>TAF15</strong></td>
<td>TATA-Box Binding Protein Associated Factor 15</td>
</tr>
<tr>
<td><strong>TBPH</strong></td>
<td>TDP-43 homolog protein in Drosophila</td>
</tr>
<tr>
<td><strong>TDP-43</strong></td>
<td>Trans-activating response region (TAR)-DNA Binding Protein 43 kDa</td>
</tr>
<tr>
<td><strong>TDPBR</strong></td>
<td>Extended Binding Region for TDP-43</td>
</tr>
<tr>
<td><strong>TERRA</strong></td>
<td>Telomeric Repeat- containing RNA</td>
</tr>
<tr>
<td><strong>TGF-β</strong></td>
<td>Transforming Growth Factor Beta</td>
</tr>
<tr>
<td><strong>TLR</strong></td>
<td>Toll-like Receptor</td>
</tr>
<tr>
<td><strong>TLR7</strong></td>
<td>Toll Like Receptor 7</td>
</tr>
<tr>
<td><strong>TNF</strong></td>
<td>Tumor Necrosis Factor</td>
</tr>
<tr>
<td><strong>TNFR</strong></td>
<td>Tumor Necrosis Factor Receptor</td>
</tr>
<tr>
<td><strong>TNFRSF9</strong></td>
<td>Tumor Necrosis Factor Receptor Superfamily, Member 9</td>
</tr>
<tr>
<td><strong>TNIK</strong></td>
<td>TRAF2 And NCK Interacting Kinase</td>
</tr>
<tr>
<td><strong>Tra2β</strong></td>
<td>Transformer-2 Protein Homolog β</td>
</tr>
<tr>
<td><strong>tRNA</strong></td>
<td>transfer RNA</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>TSC2</td>
<td>Tuberous Sclerosis 2 Protein</td>
</tr>
<tr>
<td>UPS</td>
<td>Ubiquitin Proteasome System</td>
</tr>
<tr>
<td>V1C</td>
<td>Primary Visual Cortex</td>
</tr>
<tr>
<td>WB</td>
<td>Western Blotting</td>
</tr>
<tr>
<td>YPEL</td>
<td>Yippe-like</td>
</tr>
<tr>
<td>YPEL4</td>
<td>PEL4, Yippee-like 4 (Drosophila)</td>
</tr>
<tr>
<td>ZBP1</td>
<td>Zipcode-binding Protein</td>
</tr>
<tr>
<td>ZNC</td>
<td>Nucleocytoplasmic Shuttling Signal</td>
</tr>
<tr>
<td>ZnF</td>
<td>Zinc-finger</td>
</tr>
</tbody>
</table>
ABSTRACT

Heterogeneous ribonucleoproteins (hnRNPs) are a family of RNA-binding proteins (RBPs) implicated in several steps of RNA metabolism, including transcription, pre-mRNA splicing, mRNA transport and turnover. Therefore, alteration of their physiological levels may lead to many pathological disorders, such as neurodegeneration and cancer.

In this thesis, we focused the attention on three hnRNP proteins, namely DAZAP1, hnRNP Q and hnRNP R, that we previously identified in Drosophila melanogaster as strong modulators of TAR DNA-binding Protein 43 kDa (TDP-43) activity, an RBP involved in Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Lobar Degeneration (FTLD).

First of all, we evaluated the effects of their depletion on TDP-43-controlled mRNAs within human neuroblastoma SH-SY5Y cells and in a cellular model of TDP-43 loss-of-function. We found that DAZAP1 and, to a lesser extent, hnRNP Q were the most consistent modifiers of TDP-43 activity. Therefore, to examine the connection between DAZAP1 and TDP-43, we initially characterized the interaction between these two proteins by immunoprecipitation analysis, demonstrating that DAZAP1 does not bind TDP-43 but can bind TDP-43 controlled mRNAs. Next, we decided to identify all the potentially targets regulated by TDP-43 and DAZAP1, by looking at the transcriptome status of cells silenced for these two hnRNPs and we found differently expressed genes associated with neurodegeneration (ELAVL3 NOVA2, CELF5) and inflammation (TNF, TNFRSF9, ICAM1).

Finally, we extended our analysis to the characterization of hnRNP Q and hnRNPR, since the functional rescue of TDP-43 alterations was described only for hnRNP Q, but not for hnRNPR. We investigated the subcellular distribution and profiled differentially expressed genes analysis from RNA-seq after their knockdown. Interestingly, despite their high sequence similarity, these two proteins show different cellular distribution and affect
different cellular pathways, typically associated with neurodegeneration (*PENK, NGR3, RAB26, JAG1*) and inflammatory response (*TNF, ICAMI, TNFRSF9, ICAM5*). In conclusion, our work provides insights on the involvement of the hnRNP family in controlling neuronal and inflammatory pathways, and suggests that the differential expression of these proteins could play an essential role in modulating the onset as well as the progression of neurodegenerative disorders, in particular when related to the TDP-43 proteinopathies.
1. INTRODUCTION

1.1 RNA-Binding Proteins and post-transcriptional gene regulation in brain

1.1.1 A general overview of RNA-Binding Proteins

Post-transcriptional gene regulation is an essential process involved in the cellular metabolism of coding and non-coding RNAs, that is regulated by the cooperative interaction of different ribonucleoprotein (RNP) complexes, consisting of RNA-binding proteins (RBPs) transiently or stably associated with RNAs (Dreyfuss et al., 1993).

Although some RBPs were initially identified through the biochemical method, including members of the heterogeneous ribonucleoprotein (hnRNP) family (Pinol-roma et al., 1988; Dreyfuss et al., 1984), the growing amount of data generated by whole genome sequencing of all kingdoms of life have expanded the list of known RBP proteins. In an effort to validate RBPs across human genome, in 2014 Gerstberger and collaborators analysed 800 domains from the protein family (Pfam) database known to bind RNA or to be found in RNA-related proteins, identifying 1542 RBPs (about 7.5% of all protein-coding genes in humans) (Gerstberger et al., 2014).

In this census, RBPs were mainly divided into different groups based on literature reports (Figure 1A):

- messenger RNA (mRNA)-binding;
- pre-ribosomal RNA (pre-rRNA)-binding;
- transfer RNA (tRNA)-binding;
- small nuclear RNA (snRNA)-binding;
Introduction

- small nucleolar RNA (snoRNA)-binding;
- non-coding RNA (ncRNA)-binding;
- ribosomal proteins;
- diverse targets;
- unknown targets.

Moreover, phylogenetic analysis revealed the creation of gene families in which most of RBP paralogues share 20-70% sequence identity (Gerstberger et al., 2014), probably to increase protein synthesis or their regulation across different cell types (Figure 1B).

Figure 1. Classification of RNA-binding proteins. RNA-binding proteins (RBPs) (A) and RBP families (B) are divided in groups based on RNA targets. The percentage of RBPs in each category is shown. (Source: Gerstberger et al., 2014)

The importance of RBPs in the RNA metabolism became clear from their evolutionary conservation and abundance analysis. Bioinformatics analysis of yeast and Drosophila
genome have highlighted that 5-8% and 2-3% of genes are predicted to act as RBPs, respectively (Keene, 2001). Furthermore, a considerable number of orthologs of human RBPs have been found in lower organisms, such as archaea and bacteria (Anantharaman et al., 2002).

Despite the fact that the majority of human RBPs are ubiquitously expressed, sometimes at higher levels than average cellular proteins (Vaquerizas et al. 2009; Kechavarzi & Janga 2014), 82 specialized RBPs were found in brain, germ line, bone marrow, muscle, and liver. Overall, specialised RBPs represent 2% of all RBP families (Gerstberger et al., 2014). Examples of tissue specific RBPs are proteins involved in the germ line development and maintenance, such as the Deleted in Azoospermia (DAZ) (Reynolds and Cooke, 2005), the Insulin-like Growth Factor-II mRNA-binding Proteins 1 (IGF2BP1) and 3 (IGF2BP3) (Hammer et al., 2005) and P-element Induced Wimpy Testis (PIWI) (Siomi et al., 2011). Moreover, it is interesting to note that alterations in the cellular levels of RBPs are frequently recognized following their association with accessory binding partners (Han et al., 2010), during cellular senescence (Masuda et al., 2009) or in the presence of diseases (Pino et al., 2003).

1.1.2 Structure of RNA-Binding Proteins

As a structural point of view, RBPs have a modular architecture comprising multiple RNA binding domains that create a larger binding interface to recognize much longer stretches of nucleic acids (Figure 2) (Lunde et al., 2007). Specific regions called RNA-binding domains (RBDs) are responsible for the recognition of RNA sequences or RNA structural motifs belonging to different transcripts. In eukaryotes, the RNA-recognition motif (RRM) is by far the most common and characterized RBD. This highly conserved domain was first identified in the late 1980s by biochemical analysis which revealed the association between
heterogeneous nuclear RNAs (hnRNAs) with proteins (Dreyfuss et al., 1993). RRMs are composed of 80-90 amino acids disposed in a four-stranded anti-parallel β-sheet flanked on one side by two α-helices (α1 and α2) giving the β1α1β2α2β3β4 topology (Figure 2) (Oubridge et al., 1994). In this structure the RNA binding surface is represented by an arginine or lysine residue together with two conserved aromatic residues, located on the surface of the two central β-strands, β3 and β1, in the ribonucleoprotein (RNP) motif-1 (RNP1) and ribonucleoprotein (RNP) motif-2 (RNP2), respectively (Oubridge et al., 1994). Analysis of conserved residues in the RRM-nucleic acid interface led to the definition of a common architecture of RRM-nucleic acid complex. In particular, the binding platform of RRM is based on a dinucleotide recognition by four conserved protein side chains in the centre of the β-sheet (Maris et al., 2005). Beyond these general features, the loop between secondary structure elements and the secondary structure elements themselves can be modified, in order to increase the RNA-binding affinity. For instance, the C-terminus of RRM2 and RRM3 in hnRNP I (also known as PTB) forms an extra β-strand antiparallel to β2 (Simpson et al. 2004), and the α-helix 1 in U2AF RRM is three fold longer than the canonical one (Kielkopf et al., 2004). Higher affinity can also be achieved by the combination of two or more RRM domains, as demonstrated for Polyadenylate Binding Protein (PABP) (Maris et al., 2005). Another particular feature of some RRM domains is the ability to mediate protein-protein interactions. This contact can occur between two RRMs, between an RRM associated to RNA and a non-RRM protein that acts as a cofactor, and between RRMs that do not bind RNAs and another protein. An example of the first type of protein-protein interaction is the case of hnRNP A1, in which the two RRMs contained in the N-terminal region make contacts by using their α-helix 2 (Xu et al., 1997). Examples of the second and third types are the CBP20-CBP80-RNA complex and the Y14-Magoh complex, in which Y14 binds Magoh through the β-sheet of
the RRM domain, respectively (Maris et al., 2005).

Despite the abundance of RRM, other well-established domains can be involved in the RNA recognition, including the K-homology (KH) domain, the zinc-finger (ZnF) domain and the double-stranded RNA binding domain (dsRBD) (Lunde et al., 2007). The characteristics of these domains and other RBDs are summarized in Figure 2.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Topology</th>
<th>RNA-recognition surface</th>
<th>Protein–RNA interactions</th>
<th>Representative structures (PDB ID)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRM</td>
<td>αβ</td>
<td>Surface of β-sheet</td>
<td>Interacts with about four nucleotides of siRNA through stacking, electrostatics and hydrogen bonding</td>
<td>U1A-N terminal RRM19 (1URN)</td>
</tr>
<tr>
<td>KH (type I)</td>
<td>αβ</td>
<td>Hydrophobic cleft formed by variable loop between β2, β3 and GXXG loop, Type II: same as type I except variable loop is between α2 and β2</td>
<td>Recognizes about four nucleotides of siRNA through hydrophobic interactions between non-aromatic residues and the bases; sugar-phosphate backbone contacts from the GXXG loop, and hydrogen bonding to bases</td>
<td>Nova-1 KH (type II) (1E0A), NusA (type II) (2ASB)</td>
</tr>
<tr>
<td>dsRBD</td>
<td>αβ</td>
<td>Helix α1, N-terminal portion of helix α2, and loop between β1 and β2</td>
<td>Shape-specific recognition of the minor–major groove pattern of dsRNA through contacts to the sugar-phosphate backbone; specific contacts from the N-terminal α-helix to RNA in some proteins</td>
<td>dsRBD3 from Staufen115 (1EK2)</td>
</tr>
<tr>
<td>ZnF-CCHH</td>
<td>αβ</td>
<td>Primarily residues in α-helices</td>
<td>Protein side chain contacts to bulged bases in loops and through electrostatic interactions between side chains and the RNA backbone</td>
<td>Fingers 4–6 of TFIIBα (1JUG)</td>
</tr>
<tr>
<td>ZnF-CCH</td>
<td>Little regular secondary structure</td>
<td>Aromatic side chains form hydrophobic binding pockets for bases that make direct hydrogen bonds to protein backbone</td>
<td>Stacking interactions between aromatic residues and bases create a kink in RNA that allows for the direct recognition of Watson–Crick edges of the bases by the protein backbone</td>
<td>Fingers 1 and 2 of TFIID (1KGO)</td>
</tr>
<tr>
<td>S1</td>
<td>β</td>
<td>Core formed by two β-strands with contributions from surrounding loops</td>
<td>Stacking interactions between bases and aromatic residues and hydrogen bonding to the bases</td>
<td>Ribonuclease S1 (2X1), exosome69 (2N65)</td>
</tr>
<tr>
<td>PAZ</td>
<td>αβ</td>
<td>Hydrophobic pocket formed by OB-like β-barrel and small αβ motif</td>
<td>Recognizes single-stranded 3’ overhangs of siRNA through stacking interactions and hydrogen bonds</td>
<td>PAZ2 (1S31), Aragonaute136 (1U4D), Dicer78 (2GPL)</td>
</tr>
<tr>
<td>PIWI</td>
<td>αβ</td>
<td>Highly conserved pocket, including a metal ion that is bound to the exposed C-terminal carboxylate</td>
<td>Recognizes the defining 5’ phosphate group in the siRNA guide strand with a highly conserved binding pocket that includes a metal ion</td>
<td>PIWI5 (LYTU), Aragonaute (1U4D)</td>
</tr>
<tr>
<td>TRAP</td>
<td>β</td>
<td>Edges of β-sheets between each of the 11 subunits that form the entire protein structure</td>
<td>Recognizes the GAG triplet through stacking interactions and hydrogen bonding to bases; limited contacts to the backbone</td>
<td>TRAP23 (1C9S)</td>
</tr>
<tr>
<td>Pumilio</td>
<td>α</td>
<td>Two repeats combine to form binding pocket for individual bases; helix α2 provides specificity-determining residues</td>
<td>Binding pockets for bases provided by stacking interactions; specificity dictated by hydrogen bonds to the Watson–Crick face of a base by two amino acids in helix α2</td>
<td>Pumilio14 (1M8Y)</td>
</tr>
<tr>
<td>SAM</td>
<td>α</td>
<td>Hydrophobic cavity between three helices surrounded by an electrostatic region</td>
<td>Shape-dependent recognition of RNA stem-loop, mainly through interactions with the sugar-phosphate backbone and a single base in the loop</td>
<td>Vts123 (2E5E)</td>
</tr>
</tbody>
</table>

**Figure 2. Common RNA-binding domains.** For each domain is reported a short description of the structure and properties. (Source: Lunde et al., 2007)

Together with these latter, auxiliary modules, such as regions rich in serine and arginine (S/R), arginine and glycine (R/G) or lysine and arginine (K/R)-rich basic patches (Biamonti and Riva, 1994) help increase the functional diversity of RBPs and their ability to bind multiple targets. Amongst these different auxiliary domains, the glycine-rich type is
the most frequent. This is found in most basic heterogeneous riboproteins (hnRNPs), in Nucleolin, in Fibrillarin and in several yeast RBPs (Biamonti and Riva, 1994). Basically, it contains repeats of RGG and/or RG sequence, collectively referred as RGG/RG motifs or RGG boxes, that are able to mediate RNA-binding, protein-protein interaction and cellular localization in normal and stress conditions (Thandapani et al., 2013). Moreover, the functional properties of the RGG/RG motifs can be affected by their interaction with proteins containing Tudor domains that are able to recognize arginine methylated RGG/RG motifs (Thandapani et al., 2013).

Additional layers of complexity in the structure of RBPs are achieved by post-translational modifications, including phosphorylation, arginine methylation and small ubiquitin-like (SUMO) modification (Glisovic et al., 2008). For example, hnRNP A/B, hnRNP C and hnRNP U are phosphorylated in vivo, and hnRNP A1/A2 are also methylated on arginine residues in the RGG box domain (Dreyfuss et al., 1993). The functions of these modifications seem to modulate the interactions between these proteins and RNAs or other proteins. SUMO modifications were found in hnRNP C and hnRNP M where they could contribute to the subcellular localization of these two proteins and the nucleocytoplasmic transport of associated mRNA targets (Vassileva and Matunis, 2004).

1.1.3 Different functions of RNA-binding proteins

As mentioned before, RBPs are involved in all the aspects of RNA biology, from transcription, pre-mRNA splicing and polyadenylation to modification, transport, stability and degradation of all RNAs (Figure 3) (Glisovic et al., 2008). Most importantly, RBPs can assemble in distinct multimeric complexes in order to control RNA life cycles (Romano and Buratti, 2013).
In eukaryotes, the maturation of RNAs is a process that involves RNA cleavage and ligation reactions, in which nucleotide sequences called introns and exons are removed and joined together respectively. In protein-coding mRNAs (and long non-coding RNAs) this important step of RNA metabolism occurs in a large ribonucleoprotein (RNP) machinery, called the spliceosome (Lee and Rio, 2015). Pre-mRNA splicing takes place for > 95 to 100% of human genes, and at least 74% of human genes transcribe multiple mRNAs through alternative splicing (Lee and Rio, 2015; Glisovic et al., 2008). Extrinsic nonspliceosomal RNA-binding proteins play a role in controlling alternative splicing, basically promoting or inhibiting the recognition of alternative regions by the spliceosome machinery. Among these splicing regulator proteins are Serine Arginine Rich Splicing
Factors (SR proteins), hnRNPs and tissue-specific RBPs (Baralle and Giudice, 2017). An example of the participation of RBPs in alternative splicing events is the regulation of β-tropomyosin pre-mRNA, in which two SR proteins called SF2/ASF and SC35 act in an opposite way on the intronic enhancer-dependent splicing of exon 6A (Van Der Houven Van Oordt et al., 2000).

Other two important steps in post-transcriptional gene regulation are the polyadenylation of the 3'-end of pre-mRNAs and the RNA editing. Polyadenylation involves an RBP, called Nuclear Poly(A) Binding Protein (PABPN1) that is responsible for the addition of the poly(A) tail on the extremity of each transcript after cleavage of 3'-end (Glisovic et al., 2008). The Polyadenylation Specificity Factor (CPSF) complex stimulates the poly(A) polymerase activity of PABPN1. RNA editing prevalently involves the deamination of adenosine (A) to inosine (I). This conversion requires the presence of enzymes belonging to the ADAR family, which are RNA-binding proteins. Indeed, ADAR2 combines its dsRBD with an auxiliary deaminase domain that is responsible for this conversion (Valente and Nishikura, 2005).

Once pre-mRNA processing is completed, the transcripts are exported from nucleus to the cytoplasm and this further level of RNA metabolism is promoted by a network of RBPs. These include the Zipcode-binding Protein (ZBP1), which is an RBP involved in the trafficking and translation of β-actin mRNA (Hüttelmaier et al., 2005).

Finally, RBPs are also key factors of mRNA turnover and stability, as described by the regulation of neuronal specific transcripts by RNA-binding proteins belonging to the ELAV/Hu family (Keene, 1999).
1.1.4 Implication of RNA-Binding proteins in brain development, maintenance and disease

RNA-binding proteins are key factors of neuronal development. Several studies have demonstrated that neurons have a unique system for regulating RNA metabolism, consisting of tissue-specific RBPs (De Conti et al., 2016).

In mammals, the ELAV/Hu family is involved in the stability and trafficking of specific transcripts containing AU-rich elements (ARE) predominantly in neurons. These transcripts are particularly interesting because they encode functional subsets of Early-Response Genes (ERG) or immediate gene products, including proto-oncoproteins and cytokines (Keene, 2001). The ELAV/Hu family is composed of four members: HuR, that is widely expressed and HuB, HuC and HuD which are specifics of neurons. HuR (also called HuA) has a high level of homology with a Drosophila nuclear protein namely Embryonic Lethal Abnormal Division (ELAV). Elav deletion is lethal due to abnormal development of neurons (Robinow and White, 1991). The importance of ELAV/Hu proteins is also relevant in pathology. Indeed, it is interesting to note that tumours outside the nervous system can express these proteins leading to Paraneoplastic Neurologic Syndrome (PNS) (Albert and Darnell, 2004).

The crosstalk of different RPBs controls RNA splicing patterns essential for neuronal development and maintenance. In neuronal progenitors, Polypyrimidine Tract-binding Protein 1 (PTBP1) induces the exon skipping of Polypyrimidine Tract-binding Protein 2 (PTBP2) exon 10 to form a transcript with a premature stop codon subject to non-sense mediated mRNA decay (NMD). Nevertheless, when the progenitors start to differentiate into neurons, the negative effect of PTB1 on PTBP2 exon 10 is lost by PTBP1 downregulation and at the same time upregulation of the splicing enhancer Ser/Arg Repetitive Matrix Protein 4 (SRRM4) (Quesnel-Vallières et al. 2015; Makeyev et al.
Another important RBP implicated in brain development and differentiation is the Neuro-Oncological Ventral Antigen 2 (NOVA2). This protein inhibits Disabled Homologue 1 (DAB1) exon 7bc splicing inclusion in late-generated cortical and Purkinje neurons. In turn, DAB1 regulates the Reelin signalling, an important pathway involved in neuronal development. Neuronal migration defects found in neurons with the DAB1 longer isoform, corresponding to the inclusion of exon 7b and 7c, are similar to those observed in Nova2−/− mice (Yano et al. 2010). Similarly to ELAV/Hu proteins, the abnormal expression of NOVA1/NOVA2 in breast cancer determines the establishment of a PND form called Opsoclonus Myoclonus Ataxia (POMA). In fact, the alteration of NOVA proteins leads to immune system activation and consequent formation of autoantibodies responsible for neuronal cell death in the regions of the nervous system where NOVA proteins are expressed (Lukong et al., 2008).

RNA regulation during neuronal differentiation can also involve alternative splicing of the RBP mRNAs themselves. During synaptic development, exon 19 of RNA-binding Protein FOX1 Homologue 1 (RBFOX1) is alternatively spliced by RBFOX proteins, giving rise to nuclear or cytoplasmic protein isoforms. Moreover, mutations in RBFOX proteins have been found in people suffering from autism and are related to alterations of transcripts, such as SH3 and Multiple Ankyrin Repeat Domains Protein 3 (SHANK3), Voltage-dependent L-type Calcium Channel Subunit α1C (CACNA1C) and Tuberous Sclerosis 2 Protein (TSC2) (Baralle and Giudice, 2017).

Finally, the spatial and temporal organization of gene expression is a particular feature adopted by the nervous system to control the synaptic plasticity (Zukin, Richter and Bagni 2009). In order to avoid the premature translation and degradation of specific mRNAs, silent mRNAs, RBPs and miRNAs are associated in macromolecular assemblies, including
Introduction

synaptic mRNA-silencing foci, processing bodies (P-bodies) and stress granules (SGs). Silent mRNAs wait inside the granules to reach the exact localization in which translation can start (De Conti et al., 2016). Examples of RBPs that participate to the formation of SG granules are hnRNPA2, TATA-Box Binding Protein Associated Factor 15 (TAF15), fused in Sarcoma/Translocated in Liposarcoma (FUS/TLS) and TAR DNA-binding Protein 43 (TDP-43). Stress conditions, mutations in RBPs, aging and other pathological disorders can alter the equilibrium between soluble, disperse and aggregated states leading to increased SG formation, and progressively the formation of stable and long-lived aggregates that resemble the characteristics of pathological inclusions. Indeed, several studies have identified the presence of SGs in Alzheimer’s diseases (AD), amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) (Wolozin & Apicco 2015; De Conti et al. 2016).

1.2 Transcriptome analysis using RNA sequencing and its applications in the characterization of RNA binding proteins

1.2.1 From hypothesis-driven approaches to high-throughput screening (HTS) approaches

The analysis of the complete set of all RNA molecules transcribed in a cell at a specific developmental stage or physiological condition, called its transcriptome, has given insights into how the genome is packaged and expressed, and has been useful to clarify important aspects related to genome evolution (Wang, 2001; Elliott, 2014).

In the last decade, several efforts have been made to clarify the role of RBPs in transcriptome regulation, and their involvement in neurodegeneration. Initially, the characterization of novel regulatory factors in the field of RNA metabolism was based on
traditional hypothesis-driven approaches, in which the investigation of biological effects was initiated by a prediction (Buratti, Romano and Baralle, 2013). For example, preliminary studies correlating splicing inclusion of *Cystic Fibrosis Transmembrane Conductance Regulator (CFTR)* exon 9, and the number of UG(m)/U(n) repeats close to the 3’ splice site of exon 9 (Cuppens et al., 1998; Larriba et al., 1998), identified TDP-43 as the major hnRNP involved in this splicing recognition (Buratti et al., 2001). The advantages of using this approach are the capability to set up a precise series of experiments aimed at confirm or reject the original hypothesis. However, the process that is necessary to reach an in-depth characterization is laborious and could lead to the loss of some information important for the comprehension of the initial question (Buratti, Romano and Baralle, 2013). Thanks to the advancement of system-wide approaches and bioinformatics, the analysis of gene expression changes and the identification of RNA/protein targets at a global cellular level have became more popular and have allowed more open-ended approaches that are not strictly hypothesis-driven (Buratti et al., 2013).

The term high-throughput screening (HTS) is referred to the way of rapidly assessing a large number of biological factors and chemical compounds in order to identify those which can modulate a specific cellular pathway (Buratti, Romano and Baralle, 2013). Examples of these new techniques aimed to analyse transcriptomes are microarrays (GeneChip), RNA sequencing (RNA-seq) and cross-linking immunoprecipitation sequencing (CLIP-seq) (Figure 4).
Introduction

Figure 4. High-throughput screening (HTS) approaches. Flowchart of microarray (A), RNA-seq (B) and CLIP-seq (C) experiments. (Source: Buratti et al., 2013)

Briefly, microarray screening consists of a pool of DNA sequences spotted on a solid surface that are hybridized with RNAs extracted from cells or tissues, labelled with fluorescent probes, washed and then finally analysed based on the fluorescence emitted (Dalma-Weiszhausz et al., 2006). These arrays can be designed in order to detect predicted splicing junctions, and therefore to analyse possible changes in splicing events.

Despite microarrays being extensively used for the characterization of RNA binding factors, including TDP-43 (Colombrita et al., 2012; Narayanan et al., 2013), hnRNP K (Liu and Szaro, 2011), hnRNP L (Hung et al., 2008) and hnRNP A1 (Liu et al., 2017), this approach has some limitations, such as the necessity to design/synthesize the probes that have to be attached on the solid surface. Hence, knowledge of the sequence to make the probe set is mandatory, making it impossible to detect mRNAs absent from databases or
with incomplete/incorrect genome annotations (Hurd and Nelson, 2009). Furthermore, the phenomenon of cross-hybridization makes it problematic to analyse and obtain reproducible results for highly related sequences. This restricts microarray analysis to non-repetitive genome sequences, therefore making arduous the analysis of splicing isoforms, allelic gene variants, and single-nucleotide polymorphisms (SNPs) (Shendure, 2008; Hurd and Nelson, 2009; Buratti, Romano and Baralle, 2013). Other limitations regard signal-to-noise ratio and the problems related to detect low-abundance sequences, the necessity to obtain DNA from PCR amplification with the possibility to introduce bias into samples, and finally, the wide amount of preparative methodologies, microarray formats and analytical approaches currently available (Hurd and Nelson, 2009; Ioannidis et al., 2009). Some drawbacks have been solved by the sequencing of transcriptome. Therefore, in the next paragraphs, the attention will be focalised on the RNA-seq analysis and its applications in the characterization of RNA binding proteins.

1.2.2 Next-generation sequencing (NGS): RNA sequencing

The concept of sequencing was developed in the late 70s by the pioneering works of Sanger and Coulson (Sanger, Nicklen and Coulson, 1977) and Maxam and Gilbert (Maxam and Gilbert, 1977). However, the effective power of this approach applied in the research field, has been possible later with the support of industries and the development of instruments able to produce hundreds of gigabases of nucleotide sequence output in a single run (Metzker, 2010). Nowadays, we talk about next-generation sequencing (NGS) technology to describe seq-based methods, such as RNA-seq. NGS analysis has allowed us to overcome the limitations of other HTS approached not seq-based (e.g. microarray). Indeed, the gene expression analysis carried out using RNA-seq does not require the prior knowledge of genome sequences, making possible the detection of RNA editing events and
the quantification of splicing isoforms (Hurd and Nelson, 2009; Sánchez-Pla et al., 2012; Buratti, Romano and Baralle, 2013). Moreover, the fact that the sample of interest is directly sequenced results in the removal of experimental bias from the analysis, such as cross-hybridation issue and mapping of repetitive sequences, as well as problems in the quantification of signal, because it is based on counting sequence tags rather than the relative measurement between samples. Finally, nanograms of materials are necessary and sufficient for sequencing, reducing the reliance of extensive PCR amplification of the material (Hurd and Nelson, 2009).

However, there are some shortcoming in the RNA-seq approach, including the necessity to analyse biological replicates to avoid problems related to the sample assignment to lanes or runs, and the normalization of the counts after the alignment to a reference genome. This normalization resolves issue related to the libray size, the gene length and the distribution of counts among samples (Sánchez-Pla et al., 2012).

Differential gene expression (DGE) analysis (Costa-Silva, Domingues and Lopes, 2017), characterization of alternative splicing events (Park et al., 2018) and detection of gene fusion events (Kumar et al., 2016), as well as specific RNA species (Andrés-León, Núñez-Torres and Rojas, 2016) are examples of RNA-seq applications. The type of RNA-seq application affects the choice of the sequencing platform and the relative bioinformatics analysis.

Nowadays, there are different available platforms, which were initially developed for DNA, but they can easily be applied to RNA sequencing after the conversion from RNA into more stable cDNA molecules by reverse transcriptase (Sánchez-Pla et al., 2012). Moreover, each platform uses different methodologies for template preparation, sequencing-imaging and gene alignment, determining a wide range of data output in terms of cost and quality (Metzker, 2010). Examples of platforms are Illumina (Miseq, NextSeq
500, HiSeq 2500, HiSeq X Ten), Life Technologies Ion Torrent (PGN and Proton systems), Pacific Biosciences PacBio RSII and Oxford Nanopore Technologies MinION (Reuter et al., 2015). For instance, Illumina’s and Ion Torrent’s technologies provide the possibility to generate short read length (for HiSeq 2500 and Ion PGN: reads length up to 125 bp and 200 or 400 bp, respectively) making them less suitable for particular studies, including gene isoform detection (Reuter et al., 2015). On the other hand, PacBio RII and MinION systems features to obtain long-read sequence data (> 14 kb and ~ 6-65 kb, respectively) suitabled for the identification of gene splicing events, although they are limitated by the lower coverage comparing to Illumina and Ion Torrent platforms (Reuter et al., 2015).

1.2.3. RNA sequencing applied to the characterization of RNA binding proteins

An increasing number of works described the application of sequencing analysis to explore all the possible RNA targets of a specific RBP (Buratti, Romano and Baralle, 2013), as well as to profile the transcriptomes of brain tissue affected from neurodegenerative disorders (Twine et al., 2011). In this regard, RNA-seq has been used to analyse the effect of TDP-43 knockout in the central nervous system of Drosophila (Hazelett et al., 2012) and in mouse embryonic stem cells (Chiang et al., 2010).

Nowadays, improvements in sequencing technology also provide the possibility to map the sites at which proteins are bound to the RNA, by coupling this technique with RNA immunoprecipitation (RIP) (Marchese et al., 2016) or cross-linking immunoprecipitation (CLIP) (Licatalosi et al., 2008). In RIP-seq analysis the RNA-complexes formed between the protein of interest and its interacting RNAs are co-immunoprecipitated and the bound RNAs are subsequently sequenced, while in the CLIP-seq analysis the target protein is
previously cross-linked \textit{in vitro} with its RNA targets and successively these RNAs are immunoprecipitated and sequenced.

For instance, in order to describe TDP-43 targets in TDP-43-containing protein complexes within rat cortical neurons, deep sequencing analysis was performed following RNA-immunoprecipitation (Sephton et al., 2011). RIP-seq was also used to characterize mRNA targets of hnRNP L in prostate cancer LNCaP cells (Fei et al., 2017) and to identify spliceosome-associated genes in hnRNP H1 targets (Uren et al., 2016). Interestingly, most of the hnRNP H1 targets enriched in the analysis encoded other RBPs, including TDP-43, FUS, several hnRNPs and Serine and Arginine Rich Splicing Factors (SRSFs) (Uren et al., 2016). One the other hand, CLIP-seq analysis has been used for mapping \textit{in vivo} binding positions of other RNA-binding proteins, including NOVA (Licatalosi et al., 2008), CELF1 (Le Tonquèze et al., 2016) and Transformer-2 Protein Homolog β (Tra2β) (Grellscheid et al., 2011).

\section*{1.3 The Heterogeneous Ribonucleoprotein (hnRNP) family}

\subsection*{1.3.1 General characteristics of the hnRNP family}

The most abundant RBPs belong to the heterogeneous ribonucleoprotein (hnRNP) family (Glisovic et al., 2008). hnRNPs were originally described as RBPs transiently bound to the heterogeneous nuclear RNA (hnRNA, also known as pre-mRNAs) produced by RNA polymerase II in the nucleus of eukaryotic cells (Dreyfuss et al., 1993). The 40S core particle was the first mRNA-protein complex, isolated by sucrose density gradients, and contained hnRNP A/B and C (Beyer et al., 1977). Later on, several other hnRNPs were characterized by immunoprecipitation with monoclonal antibodies and UV-cross-linking experiments leading to the identification of 20 major types of hnRNPs. These were
alphabetically named from hnRNP A1 to hnRNP U (Figure 5) (Geuens et al., 2016; Han et al., 2010). Although, the hnRNPs have many features in common with other RBPs, such as the SR proteins and ELAV-like proteins, they are generally described as a separated family. The basis for this classification arises from historical motivations, rather than phylogenetic, structural or functional evidence (Han et al., 2010).

Besides these canonical hnRNPs, several studies have classified other proteins as hnRNP-like RNA binding factors, such as CELF proteins, FOX proteins, NOVA, or TDP-43 (Busch and Hertel, 2012).

**Figure 5. Structure of the canonical hnRNP family members.** hnRNP proteins have different molecular weight ranging from 34 kDa to 120 kDa. Several structural features are shared among the hnRNP family, including RNA-binding domains (RRM, RNA-recognition motif; qRRM, quasi-RNA recognition motif and KH, K-homology domain), Arginine-glycine-glycine box (RGG-box) and auxiliary domains. (Source: Geuens et al., 2016)

Structurally, hnRNP proteins share common features, including RBDs (with the exception of hnRNP U). The majority of hnRNPs also have RGG-boxes and auxiliary domains. In addition, hnRNPs can have different splicing variants and undergo post-translation
modifications and nucleocytoplasmic shuttling (Han et al., 2010). However, regarding the latter, there are difference between the mechanisms that different hnRNPs adopt to translocate from nucleus and cytoplasm and vice-versa. For instance, hnRNP A1 and hnRNP I do not contain canonical nuclear localization signals (NLSs) and their translocation is coupled with transcription (Gama-Carvalho and Carmo-Fonseca, 2001). Interestingly, it has also been demonstrated that stress-dependent phosphorylation of hnRNP A1 is responsible for its cytoplasmic accumulation within SGs (Guil et al., 2006). On the contrary hnRNP K contains NLS and its movement is uncoupled with transcription (Gama-Carvalho and Carmo-Fonseca, 2001).

Although hnRNPs are prevalently abundant in the nucleus, it has been ascertained that hnRNP U and C are also found in cytoplasmic RNA granules (Kanai et al., 2004; Piñol-Roma and Dreyfuss, 1992), one of the hnRNP M isoforms is a membrane-bound receptor (Bajenova et al., 2003) and hnRNP Q is predominantly cytoplasmic and it is localized in mRNA granules (A Mizutani et al., 2000). hnRNP P (FUS) is also localized in cytoplasm inside stress granules (Andersson et al., 2008).

1.3.2 Multiple functions of hnRNPs

Several of the biological functions of hnRNPs involve cooperative and dynamic interactions with different proteins. Indeed, across all the steps required for mRNA biogenesis, there is a continuous remodelling of the hnRNP complex, represented by the association of particular hnRNPs with pre-mRNAs, in which the loss or acquisition of hnRNP proteins as well as other RBPs create what has been classically defined as an “mRNP code”. This “mRNP code” is unique for each transcript and the protein composition of the complex reveals its cytoplasmic fate (Singh and Valcárcel, 2005). Many of the functions of RBPs previously described in this thesis take place thanks to
hnRNP proteins and their interplay. For example, a role in transcriptional regulation was first described in 1996, when hnRNP K was found to bind in vitro to pyrimidine-rich (CT) elements in single strand DNA (ssDNA) (Michelotti et al., 1996). Successively, it has been demonstrated that hnRNP K can act as a transcriptional repressor of the *Osteocalcin* gene by inhibiting the formation of a transcriptional complex on the CT element of the *Osteocalcin* promoter (Stains et al., 2005). The transcriptional repressor activity can also take place without direct contact with DNA, as it occurs during the inhibition of *a1-Acid Glycoprotein* (*AGP*) gene (Miau et al., 1998).

Among all the processes that influence transcriptome complexity, alternative splicing is the most important event in which hnRNPs are involved. Recently, it has been demonstrated that multiple hnRNPs, among which TDP-43, hnRNP L, PTB/nPTB and hnRNPA1/A2 are able to modulate with disparate strength the inclusion of *Sortilin 1* (*SORT1*) exon 17b (Mohagheghi et al., 2016).

Interestingly, there is also substantial evidence supporting the role of RBPs in the maintenance of genome integrity and in the regulation of telomeres (Nishida et al., 2017). Indeed, hnRNP L and hnRNP C have been identified in complexes that work as DNA double strand breaks (DSB) sensors (Nishida et al., 2017). FUS/TLS is another transcriptional regulatory sensor of DNA damage, and is recruited to the *Cyclin D1* promoter after ionizing radiation exposure where it functions as a transcriptional repressor (Lagier-Tourenne et al., 2010).

Furthermore, FUS/TLS interacts through its RGG domains with a G-quadruplex consisting of telomere DNA and the functional non-coding RNAs called Telomeric Repeat-containing RNA (TERRA) to regulate histone modifications of telomeres (Takahama et al., 2015). In addition, hnRNP A/B together with other hnRNPs and telomere-binding factors, regulates the structure of telomeres, protecting them from degradation (Han et al., 2010).
1.3.3 Role of hnRNPs in neurodegenerative disorders

Albeit the equilibrium of hnRNP proteins is finely regulated and neurons are characterized by a very adaptive and dynamic architecture, perturbation of their neuronal homeostasis can lead to neurodegenerative disorders, such as ALS, FTLD, spinal muscular atrophy (SMA) and Alzheimer’s disease. In particular, these pathologies are characterized by the presence of protein aggregates, also including RBPs/hnRNP proteins (Neumann et al., 2006). The formation of these aggregates is a result of several steps consisting of protein misfolding, oligomerization and eventually formation of protofibrils or amyloid fibrils (Stefani and Dobson, 2003). A clear example of the involvement of hnRNPs in neurodegeneration is the finding of TDP-43 and FUS/TLS in ALS and FTLD inclusions. These inclusions can act as a “sink” influencing the clearance of these proteins from nucleus and it can, in turn, explain the aberrant RNA processes (e.g. alternative splicing) observed in aggregation-bearing cells. Both TDP-43 and FUS contain a prion-like domain that confers them an intrinsic propensity to aggregate (Gitler and Shorter, 2011). Together with this consideration, variations in the protein sequence can trigger the establishment of diseases. In both TDP-43 and FUS coding genes the presence of ALS-linked mutations have been described, and in the case of TDP-43 it has been demonstrated that TDP-43 ALS-linked mutations accelerate aggregation in vitro (Ugras and Shorter, 2012). More recently, the presence of ALS/FTLD-associated mutations have also been described in the prion-like domain of hnRNP A1, A2/B1 and putative mutations have also been hypothesized by a yeast functional screening in TAF15 encoding gene (Couthouis et al., 2011).

Most importantly, these hnRNPs seem to not act alone in mediating neurodegeneration. Indeed, it is thought that GGGGCC repeat expansion in the C9orf72 gene are the causative agents of associate ALS/FTLD pathologies, due to the fact that these repetitions can cause
the subtraction of a significant amount of RBPs from the protein milieu (De Conti et al., 2016).

A plethora of intracellular systems are designed to maintain this intracellular homeostasis, including the ubiquitin proteasome system (UPS) (Kriegenburg et al., 2012) and the autophagy system (Feng et al., 2014). In the UPS, proteins are primarily targeted for degradation through the covalent attachment of a small protein, called ubiquitin, and then directed to the proteasome complex for degradation (Glickman and Ciechanover, 2002). Disruption of proteasome activity is sufficient to induce cytoplasmic accumulation and aggregation of TDP-43 in primary hippocampal and cortical neurons, as well as in the immortalized motor neuron cell line NSC-34 (van Eersel et al., 2011).

In the autophagy system, harmful substances contained in the cytoplasm are degraded within lysosomes/vacuoles, and the resulting macromolecular constituents are recycled (Feng et al., 2014). Recently, it has been demonstrated that the induction of autophagy system triggered the neuronal clearance of TDP-43 with consequent enhancement of survival in human Induced Pluripotent Stem Cell (iPSC)-derived neurons and astrocytes obtained from patients with familial ALS (Barmada et al., 2014).

1.4 Trans-activating Response Region (TAR) DNA-Binding protein of 43 kDa (TDP-43)

1.4.1 Brief introduction on TDP-43

Trans-activating response region (TAR) DNA-binding protein of 43 kDa (TDP-43) is a ubiquitously expressed and highly conserved protein of 414 amino acid residues belonging to the heterogeneous nuclear ribonucleoprotein (hnRNP) family. Originally described as a transcription factor bound to the regulatory element in the Human Immunodeficiency
Virus Type 1 (HIV-1) long terminal repeat known as TAR (Ou et al., 1995), in 2001 TDP-43 was rediscovered as a splicing factor involved in the regulation of CFTR exon 9 and successively linked to different aspects of RNA metabolism, such as mRNA stability (including its own), mRNA transport, translation, and non-coding RNAs (miRNA, lncRNAs, etc.) processing (Buratti and Baralle, 2012).

Furthermore, although mostly nuclear, TDP-43 continuously shuttles from the nucleus to the cytoplasm where it forms high molecular mass complexes that include proteins and RNAs (Sephton et al., 2011; Ayala, Zago, et al., 2008).

In 2006, TDP-43 was identified as a major component of the ubiquitin-positive, α-Synuclein and Tau-negative cytoplasmic inclusions recognized in the brain and spinal cord tissues of Frontotemporal Dementia with Ubiquitin-positive Inclusions (FTLD-U) and ALS patients (Arai et al., 2006; Neumann et al., 2006). Successively TDP-43 inclusions were also reported at lower frequencies in other neurodegenerative diseases, including Alzheimer’s, and Parkinson’s (PD) disease and Lewy body (LB) disorders (Nakashima-Yasuda et al., 2007). Following the growing amount of data regarding the involvement of TDP43 protein in neurodegeneration, several other studies have focused their attention on understanding the biochemistry mechanisms behind cellular functions of TDP-43 in physiology and pathology. Some of these findings are summarized in the paragraphs below.

### 1.4.2 Characteristics of TDP-43

TDP-43 is encoded by the TARDBP gene located on the short arm (p) of human chromosome 1 at position 36.22 (1p36.22) (Figure 6), and is highly conserved among Drosophila melanogaster, Caenorhabditis elegans, mouse and human (Wang et al., 2004). The TARDBP gene consists of 6 exons. Exons 2-6 encode the functional TDP-43 protein.
However, the genomic analysis of *TARDBP* gene performed by Wang and colleagues and successively by Alton et al., have also highlighted the presence of different splicing isoforms of TDP-43 (D’Alton et al., 2015; Wang et al., 2004). Furthermore, more than 38 nonsynonymous *TARDBP* mutations have been identified in both familiar and sporadic ALS (Scotter et al., 2015).

![Figure 6. Localization of TARDBP gene on human chromosome 1.](source)

The Genomic locus encoding TDP-43 is identified by a yellow arrow. (Source: Genetic home resource)

TDP-43 expression is developmentally regulated. High levels of TDP-43 expression were found during early embryogenesis, but progressively and markedly decrease with age in post-natal brains (Sephton et al., 2010). In agreement with this finding, Huang and colleagues demonstrated the expression of TDP-43 in all post-natal tissues, such as cerebellum, forebrain, brainstem, spinal cord, skeleton muscle, heart, liver, lung, spleen and kidney, with a global decrease of TDP-43 levels with the advancement of age (Huang et al., 2010). Furthermore, TDP-43 knockout mice were also lethal during embryonic stages (Sephton et al., 2010). This observation has also been confirmed in *Drosophila* where the disruption of TDP-43 was lethal at the second larval stage (Fiesel et al., 2010), although in other *Drosophila* models, TDP-43 null flies were able to survive through adulthood (Feiguin, Vinay K. Godena, et al., 2009), suggesting the influence of modifier genes affecting TDP-43 in developmental regulation.

From a structural point of view, TDP-43 architecture reassembles the general hnRNP structural organization (Figure 7).
Figure 7. Schematic representation of TDP-43 protein domains and their associated functions. TDP-43 is composed of an N-terminal domain containing NLS, two RNA-recognition motifs (RMM1 and RRM2 in which NES is present) and a C-terminal domain containing a glycine-rich region (GRR). ALS-linked mutations are also reported in the figure. (Source: Janssens and Van Broeckhoven, 2013)

TDP43 protein is composed of:

- The **N-terminal domain**: contains the nuclear localization sequence (NLS) that is composed of NLS1 (K82RK84) and NLS2 (K95VKR98). Mutations in these two positions reduce the nuclear localization of TDP-43 (Buratti and Baralle, 2001).

- Two **RRMs**: highly evolutionary conserved 60 residue-long motifs. *D. melanogaster* and *C. elegans* homologs can functionally substitute for the human protein and vice-versa (Feiguin, Vinay K. Godena, et al., 2009; Ayala et al., 2005).

1. **RRM1**: is responsible for binding to single-stranded RNA containing UG-repeated motifs (Ayala, Zago, et al., 2008; Buratti and Baralle, 2001). These repeats are found more often in long introns and 3’ untranslated regions (3’-UTRs) rather than exons inside the human genome. Through RRM1, TDP-
43 also binds single-stranded DNA with TG repeats, as demonstrated by the electromobility shift assay (EMSA) technique (Buratti and Baralle, 2001), and double-stranded TG repeats DNA as described in the nitrocellulose filter-binding assay (Kuo et al., 2009). In particular, these interactions are mediated by two residues (Phe 147 and Phe 149) and their substitution was found to be sufficient to abolish RNA-binding properties and splicing regulatory activities of TDP-43 (Buratti & Baralle 2001; D'Ambrogio et al. 2009).

2. **RRM2**: Together with the N-terminal sequence from residues 3 to 183, RRM2 is supposed to be involved in the TDP-43 self-interaction (ability to form homo-dimers and -multimers) (Shiina et al., 2010; Kuo et al., 2009). A role in chromatin organization has also been suggested, since it may co-crystallize with TG-rich single stranded DNA in order to form highly thermal-stable dimeric assemblies (Ayala, Zago, et al., 2008). Furthermore, this region of TDP43 protein is predicted to contain a strong nuclear export signal (NES) (Strong et al., 2007).

- The **C-terminal domain**: represents almost half of the protein, spanning residues 261-414. It contains a glycine-rich region, corresponding to residues 345 to 366, that is essential for the interaction with other proteins, including hnRNP A2/B1 and hnRNP A1 (Buratti et al. 2005; D'Ambrogio et al. 2009). Furthermore, the region comprising amino acids 216–315 together with the RNA-recognition motif 1 domain are implicated in the assembly of TDP-43 into SGs (Colombrita et al., 2009). This region is also the “hot-spot” position in which the majority of TARDBP mutations have been identified in ALS and FTLD patients (Lagier-Tourenne et al., 2010). Moreover, the C-terminal domain includes a
Introduction

Glutamine/Asparagine (Q/N) spanning residues 321-369 (D’Ambrogiò et al. 2009; Fuentealba et al. 2010) that is considered the putative “prion-like domain” of TDP-43, since it has a strong tendency to aggregate (Budini et al. 2012). Deletions and mutations in the TDP-43 C-terminal domain significantly reduce the solubility of the protein and results in formation of large nuclear and cytoplasmic aggregates, thought to be a result of the lack of interaction between TDP-43 and other protein factors that promote TDP-43 solubility (Ayala, Zago, et al., 2008). Finally, the integrity of this region is extremely important for many of TDP-43’s functional properties. Indeed, TDP-43 protein lacking the C-terminal domain is no longer able to regulate the skipping of exon 9 in the CFTR gene (Wang et al., 2004).

- **TDPBR** (extended binding region for TDP-43): is present in the 3’-UTR of TDP-43 mRNA and contains several non-UG sequences that are essential for the auto-regulation of TDP-43 mRNA (Igaz et al., 2011; Polymenidou, Lagier-tourenne, et al., 2011).
1.4.3 Functions of TDP-43

As previously mentioned, TDP-43 has many functional properties that are summarised in Figure 8.

Figure 8. Schematic representation of TDP-43 functions. TDP-43 covers many functions in cells including starting from transcription to translation and degradation. TDP-43 is also involved in miRNA processing and auto-regulation. (Source: Buratti & Baralle, 2012)

TDP-43 functions include:

- **Transcriptional regulation**: TDP-43 was initially discovered as having a role in HIV-1 transcriptional repression, from which its name derives (Ou et al., 1995). However, the involvement of TDP-43 in HIV-1 regulation has not been confirmed by other studies, suggesting a restrictive role during HIV-1 replication in immune cells (Nehls et al., 2014). TDP-43 was also implicated in the regulation of spermatogenesis, because of its binding to the promoter of the mouse Spermatid-specific SP-10 (ACRV1) gene and promoting its transcriptional repression (Lagier-Tourenne et al., 2010). Moreover, consisting with is involvement in transcription,
TDP-43 was found in the human brain and in cell cultures to be associated with euchromatin (containing actively transcribed genes) through its RMM2 (Lagier-Tourenne et al., 2010).

- **Pre-mRNA splicing**: is the best-characterized function of TDP-43. An example of an alternative splicing event regulated by TDP-43 is the skipping of exon 9 of *CFTR* gene, which is promoted by the TDP-43 recognition of an UG-tract near the 3’ splice-site of this exon (Buratti and Baralle, 2001; Buratti et al., 2001). As in the case of *CFTR* gene, TDP-43 is considered to be a splicing silencer of other genes, like *Apolipoprotein AII (ApoII)* exon 3 (Mercado et al., 2005), *Eukaryotic Translation Termination Factor 1 (ETF1)* and *Retinoid X receptor gamma (RXRG)* exon 7 (Passoni et al., 2012). However, TDP43 can also promote exon splicing, as in the case of *breast cancer 1 (BRCA1)*-mutated substrate exon 12 (Passoni et al., 2012) and *polymerase delta interacting protein/S6 kinase 1 Aly/REF-like* target (*POLDIP3/SKAR*) exon 3 (Fiesel et al., 2012).

- **micro-RNA (miRNA) biogenesis**: mass spectrometry analysis has shown an interaction between TDP-43 and the Nuclear RNase III-type Protein (Drosha) that is involved in the first step of miRNA maturation (Gregory et al., 2004). A further confirmation of the involvement of TDP-43 in the miRNA biogenesis has arisen from the demonstration of TDP-43’s association with Argonaute 2 (AGO2) and DDX17 (Freibaum et al., 2010). Indeed, DDX17 directly interacts with Drosha (Shen and Hung, 2015) and AGO2 is a catalytic component of the RNA-induced Silencing Complex (RISC) (Sen and Blau, 2005).

- **RNA subcellular localization and translation**: although TDP-43 is predominantly nuclear, it has been visualized in cytoplasmic puncta strongly stained for PABPC1, a marker for cytoplasmic RNA SGs (Freibaum et al., 2010). It is thought that
cytoplasmic RNA granules may represent physiological accumulations of mRNAs and ribonucleoproteins that regulate gene expression by affecting translation, trafficking and stability (Anderson and Kedersha, 2009). This function is particularly important for neuroplasticity in which the mRNA translation is dependent on its subcellular localization. In the motoneuron-like cell line NSC-34, TDP-43 moved into SGs as a response to oxidative stress and to environmental insults (Colombrita et al., 2009). Although the link between TDP-43 and its presence into SGs is still not completely clear, it is thought that it can mediate the stabilization and localization of Low Molecular Weight Neurofilament (NEFL) mRNA during axonal repair (Lagier-Tourenne et al., 2010) by binding its 3'UTR region from 135 to 185 residues after the stop codon (Buratti and Baralle, 2008). Furthermore, in 2008 Wang and colleagues have observed the presence of TDP-43 in dendritic processing body and demonstrated that it could act as a translational repressor in vitro (Wang et al., 2008).

- **mRNA turnover**: TDP-43 is associated with the mRNA turnover from different genes, including *Cyclin-Dependent Kinase 6 (CDK6)* (Ayala, Misteli, et al., 2008) and *Histone Deacetylase 6 (HDAC6)* (Fiesel et al., 2010) through the binding to their 3'UTR within the mRNA sequences. Furthermore, TDP-43 can physically interact with the mRNA of the *Drosophila melanogaster* homologue of *Microtubule-associated Protein 1B*, called *Futsc*, preventing synaptic defects related to *Futsc* downregulation (Godena et al., 2011) and that the depletion of TDP-43 could destabilized the mRNA of *Autophagy Related 7 (ATG7)* with the consequent impairment of autophagy and accumulation of polyubiquitinated proteins (Bose, Huang and Shen, 2011).
- **Auto-regulation**: TDP-43 is able to regulate own expression by promoting RNA instability through the binding to different polyadenylation (polyA) sites localized in the TDPBR at its 3'UTR (Ayala et al., 2011). Under normal conditions, two main splicing isoforms and one isoform expressed at low levels arise from the usage of either the proximal polyA1 site or the distal polyA4 and polyA2 sites, respectively, with the efficient production of TDP-43 (Ayala et al., 2011; Eréndira Avendaño-Vázquez et al., 2012). When the nuclear levels of TDP-43 are abundant there is an increase in the binding of TDP-43 to its TDPBR. This event promotes a dual effect: the activation of a normally silent intron (intron 7) which in turn eliminates utilisation of polyA1, resulting in the preferentially usage of polyA2. At the same time, this determines the direct competition of TDP-43 with the cleavage polyA factor Cstf-64 over polyA1 with the consequence usage of polyA4 (Eréndira Avendaño-Vázquez et al., 2012). Both polyA2 and polyA4 are significantly retained in the nucleus and rapidly degraded thus reducing TDP-43 levels (Eréndira Avendaño-Vázquez et al., 2012). On the contrary, low levels of TDP-43 causes the preferentially usage of polyA1 site instead of the other possible choices (polyA2 and polyA4), resulting in an efficient TDP-43 production (Eréndira Avendaño-Vázquez et al., 2012).

### 1.4.4 TDP-43 interacting proteins

Given the importance to understand the RNA-processing functions, nuclear–cytoplasmic shuttling and aggregation properties of TDP-43, many recent studies have focused their attention on identifying TDP-43 interacting proteins. Nowadays, the most important binding partners of TDP-43 are members of the hnRNP family, such as hnRNP A2/B1, hnRNP A1, and hnRNP C1/C2 (Buratti et al., 2005). These protein interactions are
Introduction

probably essential for the modulation of the functional properties of TDP-43. For example, in the case of hnRNPA1 the alteration of mRNA processing following the loss of this RNA–binding protein correlates well with the formation of TDP-43-positive inclusions in sporadic ALS motor neurons (Honda et al., 2015). Further evidence that hnRNP proteins can play a modulatory role on TDP-43 function has been provided using cell and animal models. In particular, functional experiments of the fruit fly orthologs of the human hnRNP A1/A2 (Hrp38) overexpression in HEK293 cells have demonstrated that TDP-43 is required for the spicing inhibitor activity of this protein (Romano et al., 2014). Additionally, experiments show that the interaction between TDP-43 and fly hnRNP A2/B1 homologues Hrb87F and Hrb98DE is important to suppress CGG repeat-induced neurotoxicity in a Drosophila model of Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) (He et al., 2014).

Notwithstanding all this evidence, the relationship between pathological mislocalization of TDP-43 and general levels of hnRNP proteins, with regards to altered RNA processing events occurring in neurodegenerative diseases, still remains unclear. In an effort to answer this question, Molecular Pathology Laboratory has been recently demonstrated that the interaction of TDP-43 with hnRNP L, PTB/nPTB and hnRNP A1/A2 affects the inclusion of the exon 17b in the neurotrophic receptor SORT1 mRNA, a pathologically-relevant splicing event known to be regulated by TDP-43 (Mohagheghi et al., 2016). Interestingly in this work, Mohagheghi and colleagues also looked at the expression levels of these hnRNPs in the temporal cortex of FTLD-TDP patients, in which a significant increase in exon 17b inclusion was previously observed. A significant increase in the splicing repressors hnRNP A1/A2 and PTB in FTLD-TDP-43 was observed in patients relative to controls. This may suggest that different hnRNP expression levels, which may depend on individual or age-related differences, on top of similar TDP-43 pathologies could
Introduction

potentially account for different disease courses, with regards to both onset and progression.

Given these results, our group decided to further characterize the linkage between TDP-43 and hnRNP proteins, with the aim of finding other potential TDP-43 modulators belonging to this family. Taking advantage of *Drosophila melanogaster* and its available models of TDP-43 proteinopathies (see paragraph below), it was then possible to identify three promising hnRNPs, namely DAZAP1, hnRNP Q and hnRNP R (Appocher et al., 2017) that have subsequently become the object of my PhD thesis. Therefore, what is currently known about these factors will be reviewed in the next paragraphs.

1.4.5 TDP-43 proteinopathies

Abnormal cellular distribution and post-translational modifications of TDP-43 are present in many neurodegenerative disorders, which are now generally referred to as TDP-43 proteinopathies (Scotter et al., 2015). In pathology, TDP-43 was described as the major component of the ubiquitin-positive, α-Synuclein and Tau-negative cytoplasmatic inclusions recognized in the brain and spinal cord tissues of FTLD-U and ALS patients (Arai et al., 2006; Neumann et al., 2006). These intracellular aggregates are characterized by ubiquitination, S409/410 phosphorylation, and proteolytic cleavage forming C-terminal fragments (CTFs) of 20-25 kDa (Neumann et al., 2006). TDP-43 inclusions have also been reported to occur in other neurodegenerative disorders, including 30% of Alzheimer patients and various forms of Parkinson’s disease, although TDP-43 inclusions exist side-by-side but only partially co-localize with tau or α-Synuclein aggregates (Lagier-Tourenne et al., 2010). Two hypotheses have been suggested to explain the pathogenic role of TDP-43 inclusions: gain-of-function (GOF) and loss-of-function (LOF). It is important to note that neither hypothesis excludes the other one, and they can act at the same time to
determine the onset and progression of the disease (Lee et al., 2011). Briefly, the “gain-of-function” mechanism is based on the cytotoxicity induced by the formation of aggregates or toxic C-terminal fragments localized in the cytoplasm (Lee et al., 2011). According to this model, stress conditions can lead to redistribution of TDP-43 into the cytoplasm resulting in aggregation. Due to the fact that the auto-regulation of TDP-43 mRNA is a cytoplasmic event, it is possible that TDP-43 aggregates can still bind RNA and therefore auto-regulate themselves. This consequently may result in the decrease of newly synthesized TDP-43 proteins in affected cells and hence, nuclear clearance. The loss of normal nuclear TDP-43 may further increases cellular stress, again resulting in a vicious circle leading to cell death (Figure 9).

**Figure 9. Schematic representation of the TDP-43 gain-of-function model.** Stress conditions lead to the formation of cellular aggregates (1). These aggregates promote TDP-43 auto-regulation (2), determining a decrease of TDP-43 synthesis (3) and loss of TDP-43 nuclear functions (4), with consequent cell death (5). (Source: Lee et al., 2011)

On the other hand, the “loss-of-function” model is based on the hypothesis that cellular stress may lead to the formation of pathological stress granules or mislocalization of RNA
transport granules that are usually localized within peripheral neuritis, determining the formation of cytoplasmic inclusions that act as a “TDP-43 sink” with the clearance of this protein from the nucleus and the consequent dysregulation of many TDP-43-controlled events in RNA processing (Figure 10) (Igaz et al., 2011).

![Figure 10. Schematic representation of TDP-43 loss-of-function model. Stress conditions lead to cytoplasmic mislocalization and formation of cellular aggregates (1). These aggregates promote loss of TDP-43 from nucleus (2), determining an increase of TDP-43 synthesis (3) and aggregation (4), with consequent cell death (5). (Source: Lee et al., 2011)](image.png)

1.4.6 Drosophila as a model for understanding TDP-43 proteinopathies

The interest in Drosophila melanogaster as a powerful organism for studying cell biology became evident in the early twentieth century. Initially, Drosophila was introduced as an experimental model for investigating the Mendelian theory of heredity (Muqit and Feany, 2002), but successively it has been widely used in the neurodegeneration field. Indeed, Drosophila presents many advantages, including a short life cycle (from 40 to 120 days), a compact genome, the possibility to screen substances/drugs very quickly and a high
neuronal complexity due to receptors, ion channels and neurotransmitters also found in humans (Romano et al., 2012).

There are different species of *Drosophila* with differences in morphology, genetic background and habitat. Interestingly, *Drosophila melanogaster* has approximately 75% of genes whose the equivalent human homologue is associated with disease (Romano et al., 2012).

Regarding neurodegenerative disorders, various fly models have been used to understand the pathological mechanisms behind Parkinson’s, Alzheimer’s and Huntington’s disease, well as spinocerebellar ataxia, tauopathies and TDP-43 proteinopathies (Muqit and Feany, 2002).

Here, I am going to present some of the characteristics of the main homologue of human TDP-43 in *Drosophila*, called TBPH, and some of the findings obtained using this organism as an experimental model.

Briefly, TDP-43 and TBPH have a strong homology at the level of the region containing RRM1 and RRM2 (Figure 11A) (Wang et al., 2004). Another characteristic in common with TDP-43 is the presence of different splicing isoforms (Figure 11B) (Wang et al., 2004), as well as, the preferential binding to UG-rich sequences (Ayala et al. 2005).
**Figure 11. Comparison of human TDP-43 and Drosophila TBPH proteins.** Schematic representation of the protein structure in both (A) and alignment of the human (TDP-43) and fly (TBPH) proteins (B) are reported. (Source: Modified from Romano et al., 2012)

Functionally, TBPH shows same splicing inhibitory effects both in vitro and in vivo (Ayala et al., 2005), and although the C-terminal tail of TBPH represents the region with less...
homology with TDP-43 it can still bind to the same nuclear protein interactors as TDP-43 (D’Ambrogio et al. 2009).

Most importantly, it is possible to achieve different fly phenotypes, including null allele flies (loss-of-function model), transgenic flies overexpressing wild-type TDP-43 or ALS-related mutations (gain-of-function model), and transgenic flies silenced for this factor (RNA interference transgenic flies).

These approaches are described as follows:

- **Loss-of-function model (null allele flies):** were created by the disruption of *TBPH* gene to test if an ALS-like phenotype could be generated. Different approaches were used to achieve this aim, including: knockout of *TBPH* through the partial removal of its coding and regulatory sequence (Feiguin et al., 2009); point mutation (G>A) in order to create a stop codon (Lu et al., 2009); deletion of 932 bp at the 5’ end of *TBPH* (Lin et al., 2011); and the generation of null allele in which the entire coding sequence was deleted (Fiesel et al., 2010). Interestingly, in this latter approach, Fiesel and collaborators observed lethality at the second larval stage. This is in contrast to the *TBPH* knockout approach of Feiguin and collaborators that has reported viability of some larvae after embryogenesis, with most of them able to reach the pupal stage and undergo metamorphosis but do not eclose, remaining trapped inside their pupal cages. However, the knockout flies able to get rid the external cuticle presented uncoordinated movements, reduction of axonal branches and synaptic boutons present inside muscles (Feiguin et al., 2009). Nevertheless, the overexpression of human TDP-43 (hTDP-43) was able to rescue the aberrant phenotype (Feiguin et al., 2009; Lu et al., 2009).

- **Transgenic flies:** were achieved by the creation of transgenic flies using the Gal4-UAS system (Figure 12) (Brand and Perrimon, 1993). This system permits the
ectopic expression of a human transgene (gain-of-function model), as well as the knockdown of an endogenous gene (RNA interference transgenic flies) in a specific tissue or cell type. For instance, it has been widely demonstrated that the overexpression of hTDP-43 or TBPH is able to cause eye degeneration associated with cell death (Hanson et al., 2010; Estes et al., 2011). In a similar way, the overexpression of wild-type TDP-43 in mushroom bodies and in motor neurons caused neuronal death and axonal loss, whereas the overexpression of TDP-43 in motor neurons lead to cytoplasmic and axonal aggregation (Li et al., 2010). Loss of motor neurons and reduction in life span were also identified after the overexpression of ALS/FTLD-linked and synthetic mutations (Voigt et al., 2010). Furthermore, it has been demonstrated that the ectopic expression of wild-type TDP-43 or its variants was associated with age- and dose-dependent toxicity (Hanson et al., 2010). TDP-43 toxicity was also found to correlate with its subcellular localization (Ritson et al., 2010; Miguel et al., 2011). Indeed, cytoplasmic and nuclear inclusions of TDP-43 were found toxic in different tissues, such as eyes, muscles and glia (Miguel et al., 2011). However, the formation of inclusions was not a prerequisite for TDP-43 toxicity, since increased levels of both nuclear and cytoplasmic TDP-43 were found toxic even in absence of inclusions (Miguel et al., 2011). In contrast, the generation of transgenic flies silenced for TBPH exclusively in neuronal tissue was found to induce similar locomotive defects observed in the null allele flies (Feiguin, et al., 2009). Moreover, Li and colleagues showed that the depletion of TBPH in the central nervous system and the mushroom bodies of Drosophila caused axonal loss and neuronal death (Li, et al., 2010).
Transgenic flies were not used only to assay the TDP-43 misregulation in the context of ALS-linked mutations, but also to clarify the role of TDP-43 in other degenerative disorders, as in the case of the Inclusion Body Myopathy associated with Paget’s Disease of Bone and Frontotemporal Dementia (IBMPFD) (Ritson et al., 2010).

**Figure 12. Gal4-UAS system.** This system is based crossing two transgenic fly lines. The first fly line (UAS–human-transgene fly) harbours the transgene of interest that is placed downstream of a UAS activation domain consisting of GAL4-binding sites. The second fly line (enhancer-trap GAL4 fly) harbours a yeast transcriptional activator called GAL4 located downstream of a cell- or tissue- specific promoter. These latter flies are known as “drivers”. In this system, the ectopic expression of the transgene of interest is obtained by crossing the two fly lines. There is a wide array of promoters available for the ectopic expression of a transgene, such as the pan-neural promoter *elav* (embryonic lethal, abnormal vision) or the eye-specific promoter GMR (Glass Multimer Reporter). (Source: Muqit & Feany, 2002)
1.5 DAZ Associated Protein 1 (DAZAP1)

1.5.1 Characteristics of DAZAP1

DAZ associated protein 1 (DAZAP1) was originally identified as an RBP associated with the DAZ proteins, DAZ and DAZ-Like (DAZL), using a yeast two-hybrid system (Tsui et al., 2000). DAZAP1 is encoded by the DAZAPI gene located on the short arm (p) of human chromosome 19 at position 13.3 (19p13.3) (Figure 13). DAZAPI consists of 12 exons spanning about 28 kb (Dai et al., 2001; Tsui et al., 2000).

Figure 13. Localization of DAZAPI gene on human chromosome 19. The Genomic locus encoding DAZAPI is identified by a yellow arrow. (Source: Modified from Genetic home resource)

Two major DAZAPI transcript variants of 1.75 and 2.4 kb are generated through alternative polyadenylation (Yang and Yen, 2013; Tsui et al., 2000). These transcripts are highly abundant in human testis (Dai et al., 2001; Tsui et al., 2000) and ovaries (Pan et al., 2005), with a low expression level in non reproductive organs including liver, brain, placenta, heart, skeletal muscle, kidney and pancreas (Dai et al., 2001; Tsui et al., 2000). Furthermore, DAZAP1 is extremely conserved among the evolution. Human and mouse DAZAP1 share 89% similarity in the coding sequence, although they diverge in protein sequence for 9 amino acid residues: 7 substitutions and 2 deletions/insertions. Furthermore, both of the mammalian proteins share 89% similarity and 81% identity with the Proline-rich RNA Binding Protein (Prrp) of Xenopus (Dai et al., 2001).
From a structural point of view, the full-length protein represents the major splicing isoform that is composed of 407 amino acid residues (≈ 45 kDa), with two RRMs in the N-terminal portion, and a proline-rich (Pro-rich) C-terminal region (Figure 14) (Dai et al., 2001; Tsui et al., 2000). The high level of similarity of these RRMs with those present in many other hnRNPs (e.g. Drosophila hnRNP 27C) makes DAZAP1 a member of the hnRNP protein family (Dai et al., 2001; Tsui et al., 2000).

**Figure 14. Schematic representation of DAZAP1 domains.** The full-length DAZAP1 (407 amino acid residues) is composed of an N-terminal domain containing two RNA-recognition motifs (RRM1 and RRM2) and a C-terminal domain containing a proline-rich region (Pro-rich).

Furthermore, DAZAP1 was found to shuttle between nucleus and cytoplasm during different stage of mouse spermatid maturation, as well as in non germinal cells in response to external stimuli (Vera et al., 2002; Lin and Yen, 2006). In 2006, a study carried out in somatic cells demonstrated DAZAP1’s association with nuclear structures as a component of hnRNP particles, and mapped the region comprising the last 25 amino acids at the C-terminus (C25) as important residues for the nuclear localization (Lin and Yen, 2006). This C25 region does not share sequence similarity to other NLS or shuttling signals, and it also contains the nuclear export signal for cytoplasmic localization. Therefore, it was called the nucleocytoplasmic shuttling signal (ZNC) (Lin and Yen, 2006).

The sublocalization of DAZAP1 was also found to be associated with the transcription state of its mRNA targets (Lin et al., 2012) and the acetylation of the Lysine 150 residue within its second RRM (Sasaki et al., 2012). In particular, the signal necessary and
sufficient for the transcription-dependent nuclear localization of DAZAP1 was represented by a sequence of 42 amino acid residues located at the N-terminal (N42) of DAZAP1 (Lin et al., 2012), and the mutation of Lysine 150 was correlated with its cytoplasmic retention and association with mitochondria (Sasaki et al., 2012).

1.5.2 Functions of DAZAP1

Nowadays, little information is available in literature regarding the natural RNA substrates and cellular functions of DAZAP1, albeit, as a member of hnRNP family, its cellular activity can be related to the regulation of RNA metabolism.

The first evidence supporting this hypothesis came out when DAZAP1 was found to be associated with RNA homopolymers, preferentially polyU and polyG tracts, in vitro (Dai et al., 2001; Tsui et al., 2000); and later when it was identified as an AU-rich element (ARE) binding protein bound to the ARE of the Tumor Necrosis Factor-alpha (TNF-α) (Rousseau et al., 2002) and the Cellular Oncogene Fos (c-Fos) (Morton et al., 2006) mRNAs. Two consensus sequences (AAAUAG and GU1-3AG) and two secondary structures (loops L1 and L2) were further characterized in the RNA ligands of the mouse DAZAP1 orthologous and in the 3'UTR of the all known target mRNAs of Xenopus Prrp (Hori et al., 2005). Interestingly, the AUG sequence contained in both the DAZAP1 consensus sequences was also reported in the mRNA targets of other RBPs, including hnRNP A1 (Burd and Dreyfuss, 1994), hnRNP D (Kajita et al., 1995) and Musashi1 (Msi1) (Imai et al., 2001). In this context, it has been proposed the role of DAZAP1 in the transport/localization (Lin et al., 2012; Lin and Yen, 2006; Zhao et al., 2001) and translation of different mRNAs is through the binding to their 3'UTRs (Smith et al., 2011; Yang et al., 2009; Morton et al., 2006). In particular, DAZAP1 was found to be accumulated in the cytoplasm when transcription was inhibited, suggesting that this protein
could translocate from nucleus to cytoplasm with exported mRNAs (Lin and Yen, 2006) as described for hnRNP A1 (Piñol-Roma and Dreyfuss, 1992). Moreover, DAZAP1 was able to stimulate translational initiation of mRNAs in a manner that was sensitive to the adenylation status of those mRNAs (Smith et al., 2011).

Other interesting findings demonstrate that the nuclear fraction of DAZAP1 is able to affect the splicing of different transcripts. It was identified, together with hnRNP A1, as a splicing-inhibitor factor of the mutant T6 of BRCA1 exon 18 (Goina et al., 2008). DAZAP1 was also implicated in the mRNA splicing of the Ataxia Teleangiectasia Mutated (ATM) gene through the interaction with Alu-derived Intron Splicing Enhancer (ISE) and the consequent inclusion of a cryptic exon (Pastor and Pagani, 2011). DAZAP1 was also found to regulate the intron inclusion of different mRNA targets in a minigene system (Chen et al., 2013), and was described as binding to an intronic splicing enhancer or silencer in human cells (Wang et al., 2012, 2013). Moreover, a recent publication has highlighted the role of DAZAP1 in cell proliferation and migration. Indeed, mRNA-seq analysis of HEK293 cell line identified several genes, most of them related to the maintenance of cell proliferation, regulated by the splicing activity of DAZAP1. This activity was also found to be affected by the Extracellular Signal-regulated Protein Kinase (MEK/Erk) signalling pathway through the phosphorylation of the DAZAP1 proline-rich domain (Choudhury et al., 2014).

1.5.3 DAZAP1 interacting proteins

Few DAZAP1 binding partners are described in the literature. The first well-characterized interactors are certainly DAZ and DAZL (Dai et al., 2001; Tsui et al., 2000), but other putative interacting proteins were established later. In most of the cases, the DAZAP1-protein contacts occurred between the RRM5 of DAZAP1 and the C-terminus of its
partners, as described for hnRNP U like 1, hnRNP A/B, hnRNP A1, hnRNP D and DEAD-Box Helicase 20 (DDX20) (Yang et al., 2009). However, more recently it has been reported that the transcription-dependent nuclear localization of DAZAP1 could be involved in the association of DAZAP1 to its partners, as described for the binding to the Steroid Receptor RNA Activator Stem-loop Interacting RNA Binding Protein (SLIRP) (Lin et al., 2012).

Moreover, the phosphorylation state of DAZAP1 was demonstrated to disrupt the interaction with binding partners, affecting its cellular functions. Indeed, the phosphorylation of Thr269 and Thr315 in its Proline-rich C-terminus by the Extracellular Signal Regulated Kinase 2 (ERK2) could promote the dissociation of DAZAP1 from DAZ, allowing the interaction of DAZ with other proteins involved in mRNA translation such as PABP (Morton et al., 2006).

1.5.4 DAZAP1 and pathologies

A link between alteration of DAZAP1 homeostasis and the development of disorders is still not clear. However, two studies have described how the t(1;19)(q23;p13.3) translocation present in TS-2 Acute Lymphoblastic Leukemia (ALL) cells could lead to the expression of two in-frame fusion proteins between DAZAP1 gene and Myocyte Enhancer Factor 2D (MEF2D) gene with oncogenic properties: DAZAP1/MEF2D and MEF2D/DAZAP1. Indeed, DAZAP1/MEF2D and MEF2D/DAZAP1 efficiently activated the transcription of genes important for lymphocyte growth and retained RNA-binding properties, respectively, indicating for the first time DAZAP1 as a putative factor implicated in cancer (Prima and Hunger, 2007; Prima et al., 2005).
1.6 Heterogeneous Nuclear Ribonucleoprotein Q (hnRNP Q) and Heterogeneous Nuclear Ribonucleoprotein R (hnRNP R): two closely related hnRNP proteins

1.6.1 Characteristics of hnRNP Q and hnRNP R

The RNA binding protein hnRNP Q, also known as SYNCRIP, was first described in 1997 as a nucleocytoplasmic protein interacting with the synaptotagmin isoform II (Syt-II) C2AB domain in mouse brain lysate (Mizutani et al., 1997). hnRNP Q was subsequently found in association with human Survival of Motor Neurons (SMN) (Mourelatos et al., 2001). The hnRNP Q protein is encoded by the SYNCRIP gene located on the long arm (q) of human chromosome 6 at position 14.3 (6q14.3) (Figure 15). SYNCRIP consists of 12 exons that are subjected to alternative splicing resulting in different isoforms, three of which are the most representative: hnRNP Q3, hnRNP Q2 and hnRNP Q1 (Figure 16A-C) (Mourelatos et al., 2001).

Figure 15. Localization of SYNCRIP gene on human chromosome 6. The Genomic locus encoding hnRNP Q is identified by a yellow arrow. (Source: Modified from Genetic home resource)

Regarding these isoforms, the nucleic acid sequence of hnRNP Q3 (also known as GRY-RBP) (Neubauer et al., 1998) encodes the longest variant with the closest protein sequence
identity (83% homology) to hnRNP R (Mourelatos et al., 2001). hnRNP Q3 is composed of 623 amino acid residues (≈ 70 kDa) containing an acidic domain (AcD), three RRM s, two NLS and a RGG-box (Figure 16B). The Q2 variant (≈ 65 kDa) lacks of 36 amino acids (302-336) between RRM2 and RRM3 comparing to the longest variant hnRNP Q3 (Figure 16A), while hnRNP Q1 (≈ 62 kDa) lacks of the second NLS and RGG-box region (549-623) with respect to hnRNP Q3 and contains a unique C-terminal domain (VKGVEAGPDLQQ) (Figure 16C) (Mourelatos et al., 2001). Furthermore, hnRNP Q is expressed in different tissues, such as spinal cord, muscle, brain, heart, kidney, liver and spleen (Rossoll et al., 2002).

![Figure 16. Schematic representation of protein domains and major isoforms of hnRNP Q.](image)

On the other hand, a 2.5 kb cDNA clone corresponding to hnRNP R was first described in 1998 in the serum of patients with autoimmunity symptoms (Hassfeld et al., 1998) and
subsequently identified, like hnRNP Q, as a factor bound to the SMN protein (Rossoll et al., 2002). The protein hnRNP R is encoded by the *HNRNPR* gene located on the p arm of human chromosome 1 at position 36.12 (p36.12) (Figure 17) and consists of 11 exons. Different isoforms are generated from alternative splicing events.

**Figure 17. Localization of *HNRNPR* gene on human chromosome 1.** The Genomic locus encoding hnRNP R is identified by a yellow arrow. (Source: Modified from Genetic home resource)

The isoform 1 of hnRNP R (hnRNP R1) is the most highly expressed and it is composed of 633 amino acid residues (≈ 82 kDa) containing an AcD domain, three RRM s, two NLS, a RGG-box and a Q/N-rich domain at C-terminus (Figure 18A) (Hassfeld et al., 1998). Moreover, this variant is ubiquitously expressed in all tissues (Huang et al., 2005; Rossoll et al., 2002). Recently a low level expressed and neuronal-specific isoform (hnRNP R2, ≈ 75 kDa) has been characterized. This isoform differs from the R1 because of the exclusion of exon 5 (114 bp) from the cDNA sequence, resulting in the absence of 41 amino acids (129-166) between the AcD domain and the first RRM (Figure 18B) (Huang et al., 2005).
Interestingly, hnRNP Q and hnRNP R are extremely conserved through evolution, although in lower organisms (e.g. *D. melanogaster* and *C. elegans*) they are represented by just one well-conserved orthologue (Cappelli et al., 2018; McDermott et al., 2012; Kabat et al., 2009). This suggests the occurrence of a progressive functional divergence into two distinct paralogues in higher organisms, such as mammals. In particular, the *Drosophila* homologue CG17838 (*Syp*) shares 46-47% sequence identity and ca. 57-60% of similarity to mammalian hnRNP Q and hnRNP R (Cappelli et al., 2018; McDermott et al., 2012). The human counterparts share almost 50% amino acid identity in their three RNA-binding domains with HRP-2, the nematode homologue of the mammalian hnRNP Q and R proteins (Kabat et al., 2009).

Regarding the subcellular localization of hnRNP Q and hnRNP R, the presence of a NLS motif in the protein sequence indicates the prevalent retention of these RBPs in the nucleus, even if translocation in the cytoplasm was observed in the case of hnRNP Q (McDermott et al., 2014; Quaresma et al., 2009; Akihiro Mizutani et al., 2000). Indeed, the transport of hnRNP Q within the dendrites of cultured rat hippocampal neurons has been

---

**Figure 18. Schematic representation of protein domains and major isoforms of hnRNP R.** The *HNRNPR* gene encodes two main splicing isoforms: hnRNP R1 (A) and hnRNP R2 (B). In this figure AcD (acidic domain), RRM1s (RNA-recognition motifs), NLS (nuclear localization signal), RGG (Arg-Gly-Gly)-box and Glutammine/Asparagine (Q/N)-rich domain are highlighted in colored boxes; relative sequence position and amino acid length of each isoform is also reported. (Source: Modified from Cappelli, Romano and Buratti, 2018)
described as a component of mRNA granules (Bannai et al., 2004). hnRNP Q was also found re-localized to cytoplasmic granules after stress condition (Quaresma et al., 2009). Moreover, in a study performed in 2006, the methylation status of the RGG box of hnRNP Q was reported to be associated with its shuttling (Passos et al., 2006).

Contrary to hnRNP Q, Mizutani and collaborators proposed the existence of a second putative bipartite-type NLS motif (KRKADGYNPDSKRR), similar to that observed for nucleoplasmin (Robbins et al., 1991), located at the C-terminal region of hnRNP R (572–586 amino acids) to explain the prevalently nuclear localization of hnRNP R (Mizutani et al., 2000). However, the expression of hnRNP R was not only detected in the nucleus, but was also found in cytoplasm and neurites of cultured embryonic motor neurons (Dombert et al., 2014; Rossoll et al., 2002) and in the cytoplasm, dendrites and axon terminals of rat bipolar cells (Peng et al., 2009).

### 1.6.2 Functions of hnRNP Q and hnRNP R

Functionally, hnRNP Q and hnRNP R are known to regulate different aspects of RNA maturation. For example, the C. elegans orthologue of hnRNP Q and hnRNP R was found to be involved in the alternative splicing of genes containing cassette exons flanked by an intronic UCUAUC motif (Kabat et al., 2009).

More specifically to the mammal hnRNP Q counterpart, the Q3 isoform was identified as a component of the spliceosome in 1998 (Mourelatos et al., 2001; Neubauer et al., 1998). Subsequently the capability of all the other variants to bind in vitro pre-mRNA and intron-containing RNA intermediates has been discovered (Mourelatos et al., 2001). Recently, hnRNP Q1 was found to promote inclusion of SMN2 exon7, whereas the other two variants (hnRNP Q2/Q3) were found to antagonize the activity of hnRNP Q1 by inducing exon 7 exclusion (Chen et al., 2008).
Controlling RNA editing was described as another hnRNP Q function, since the Q3 isoform was discovered to inhibit C-to-U RNA editing of the *Apolipoprotein B* (*apoB*) mRNA (Blanc et al., 2001). In addition, hnRNP Q3 can also affect mRNA transport, as demonstrated by its colocalization with ribosomal proteins and other RBPs in neuronal mRNA granules (Bannai et al., 2004; Kanai et al., 2004), and translation by competing with poly(A) binding proteins (Svitkin et al., 2013) or as in the case of mouse Cytochrome 1 (*Cry*) gene by binding its 5'UTR (Lim et al., 2016).

On the other hand, hnRNP R was identified as a splicing factor involved in the assembly of the pre-catalytic spliceosome during stress conditions (Tonevitsky et al., 2009). Depletion of hnRNP R impairs translocation of β-actin mRNA to the axonal growth cone in isolated embryonic mouse motor neurons (Glinka et al., 2010). Moreover, hnRNP R is involved in the transcription and degradation process of *c-Fos* mRNA in retinal cells (Huang et al., 2008) and in the expression of immunity factors (Meiningen et al., 2016; Reches et al., 2016).

Finally, hnRNP Q and hnRNP R regulate the morphology and the development of neural structures (McDermott et al., 2012, 2014; Chen et al., 2012). For instance, *Drosophila Syp* is required to regulate the postsynaptic morphology of neuromuscular junction (NMJ) through the maintenance of protein levels of important transcripts associated with neuron/cell projection morphogenesis or development (McDermott et al., 2014). *Drosophila Syp* also coordinates the translation in post-synapses with retrograde signalling to pre-synapses (Halstead et al., 2014). Moreover, hnRNP Q1 is also able to modulate neuronal morphogenesis and neurite branching in a mouse neuroblastoma cell line by interacting with different mRNAs related to Cdc42 signalling (Chen et al., 2012). Depletion of hnRNP R can reduce axonal 7SK, a nuclear RNA that modulates
transcription, and in turn led to defective axon growth in NSC-34 cells (Briese et al., 2018).

1.6.3 hnRNP Q and hnRNP R interacting proteins

The interest in hnRNP Q and hnRNP R characterization was initially promoted by the identification of SMN protein as a potential binding partner (Rossoll et al., 2002; Mourelatos et al., 2001). The RGG motif in the C-terminus of both hnRNP Q and hnRNP R was proposed to mediate this association (Rossoll et al., 2002).

At the present, however, few studies have described additional interacting proteins. Among them, hnRNP Q and hnRNP R were identified as putative TDP-43-associated proteins (Ling et al., 2010), as well as it has been demonstrated their interaction with Rev protein of HIV (Hadian et al., 2009). Moreover, hnRNP Q was found to bind specifically to galectin-3, a protein that stimulates proliferation in various cell types (Yoo et al., 2009) and the Q1 isoform was found to be associated through its Gly-rich C-terminus with the multifunctional nuclear phosphoprotein NS1 required for viral replication of the minute virus of mice (Harris et al., 1999).

1.6.4 Implication of hnRNP Q and hnRNP R in pathology

Nowadays, information about the development and progression of diseases associated with aberrant levels of hnRNP Q and hnRNP R is still incomplete. One study has investigated the role of hnRNP Q1 in tumorigenesis of colorectal cancer (Lai et al., 2017). Both hnRNP Q1 protein and mRNA levels were found to be upregulated in human tumor tissues compared with adjacent normal ones. Moreover, hnRNP Q1 was identify to bind to the 5'UTR of Aurora-A mRNA regulating its translation and in turn increasing cell proliferation and tumor progression (Lai et al., 2017).
Replication of the minute virus of mice (Harris et al., 1999) and HIV (Hadian et al., 2009) were also affected by hnRNP Q. In the first case, hnRNP Q1 was proposed as a cofactor in the double-stranded DNA-binding activity of NS1, whereas in the second study, the overexpression of hnRNP Q in concert with other hnRNP proteins, including hnRNP R, influenced the HIV-1 production by persistently infected astrocytes.
2. AIM OF THE PROJECT

As mentioned in the previous Chapters, several evidences have suggested an interesting correlation between TDP-43 pathology and hnRNP misregulation. In order to clarify the tangled scenario underlying TDP-43 metabolism and identify physiological partners of TDP-43 among the hnRNP family, the Molecular Pathology laboratory at ICGEB has recently taken advantage of *Drosophila melanogaster* as a TDP-43 proteinopathy model.

First of all, we evaluated how many human hnRNP proteins have homologs in *Drosophila*. In total, we found 11 human hnRNPs having a counterpart conserved in *Drosophila*. Then, in order to determine which of these genes were able to modulate TDP-43 activity *in vivo*, we performed a systematic gene suppression screening crossing flies silenced for these hnRNPs either with TBPH/TDP-43 hypomorphic flies (loss-of-function model) or transgenic flies overexpressing TBPH/TDP-43 in the eye compartment (gain-of-function model). We found that three hnRNPs, namely Hrb27c (DAZAP1) and Syncrip/Syp (hnRNP Q and R) had an extremely negative effect when knocked out in combination with TBPH. However, at the same time, they seemed to be protective in the TBPH gain-of-function expression model. Taken together, these results suggested that alteration of the expression levels of hnRNP proteins could represent a response against TDP-43 gain and loss-of-function events and could enable an *in vitro* characterization by using cell cultures.

Therefore, the main purpose of my PhD project was to clarify the bases of the opposite results observed in both the *Drosophila* models of TDP-43 proteinopathies after DAZAP1, hnRNP Q and hnRNP R depletion. This was achieved through the investigation of their functional relationship with TDP-43 within human neuronal-like cells, in order to highlight differences and similarities in regulation of RNA metabolism.
Initially, I studied the influence of each hnRNP protein on well-known TDP-43 regulated events and subsequently analysed the overall transcriptome status of RNA samples collected from human neuroblastoma (SH-SY5Y) cells treated with siRNA against TDP-43 and each hnRNP, in order to identify commonly regulated RNA targets involved in neuron development and neurodegeneration disease.
3. MATERIALS AND METHODS

3.1 Cell cultures, gene knockdown and cell differentiation

3.1.1 Cell cultures

Human neuroblastoma SH-SY-5Y cells (ATCC), HeLa cells (ATCC) and Flp-In HEK293 cells (Invitrogen, ThermoFisher, Waltham, MA, USA) expressing inducible Flag-tagged TDP-43 12XQ/N aggregates were cultured in DMEM–Glutamax-I (Gibco-BRL, Life Technologies Inc., Frederick, MD, USA) supplemented with 10% fetal bovine serum (Gibco-BRL, Life Technologies Inc., Frederick, MD, USA) and 1% Antibiotic-Antimycotic-stabilized suspension (SigmaAldrich, St. Louis, MO, USA) at 37°C in an incubator with humidified atmosphere of 5% CO2.

The mouse motor neuron-like NSC-34 cell line (a kind gift of Antonia Ratti, University of Milan, Milan, Italy) was cultured in Dulbecco’s modified Eagle’s medium (DMEM)–Glutamax-I (Gibco- BRL, Life Technologies Inc., Frederick, MD, USA) supplemented with 5% fetal bovine serum (FBS) (SigmaAldrich, St Louis, MO, USA) and 1% Antibiotic-Antimycotic-stabilized suspension (Sigma-Aldrich, St Louis, MO, USA) at 37°C in an incubator with humidified atmosphere of 5% CO2.

All the cell cultures were used up to 20 passages.

3.1.2 Gene knockdown of SH-SY5Y neuroblastoma cell line

To achieve optimal knockdown efficiency, three rounds of silencing were performed on day 1, 3 and 5 using HiPerFect Transfection Reagent (Qiagen Inc, Gaithersburg, MD, USA), according to the manufacturer’s instructions. Briefly, 50X10^4 cells were seeded in 6-well plates and silenced with a mix composed of 100 µl Opti-MEM (Life-Technologies,
CA, USA), 3 µl of 40 µM siRNA and 6 µl hyperfectamine (day 1). At the third day (day 3) cells were trypsinised, plated in a p60 plate and silenced for the second time using a double volume of mix. The same procedure was repeated at the third time (day 5) plating cells in a p100 plate and using a double volume of mix with respect to the second round of silencing. The final siRNA concentration for each time of silencing was 80 nM. After 48 hours cells were analysed by western blot (WB), pre-mRNA splicing, gene expression or used for immunofluorescence (IF) analysis.

For experiments with double knockdown (siTDP-43/siDAZAP1, siTDP-43/sihnRNP Q and siTDP-43/sihnRNP R) the amount of siRNA used for depletion of each gene was the half of the single knockdown.

The siRNA sense sequences used in this study are described in Table 1.

Table 1. siRNA sequence from 5'-prime to 3'-prime are listed.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>siRNA sense sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fire-fly luciferase (siLUC-control)</td>
<td>5'-uaaggcuauagagauac-3'</td>
</tr>
<tr>
<td>TDP-43 (siTDP-43)</td>
<td>5'-gcaaagccaaugagccu-3'</td>
</tr>
<tr>
<td>DAZAP1 (siDAZAP1)</td>
<td>5'-gagacucugcgcagcuacu-3'</td>
</tr>
<tr>
<td>hnRNP Q (sihnRNP Q)</td>
<td>5'-agacagucucuucucu-3'</td>
</tr>
<tr>
<td>hnRNP R (sihnRNP R)</td>
<td>5'-cauuuggauguacgucuu-3'</td>
</tr>
</tbody>
</table>

3.1.3 Gene knockdown of Flp-In HEK293 cell line expressing TDP-43 aggregates

Flp-In HEK293 cells with inducible siRNA resistant FLAG-tagged TDP-43-12XQ/N aggregates (75X10⁴) were seeded in 6-well plates in 1.5 ml of culture medium containing serum and antibiotics. A total of 4 µl of 40 µM siRNA against DAZAP1, hnRNP Q and
hnRNP R was diluted in 90 µl of Opti-MEM (Life-Technologies, CA, USA) and 6 µl of HiPerFect Transfection Reagent (Qiagen Inc, Gaithersburg, MD, USA) (day 1). Second and third siRNA treatment were performed at day 2 and day 3, respectively. For each round of silencing, cells were tripinised, plate in a new 6-well and silenced with the same amount of siRNA mixture as described for day 1. Importantly, for all the conditions, the silencing was performed in duplicates because of one plate was used for WB analysis and the second one for IF analysis. For this latter, the third round of siRNA treatment was performed in 6-well plates containing microscope slides treated with 0.1 mg/ml Poly-L-Lysine (Sigma-Aldrich, St Louis, MO, USA).

The list of siRNA used for silencing Flip-In HEK293 is the same described in section 3.1.2 and reported in Table 1.

At day 4, culture medium was changed and the TDP-43 12XQ/N aggregation was induced by addition of 1 µg/ml tetracycline solution (Sigma-Aldrich, St Louis, MO, USA). After 48 hours, cells were prepared for WB or IF analysis.

### 3.1.4 NSC-34 differentiation

NSC-34 cells were seeded to reach 70% confluence the day after and the proliferation medium was exchanged 24 hours later to fresh differentiation medium containing 1:1 DMEM/F-12 Ham (Sigma-Aldrich, St Louis, MO, USA), 1% FBS (Sigma-Aldrich, St Louis, MO, USA), 1% modified Eagle’s medium nonessential amino acids (NEAA) (Sigma-Aldrich, St Louis, MO, USA), 1% Antibiotic-Antimycotic-stabilized suspension (Sigma-Aldrich, St Louis, MO, USA) and 1 µM all-trans retinoic acid (RA) (Sigma-Aldrich, St Louis, MO, USA). Differentiation medium was changed every two days and cells were allowed to differentiate for up to 5 days. NSC-34 cells, maintained on proliferation medium (DMEM, 5% FBS, 1% Antibiotic-Antimycotic suspension)
Materials and Methods

represented the undifferentiated control group. Images of undifferentiated (control, CNTRL) and differentiated NSC-34 cells were acquired by light microscopy using a Leica DMIL LED microscope equipped with a 20X objective, a Leica DFC450 C camera (Leica Microsystems, Cambridge, UK) and LAS v.4.4.0 Software (Leica application suit). The average length of neurites in the differentiation media was compared to that in the proliferation media and quantified using Fiji NeuronJ (Meijering et al., 2004). Neurite length was analyzed by imaging a minimum of 5 cells per field. The mean of neurite length ± standard error is reported. Statistical significance was calculated using t-test (indicated as *** for P ≤ 0.001).

3.2 DNA plasmids for protein overexpression

3.2.1 Generation of competent DH5α Escherichia coli cells

To prepare competent DH5α E. coli cells, the protocol of Chung and collaborators (Chung et al., 1989) was followed introducing few modifications. DH5α stored in glycerol (1:1) at -80 °C were defrosted in ice and then plate in Petri dishes containing Luria-Bertani broth (LB) without antibiotics. The day after, one colony was picked up and grown overnight at 37 °C with 10 ml of LB (pre-inoculum). The third day 1-2 ml of pre-inoculum was diluted into 50 ml of LB and grown at 37 °C until its optical density reached an OD590 value of 0.3-0.4. Once the OD desired was obtained, bacteria were collected by centrifugation at 1000 xg for 10 minutes at 4 °C, resuspended into 5 ml of ice-cold diluted TSS 1X pH 6.5 (Table 2) and aliquoted in 1.5 ml sterile tubes and instantly frozen in liquid NO2 for storage at -80 °C. To evaluate the competence of cells, 0.1 ng/µl of pUC18 plasmid was used to transform 100 µl of DH5α cells (see section below). Competence was calculated as following:
Materials and Methods

\[
\text{Competence (10}^6\text{)} = \frac{\text{Colony Forming Units}}{\text{Concentration of } p\text{UC18}}
\]

Table 2. Composition of TSS 1X.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Amount for 50 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyethylene glycol (PEG 400) (SERVA Electrophoresis GmbH, Heidelberg, Germany)</td>
<td>5 g</td>
</tr>
<tr>
<td>MgCl(_2)</td>
<td>35 mM</td>
</tr>
<tr>
<td>DMSO</td>
<td>5%</td>
</tr>
<tr>
<td>Up to 50 ml with LB</td>
<td>35 ml ca.</td>
</tr>
</tbody>
</table>

### 3.2.2 Transformation of competent DH5α cells

The amplification of desired plasmid DNA was obtained through bacteria transformation. Competent bacteria (30 µl) were transformed with 0.1 ng/µl plasmid DNA as follows:

- Incubation 30 minutes on ice
- Heat-shock 2 minutes at 42 °C
- Incubation 1 minute on ice
- Incubation 30 minutes at 37 °C (shaking) in the presence of 60-100 µl of LB without antibiotic
- Plate in Petri dish supplemented with antibiotic and incubate over-night at 37 °C

The next day, one colony was picked up and grown overnight at 37 °C in the presence of 50 ml (inoculum for Midiprep) or 6 ml (inoculum for Miniprep) of LB with antibiotic. After that, cells were collected through centrifugation for 30 minutes at 1937 xg and the resulting pellets were stored at -20 °C.
3.2.3 Small scale preparation of plasmid DNA (Miniprep)

Small scale preparation of plasmid DNA was performed using Wizard Plus SV Minipreps DNA purification system (Promega, Fitchburg, WI, USA) following the manufacturer’s instructions. Basically, bacteria pellets were resuspended with 250 µl of cell resuspension solution by vortexing. Successively, 250 µl of Cell Lysis Solution and 10 µl of Alkaline Protease Solution was added to sample, mixed by inversion 4 times and incubated for 5 minutes at room temperature. To block the reaction, the Neutralization Solution (350 µl) was added into the bacteria lysate, mixed by inversion 4 times and centrifuged at maximum speed for 10 minutes at room temperature. After that, the cleared lysate was decanted into a spin column and centrifuged at maximum speed for 1 minute at room temperature. The flow-through was discarded and the spin column was washed two times with 750 µl and 250 µl of Wash Solution, respectively. Each step of washing was spaced out by centrifugation at maximum speed for 1 and 2 minutes at room temperature. Finally, the plasmid DNA was eluted by adding 100 µl of Nuclease-Free water into the spin column and centrifuging at maximum speed for 1 minute at room temperature. The obtained DNA was stored at -20 °C.

3.2.4 Middle scale preparation of plasmidic DNA (Midiprep)

Middle scale preparation of plasmid DNA was performed using JETSTAR Plasmid Purification kit (Genomed, Brussels, Belgium) following manufacture’s instructions. Before starting with the solution within the kit, 50 µl of RNAse A was added to the cell pellet. Successively, 4 ml of Cell Resuspension Buffer E1 (supplemented with RNAse A) was used for resuspending the pellet until its homogenization. After that, 4 ml of Lysis Buffer E2 was added into the cell resuspension, mixed by inverting the capped tube and incubated at room temperature for 15-20 minutes to facilitate the lysis of bacteria. To
precipitate the nucleic acids, 4 ml of Precipitation Buffer (E3) was added to the lysate, mixed by inverting the tube and centrifugated at 1937 xg for 30 minutes at 4 °C. The supernatant was then loaded onto a pre-equilibrated column with 10 ml of Equilibration Buffer (E4) and the solution was allowed to drain by gravity flow. The column was then washed twice with 10 ml Wash Buffer (E5) and the flow-through was discarded. To elute plasmid DNA, 5 ml of Elution Buffer (E6) was added to the column previously placed in a new sterile 50 ml centrifuge tube. 3.5 ml of isopropanol was then added to the eluate, mixed well and centrifuged at 1937 xg for 30 minutes at 4°C. Finally, after air-drying the pellet for 10 minutes, the DNA pellet was resuspended in 200-300 µl of Nuclease-Free water and stored at -20 °C.

3.2.5 Overexpression of Flag-tagged siRNA resistant TDP-43 and Flag-tagged DAZAP1 in HeLa cells

In this thesis, I used two plasmid DNA vectors (pFLAG-CMV4 and pFLAG-CMV2) containing a cloned sequence for the overexpression of Flag-tagged siRNA resistant TDP-43 and Flag-tagged DAZAP1, respectively (Figure 19A-B). The expression of each cloned DNA was performed by transfection of HeLa cells with Effectene Transfection Reagent (Qiagen Inc, Gaithersburg, MD, USA) following the manufacturer’s instructions. Briefly, cells were plated in a p100 tissue culture dish and grown under their normal conditions (see section 3.1.1) to reach 70% confluence on the day of transfection. The day of transfection, 3 µg of plasmid was diluted with 300 µl of DNA-condensation buffer (Buffer EC). After that, 16 µl of Enhancer was added into the solution, mixed by vortexing, spun down and incubated 5 minutes at room temperature to allow the condensation of DNA. Effectene reagent (20 µl) was then added to the condensed DNA, mixed by vortexing 10 seconds, spun down and incubated 10 minutes at
Materials and Methods

room temperature to produce condensed Effectene-DNA complexes. Cells were washed with 1X PBS and the culture medium was exchanged with new one. After 10 minutes of incubation, the Effectene-DNA complexes were mixed with medium and directly added to the cells.

Figure 19. plasmid DNA for overexpression. (A) Map of pFLAG-CM4-TDP-43 (siRNA resistant). (B) Map of pFLAG-CMV2-DAZAP1.
3.3 Protein detection analysis

3.3.1 Western blot analysis for checking gene knockdown

Total protein extracts were obtained by sonicating cells two times for 5 minutes, high power with a BioRuptor UCD-200 (Diagenode, Belgium) in a mild lysis buffer composed of 1X Phosphate Saline Buffer (PBS) and 1X Complete Protease Inhibitor Cocktail (Roche Diagnostics, Mannheim, Germany). Disrupted cells were centrifuged 5 minutes at 16100 xg and the supernatant (protein extract) was collected in a new tube. Protein extract (20 µg) from each sample was resuspended in 1X sample buffer (Table 3), boiled at 95 °C for 5 minutes and loaded on a 10% Sodium Dodecyl sulfate (SDS)–polyacrylamide denaturing gel (Table 4). The gel was then electroblotted onto a PVDF blotting membrane (Immobilion-P, Millipore, Billerica, USA) pre-activated with 100% methanol by using a semi-dry protein blotter (Lighting Blotter, PerkinElmer, Massachusetts, USA) for 30 minutes keeping the voltage constant at 27 V. After blotting, the membrane was blocked with 4 % skimmed milk (Non-fat dry milk) or 4 % Bovine Serum Albumin (BSA) (Sigma-Aldrich, St Louis, MO, USA) in 1x PBS and 0.1% Tween-20 as described in the Table 5. Proteins were probed with incubation of primary (AbI) and subsequently HRP-conjugated secondary (AbII) antibodies in order to confirm the gene knockdown efficacy. The list of primary (AbI) and secondary (AbII) antibodies used for western blot are shown in Table 5. The chemiluminescence of the target proteins were detected with ECL Luminata Classico Western HRP substrate (Merck Millipore, Burlington, MA, USA) and the images were acquired using Alliance 9.7 Western Blot Imaging System (UVItec Limited, Cambridge, UK) or using X-ray film. Tubulin, available in-house, was used as total protein loading control. In order to quantify the percentage of remaining protein after siRNA knockdown, densitometric analysis of the Western Blot bands was performed using ImageJ software.
Table 3. Composition of stock 4X sample buffer.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCL pH 6.8</td>
<td>240 mM</td>
</tr>
<tr>
<td>Glycerol</td>
<td>40%</td>
</tr>
<tr>
<td>SDS</td>
<td>8%</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>0.04%</td>
</tr>
<tr>
<td>2-mercaptoethanol</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 4. Composition of 10% SDS-acrylamide gel.

<table>
<thead>
<tr>
<th>Running gel 10%</th>
<th>Stacking gel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reagent</strong></td>
<td><strong>Volume</strong></td>
</tr>
<tr>
<td>Tris-HCL 1.5 M pH 8.8</td>
<td>5 ml</td>
</tr>
<tr>
<td>Protogel 30% (w/v) Acrylamide: 0.8 %</td>
<td>6.66 ml</td>
</tr>
<tr>
<td>(w/v) bis-acrylamide Stock Solution</td>
<td>200 µl</td>
</tr>
<tr>
<td>10% SDS</td>
<td>200 µl</td>
</tr>
<tr>
<td>Ammonium persulfate (APS)</td>
<td>200 µl</td>
</tr>
<tr>
<td>Tetramethylethylenediamine (TEMED)</td>
<td>20 µl</td>
</tr>
<tr>
<td>H₂O</td>
<td>8.34 ml</td>
</tr>
<tr>
<td><strong>Reagent</strong></td>
<td><strong>Volume</strong></td>
</tr>
<tr>
<td>Tris-HCL 0.5 M pH 6.8</td>
<td>2.50 ml</td>
</tr>
<tr>
<td>Protogel 30% (w/v) Acrylamide: 0.8 %</td>
<td>560 µl</td>
</tr>
<tr>
<td>(w/v) bis-acrylamide Stock Solution</td>
<td>50 µl</td>
</tr>
<tr>
<td>10% SDS</td>
<td>50 µl</td>
</tr>
<tr>
<td>Ammonium persulfate (APS)</td>
<td>50 µl</td>
</tr>
<tr>
<td>Tetramethylethylenediamine (TEMED)</td>
<td>10 µl</td>
</tr>
<tr>
<td>H₂O</td>
<td>2 ml</td>
</tr>
</tbody>
</table>
Table 5. Specifics of AbI and AbII used for checking gene knockdown.

<table>
<thead>
<tr>
<th>Blocking solution</th>
<th>AbI</th>
<th>Company</th>
<th>Dilution</th>
<th>AbII</th>
<th>Company</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>Polyclonal rabbit α-TDP-43</td>
<td>Proteintech</td>
<td>1:2000</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Polyclonal rabbit α-DAZAP1</td>
<td>Homemade</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>BSA</td>
<td>Polyclonal rabbit α-hnRNP Q</td>
<td>Sigma-Aldrich</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>BSA</td>
<td>Polyclonal rabbit α-hnRNP R</td>
<td>Abcam</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Polyclonal mouse α-Tubulin</td>
<td>Homemade</td>
<td>1:10000</td>
<td>Polyclonal goat α-mouse</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
</tbody>
</table>

3.3.2 TDP-43 and DAZAP1 co-immunoprecipitations (CO-IP) analysis

HeLa cells (1.2X10^6) were grown overnight in a p100 tissue culture dish to achieve 70% of confluence for transfection with 3 μg of pFLAG-TDP-43 wild-type using the Effectene reagent as described in section 3.2.5. After 24 hours, cell culture medium was removed and cells were washed with cold 1X PBS and harvested. Cells were lysed in 500 μL of IP buffer (Table 6) by sonication (3 min., mid power) in an ice-cooled sonicating bath (BioRuptor UCD-200, Diagenode, Belgium). The cell lysate was pre-cleared by incubation with 30 μl Protein A/G PLUS agarose beads (Santa Cruz Biotechnology Inc., Dallas, Texas, USA) in IP buffer for 1.5 hours at 4 °C. The pellet was discarded, and the supernatant was used for immunoprecipitation: the cell lysate was incubated with 2 μg of mouse monoclonal anti-FLAG M2 antibody (Sigma-Aldrich, St Louis, MO, USA) on a
rotating device for 1 hour at 4 °C. Then, 30 μl of Protein A/G PLUS agarose beads were added to each sample and incubated overnight at 4°C. The pellet was then washed three times in ice-cold IP buffer. The supernatant was discarded, and the pellet was re-suspended in 30 μl of 4X loading buffer. The samples were fractionated by SDS-PAGE (10%) and analysed by immunoblotting as described in section 3.3.1 using the antibodies reported in Table 7.

Table 6. Composition of IP buffer.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCL pH 7.5</td>
<td>20 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>110 mM</td>
</tr>
<tr>
<td>Triton-X</td>
<td>0.5%</td>
</tr>
<tr>
<td>Complete Protease Inhibitor Cocktail</td>
<td>1X</td>
</tr>
</tbody>
</table>

Table 7. List of AbI and AbII used for CO-IP experiment.

<table>
<thead>
<tr>
<th>Blocking solution</th>
<th>AbI</th>
<th>Company</th>
<th>Dilution</th>
<th>AbII</th>
<th>Company</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>Polyclonal rabbit α-TDP-43</td>
<td>Proteintech</td>
<td>1:2000</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Polyclonal rabbit α-DAZAP1</td>
<td>Homemade</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Polyclonal rabbit α-hnRNP H</td>
<td>Homemade</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
</tbody>
</table>

3.3.3 Nuclear and cytoplasmic extraction using NER-PER kit

SH-SY5Y cells were seeded in p100 dishes to reach 90% confluence the day of nuclear and cytoplasm extraction. Cells from two dishes were pooled together and the resulting
pellet was treated using NER-PER Nuclear and Cytoplasmic Extraction Reagents (ThermoFischer, Waltham, MA, USA) as described in the manufacturer’s instructions. Briefly, 24 hours after seeding cell culture medium was removed and cells were washed with 1X PBS and harvested. The cell pellet was resuspended with ice-cold CER I buffer according to the packed cell volumes (Table 8) and vortexed on the highest setting for 15 seconds to fully suspend the cell pellet. The lysate was incubated on ice for 10 minutes and successively ice-cold CER II buffer according to Table 8 was added to tube. After that, the lysate was vortexed twice for 5 seconds on the highest setting with 1 minute incubation on ice between each round and then centrifuged for 5 minutes at maximum speed. The obtained supernatant (cytoplasmic fraction) was transferred to a pre-chilled tube and stored at -80°C until WB analysis. The insoluble pellet, containing nuclei, was suspended with ice-cold NE buffer according to Table 8, vortexed in the highest setting for 15 seconds every 10 minutes for a total of 40 minutes. After that, the nuclear fraction was centrifuged at maximum speed for 10 minutes and moved in a pre-chilled tube stored at -80 °C until the WB analysis.

For each sample fractionated with NER-PER kit, 15 µg of cytoplasmic and nuclear extracts were analyzed by WB as described in section 3.3.1 using the antibodies reported in Table 9.

Table 8. Reagent volumes for different packed cell volumes.

<table>
<thead>
<tr>
<th>Packed cell volume (µl)</th>
<th>CER I (µl)</th>
<th>CER II (µl)</th>
<th>NER (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
<td>5.5</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>200</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>50</td>
<td>500</td>
<td>27.5</td>
<td>250</td>
</tr>
<tr>
<td>100</td>
<td>1000</td>
<td>55</td>
<td>500</td>
</tr>
</tbody>
</table>
Table 9. List of AbI and AbII antibodies used for nuclear and cytoplasmic fractionation.

<table>
<thead>
<tr>
<th>Blocking solution</th>
<th>AbI</th>
<th>Company</th>
<th>Dilution</th>
<th>AbII</th>
<th>Company</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA</td>
<td>Polyclonal rabbit α-hnRNP Q</td>
<td>Sigma-Aldrich</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>BSA</td>
<td>Polyclonal rabbit α-hnRNP R</td>
<td>Abcam</td>
<td>1:500</td>
<td>Polyclonal goat α-rabbit</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Polyclonal mouse α-Tubulin</td>
<td>Homemad e</td>
<td>1:1000</td>
<td>Polyclonal goat α-mouse</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
<tr>
<td>Milk</td>
<td>Monoclonal mouse α-p84</td>
<td>Abcam</td>
<td>1:1000</td>
<td>Polyclonal goat α-mouse</td>
<td>Dako Denmark A/S</td>
<td>1:2000</td>
</tr>
</tbody>
</table>

3.4 pre-mRNA splicing, gene expression and RNA-seq analysis

3.4.1 RNA extraction

RNA was obtained using EuroGold Trifast (Euroclone, Milan, Italy), following the manufacturer’s instructions. Briefly, cell culture medium was removed and cells were washed with cold 1X PBS and harvested. According to the volumes reported in Table 10, cell pellet was resuspended with EuroGold Trifast by pipetting and incubated for 5 minutes at room temperature. Chloroform (Table 10- chloroform I) was then added to each sample and the mixture was strongly shacked and incubated 10 minutes at room temperature.
Subsequently, samples were centrifuged at maximum speed for 20 minutes at 4 °C and the supernatant fraction (transparent upper phase) was transferred in a new sterile tube and mixed with an equal volume of chloroform for a second cleaning step (chloroform II). Samples were then centrifuged at maximum speed for 10 minutes at 4 °C and the supernatant was transferred in a new sterile tube. The RNA was precipitated adding isopropanol to this latter (Table 10), incubated for 10 minutes on ice, and then centrifuged at maximum speed for 20 minutes at 4 °C. At the end of the process, the resulting pellet (RNA) was washed with 70% ethanol, centrifuged at maximum speed for 5-10 minutes at 4 °C, resuspended in 30-60 µl of RNase-free water and quantified with Nanodrop 1000 (ThermoFisher, Waltham, MA, USA).

**Table 10.** Volume of reagent using for RNA extraction according to cell dish.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>EuroGold Trifast</td>
<td>500 ml</td>
<td>EuroGold Trifast</td>
<td>1 ml</td>
</tr>
<tr>
<td>Chloroform I</td>
<td>100 µl</td>
<td>Chloroform I</td>
<td>200 µl</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>250 µl</td>
<td>Isopropanol</td>
<td>500 µl</td>
</tr>
<tr>
<td>70 % Ethanol</td>
<td>400 µl</td>
<td>70 % Ethanol</td>
<td>800 µl</td>
</tr>
<tr>
<td>RNase-free water</td>
<td>30 µl</td>
<td>RNase-free water</td>
<td>60 µl</td>
</tr>
</tbody>
</table>

### 3.4.2 Reverse transcription reaction

Complementary DNA (cDNA) was retrotranscribed by using 1 µg of total RNA in the presence of random primers (Sigma-Aldrich, St Louis, MO, USA) and Moloney murine leukemia virus (M-MLV) Reverse Transcriptase (Invitrogen, ThermoFisher, Waltham, MA, USA). RNA samples prepared as described in Table 11 were denaturated at 85 °C for
5 minutes and chilled 1 minutes on-ice to avoid renaturation of RNA. After that, 20 µl of mixture containing M-MLV enzyme (Table 12) were added to RNA sample mixture (20 µl) and incubated 1-2 hours at 37 °C for reverse transcription reaction. cDNA samples were then stored at -20 °C or placed on-ice for pre-mRNA splicing analysis or gene expression analysis.

Table 11. Preparation of RNA sample mixture.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration (20 µl of reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA sample</td>
<td>50 ng/µl</td>
</tr>
<tr>
<td>Random primers</td>
<td>1 ng/µl</td>
</tr>
</tbody>
</table>

Table 12. Mixture for reverse transcription reaction.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration (20 µl of reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X First strand buffer</td>
<td>2X</td>
</tr>
<tr>
<td>Dithiothreitol (DTT) 0.1 M</td>
<td>0.02 M</td>
</tr>
<tr>
<td>Deoxynucleotide (dNTP) Solution Mix 5mM</td>
<td>1 mM</td>
</tr>
<tr>
<td>M-MLV 200 U/µl</td>
<td>5 U/µl</td>
</tr>
</tbody>
</table>

3.4.3 pre-mRNA splicing analysis

To evaluate the pre-mRNA splicing, polymerase chain reaction (PCR) was performed from cDNA obtained as described in section 3.4.2, and derived from SH-SY5Y cells depleted for TDP-43, DAZAP1, hnRNP Q and hnRNP R, or co-silenced for TDP-43/DAZAP1, TDP-43/hnRNP Q and TDP-43/hnRNP R. For each sample, 3 µl of cDNA was mixed with 47 µl of PCR mix solution prepared as described in Table 13. Taq DNA polymerase (New England Biolabs, Ipswich, MA, USA) was used for reaction. Forward and reverse primer
sequences for each gene tested and the relative PCR conditions are reported in Table 14 and figure 20A-B, respectively. For each target gene, PCR amplification was performed for 35 cycles and products (10 µl of total volume of reaction) were routinely fractionated in 1.5% (w/v) agarose gels. The percentage of exon inclusion for each tested gene was evaluated by densitometric analysis using ImageJ software. For MADD gene was calculated the percentage of pseudoexon inclusion. The statistical significance was calculated using an unpaired t-test. P<0.05 was considered significant (n = 3) (* P < 0.05, ** P < 0.01, and *** P<0.001).

Table 13. PCR reaction mixture.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration (50 µl of reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X ThermoPol reaction Buffer</td>
<td>1X</td>
</tr>
<tr>
<td>Deoxynucleotide (dNTP) Solution Mix 5mM</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>Primer Forward 10 ng/µl</td>
<td>0.3 ng/µl</td>
</tr>
<tr>
<td>Primer Reverse 10 ng/µl</td>
<td>0.3 ng/µl</td>
</tr>
<tr>
<td>Taq Biolabs 5 U/µl</td>
<td>0.02 U/µl</td>
</tr>
</tbody>
</table>

Table 14. List of genes and relative primers (forward and reverse sequences) used for pre-mRNA splicing analysis.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Forward sequence (5'-3')</th>
<th>Reverse sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLDIP3</td>
<td>5'-gcttaatgccagacggggagttg-3'</td>
<td>5'-tcatctcactccagttcatatttt-3'</td>
</tr>
<tr>
<td>TNIK</td>
<td>5'-caagggagagaaggagctg-3'</td>
<td>5'-ctgatgctgaagggaaactaag-3'</td>
</tr>
<tr>
<td>STAG2</td>
<td>5'-gtatgatttaactgtgtaaaagttcatg-3'</td>
<td>5'-tgattcatccataattgaagctgga-3'</td>
</tr>
<tr>
<td>MADD</td>
<td>5'-gacctgaattgttggtgcagttccct-3'</td>
<td>5'-catgggtgtctgttaactttgtggtgct-3'</td>
</tr>
</tbody>
</table>
Figure 20. PCR conditions used in the pre-mRNA splicing assay. (A) Protocols for PCR amplification of POLDIP3 and TNIK gene. (B) Protocols for PCR amplification of STAG2 and MADD gene.

3.4.4 RNA immunoprecipitation (RNA-IP) analysis

HeLa cells (1.2X10⁶) were grown overnight in a p100 tissue culture dish to achieve 70% of confluence for transfection with 3 μg flag-DAZAP1 plasmid using Effectene reagent as described in section 3.2.5. After 24 hours, the cell culture medium was removed and cells were washed with cold 1X PBS and harvested. Cells were lysated with HEGN buffer (Table 15) and sonicated (5 minutes, high power) in an ice-cooled sonicating bath (BioRuptor UCD-200, Diagenode, Belgium). HeLa lysate (40 μg) was incubated for 1 hour at 4 °C in HEGN buffer supplemented with 7.5 μg/μl heparin (Sigma-Aldrich, St Louis, MO, USA) and Protein A/G Agarose beads (Santa Cruz Biotechnology Inc., Dallas, Texas, USA), pre-coated for 1 hour, 4 °C in rotation with 5 μg of anti-Flag antibody (IP-Flag) (Sigma-Aldrich) or with uncoated beads as controls (IP-Beads). After washes with HEGN + DOC 0.2% + Urea 0.5 M (four-times separated with centrifugation of 5 minutes at 15000 xg), mRNA was extracted from immunoprecipitated RNPs using EuroGold Trifast (see section 3.4.1; volumes for p35 dish). The abundance of possible DAZAP1 target genes was measured by quantitative Real-Time PCR on three independent experiments as described below.
Materials and Methods

Table 15. Composition of HEGN buffer.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepes pH 7.5</td>
<td>20 mM</td>
</tr>
<tr>
<td>NaCl</td>
<td>150 mM</td>
</tr>
<tr>
<td>EDTA</td>
<td>0.5 mM</td>
</tr>
<tr>
<td>Glycerol</td>
<td>10%</td>
</tr>
<tr>
<td>Triton-X</td>
<td>0.1%</td>
</tr>
<tr>
<td>DTT</td>
<td>1 mM</td>
</tr>
<tr>
<td>Complete protease inhibitor cocktail</td>
<td>1X</td>
</tr>
</tbody>
</table>

3.4.5 Gene expression analysis

Total RNA was extracted from SH-SY5Y cells transfected with siRNA against TDP-43, DAZAP1, hnRNP Q and hnRNP R or co-transfected with siRNA against TDP-43 in the presence of siRNA against each of the three tested hnRNPs as describe in section 3.1.2. RNA obtained from SH-SY5Y cells treated with siRNA against luciferase (siLUC) was used as a negative control. For each RNA sample, reverse transcription was performed in order to synthetize cDNA as described in section 3.4.2. The quantification of gene expression levels was carried out by quantitative Real-time PCR using iQ SYBR green supermix (Bio-Rad, Hercules, CA, USA). For each sample, 4 µl of cDNA (diluted 1:10-1:5 in sterile water) was mixed with 6 µl of iQ mix solution prepared as described in Table 16. Quantitative Real-time PCRs were performed on a MiniOpticon real-time PCR and on a CFX96 Real-time systems (Bio-Rad, Hercules, CA, USA). The expression levels were determined using the 2^(-ΔΔCT) method (Schmittgen and Livak, 2008). For primer designer two online tools were used: PrimerBank (https://pga.mgh.harvard.edu/primerbank/) and qPrimerDepot (http://primerdepot.nci.nih.gov/). The list of primers (forward and reverse
sequence) and the qPCR conditions are reported in Table 17 and Figure 21A-C, respectively. To normalize the results, the housekeeping gene *Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH), Ribosomal Protein L13a (RPL13A), Hypoxanthine Phosphoribosyltransferase 1 (HPRT1)* and *RNA Polymerase II Subunit A (POLR2A)* were also used (Table 18). In particular, GAPDH and RPL13A were used to normalize simultaneously the expression of *MADD, BRD8, TNIK, ELAVL3, NOVA2, RELN, STX3, CELF5, YPEL4, ACHE, TNF, TNFRSF9* and *ICAM1* gene in SH-SY5Y cells. Whereas, HPRT1 and POLR2A were used to normalize simultaneously the expression of *PENK, KLF4, KLHL4, NRG3, RAB26, ARHGAP36, CT55, CARTPT, FOSB, JAG1, ICAM5, DUOXA1, HMOX1, KCNAB1, ACP5, SDCBP2, EFEMP1, SYNCRIP (hnRNP Q) and HNRNPR* gene in SH-SY5Y cells. Moreover, *Succinate Dehydrogenase Complex Flavoprotein Subunit A (SDHA)* was used as a negative control for RNA-immunoprecipitation analysis performed in HeLa cells.

The mean of relative expression levels ± standard deviation/standard error of three independent experiments are reported. Statistical significance was calculated using an unpaired t-test. P<0.05 was considered significant (n = 3) (* P < 0.05, ** P < 0.01, and *** P<0.001).

**Table 16.** Composition of iQ mix solution.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>iQ SYBR green supermix</td>
<td>Dilution 1:2</td>
</tr>
<tr>
<td>20 µM Primer forward</td>
<td>0.05 µM</td>
</tr>
<tr>
<td>20 µM Primer forward</td>
<td>0.05 µM</td>
</tr>
</tbody>
</table>
Table 17. List of target genes and relative primers (forward and reverse sequences) used for gene expression analysis.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Forward sequence (5'-3')</th>
<th>Reverse sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>MADD</td>
<td>5'-ttgagaccaacttgcacca-3'</td>
<td>5'-agactcgctggctcaaatct-3'</td>
</tr>
<tr>
<td>BRD8</td>
<td>5'-gcagcgtgtttacagatgac-3'</td>
<td>5'-aatagtgcaanaaacaaatagc-3'</td>
</tr>
<tr>
<td>TNIK</td>
<td>5'-tggaaatcaagggcagttt-3'</td>
<td>5'-tctctctaatcccctggctgac-3'</td>
</tr>
<tr>
<td>POLDIP3</td>
<td>5'-tactacccctgaagggagc-3'</td>
<td>5'-tactacccctgaagggagc-3'</td>
</tr>
<tr>
<td>STAG2</td>
<td>5'-aaatcaaaacacttacaggg-3'</td>
<td>5'-tactacccctgaagggagc-3'</td>
</tr>
<tr>
<td>ELAVL3</td>
<td>5'-cctaataacaggaagacca-3'</td>
<td>5'-getgaagctcaggtcttact-3'</td>
</tr>
<tr>
<td>NOVA2</td>
<td>5'-aaggggaaaactcctgct-3'</td>
<td>5'-getgaagctcaggtcttact-3'</td>
</tr>
<tr>
<td>RELN</td>
<td>5'-ttgaggttcctagtttttc-3'</td>
<td>5'-tctctctaatcccctgaagggagc-3'</td>
</tr>
<tr>
<td>STX3</td>
<td>5'-tcggccagaccttctcaggt-3'</td>
<td>5'-tctctctaatcccctgaagggagc-3'</td>
</tr>
<tr>
<td>CELF5</td>
<td>5'-aaacttctctgggcagat-3'</td>
<td>5'-gacagttgagaagacca-3'</td>
</tr>
<tr>
<td>YPEL4</td>
<td>5'-cggagatctctgaggagaca-3'</td>
<td>5'-gtgttcaggggca gagaag-3'</td>
</tr>
<tr>
<td>ACHE</td>
<td>5'-agaaaaagctccttcgggtc-3'</td>
<td>5'-tagacccctccctctccac-3'</td>
</tr>
<tr>
<td>TNF</td>
<td>5'-actaccaatgactccttctg-3'</td>
<td>5'-gagacctggagagtagatggag-3'</td>
</tr>
<tr>
<td>TNFRSF9</td>
<td>5'-ttggatgagcttgttcttg-3'</td>
<td>5'-agagaggtatctgcggagagl-3'</td>
</tr>
<tr>
<td>ICAM1</td>
<td>5'-gscgcccagcttttttaac-3'</td>
<td>5'-tagacccctccctctccac-3'</td>
</tr>
<tr>
<td>PENK</td>
<td>5'-gtgacgcttacctttctg-3'</td>
<td>5'-gagacctggagagtagatggag-3'</td>
</tr>
<tr>
<td>KLF4</td>
<td>5'-gcggccaaacaccttacaaag-3'</td>
<td>5'-cactgtttgtgcggag-3'</td>
</tr>
<tr>
<td>KLHL4</td>
<td>5'-ttggagatgatggtctgtg-3'</td>
<td>5'-aagagttgtgctctggtg-3'</td>
</tr>
<tr>
<td>NRG3</td>
<td>5'-tatcctacacaggtggtac-3'</td>
<td>5'-tgaagcattcctctatggagca-3'</td>
</tr>
<tr>
<td>RAB26</td>
<td>5'-tactacccctacagttggt-3'</td>
<td>5'-ccgcaggtgctctgtggtg-3'</td>
</tr>
<tr>
<td>ARHGAP36</td>
<td>5'-ttaaatcgcagaccaaag-3'</td>
<td>5'-gacagctcactcagagac-3'</td>
</tr>
<tr>
<td>CT55</td>
<td>5'-atgttgtggaactggcaatgtg-3'</td>
<td>5'-gacagctcactcagagac-3'</td>
</tr>
<tr>
<td>CARTPT</td>
<td>5'-ccgagcctcggggactact-3'</td>
<td>5'-atgggaacaagctctcttggag-3'</td>
</tr>
<tr>
<td>FOSB</td>
<td>5'-acccttcgacagttctcatt-3'</td>
<td>5'-gaagaacctggtggtgctc-3'</td>
</tr>
<tr>
<td>JAG1</td>
<td>5'-atcttgttctcagttttct-3'</td>
<td>5'-gacagctcactcagagac-3'</td>
</tr>
<tr>
<td>ICAM5</td>
<td>5'-gctctctgtgctctctcag-3'</td>
<td>5'-gacagctcactcagagac-3'</td>
</tr>
</tbody>
</table>
Table 18. List of housekeeping genes and relative primers (forward and reverse sequences) used for normalization of gene expression analysis.

<table>
<thead>
<tr>
<th>Housekeeping gene</th>
<th>Forward sequence (5'-3')</th>
<th>Reverse sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>5'-cgtctctgtctctctcttt-3'</td>
<td>5'-cccgtgtgtgtgtacgatgt-3'</td>
</tr>
<tr>
<td>RPL13A</td>
<td>5'-cctggaggagaaggaaggaga-3'</td>
<td>5'-ttgagggctctgtgttatatttctaa-3'</td>
</tr>
<tr>
<td>SDHA</td>
<td>5'-tgagggagaagggacatctg-3'</td>
<td>5'-ccacactgttcaaatctgtc-3'</td>
</tr>
<tr>
<td>HPRT1</td>
<td>5'-tgacactggaagcaaaaatgca-3'</td>
<td>5'-gttgtttttcaccagcaagct-3'</td>
</tr>
<tr>
<td>POLR2A</td>
<td>5'-gcaccgttcaatgca-3'</td>
<td>5'-gtgtggtgttccataa-3'</td>
</tr>
</tbody>
</table>
Materials and Methods

Figure 21. Real-Time PCR conditions for gene expression analysis. (A) Standard protocol for amplification of target genes. (B) Specific Real-Time PCR protocol for amplification of MADD, BRD8 and TNIK gene. (C) Specific Real-Time PCR protocol for amplification of YPELA, CELF5, KLF4, KLHL4 gene.

3.4.6 RNA-seq and analysis of differentially expressed genes (DEGs)

Total RNA was extracted from luciferase (control), TDP-43, DAZAP1, hnRNP Q and hnRNP R depleted SH-SY5Y cells, as described previously. RNA sequencing was performed by Eurofins (www.eurofins.com) using an Illumina HiSeq 2500 instrument on three independent experiments obtained for each tested hnRNPs. Data processing was carried out with the following software: HiSeq Control Software v2.0.12.0, RTA v1.17.21.3 and bcl2fastq-1.8.4. Alignment to human reference sequence was performed by
Materials and Methods

BWA-MEM (version 0.7.12-r1039, http://bio-bwa.sourceforge.net/) and the raw read counts were created using featureCounts (http://bioinf.wehi.edu.au/featureCounts/). Only reads with unique mapping positions and a mapping quality score at least 10 were considered for read counting. Raw read counts were converted to Counts per million (CPM) values by Trimmed mean of M-values (TMM) normalization (edgeR package http://bioconductor.org/packages/release/bioc/html/edgeR.html, (Robinson and Oshlack, 2010). Features had to have a counts-per-million value of more than one in at least three samples or were removed, resulting in the removal of 47622 of the 64769 features. Differential expression analysis was performed on the remaining 17417 genes using edgeR package.

Regarding the interpretation of the differential expression genes data and pathway mapping related to TDP-43 and DAZAP1, PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System and the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 tools were used. Venn diagrams were generated using the VENN diagram generator at http://bioinfogp.cnb.csic.es/tools/venny/. GOseq package from R (Young et al., 2010) was also used for Gene ontology (GO) and KEGG pathway analysis of RNA-seq data. Categories significantly enriched (P< 0.05) were considered.

3.4.7 Statistical analysis of data

Statistical analysis of gene expression and splicing analysis were carried out by an unpaired two-sample \( t \) test on three independent experiments. \( P<0.05 \) was considered significant (* \( P < 0.05 \), ** \( P < 0.01 \), and *** \( P<0.001 \)). Regarding the RNA-seq analysis, DEGs were identified using the fold change cut-off of 30% with respect to a control sample treated with siRNA against fire-fly luciferase (siLUC) (upregulation cut-off: >1.3; downregulation cut-off: <0.7-FC) and the cut-off \( P<0.05 \) for statistical significance.
3.5 Immunofluorescence (IF) analysis

3.5.1 Immunofluorescence analysis of Flp-In HEK293 cell line expressing TDP-43 aggregates

Flp-In HEK293 cells expressing inducible TDP-43 aggregates were plated and treated as described in section 3.1.3. For immunofluorescence analysis the cells were washed 3 times with PBS, fixed in 3.2% paraformaldehyde in PBS for 1 hour at room temperature and permeabilized by using 0.3% Triton in PBS for 5 minutes on ice. Cells were then blocked with 2% BSA/PBS for 20 minutes at room temperature and incubated overnight at 4 °C in a humidified chamber with 60 µl of primary antibody against TDP-43 and the FLAG-tag prepared as described in Table 19. The day after, cells were washed 3 times with PBS, incubated with secondary antibodies conjugated with probes (Table 19) for 1 hour at room temperature and coverslipped with Vectashield-DAPI mounting medium (Vector Laboratories, Burlingame, CA, USA). Each slide was analyzed using Nikon Ti-E confocal microscope with a 60X oil objective.

<table>
<thead>
<tr>
<th>AbI</th>
<th>Company</th>
<th>Dilution</th>
<th>AbII</th>
<th>Company</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyclonal rabbit α-TDP-43</td>
<td>Proteintech</td>
<td>1:200</td>
<td>AlexaFluor 488 donkey α-rabbit</td>
<td>Invitrogen</td>
<td>1:500</td>
</tr>
<tr>
<td>Monoclonal mouse α-FLAG</td>
<td>Sigma-Aldrich</td>
<td>1:200</td>
<td>AlexaFluor 594 donkey α-mouse</td>
<td>Invitrogen</td>
<td>1:500</td>
</tr>
</tbody>
</table>

Table 19. List of AbI and AbII for IF analysis of Flp-In HEK293 expressing TDP-43 12XQ/N.
3.5.2 Immunofluorescence analysis of SH-SY5Y cells and differentiated NSC-34 cells

SH-SY5Y cells (40×10⁴), SH-SY5Y treated with siRNA against DAZAP1, hnRNP Q and hnRNP R (40×10⁴) and differentiated NSC-34 cells (35×10⁴) were plated in 6-well plates containing coverslips. For SH-SY5Y depleted cells and differentiated NSC-34 cells coverslips were coated with poly-L-lysine solution at a final concentration of 0.01% (w/v) in H2O (SigmaAldrich, St Louis, MO, USA). After 24 hours, cells were washed three times with PBS, fixed in 3.2% paraformaldehyde in PBS for 1 hour at room temperature and permeabilized by using 0.3% Triton in PBS for 5 minutes on ice. Cells were then blocked with 2% BSA/PBS for 20 minutes at room temperature and incubated overnight at 4 °C in a humidified chamber with 60 µl of primary antibodies prepared as described in Table 20. The day after, cells were washed 3 times with PBS, incubated with secondary antibodies conjugated with a fluorescent dye (Table 20) for 1 hour at room temperature and coverslipped with Vectashield-DAPI mounting medium (Vector Laboratories, Burlingame, CA, USA). Each slide was analysed using Nikon Ti-E confocal microscope with a 60X oil objective available in ICGEB institute or at the microscopy facility of University of Trieste, using a Nikon Eclipse C1si confocal microscope system mounted on a Nikon TE-2000U inverted microscope with a 60X objective.
Materials and Methods

Table 20. List of AbI and AbII for IF analysis of SH-SY5Y cells and differentiated NSC-34 cells.

<table>
<thead>
<tr>
<th>AbI</th>
<th>Company</th>
<th>Dilution</th>
<th>AbII</th>
<th>Company</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyclonal rabbit α-TDP-43</td>
<td>Proteintech</td>
<td>1:200</td>
<td>AlexaFluor 594 donkey α-rabbit</td>
<td>Invitrogen</td>
<td>1:500</td>
</tr>
<tr>
<td>Polyclonal rabbit α-hnRNP Q</td>
<td>Sigma-Aldrich</td>
<td>1:200</td>
<td>AlexaFluor 488 donkey α-rabbit</td>
<td>Invitrogen</td>
<td>1:500</td>
</tr>
<tr>
<td>Polyclonal rabbit α-hnRNP R</td>
<td>Abcam</td>
<td>1:200</td>
<td>AlexaFluor 488 donkey α-rabbit</td>
<td>Invitrogen</td>
<td>1:500</td>
</tr>
</tbody>
</table>
4. RESULTS

4.1 Effects of DAZAP1, hnRNP Q and hnRNP R suppression on TDP-43 activity in SH-SY5Y cell line

4.1.1 DAZAP1, hnRNPD and hnRNPR can regulate and rescue TDP-43 controlled splicing events without altering the expression of TDP-43

Thanks to the results obtained from the *Drosophila melanogaster* model of TDP-43 loss- or gain-of-function, in collaboration with the Neurobiology group in ICGEB, we decided to investigate the effects of DAZAP1, hnRNP Q and hnRNP R within human neuroblastoma (SH-SY5Y) cells. First of all, we measured the efficiency of siRNA treatment against each hnRNP and subsequently, we evaluated if the silencing of DAZAP1, hnRNP Q and hnRNP R was capable to affect the expression of TDP-43 at the protein level, by using Western Blot analysis. As shown in Figure 22A-B the efficiency of siRNA treatment against TDP-43, DAZAP1, hnRNP Q and hnRNP R was sufficient to knockdown ~80-70% of each protein in SH-SY5Y cells. Moreover, the depletion of these hnRNP proteins was not able to modify TDP-43 expression (Figure 22B), as well as the protein expression of DAZAP1, hnRNP Q and hnRNP R did not change after siRNA depletion of TDP-43 (Figure 22C-E).
Results

Figure 22. Western Blot analysis of siRNA treatments in SH-SY5Y cells. (A) Protein expression level of DAZAP1, hnRNP Q and hnRNP R after siRNA treatment comparing to cells treated with siRNA against firefly luciferase (siLUC) used as a control. The percentage of residual protein is also reported. (B) TDP-43 expression level following siRNA silencing of DAZAP1, hnRNP Q, and hnRNP R comparing to siLUC treated cells. Silencing of TDP-43 is also reported (the percentage of residual protein is 27%). (C-E) Protein expression level of DAZAP1 (C), hnRNP Q (D) and hnRNP R (E) were measured in SH-SY5Y cells treated with siRNA against TDP-43 comparing to siLUC treated cells. These experiments were performed in collaboration with Dr. Fatemeh Mohagheghi.

Successively, we analysed the ability of these hnRNP proteins to modulate pre-mRNA splicing events known to be controlled by TDP-43. Indeed, recent studies have demonstrated that the knockdown of TDP-43 causes exon 3 skipping of the POLDIP3/SKAR gene (Fiesel et al., 2012), as well as an increase in TNIK exon 15 (Colombrita et al., 2015) and STAG2 exon 30b inclusion (De Conti et al., 2015) and the exclusion of exon 31, in addition to the inclusion of a pseudoxon, in the MADD gene (De Conti et al., 2015)
We silenced each tested hnRNP protein (including TDP-43) in SH-SY5Y cells by using siRNA, and then we evaluated the different splicing pattern of SKAR/POLDIP3 exon3, TNIK exon 15, STAG2 exon 30b and MADD exon 31. According to previous studies, the silencing of TDP-43 was able to reproduce the splicing profile observed for each tested gene (Figure 23). As expected, the silencing of TDP-43 was associated with an increased exclusion of POLDIP3 exon 3 (Figure 23A), with an increased inclusion of TNIK exon 15 (Figure 23B), and of STAG2 exon 30b (Figure 23C), as well as with a decrease of exon 31 inclusion and simultaneous activation of a pseudoexon in the MADD gene (Figure 23D).

On the other hand, knocking down expression of DAZAP1 boosted POLDIP3 exon 3 inclusion (Figure 23A), and reduced both TNIK exon 15 (Figure 23B) and of STAG2 exon 30b inclusion below normal levels (Figure 23C). Likewise, removal of hnRNP Q was associated with a drop in the inclusion of both POLDIP3 exon 3 (Figure 23A) and TNIK exon 15 (Figure 23B). Finally, knocking down expression of hnRNP R was able to alter only the splicing of the TNIK gene, by causing an enhancement of exon 15 inclusion (Figure 23B). Notably, the splicing profile of MADD gene was modified only by TDP-43 silencing (Figure 23D).
Figure 23. Effects of different siRNA treatment on TDP-43-regulated events. SH-SY5Y cells were treated with different siRNA against TDP-43, DAZAP1, hnRNP Q and hnRNP R and RT-PCR was used to evaluate the effects of this depletion on POLDIP3 exon 3 (A), TNIK exon 15 (B), STAG2 exon 30b (C) and MADD exon 31 (D). The agarose gel was loaded with the following samples: control siRNA Luciferase transfected cells (lane 1, siLUC), depleted of TDP-43 (lane 2, siTDP-43), depleted of DAZAP1 (lane 3, siDAZAP1), depleted of hnRNP Q (lane 4, sihnRNQ) and depleted of hnRNP R (lane 5, sihnRNPR). The identity of the various transcripts is reported on the right. For the MADD gene, the appearance of a pseudoexon is also reported (Ps.Ex.). *P < 0.05, **P < 0.01, ***P < 0.001 (n = 3), calculated by an unpaired t-test. These experiments were performed in collaboration with Dr. Fatemeh Mohagheghi.
As a result of these changes, we were also interested in investigating if the co-silencing of TDP-43 and DAZAP1, hnRNP Q or hnRNP R could rescue the splicing alteration induced by TDP-43 depletion. Therefore, we silenced SH-SY5Y cells with both siRNA against TDP-43 and each hnRNP, and we analysed the resulting splicing profile of \textit{POLDIP3}, \textit{TNIK}, \textit{STAG2} and \textit{MADD} gene. Co-silencing of DAZAP1 and TDP-43, on one hand, restored the splicing patterns of the \textit{TNIK} (Figure 24B) and the \textit{MADD} gene (Figure 24D) as observed in the control transfections. On the other hand, it was able to counteract partially the \textit{POLDIP3} exon3 skipping (Figure 24A), and completely \textit{STAG2} exon30b inclusion (Figure 24C), caused by TDP-43 silencing alone.
Figure 24. Rescue of TDP-43 controlled events. SH-SY5Y cells were co-silenced with siRNA against TDP-43 in the presence of siRNA against DAZAP1, hnRNP Q and hnRNP R. RT-PCR was used to evaluate the effects of this depletion on POLDIP3 exon 3 (A), TNIK exon 15 (B), STAG2 exon 30b (C) and MADD exon 31 (D). The agarose gel was loaded with the following samples: control siRNA Luciferase transfected cells (lane 1, siLuc), depleted of TDP-43 (lane 2, siTDP-43), depleted of TDP-43 and DAZAP1 (lane 3, siTDP-43/siDAZAP1), depleted of TDP-43 and hnRNP Q (lane 4, siTDP-43/sihhnRNP Q) and depleted of TDP-43 and hnRNP R (lane 5, siTDP-43/sihhnRNP R). The identity of the various transcripts is reported on the right. For the MADD gene, the appearance of a pseudoexon is also reported (Ps.Ex.). *P < 0.05, **P < 0.01, ***P < 0.001 (n = 3), calculated by an unpaired t-test. These experiments were performed in collaboration with Dr. Fatemeh Mohagheghi.
Figure 25 summarizes these results:

Figure 25. Schematic diagram of TDP-43, DAZAP1, hnRNPQ and hnRNPR activity on RNA splicing of *POPLDIP3, STAG2, TNIK* and *MADD* gene. Exon inclusion and exon exclusion are highlighted in green and red, respectively. The colours also indicate the relative effects (enhancer or silencer activity) of each hnRNP on the target genes. Moreover, the figure reports the rescue of TDP-43 depletion together with silencing of each hnRNP.

The second important event that we decided to test was the effect of these hnRNPs on gene expression levels of the TDP-43-controlled genes, previously published in our laboratory. Once again, we treated SH-SY5Y cells with siRNA against TDP-43 and each hnRNP of interest and then we analysed the expression of *MADD* and *BRD8* genes, since their expression was found to be altered in HEK293 cells depleted for TDP-43 (De Conti et al., 2015). We also included the evaluation of *TNIK* gene expression levels as a control. As shown in Figure 26A-C, a significant reduction of *MADD* gene expression levels was associated with siRNA depletion of TDP-43, and a slight but significant reduction of *TNIK* gene expression levels was observed after siRNA depletion of hnRNPQ, suggesting that
the other hnRNPs were not able to modify the expression of these TDP-43 targets. Afterwards, the co-silencing effects of TDP-43 and DAZAP1, hnRNP Q or hnRNP R on the gene expression were tested. Apparently, only the simultaneous depletion of TDP-43 and DAZAP1 was able to rescue the reduction of *MADD* gene expression levels and induce a significant increment of *BRD8* gene expression levels (Figure 26D). No statistically significant effect was detected for *TNIK* gene expression levels (Figure 26F).

Figure 26. Effects of DAZAP1 and hnRNP Q/R depletion in the absence (A-C) or in the presence (D-F) of TDP-43 co-silencing on gene expression events controlled by TDP-43. Real-time PCR quantification analysis of *MADD* (A, D), *BRD8* (B, E) and *TNIK* (C, F) endogenous transcript levels following siRNA transfection in SH-SY-5Y cells from three independent experiments. Each bar reports the mean ± standard deviation of three independent experiments. *P < 0.05, **P < 0.01 (n = 3), calculated by an unpaired t-test.
4.1.2 DAZAP1, hnRNP Q and hnRNP R do not alter TDP-43 nuclear localization

Since modifications of mRNA splicing/expression were observed in TDP-43 controlled genes after hnRNP depletion, we were also interested in studying the localization of endogenous TDP-43, as its nucleo-cytoplasmic shuttling usually occurs after protein homeostasis alterations within the cells. For this reason, we performed immunofluorescence assays on SH-SY5Y cells treated with siRNA against each hnRNP protein of interest. As shown in Figure 27A-D, none of these treatments caused appreciable changes in TDP-43 localization, that remained mostly nuclear.

![Immunofluorescence images](image-url)

Figure 23. Localization of endogenous TDP-43 in SH-SY5Y cells. Cells were transfected with siRNA against fire-fly Luciferase (control) (A), DAZAP1 (B), hnRNP Q (C) and hnRNP R (D). Cell nuclei were
Results

visualized using DAPI staining. Scale bars: 32 µm. For each image, b/w inserts (scale bars: 10 µm) of a selected area (white dashed square) indicating DAPI (on the left) and α-TDP-43 (on the right) staining are also provided.

4.2 Cellular model of TDP-43 pathology

4.2.1 Analysis of siRNA-mediated depletion of DAZAP1, hnRNP Q and hnRNP R using a cellular model of TDP-43 loss-of-function

To examine in depth the effects of siRNA-mediated depletion of DAZAP1, hnRNP Q and hnRNP R in a context more close to the scenario of TDP-43 pathology, we have taken advantage of a cellular model of TDP-43 loss-of-function. This model is already available in our laboratory and established in Flp-In HEK293 cell line. Recently, it has been described a TDP-43 aggregation system, consisting of 12 tandem repeats carrying the 331–369-residue Gln/Asn region of TDP-43 fused to FLAG-tag or green fluorescentprotein (EGFP), reproduces many features of the TDP-43 inclusions observed in patients suffering from ALS and FTLD (Budini et al., 2012, 2014) (Figure 28A). Therefore, in order to evaluate the localization of the endogenous TDP-43 in this cellular model, we used Flp-In HEK293 cells producing FLAG-TDP43-12XQ/N aggregates (upon tetracycline induction) to perform immunofluorescence assays, after DAZAP1, hnRNP Q and hnRNP R knockdown. Despite it is difficult to quantify the amount of TDP-43 aggregation, the results obtained in this experiment show that the endogenous TDP-43 seems to co-localize with the inducible aggregates, and remains mostly nuclear either in the presence or absence of siRNA treatment (Figure 28). Therefore, the depletion of DAZAP1, hnRNP Q and hnRNP R appears not able to affect the aggregation propency of TDP-43, suggesting why, at the present, there is no evidence of the presence of these hnRNPs in the TDP-43 inclusions of ALS and FTLD patients.
Nonetheless, the silencing of both DAZAP1 and hnRNP Q seems to restore *POLDIP3* exon 3 inclusion (Figure 28C, left panel) and *STAG2* exon 30b skipping (Figure 28C, right panel) induced by the formation of the aggregates, as previously observed in the RNA splicing experiments performed in the SH-SY5Y cells.

**Figure 24.** Effects of hnRNP silencing in Flp-In HEK293 cells expressing FLAG-TDP43-12XQ/N aggregates. (A) Schematic representation of TDP-43 12XQ/N. (B) Immunofluorescence analysis showing the localization of endogenous (green) and aggregated (red) TDP-43 in the absence of hnRNP silencing (TDP-12XQ/N) and in the presence of siRNA against fire-fly Luciferase (siLUC, control), DAZAP1, hnRNP
Q and hnRNP R. Co-localization of endogenous TDP-43 with 12XQ/N aggregates is observed as yellow dots. Scale bars: 16 μm. Western Blot for checking hnRNP silencing is also reported. (C) RT-PCR of POLDIP3 and STAG2 genes in the presence of TDP-12XQ/N aggregates (lane 1, Tet) and silencing of siLUC (lane 2, control), siDAZAP1 (lane 3), sihnRNP Q (lane 4) and sinhRNP R (lane 5). Splicing analysis of POLDIP3 and STAG2 genes were performed in collaboration with Dr. Fatemeh Mohagheghi.

4.3 Characterization of TDP-43 and DAZAP1 interaction

4.3.1 DAZAP1 does not bind to TDP-43 but can bind in vivo to TDP-43 controlled mRNAs

Another possible connection between TDP-43 and hnRNP proteins able to explain the results obtained until now could be represented by their ability to bind TDP-43 and directly affect its functional properties. To this regard, proteomic studies on TDP-43 have confirmed the presence of hnRNP Q and hnRNP R in complexes that contain TDP-43 (Blokhuis et al., 2016; S. Ling et al., 2010), but no previous studies have uncovered a possible interaction between TDP-43 and DAZAP1. Therefore, we examined this interesting aspect by transfecting HeLa cells with flagged-TDP-43 and testing the presence of DAZAP1 by a co-immunoprecipitation (Co-IP) assay. This experiment shows that TDP-43 and DAZAP1 do not seem to co-IP together (Figure 29A).

Considering that data, we tested if DAZAP1 could be capable of binding to TDP-43 mRNA targets analysed in Figure 23, 24 and 26. To do this, we performed an RNA-immunoprecipitation (RNA-IP) assay by transfecting HeLa cells with flagged-DAZAP1, immunoprecipitating the RNA bound to DAZAP1 and testing by RT-qPCR the binding of DAZAP1 to BRD8, TNIK, STAG2, POLDIP3 and MADD transcripts. As shown in Figure 29B, all these transcripts were enriched after RNA-IP analysis, although BRD8 and TNIK in smaller amount. This could be dependent on their lower abundance in HeLa cells and/or
it could be associated with their weaker/unstabler interaction with TDP-43 than the other tested transcripts. *GAPDH* and *SDHA* were also tested as controls.

Importantly, in both the experiments of immunoprecipitation we decided to take advantage of HeLa cells instead of using SH-SY5Y cells in order to maximize the transfection efficiency with plasmid DNA.

**Figure 25. Probing the DAZAP1 - TDP-43 interaction.** (A) Co-immunoprecipitation experiment was performed by overexpressing FLAG-TDP-43 in HeLa cells and the immunoprecipitation was checked by Western Blot against DAZAP1 (upper panel). The presence of hnRNP H1 in the immunoprecipitated sample was used as a positive control (middle panel). The lower panel shows the levels of flagged and endogenous TDP-43 in the Input and immunoprecipitated sample. (B) RNA immunoprecipitation experiments (n=3) to monitor for DAZAP1 binding to *BRD8, TNIK, STAG2, POLDIP3, MADD* transcripts and also two housekeeping genes, *GAPDH* and *SDHA* (used as controls). RNA immunoprecipitation was performed in collaboration with Dr. Fatemeh Mohagheghi.
4.3.2 TDP-43 and DAZAP1 can alter the expression of neuronal related mRNAs in SH-SY5Y cells

Based on the results obtained from mRNA analysis of TDP-43 targets, in order to determine the wide effects of the hnRNPs of interest, we decided to obtain the whole transcriptome status of SH-SY5Y cells depleted by TDP-43, DAZAP1, hnRNP Q or hnRNP R.

In this regard, because DAZAP1 seems to be the best modifier of the TDP-43 controlled events, we initially focused the attention on the RNA-seq data achieved following silencing of DAZAP1 and TDP-43. Three independent experiments were performed. Notably, prior to the differential expression analysis, weak expression features with counts-per-million value (CPM) of less than one were excluded from the analysis, resulting in the removal of 47622 of the 64769 features initially generated by RNA-seq analysis. Therefore, a total of 17147 final features underwent to differential expression analysis. The putative differentially expressed genes (DEGs) were identified using a fold change (FC) cut-off of 30% with respect to a control sample treated with siRNA against fire-fly luciferase (siLUC) (upregulation cut-off: >1.3; downregulation cut-off: <0.7-FC).

Among the total 17147 genes analysed, 3533 genes were differentially expressed in siTDP-43 treated cells, in comparison to siLUC treated cells. Of these 3533 genes, 1173 (33%) genes were downregulated and 2360 (67%) genes were upregulated.

In the same way, we analysed the results obtained from cells silenced for DAZAP1 and we found that among the total 17147 genes analysed, 7571 genes resulted to be differentially expressed in siDAZAP1 treated cells in comparison to siLUC treated cells. Of these 7571 genes, 3244 (43%) genes were downregulated and 4327 (57%) genes were upregulated.

Interestingly, enough 1924 genes of the differentially expressed genes in siTDP-43 and siDAZAP1 treated cells were commonly regulated (Figure 30A).
To validate the RNA-seq data, we selected 10 genes belonging to the top 100 differentially expressed genes by siTDP-43 and siDAZAP1 treatment, based on their involvement in neuron development and brain inflammation (Figure 30B). The fold changes obtained from RT-qPCR were compared with RNA-seq expression analysis results (Figure 30C). *CUGBP Elav-like Family Member 5 (CELF5), Syntaxin 3 (STX3), Acetylcholinesterase (ACHE), Tumor Necrosis Factor (TNF), Tumor Necrosis Factor Receptor Superfamily, Member 9 (TNFRSF9), Intercellular Adhesion Molecule 1 (ICAM1) and Yippee-like 4 (YPEL4)*, were prominently found to be upregulated both by siDAZAP1 treatment and by siTDP-43 treatment. *ELAV like Neuron-Specific RNA binding protein 3 (ELAVL3) and Neuron-Oncological Ventral Antigen 2 (NOVA2)* were found to be downregulated in both siTDP-43 and siDAZAP1 treatment, whereas *Reelin (RELN)* was found to be downregulated only in siDAZAP1 treated cells and was chosen as a neutral gene (for TDP-43 treated cells) and it was also confirmed by qPCR validation. RNA-seq analysis for both TDP-43 and DAZAP1 was also represented by volcano plot and MAPlot (Figure 31 and 32, respectively), in which \( P < 0.05 \) was used as a statistical hypothesis test. In both graphs, differentially expressed genes (DEGs) are highlighted in red (downregulated) and in green (upregulated) based on the Pvalue and the cut-off of 30% variation with respect to the control treated cells (siLUC). For both siTDP-43 and siDAZAP1 treated cells in the volcano plot is reported the position of the DEGs validated by using qPCR, whereas the position of *TARDBP* and *DAZAPI* genes is reported in TDP-43-related MAPlot and DAZAP1-related MAPlot, respectively.

Furthermore, we performed a pathway analysis using PANTHER Pathway and DAVID annotations to investigate functional associations of gene expression changes and pathways in which differentially expressed genes were involved and enriched. As shown in Figure 30D 51% of 433 pathways identified in siTDP-43 treated cells and 57% of 889 pathways
identified in siDAZAP1 treated cells were related to brain function and neurodegeneration. Interestingly, the greater amount of common differentially regulated genes is related to inflammation (Figure 33B), suggesting an important role of this pathway in the presence of hnRNP dis-regulation.

Finally, we also looked at the Gene ontology (GO) functional enrichment analysis using the UniProt-KB Retrieve/ID mapping tool (Figure 33A). This analysis showed that siTDP-43 and siDAZAP1-treated cells share a similar amount of differentially expressed genes in the subgroups of "Biological process” and "Molecular function” categories. For both TDP-43 and DAZAP1, the "molecular function" category of GO terms associated with regulated transcripts includes a large percentage of “transcription factor activity, transcription factor binding, nucleic acid binding transcription factor activity, signal transducer activity, transporter activity, molecular function regulator and receptor activity”. On the other hand, in the “biological process” category, a similar percentage in genes involved in “cellular process, biological regulation, single-organism process and metabolic processes” groups were notably represented for both siTDP-43 and siDAZAP1.
Results

A

<table>
<thead>
<tr>
<th>Total genes analyzed</th>
<th>17147</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Downregulated (&lt;0.7 vs siLuc):</strong></td>
<td></td>
</tr>
<tr>
<td>siTDP-43</td>
<td>1173</td>
</tr>
<tr>
<td>siDAZAP1</td>
<td>3244</td>
</tr>
<tr>
<td>siTDP-43/siDAZAP1 common</td>
<td>489</td>
</tr>
<tr>
<td><strong>Upregulated (&gt;1.3 vs siLuc):</strong></td>
<td></td>
</tr>
<tr>
<td>siTDP-43</td>
<td>2360</td>
</tr>
<tr>
<td>siDAZAP1</td>
<td>4327</td>
</tr>
<tr>
<td>siTDP-43/siDAZAP1 common</td>
<td>1435</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>ENSEMBL</th>
<th>Gene Symbol</th>
<th>Description</th>
<th>RNAseq</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000196361</td>
<td>ELAVL3</td>
<td>ELAV-like neuron-specific RNA binding protein 3</td>
<td>0.28</td>
</tr>
<tr>
<td>ENSG00000104967</td>
<td>NOVA2</td>
<td>Neuro-ocular ventral antigen 2</td>
<td>0.38</td>
</tr>
<tr>
<td>ENSG00000189056</td>
<td>RELN</td>
<td>Relien</td>
<td>1.06</td>
</tr>
<tr>
<td>ENSG00000166900</td>
<td>CELF5</td>
<td>CELF5, CUGBP, Elav-like family member 5</td>
<td>4.59</td>
</tr>
<tr>
<td>ENSG00000087085</td>
<td>STX3</td>
<td>Syntaxin 3</td>
<td>5.49</td>
</tr>
<tr>
<td>ENSG00000233210</td>
<td>ACHE</td>
<td>Acetylcholinesterase</td>
<td>5.83</td>
</tr>
<tr>
<td>ENSG00000049249</td>
<td>TNF</td>
<td>Tumor necrosis factor</td>
<td>20.54</td>
</tr>
<tr>
<td>ENSG00000090339</td>
<td>TNFRSF9</td>
<td>Tumor necrosis factor receptor superfamily, member 9</td>
<td>35.97</td>
</tr>
<tr>
<td>ENSG00000161082</td>
<td>ICAM1</td>
<td>Inter cellular adhesion molecule 1</td>
<td>50.03</td>
</tr>
<tr>
<td>ENSG00000166793</td>
<td>YPEL4</td>
<td>YPEL4, vimentin-like 4 (Drosophila)</td>
<td>81.11</td>
</tr>
</tbody>
</table>

B

Figure 26. Validation of the TDP-43 or DAZAP1 silencing and comparison between RNA-seq and qRT-PCR results. (A) Summary of downregulated (<0.7× versus siLuc) and upregulated (>1.3× versus siLuc) genes after siTDP-43 or siDAZAP1 treatments. The number of common (between siTDP-43 and siDAZAP1) downregulated and upregulated genes is also shown. (B) List of genes associated with brain functions (**ELAV3, NOVA2, RELN, STX3, ACHE, YPEL4, CELF5**) or inflammation (**TNF, TNFRSF9, ICAM1**) selected for validation of the RNA-seq analysis. The expression levels of genes following siTDP-43
or siDAZAP1 treatments versus the control condition (siLUC) is indicated. (C) Validation of RNA-seq by real time PCR of the ten selected transcripts. The results are represented as relative expression compared with the control (siLUC). Each bar reports the mean ± standard deviation of three independent experiments. The single asterisks indicate significant differences (P< 0.05) between the indicated measurements. (D) Pathways analysis of differentially expressed genes following TDP-43 and DAZAP1 depletion as determined by PANTHER, DAVID and UniProt analyses.

Figure 27. Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for TDP-43. (A) Volcano plot reports in x-axis the magnitude (Log2FoldChange) and in y-axis the significance (–log10(Pvalue)) of siLUC versus siTDP-43 treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively, and validated genes are signalled by arrows. The horizontal blue line signals statistical significance threshold (P= 0.05). (B) MAPlot reports in x-axis the normalized library
size (Log2CPM, counts per million) and in y-axis the magnitude (Log2FoldChange) of siLUC versus siTDP-43 treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively. The position of TARDBP gene (encoding TDP-43 protein) is highlighted by an arrow.

---

**Figure 28.** Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for DAZAP1. (A) Volcano plot reports in x-axis the magnitude (Log2FoldChange) and in y-axis the significance ($-\log_{10}(P\text{value})$) of siLUC versus siDAZAP1 treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively, and validated genes are signalled by arrows. The horizontal blue line signals statistical significance threshold ($P=0.05$). (B) MAPlot reports in x-axis the normalized library size (Log2CPM, counts per million) and in y-axis the magnitude (Log2FoldChange) of siLUC versus siDAZAP1 treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively. The position of DAZAP1 gene (encoding DAZAP1 protein) is highlighted by an arrow.
Results

Figure 29. GO functional enrichment analysis of the genes significantly co-regulated by TDP-43 and DAZAP1. (A) For TDP-43, 2269 out of 3535 (1173 downregulated <0.7 and 2360 upregulated >1.3 vs. siLUC) Ensembl identifiers were successfully mapped to 8481 UniProtKB IDs. For DAZAP1, 5137 out of
7571 (3244 downregulated <0.7 and 4327 upregulated>1.3 vs. siLUC) Ensembl identifiers were successfully mapped to 19541 UniProtKB IDs. The clustered columns charts show the percentage of the successfully mapped Ensembl IDs. (B) Pathway analysis of the differentially expressed genes common between siTDP-43 and siDAZAP1 depletions, as determined by PANTHER analysis.

4.4 Characterization of hnRNP Q and hnRNP R

4.4.1 Analysis of protein sequence identity and similarity

To further refine the characterization of the hnRNPs in this study, we decided to extend our investigation to hnRNP Q and hnRNP R, in particular because they share similar protein sequence (Zissimos Mourelatos et al., 2001) and because in lower organisms, such as Drosophila melanogaster (McDermott et al., 2012) and Caenorhabditis elegans (Kabat et al., 2009), they were found in one well conserved ortholog, suggesting the occurrence of a progressive functional divergence during the evolution.

Concerning this aspect, we decided to compare the protein sequence of the major splicing isoforms of each human counterpart with their D. melanogaster ortholog (Table 21). To this aim, we used EMBOSS needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) and we calculated the percentage of identity and similarity of human hnRNP Q1, hnRNP Q2, hnRNP Q3, hnRNP R1 and hnRNP R2 with respect to the Drosophila variant CG17838 (isoform F)/Syp (Figure 34), that we previously used in our TBPH loss- and gain-of-function model and it has been well-described in larvae, pupae, and in all adult tissues (McDermott et al., 2012). The obtained results confirmed the high level of homology of each human splicing variant compared to the Drosophila one.
Figure 30. Structure of *Drosophila melanogaster* CG17838 (isoform F)/Syp variant. This variant contains a conserved AcD, three RRM (RRM1, RRM2 and RRM3) and two NLS. Each domain is highlighted in colored boxes and the relative sequence position and amino acids length is also reported.

Table 21. Percentage of identity and similarity of human hnRNP Q and hnRNP R with respect to *Drosophila* CG17838 (isoform F)/Syp variant.

<table>
<thead>
<tr>
<th>Isoform</th>
<th>% identity</th>
<th>% similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnRNP Q1</td>
<td>47.1</td>
<td>59.6</td>
</tr>
<tr>
<td>hnRNP Q2</td>
<td>42.9</td>
<td>53.1</td>
</tr>
<tr>
<td>hnRNP Q3</td>
<td>45.9</td>
<td>57.1</td>
</tr>
<tr>
<td>hnRNP R1</td>
<td>46.0</td>
<td>58.2</td>
</tr>
<tr>
<td>hnRNP R2</td>
<td>43.0</td>
<td>53.8</td>
</tr>
</tbody>
</table>

4.4.2 Subcellular localization of hnRNP Q and hnRNP R in SH-SY5Y cells

In order to investigate the reasons underlying a possible functional divergence between hnRNP Q and hnRNP R, we initially decided to look at their subcellular distribution in SH-SY5Y cells through an immunofluorescence analysis. As shown in Figure 35, hnRNP Q was spread amongst cytoplasm and nucleus, while hnRNP R seemed to be mostly nuclear.
Results

Figure 31. Cellular localization of endogenous hnRNP Q and hnRNP R in SH-SY5Y cells. Immunofluorescence analysis of the endogenous hnRNP Q and hnRNP R (shown in green) in SH-SY5Y cells. Nuclei were visualized using DAPI staining. Scale bars: 17 µm.

These results were also confirmed by nuclear cytoplasmic fractionation, as reported in Figure 36. Regarding hnRNP Q, three major splicing isoforms have been so far previously reported (Mourelatos et al. 2001). These isoforms are characterized by the presence of two NLS in the hnRNP Q3 and hnRNP Q2 variants (with a molecular weight of 70 kDa and 65 kDa, respectively) and one NLS in the hnRNP Q1 (with a molecular weight of 62 kDa). Our experiments detected three immunoreactive bands (∼58 kDa, ∼66 kDa and ∼75 kDa), differentially distributed between nucleus and cytosol (Figure 36A). The molecular weight of ∼75 kDa was consistent with that of hnRNP Q2/hnRNP Q3 isoforms and the molecular weight of ∼66k Da with that of hnRNP Q1 isoform. The lower band (∼58 kDa) cannot be associated to any known hnRNP Q isoform and could correspond to a further variant that still remains to be characterized, or non-specific cross-reaction. This observation could be explained by the fact that the latter variant gives a weak signal and is therefore not always
detectable. For this, it was not detected in our previous Western Blot analysis (Figure 22, 28A).

The same analysis was repeated for hnRNP R, confirming its presence predominantly in the nuclear fraction (Figure 36B). The antibody used for staining the membrane (ab30930) detected three bands: ~71 kDa, ~75 kDa and ~80 kDa. In the first set of Western Blot experiments performed using hnRNP R antibody (Figure 22, 28A), we only reported the band corresponding to the predicted molecular size described in the datasheet provided by the company. Due to the fact that the literature regarding different hnRNP R isoforms (and consequently variable size of this protein) is still incomplete, in the second part of the current project, we decided to additionally consider the other two hnRNP isoforms recognized by the same antibody. These two isoforms were not taken into account in the first part of the experiments, since they were not in agreement with the predicted size reported by the antibody producer.

Therefore, according to the literature, we assigned the higher molecular weight band (~ 80 KDa) to the major hnRNPR isoform, also known as R1 (NP_005817.1. NM_005826.4. [O43390-1]); and the second higher band (~75 kDa) to the more neuronal variant, namely R2 (NP_001284549.1, NM_001297620.1 [O43390-3]). On the other hand, we supposed that the lower band of ~71 kDa was associated with another splicing variant of hnRNP R that still needs to be characterized.

Overall, the different localization of hnRNP Q and hnRNP R suggested that their nuclear-cytoplasmic distribution is differentially regulated and this could result in a differential control of the regulated cellular pathways.
4.4.3 Subcellular localization of hnRNP Q and hnRNP R is not altered by their reciprocal silencing

Because of the close relationship between these two factors, we decided to investigate if the silencing of hnRNP Q was able to affect the gene expression and cellular localization of hnRNP R and vice-versa (Figure 37). We observed no significant differences in the mRNA levels, as well as in the endogenous localization of hnRNP Q after siRNA treatment against hnRNP R and of hnRNP R after siRNA treatment against hnRNP Q.
Figure 33. mRNA levels and cellular localization of endogenous hnRNP Q and hnRNP R after siRNA treatment in SH-SY5Y cells. (A) Relative expression of hnRNP Q after siRNA against hnRNP R (sihnRNP R) and vice-versa. All the samples were compared to siLUC treated cells. Each bar reports mean ± standard error (n=3). Statistical differences were evaluated using Student’s t-test (***, P<0.001). (B) Immunofluorescence analysis of the endogenous human hnRNP Q and hnRNP R (shown in green) in SH-SY5Y cells after sihnRNP R and sihnRNP Q, respectively. Nuclei were visualized using DAPI staining. Scale bars: 17 μm.
4.4.4 Subcellular localization of hnRNP Q and hnRNP R is not altered by retinoic acid (RA)-induced neuronal differentiation in NSC-34 cells

It is well-known that the activation of specific signaling pathways leading to cellular differentiation can be differently regulated depending on the subcellular localization of specific controlling factors (Robinow and White, 1991). It is also known that both hnRNP Q and hnRNP R can regulate neuronal pathways during development (Chen et al., 2012; McDermott et al., 2012, 2014).

Therefore, we tested if hnRNP Q or hnRNP R might change their cellular localization under neuronal differentiation. To this aim, immunofluorescence experiments were carried out after inducing differentiation up to 5 days with RA at 1 µM (Figure 38A) of the murine motor neuron-like NSC-34 cell line. Considering the high conservation of hnRNP Q and hnRNP R in mouse (more than 99% identity and similarity), we decided to use this cell line due to its ability to differentiate in neuron-like cells better and faster than the SH-SY5Y neuroblastoma cells.

The results indicated that the differentiation did not affect the subcellular distribution of both hnRNP Q and hnRNP R (Figure 38B).
Results

Figure 34. Endogenous localization of hnRNP Q and hnRNP R is not affected by differentiation with 1 µM of RA in NSC-34 cells. (A) Light microscopy images of undifferentiated (CNTRL) and differentiated cells treated up to 5 days with medium containing 1 µM RA. Scale bars: 58 µm, 20X magnification. Neurite length comparison (grey bars) of CNTRL and RA is also reported as mean ± standard error (n=5). Statistical differences were evaluated using Student’s t-test (* P<0.001). (B) Immunofluorescence analysis of the endogenous mouse hnRNP Q and hnRNP R (shown in green) in NSC-34 cells with or without 1 µM RA treatment. Nuclei were visualized using DAPI staining. Scale bars: 17 µm.
4.4.5 Depletion of hnRNP Q and hnRNP R alters neuronal-related genes

To define functional characteristics of human hnRNP Q and hnRNP R within cells, we decided to perform gene expression profiling of SH-SY5Y silenced for these hnRNPs. First of all, we demonstrated the reduction of mRNA levels of hnRNP Q and hnRNP R using qPCR after their knockdown (Figure 39A-40A).

Then, we performed an RNA-seq analysis by comparing data obtained from sihnRNP Q or sihnRNP R treated cells, relative to siLUC control samples. Three independent experiments were performed.

As described for RNA-seq analysis of DAZAP1 and TDP-43, prior to the differential expression analysis, weak expression features with counts-per-million value (CPM) of less than one were excluded from the analysis, resulting in the removal of 47622 of the 64769 features initially generated by RNA-seq analysis. Therefore, a total of 17147 final features underwent to differential expression analysis. Up- and downregulated genes were identified applying a FC cut-off of 30% with respect to control siLUC samples (upregulation cut-off: >1.3; downregulation cut-off: <0.7-FC) and the P value < 0.05.

Regarding sihnRNP Q, the total number of DEGs was 2819 (out of the 17147 analyzed genes), among which 1380 (49%) genes were upregulated and 1439 (51%) genes were downregulated (Figure 39A). On the other hand, after treatment with siRNA against hnRNP R, the total number of DEGs was 1517 genes (out of the 17147 analyzed genes) among which 957 (63%) genes were upregulated and 560 (37%) genes were downregulated (Figure 40A).

Validation of RNA-seq data was carried out by qPCR expression analysis of 10 genes selected among the top 100 DEGs for their potential involvement in brain functions and neuroinflammation (Figures 39B-40B).

Regarding cells silenced for hnRNP Q, Tumor Necrosis Factor (TNF), Intercellular
Adhesion Molecule 1 (ICAM1), Proenkephalin (PENK), Tumor Necrosis Factor Receptor Superfamily, Member 9 (TNFRSF9), Kruppel-like Factor 4 (gut) (KLF4), Kelch-like Family Member 4 (KLHL4) and Neurogulin 3 (NRG3) were found to be upregulated, while RAB26, Member RAS Oncogene Family (RAB26), Rho GTPase Activating Protein 36 (ARHGAP36) and Cancer/testis Antigen 55 (CT55) were found downregulated (Figure 39C). On the other hand, concerning cells silenced for hnRNP R, CART prepropeptide (CARTPT), FBJ Murine Osteosarcoma Viral Oncogene Homolog B (FOSB), Jagged 1 (JAG1), Intercellular Adhesion Molecule 5, Telencephalin (ICAM5), Dual Oxidase Maturation Factor 1 (DUOX1) and Heme Oxygenase (decycling 1) (HMOX1) were found to be upregulated, while Potassium Voltage-gated Channel, Shaker-related Subfamily Beta (KCNAB1), Acid Phosphatase 5, Tartrate Resistant (ACP5), Syndecan Binding Protein (Syntenin) 2 (SDCBP2) and EGF containing Fibulin-like Extracellular Matrix Protein 1 (EFEMP1) were found to be downregulated (Figure 40C). In conclusion, the results of our qPCR validation are consistent with those obtained with the RNA-seq analysis.
Results

Figure 35. Validation of hnRNP Q silencing and comparison between RNA-seq and RT-qPCR results. (A) Summary of differentially expressed genes (DEGs) with number of downregulated (<0.7x versus siLUC) and upregulated (>1.3x versus siLUC) genes after sihnRNP Q treatment. Number of DEGs with P<0.05 is also reported. (B) List of DEGs validated by RT-qPCR and associated with brain functions/neurodegeneration (PENK, KLF4, KLHL4, NRG3, RAB26, and ARHGAP36), inflammation (TNF, ICAM1, TNFRSF9) and other functions (CT55). (C) RT-qPCR validation of ten selected transcripts. Each bar reports mean ± standard error of three independent experiments. Statistical differences were evaluated using
Student’s t-test (* P<0.05, ** P<0.01, *** P<0.001). (D) Pathway analysis of differentially expressed genes following hnRNP Q depletion as determined by PANTHER, DAVID and UniProt analyses.

Figure 36. Validation of hnRNP R silencing and comparison between RNA-seq and RT-qPCR results. (A) Summary of DEGs with number of downregulated (<0.7x versus siLUC) and upregulated (>1.3x versus siLUC) genes after sihnRNP R treatment. Number of DEGs with P<0.05 is also reported. (B) List of DEGs validated by RT-qPCR and associated with brain functions/neurodegeneration (CARTPT, FOSB, JAG1, ...
DUOX1, HMOX1, KCNAB1, SDCBP2, EFEMP1), inflammation (ICAM5, ACP5). (C) RT-qPCR validation of ten selected transcripts. Each bar reports mean ± standard error of three independent experiments. Statistical differences were evaluated using Student’s t-test (*P<0.05, ** P<0.01, ***P<0.001). (D) Pathway analysis of differentially expressed genes following hnRNP R depletion as determined by PANTHER, DAVID and UniProt analyses.

An overview of these RNA-seq data was also provided by volcano plots (Figure 41) and MAPlots (Figure 42). Upregulated and downregulated genes were highlighted in green and red, respectively. Moreover, in the volcano plot, we reported the position of the DEGs validated in Figure 39C and 40C using qPCR. On the other hand, in the MAPlot, we highlighted the position of SYNCRIP (hnRNP Q, Figure 42A) and HNRNPR (Figure 42B) gene, respectively.
Figure 37. Volcano plot of hnRNP Q and hnRNP R. Schematic representation of RNA-seq data of sihnRNP Q (A) and sihnRNP R (B) treated cells. Up- and downregulated genes are reported as green and red dots, respectively. Not DEGs are represented as grey dots. Black arrows show qPCR validated genes. The horizontal blue line signals statistical significance threshold (P= 0.05).
Figure 38. Graphic representation of RNA-seq analysis performed on SH-SY5Y cells depleted for hnRNP Q and hnRNP R. (A) MAPlot reports on the x-axis the normalized library size (Log2CPM, counts per million) and on the y-axis the magnitude (Log2FoldChange) of siLUC versus sihnRNP Q treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively. The position of SYNCRIP gene (encoding hnRNP Q protein) is highlighted by an arrow. (B) MAPlot reports on the x-axis the normalized library size (Log2CPM, counts per million) and on the y-axis the magnitude (Log2FoldChange) of siLUC versus sihnRNP R treated cells. Upregulated and downregulated genes are highlighted in green and red, respectively. The position of HNRNPR gene (encoding hnRNP R protein) is highlighted by an arrow.
4.4.6 hnRNP Q and hnRNP R show different and common features after GO enrichment and KEGG pathway analysis

GO enrichment was performed to show up functional similarities and differences in DEGs regulated by hnRNP Q and hnRNP R. To this aim, we took advantage of the GOseq R Bioconductor package (Young et al., 2010) and considered for the final analysis only GO term of the “biological process” (BP), “molecular function” (MF), and “cellular component” (CC) categories reaching the P value threshold < 0.05 for significance. In particular, top 15 GO terms of each category were selected to find categories specific for hnRNP Q or hnRNP R and categories commonly present in both two proteins (Figure 43). In hnRNP Q data set, we found 1152 significantly enriched terms out of 2819 DEGs used as input for GO analysis. The top enriched GO categories were “membrane” (Pvalue = 3.17E-08), “bounding membrane to organelle” Pvalue = 7.29E-07), “cell morphogenesis involved in differentiation” (Pvalue = 1.10E-06), “cell development” (Pvalue = 1.42E-06) and “cell adhesion” (Pvalue = 3.17E-06). On the other hand, in hnRNP R data set, we found 955 significantly enriched terms out of 1517 DEGs used as input for GO analysis. The top enriched GO categories were “system development” (Pvalue = 2.34E-08), “tissue development” (Pvalue = 4.21E-08), “multicellular organism development” (Pvalue = 3.39E-07), “cell differentiation” (Pvalue = 4.28E-07) and “signal transduction” (Pvalue =1.13E-06). Notably, when we looked at GO terms differentially enriched in hnRNP Q and hnRNP R DEGs, we found that “intrinsic component of membrane”, “plasma membrane”, “cell periphery”, “integral component of membrane” and “membrane part” were particularly enriched in hnRNP Q depleted cells, while “signal receptor activity”, “transmembrane receptor activity”, “transmembrane signalling receptor activity”, “receptor activity” and “molecular transducer activity” were more enriched in hnRNP R depleted treated cells.
Regarding KEGG analysis, we found 29 and 16 terms with significant gene enrichment (P < 0.05) for hnRNP Q and hnRNP R, respectively (Figure 44) and the majority of them were inflammation-related pathways. Indeed, “toll-like receptor signalling pathway” (P-value = 0.002), “ECM-receptor interaction” (P-value = 0.002), “adipocytokine signaling pathway” (P-value = 0.003), “toxoplasmosis” (P-value = 0.004) and “rheumatoid arthritis” (P-value = 0.007) were particularly enriched in DEGs obtained by sihnRNP Q silencing, whereas “cell adhesion molecules (CAMs)” (P-value = 4.58E-05), “ECM-receptor interaction” (P-value = 0.001), “cytokine-cytokine receptor interaction” (P-value = 0.01), “T cell receptor signaling pathway” (P-value = 0.02) and “malaria” (P-value = 0.02) were particularly enriched in DEGs obtained by sihnRNP R silencing.

In conclusion, this analysis suggested a specific specialization acquired from hnRNP Q and hnRNP R during evolution. In particular, GO terms linked to the organization of lipid layers and in the regulation of cell-cell and cell-extracellular matrix contacts seemed to be influenced by sinhnRNP Q treatment, while, GO terms associated with cell/tissue development and cell signalling was found to be regulated by depletion of hnRNP R.
R. the first 15 molecular function (MF) and cellular component (CC) Figure oxidoreductase activity, acting on the CH oxygen (*) Positive sub acting morphogenesis - Regulation as G- Cell regulation categories (Transmembrane protein Transmembrane Intrinsic Integral component of plasma membrane Intrinsic Integral component of membrane Intracellular component of membrane Multicellular organism development Cellular developmental process Cell differentiation Multicellular organism development Tissue development System development Neurotransmitter transporter activity Beta-galactosidase activity Oxioreductase activity (*) Transforming growth factor beta binding Heparin binding Transmembrane receptor protein kinase activity Growth factor binding Glycosaminoglycan binding Cell projection membrane Extracellular matrix Extracellular region part

Results

Figure 39. GO enrichment analysis of hnRNP Q and hnRNP R DEGs. Biological process (BP), molecular function (MF) and cellular component (CC) GO categories are represented. Each category reported the first 15 sub-categories (P<0.05) of DEGs for cells treated with siRNA against both hnRNP Q and hnRNP R.

(*) acting on the CH-NH2 group of donors, oxygen as acceptor
Figure 40. KEGG pathway analysis of hnRNP Q and hnRNP R. (A) First 10 pathways identified in DEGs of SH-SY5Y cells treated with siRNA against hnRNP Q (P<0.05). (B) First 10 pathways identified in DEGs of SH-SY5Y cells treated with siRNA against hnRNP R (P<0.05).
4.5 An overview of the entire transcriptome status after silencing of TDP-43, DAZAP1, hnRNP Q and hnRNP R

4.5.1 TDP-43, DAZAP1 and hnRNP Q share common regulated-transcripts associated with ALS pathology

Following the RNA-seq analysis of cells depleted for TDP-43, DAZAP1, hnRNP Q or hnRNP R we decided to combine all the DEGs (FC cut-off of 30% vs siLUC; P <0.05) of each hnRNP silencing in order to find commonly regulated transcripts (Figure 45).

![Venn diagram](image_url)

**Figure 41. Venn diagramm of RNA-seq data.** Comparison of the number of transcripts altered in TDP-43, hnRNP Q and hnRNP R depleted SH-SY5Y cell. Genes co-regulated by TDP-43, DAZAP1 and hnRNP Q but not hnRNP R are highlighted with red circle.

As described previously, our results in SH-SY5Y cells and Flp-In HEK293 expressing TDP-43 12XQ/N demonstrated that DAZAP1 and hnRNP Q were particularly effective in modulating the activity of TDP-43 targets. In this way, we were interested in looking for
all the possible mRNA targets regulated by TDP-43, DAZAP1, hnRNP Q but not hnRNP R and we identified 214 genes reflecting these criteria (Figure 45, Table 22). This list of genes shows that more than 50% of the transcripts affected by TDP-43, DAZAP1, and hnRNP Q play an important role in neuronal maintenance/functions (such as \textit{NOVA2}), immune system and inflammation (\textit{TNF, TNFRSF9, ICAM-1}), and apoptosis (Figure 46). These observations lend further support to recent studies that hypothesize a role of immune system and apoptosis in the ALS pathogenesis. Furthermore, in this list of 214 co-regulated genes there are already several candidates that are extremely promising for the understanding the RNA misregulation occurring during neurodegenerative disorders and that will be discussed in the next chapter.

\textbf{Table 22.} Co-regulated transcripts between TDP-43, DAZAP1 and hnRNP Q

<table>
<thead>
<tr>
<th>ENSEMBL</th>
<th>Gene Symbol</th>
<th>RNA-seq Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TDP-43</td>
</tr>
<tr>
<td>ENSG00000181201</td>
<td>\textit{HIST3H2BA}</td>
<td>0.1</td>
</tr>
<tr>
<td>ENSG00000205625</td>
<td>\textit{RP11-15G8.1}</td>
<td>0.2</td>
</tr>
<tr>
<td>ENSG00000157703</td>
<td>\textit{SVOPL}</td>
<td>0.3</td>
</tr>
<tr>
<td>ENSG00000042832</td>
<td>\textit{TG}</td>
<td>0.3</td>
</tr>
<tr>
<td>ENSG00000100078</td>
<td>\textit{PLA2G3}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000258472</td>
<td>\textit{RP11-192H23.4}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000104967</td>
<td>\textit{NOVA2}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000230138</td>
<td>\textit{RP11-117D22.2}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000206838</td>
<td>\textit{SNORA5A}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000162188</td>
<td>\textit{GNG3}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000154997</td>
<td>\textit{SEPT14}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000167244</td>
<td>\textit{IGF2}</td>
<td>0.4</td>
</tr>
<tr>
<td>ENSG00000130600</td>
<td>\textit{H19}</td>
<td>0.4</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Fold Change 1</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td>ENSG00000132141</td>
<td>CCT6B</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000130176</td>
<td>CNN1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000175785</td>
<td>PRIMA1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000165424</td>
<td>ZCCHC24</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000265971</td>
<td>RP11-269G24.6</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000179921</td>
<td>GPBAR1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000167889</td>
<td>MGAT5B</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000142197</td>
<td>DOPEY2</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000163536</td>
<td>SERPINI1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000115756</td>
<td>HPCAL1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000173517</td>
<td>PEAK1</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000141753</td>
<td>IGFBP4</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000092929</td>
<td>UNC13D</td>
<td>0.5</td>
</tr>
<tr>
<td>ENSG00000123091</td>
<td>RNF11</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000198570</td>
<td>RD3</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000267980</td>
<td>AC007292.6</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000271155</td>
<td>RP11-435O5.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000179242</td>
<td>CDH4</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000185900</td>
<td>POMK</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000100092</td>
<td>SH3BP1</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000135097</td>
<td>MSII</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000198929</td>
<td>NOS1AP</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000177830</td>
<td>CHID1</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000144677</td>
<td>CTDSPL</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000183044</td>
<td>ABAT</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000270015</td>
<td>RP11-540B6.6</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000069431</td>
<td>ABCC9</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000164125</td>
<td>FAM198B</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000175093</td>
<td>SPSB4</td>
<td>0.6</td>
</tr>
<tr>
<td>ID</td>
<td>Gene</td>
<td>Value1</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>ENSG00000084710</td>
<td>EFR3B</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000167524</td>
<td>SGK494</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000101115</td>
<td>SALL4</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000088280</td>
<td>ASAP3</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000171303</td>
<td>KCNK3</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000159873</td>
<td>CCDC117</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000175229</td>
<td>GAL3ST3</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000101986</td>
<td>ABCD1</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG0000003249</td>
<td>DBNDD1</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000148700</td>
<td>ADD3</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000197122</td>
<td>SRC</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000174672</td>
<td>BRSK2</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000176749</td>
<td>CDK5R1</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000239887</td>
<td>C1orf226</td>
<td>0.6</td>
</tr>
<tr>
<td>ENSG00000146094</td>
<td>DOK3</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000125520</td>
<td>SLC2A4RG</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000149295</td>
<td>DRD2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000198168</td>
<td>SVIP</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000116981</td>
<td>NT5C1A</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000122254</td>
<td>HS3ST2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000197586</td>
<td>ENTPD6</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000173621</td>
<td>LRFN4</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000165238</td>
<td>WNK2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000108433</td>
<td>GOSR2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000261762</td>
<td>RP11-650L12.2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000121039</td>
<td>RDH10</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000073670</td>
<td>ADAM11</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000128482</td>
<td>RNF112</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000164742</td>
<td>ADCY1</td>
<td>0.7</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Log2 Fold Change 1</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>ENSG00000172794</td>
<td>RAB37</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000278619</td>
<td>MRM1</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000149260</td>
<td>CAPN5</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000259291</td>
<td>RP11-617F23.1</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000141750</td>
<td>STAC2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000100034</td>
<td>PPM1F</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000007402</td>
<td>CACNA2D2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000133997</td>
<td>MED6</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000139200</td>
<td>PIANP</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG0000015532</td>
<td>XYL1T2</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000171310</td>
<td>CHST11</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000162959</td>
<td>MEMO1</td>
<td>0.7</td>
</tr>
<tr>
<td>ENSG00000130772</td>
<td>MED18</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000171295</td>
<td>ZNF440</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000158373</td>
<td>HIST1H2BD</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000185658</td>
<td>BRWD1</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000171992</td>
<td>SYNOPO</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000155729</td>
<td>KCTD18</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000136895</td>
<td>GARNL3</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000177239</td>
<td>MAN1B1</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000158470</td>
<td>B4GALT5</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000112773</td>
<td>FAM46A</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000169981</td>
<td>ZNF35</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000139117</td>
<td>CPNE8</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000152518</td>
<td>ZFP36L2</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000181523</td>
<td>SGSH</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000033100</td>
<td>CHPF2</td>
<td>1.3</td>
</tr>
<tr>
<td>ENSG00000198046</td>
<td>ZNF667</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000090263</td>
<td>MRPS33</td>
<td>1.4</td>
</tr>
<tr>
<td>Ensembl ID</td>
<td>Gene Symbol</td>
<td>Ratio</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>ENSG00000188042</td>
<td>ARL4C</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000126947</td>
<td>ARMCX1</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000162396</td>
<td>PARS2</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000136167</td>
<td>LPC1</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000138760</td>
<td>SCARB2</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000067113</td>
<td>PPAP2A</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000106617</td>
<td>PRKAG2</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000196437</td>
<td>ZNF569</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000248383</td>
<td>PCDHAC1</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000143469</td>
<td>SYT14</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000273820</td>
<td>USP27X</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000185947</td>
<td>ZNF267</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000250091</td>
<td>DNAH10OS</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000170035</td>
<td>UBE2E3</td>
<td>1.4</td>
</tr>
<tr>
<td>ENSG00000146966</td>
<td>DENND2A</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000272947</td>
<td>RP11-71H17.9</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000116574</td>
<td>RHOU</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000153214</td>
<td>TMEM87B</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG000000003436</td>
<td>TFPI</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000238105</td>
<td>GOLGA2P5</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000138078</td>
<td>PREPL</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000132824</td>
<td>SERINC3</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000167740</td>
<td>CYB5D2</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000186866</td>
<td>POFUT2</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000031691</td>
<td>CENPQ</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000088727</td>
<td>KIF9</td>
<td>1.5</td>
</tr>
<tr>
<td>ENSG00000210107</td>
<td>MT-TQ</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000106853</td>
<td>PTGRI</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000153363</td>
<td>LINC00467</td>
<td>1.6</td>
</tr>
<tr>
<td>Accession</td>
<td>Gene Symbol</td>
<td>Log2Fold</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>ENSG00000279192</td>
<td>PWAR5</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000108733</td>
<td>PEX12</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000171365</td>
<td>CLCN5</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000157680</td>
<td>DGKI</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000100906</td>
<td>NFKBIA</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000135776</td>
<td>ABCB10</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000105708</td>
<td>ZNF14</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000107165</td>
<td>TYRP1</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000256304</td>
<td>CCDC150P1</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000171451</td>
<td>DSEL</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000177272</td>
<td>KCNA3</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000166192</td>
<td>SENP8</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000128923</td>
<td>FAM63B</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000164305</td>
<td>CASP3</td>
<td>1.6</td>
</tr>
<tr>
<td>ENSG00000114923</td>
<td>SLC4A3</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000134326</td>
<td>CMPK2</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000118976</td>
<td>OTUD4P1</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000280287</td>
<td>RP13-554M15.7</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG000000253869</td>
<td>PIGFP1</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000085449</td>
<td>WDFY1</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000090612</td>
<td>ZNF268</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000149243</td>
<td>KLHL35</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000087074</td>
<td>PPP1R15A</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000117586</td>
<td>TNFSF4</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000162616</td>
<td>DNAJB4</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000228876</td>
<td>AC010745.2</td>
<td>1.7</td>
</tr>
<tr>
<td>ENSG00000163820</td>
<td>FYCO1</td>
<td>1.8</td>
</tr>
<tr>
<td>ENSG00000128016</td>
<td>ZFP36</td>
<td>1.8</td>
</tr>
<tr>
<td>ENSG00000166432</td>
<td>ZMAT1</td>
<td>1.8</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Symbol</td>
<td>Fold Change</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>ENSG00000127561</td>
<td>SYNGR3</td>
<td>1.8</td>
</tr>
<tr>
<td>ENSG00000135318</td>
<td>NT5E</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000223427</td>
<td>AC016716.1</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000107201</td>
<td>DDX58</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000168398</td>
<td>BDKRB2</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000080709</td>
<td>KCNN2</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000135926</td>
<td>TMBIM1</td>
<td>1.9</td>
</tr>
<tr>
<td>ENSG00000168461</td>
<td>RAB31</td>
<td>2.0</td>
</tr>
<tr>
<td>ENSG00000142102</td>
<td>ATHL1</td>
<td>2.0</td>
</tr>
<tr>
<td>ENSG00000254473</td>
<td>RP11-522I20.3</td>
<td>2.0</td>
</tr>
<tr>
<td>ENSG00000142303</td>
<td>ADAMTS10</td>
<td>2.0</td>
</tr>
<tr>
<td>ENSG00000115665</td>
<td>SLC5A7</td>
<td>2.0</td>
</tr>
<tr>
<td>ENSG00000156966</td>
<td>B3GNT7</td>
<td>2.1</td>
</tr>
<tr>
<td>ENSG00000259381</td>
<td>RP11-192M23.1</td>
<td>2.2</td>
</tr>
<tr>
<td>ENSG00000248323</td>
<td>LUCAT1</td>
<td>2.2</td>
</tr>
<tr>
<td>ENSG00000259834</td>
<td>KCNA3</td>
<td>2.2</td>
</tr>
<tr>
<td>ENSG00000166246</td>
<td>C16orf71</td>
<td>2.2</td>
</tr>
<tr>
<td>ENSG00000135048</td>
<td>TMEM2</td>
<td>2.2</td>
</tr>
<tr>
<td>ENSG00000050628</td>
<td>PTGER3</td>
<td>2.3</td>
</tr>
<tr>
<td>ENSG00000248905</td>
<td>FMN1</td>
<td>2.3</td>
</tr>
<tr>
<td>ENSG00000168026</td>
<td>TTC21A</td>
<td>2.3</td>
</tr>
<tr>
<td>ENSG00000276603</td>
<td>RP11-425M5.7</td>
<td>2.3</td>
</tr>
<tr>
<td>ENSG00000166401</td>
<td>SERPINB8</td>
<td>2.4</td>
</tr>
<tr>
<td>ENSG00000168542</td>
<td>COL3A1</td>
<td>2.4</td>
</tr>
<tr>
<td>ENSG00000267682</td>
<td>CTD-3220F14.2</td>
<td>2.5</td>
</tr>
<tr>
<td>ENSG00000183696</td>
<td>UPP1</td>
<td>2.5</td>
</tr>
<tr>
<td>ENSG00000130762</td>
<td>ARHGEF16</td>
<td>2.5</td>
</tr>
<tr>
<td>ENSG00000148180</td>
<td>GSN</td>
<td>2.7</td>
</tr>
<tr>
<td>ENSG00000175592</td>
<td>FOSL1</td>
<td>2.8</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Log2 Ratio 1</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>ENSG00000161638</td>
<td>ITGA5</td>
<td>2.8</td>
</tr>
<tr>
<td>ENSG00000146592</td>
<td>CREB5</td>
<td>2.8</td>
</tr>
<tr>
<td>ENSG000000247157</td>
<td>LINC01252</td>
<td>2.8</td>
</tr>
<tr>
<td>ENSG000000136052</td>
<td>SLC41A2</td>
<td>2.8</td>
</tr>
<tr>
<td>ENSG000000224165</td>
<td>DNAJC27-AS1</td>
<td>2.9</td>
</tr>
<tr>
<td>ENSG000000137331</td>
<td>IER3</td>
<td>3.1</td>
</tr>
<tr>
<td>ENSG000000125246</td>
<td>CLYBL</td>
<td>3.1</td>
</tr>
<tr>
<td>ENSG00000063438</td>
<td>AHR</td>
<td>3.1</td>
</tr>
<tr>
<td>ENSG00000023839</td>
<td>ABCC2</td>
<td>3.3</td>
</tr>
<tr>
<td>ENSG000000140511</td>
<td>HAPLN3</td>
<td>3.3</td>
</tr>
<tr>
<td>ENSG000000114670</td>
<td>NEK11</td>
<td>3.3</td>
</tr>
<tr>
<td>ENSG000000108846</td>
<td>ABCC3</td>
<td>3.5</td>
</tr>
<tr>
<td>ENSG000000266921</td>
<td>RP11-15A1.7</td>
<td>3.7</td>
</tr>
<tr>
<td>ENSG000000163131</td>
<td>CTSS</td>
<td>3.9</td>
</tr>
<tr>
<td>ENSG00000172752</td>
<td>COL6A5</td>
<td>4.0</td>
</tr>
<tr>
<td>ENSG000000159450</td>
<td>TCHH</td>
<td>4.2</td>
</tr>
<tr>
<td>ENSG000000106366</td>
<td>SERPINE1</td>
<td>4.4</td>
</tr>
<tr>
<td>ENSG000000240065</td>
<td>PSMB9</td>
<td>4.4</td>
</tr>
<tr>
<td>ENSG000000106070</td>
<td>GRB10</td>
<td>5.4</td>
</tr>
<tr>
<td>ENSG000000204248</td>
<td>COL11A2</td>
<td>5.8</td>
</tr>
<tr>
<td>ENSG000000134070</td>
<td>IRAK2</td>
<td>6.5</td>
</tr>
<tr>
<td>ENSG000000196664</td>
<td>TLR7</td>
<td>7.0</td>
</tr>
<tr>
<td>ENSG000000187164</td>
<td>KIAA1598</td>
<td>7.6</td>
</tr>
<tr>
<td>ENSG000000122861</td>
<td>PLAU</td>
<td>8.7</td>
</tr>
<tr>
<td>ENSG000000232810</td>
<td>TNF</td>
<td>20.5</td>
</tr>
<tr>
<td>ENSG00000049249</td>
<td>TNFRSF9</td>
<td>36.0</td>
</tr>
<tr>
<td>ENSG00000090339</td>
<td>ICAM1</td>
<td>50.0</td>
</tr>
</tbody>
</table>

Results
**Results**

**Figure 42. Pathways of genes co-regulated only by TDP-43, DAZAP1 and hnRNP Q.** Classification of 214 genes whose expression is co-regulated in TDP-43, DAZAP1, hnRNP Q, but not hnRNP R depleted SH-SY5Y cells according to the GO categorization system.

<table>
<thead>
<tr>
<th>Pathways</th>
<th>TDP-43/DAZAP1/hnRNP Q co-regulated genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurological function</td>
<td>54 (25%)</td>
</tr>
<tr>
<td>Immune-related</td>
<td>28 (13%)</td>
</tr>
<tr>
<td>Apoptosis-related</td>
<td>23 (11%)</td>
</tr>
<tr>
<td>Other pathways</td>
<td>109 (51%)</td>
</tr>
</tbody>
</table>
5. DISCUSSION

5.1 TDP-43 activity is modulated by the interaction with other hnRNP proteins

Since TDP-43 was discovered as a major component of ubiquitinated inclusions in ALS and FTLD patients (Neumann et al., 2006) the interest in better characterizing the pathological role of this protein has considerably increased. In particular, in past years a lot of effort was put into better defining TDP-43 protein-protein interactions and in the identification of all the possible transcripts regulated by TDP-43 that could be altered in pathology. In this context, it has been well-characterized that TDP-43 can interact with other members of the hnRNP family and that this binding can affect the functions of both factors (Buratti and Baralle, 2012). Examples of TDP-43 interacting proteins are several members of hnRNP A/Bs family (D’Ambrogio et al., 2009; Romano et al., 2014), hnRNP H, and FUS/TLS (Freibaum et al., 2010).

In this thesis, I have investigated the relationship between TDP-43 and three hnRNP proteins (DAZAP1, hnRNP Q and hnRNP R) not yet completely characterized, although all these proteins were previously found to be associated with neuronal functions (Bannai et al., 2004; Glinka et al., 2010; Pastor and Pagani, 2011). We had also demonstrated, together with the Neurobiology group, that they represented powerful modifiers of TDP-43/TBPH in Drosophila melanogaster.

At the functional level, we first demonstrated that DAZAP1, hnRNP Q and hnRNP R can with different ability modify the splicing pattern or gene expression of well-known TDP-43 targets, such as POLDIP3, TNIK, STAG2, MADD and BRD8. These genes are involved in several aspects of cell proliferation (POLDIP3, STAG2), apoptosis (MADD, BRD8) and neuronal functions (TNIK). In particular, alteration in the POLDIP3 (also known as SKAR)
pre-mRNA splicing profiles were previously reported to be associated with depletion of TDP-43 in vitro and in motor cortex, spinal cord and spinal motor neurons of patients affected by ALS (Shiga et al., 2012). This previous work observed that knockdown of TDP-43 regulated the production of an alternative splicing isoform, namely variant-2, through the skipping of exon 3 and therefore influenced POLDIP3 function in mediating cell proliferation and growth (Shiga et al., 2012). In our experiments, we demonstrated that the inclusion of POLDIP3 exon 3 was also significantly altered in the presence of DAZAP1 and hnRNP Q knockdown (Figure 23A), indicating that splicing regulation is not exclusive for TDP-43 but it can also depend on the participation of other factors. This hypothesis was also corroborated by the fact that the co-silencing of TDP-43 with DAZAP1 was able to recover the POLDIP3 splicing profile (Figure 24A), indicating that the engagement of DAZAP1 is able to modulate the TDP-43 activity.

Other interesting observations supporting the role of DAZAP1, hnRNP Q and hnRNP R in regulating TDP-43 activity were obtained by studying the pre-mRNA splicing analysis of TNIK and STAG2 genes, as well as investigating changes in mRNA expression levels of TNIK.

TNIK is a serine/threonine kinase belonging to the Ste20 family that was originally cloned from a human cDNA library and successively identified in mouse brain (Fu et al., 1999). It was also described as an important factor for synaptic functions and cytoskeleton organization (Traikov et al., 2014; Fu et al., 1999), as well as, as a potential risk factor for several mental disorders (Matigian et al., 2007; Glatt et al., 2005). The association between TDP-43 and TNIK arose from a transcriptomic study in 2015 where the analysis of TDP-43 and FUS depletion led to the identification of several genes related to neuronal pathways and RNA metabolism, whose expression or pre-mRNA splicing profile was modified after siRNA treatment (Colombrita et al., 2015). In our study, we observed that
DAZAP1, hnRNP Q and even hnRNP R were able to alter TNIK’s splicing profile (Figure 23B) and, in the case of DAZAP1 and hnRNP Q, also to recover it after TDP-43 depletion (Figure 24B). Furthermore, we showed a decrease of TNIK mRNA levels after treatment of cells with an siRNA against hnRNP Q (Figure 26C).

Along the same lines, we looked at the STAG2 splicing profile. STAG2 is a component of the cohesin complex involved in separation of sister chromatids during cell division (Berkowitz, 2014). In 2015, our laboratory showed that STAG2 was involved in the accumulation of cells in G2/M phase due to upregulation of its exon 30b+ isoform after siTDP-43 treatment (De Conti et al., 2015). In this thesis I show that STAG2 exon 30b+ inclusion was also affected by DAZAP1 silencing (Figure 23C), and the depletion of both proteins was able to rescue the STAG2 profile (Figure 24C).

To complete the list of known TDP-43 targets, we further tested our hnRNPs for mediating changes in MADD and BRD8 gene expression regulation. The MADD gene is expressed in low levels in all tissues and can act as a pro-apoptotic factor when its phosphorylation is reduced by the Tumour Suppressor Phosphatase and Tensin Homolog (PTEN) (Jayarama et al., 2014). Interestingly, the MADD gene was also found to be important for neuronal survival since alteration in its splicing isoforms were found in in neuronal cell lines exposed to high concentrations of oligomeric Aβ peptides (Mo et al., 2012). Tollervey et al. and subsequently De Conti et al. demonstrated that siTDP-43 treatment could affect both the splicing process of this gene, in particular by inducing exon 31 exclusion plus activation of a pseudoexon, but also its mRNA levels (De Conti et al., 2015; Tollervey et al., 2011). Here, we demonstrated that although none of the tested hnRNPs was able to modify the MADD splicing profile (Figure 26A), there was significant rescue of its mRNA expression after the contemporary silencing of TDP-43 with DAZAP1 using siRNA (Figure 26D).
Regarding BRD8, little information is available in the literature related to its cellular function. However, this protein is known to be a member of bromodomain (BRD)-containing proteins required for genome maintenance, since its depletion was demonstrated to trigger p53-mediated apoptosis (Lashgari et al., 2018). Interestingly, although BRD8 was identified as a TDP-43 target gene in HEK293 cells (De Conti et al., 2015), in the neuroblastoma cell line the silencing of TDP-43 was not able to induce an effect on BRD8 expression levels, suggesting a cell-specific effect for this regulation (Figure 26B). Nevertheless, we could observe a slight and not significant overexpression in siDAZAP1 treated cells (Figure 26B).

Taken together these results suggested a strong correlation between TDP-43 function and hnRNPs, especially with regards to DAZAP1 and to hnRNP Q but not for hnRNP R. These data therefore represented the starting point of my further investigations.

First of all, to complete the analysis of the effects of DAZAP1, hnRNP Q and hnRNP R on TDP-43, we assessed whether these could cause changes in TDP-43 protein levels after treatment with siRNAs against each individual hnRNP and vice-versa (Figure 22B-E). This was the first item that had to be tested because it has been shown that appropriate expression of TDP-43 is critical for carrying out its cellular functions. In fact, overexpression of TDP-43 was found to be toxic to neurons (Swarup et al., 2011; Li, Ray, Rao, Shi, Guo, Chen and Woodruff, 2010), and also its complete depletion to be lethal for mouse development (Sephton et al., 2010) and Drosophila survival (Feiguin et al., 2009). Similarly, the alteration of DAZAP1, hnRNP Q and hnRNP R expression could be linked to tumorigenesis (Lai et al., 2017; Prima and Hunger, 2007). In my experiments, however, I saw that DAZAP1, hnRNP Q and hnRNP R were not able to change TDP-43 protein levels and the expression of each hnRNPs was not influenced after siRNA depletion of TDP-43.
We also looked at the TDP-43 subcellular localization in SH-SY5Y cells following sihnRNP treatment (Figure 27), since one of the most important features of TDP-43 is the ability to shuttle between nucleus and cytoplasm (Ayala et al., 2008). In fact, although most of the protein is physiologically present in nucleus, approximately 30% of TDP-43 is normally found to be localized in cytosol where it participates in regulating mRNA stability, export and localization of newly transcripts (Lagier-Tourenne et al., 2010). However, modification of TDP-43 homeostasis can lead to re-distribution within cells, as described in neurodegeneration processes in which TDP-43 undergoes cytoplasmic accumulation and consequent aggregation (Neumann et al., 2006). In our experiments, however, TDP-43 localization was not disrupted by depletion of DAZAP1, hnRNP Q and hnRNP R.

Therefore, the results achieved so far indicated that the variations in splicing profile and gene expression levels of TDP-43 targets were not dependent on aberrant modifications of basic TDP-43 features, such as protein expression and subcellular localization. In support of these results, I also examined hnRNP effects using a more realistic scenario of TDP-43 pathology and we analysed the capability of endogenous TDP-43 to aggregate in the presence of FLAG-TDP43-12XQ/N aggregates (Figure 28). In this LOF model, the endogenous TDP-43 can fully interact with these aggregates and undergo sequestration into nuclear and cytoplasmic insoluble aggregates (Budini et al., 2014). Basically, we concluded that in cells treated with siRNA against hnRNPs the aggregation propency of TDP-43 was not affected. Furthermore, we were able to confirm our previous results about the rescue of MADD and STAG2 splicing profile, following siDAZAP1 and sihnRNP Q treatment.

Since DAZAP1 was the most consistent modifier of TDP-43 activity and no previous literature has undercovered a possible connection between these two proteins, our interest
then moved into the characterization of the relationship between TDP-43 and DAZAP1. After ascertaining that DAZAP1 was not directly binding to TDP-43 I saw that it could bind all its mRNA targets (Figure 29). Therefore, we decided to identify all the possible targets commonly regulated by knockdown of both proteins through RNA-seq analysis. We found that the depletion of TDP-43 and DAZAP1 prevalently determined an upregulation of genes with respect to the number of downregulated genes (Figure 30A) and the amount of DEGs grouped in the same categories of GO analysis (Figure 33A) were similar in both conditions, indicating therefore the possibility that both proteins could influence the same cellular pathways with an analogous mechanism.

Moreover, these pathways mostly affected genes related to neuronal and inflammation pathways (Figure 30D and Figure 33B), confirming the importance of hnRNPs in the regulation of RNA metabolism in SH-SY5Y cells. In fact, taking a closer look at the regulated genes chosen for qPCR validation and belonging to the top 100 DEGs, we found nine interesting genes associated to brain functions and inflammatory response: \textit{ACHE}, \textit{STX3}, \textit{ELAVL3}, \textit{NOVA2}, \textit{CELF5}, \textit{TNF}, \textit{TNFRSF9}, \textit{YPEL4} and \textit{ICAM1} (Figure 30B-C). Notably, these genes were found to be upregulated (\textit{ACHE}, \textit{STX3}, \textit{CELF5}, \textit{TNF}, \textit{TNFRSF9}, \textit{ICAM1}, \textit{YPEL4}) and downregulated (\textit{ELAVL3}, \textit{NOVA2}) in the same direction in both siTDP-43 and siDAZAP1 treated cells, supporting again the idea of a common mechanism of action of these two hnRNPs (Figure 30B).

In this respect, it is therefore worthwhile to better describe the role of these factors in the nervous system:

\textit{ELAVL3} (also known as HuC) is a member of the Elav-like family and one of the mammalian homologous of \textit{Drosophila} Elav (Ogawa et al., 2018). The ELAV-like family is composed of four members: three specifically expressed in neurons known as neuronal \textit{ELAVL} (n\textit{ELAVL}: \textit{ELAVL2}, \textit{ELAVL3} and \textit{ELAVL4}) and one ubiquitously expressed
(ELAVL1). nELAVL proteins are RNA-binding proteins involved in RNA metabolism, promoting neuronal differentiation. They are also implicated in neurodegenerative diseases, including Alzheimer’s and Parkinson’s disease (Ogawa et al., 2018). ELAVL3 has been demonstrated to regulate proteins implicated in amino acid biosynthesis, and it plays a role in mRNA stability and in alternative splicing of genes important for synaptic cytoskeletal dynamics (Ince-Dunn et al., 2012). In particular, ELAVL3 affects the alternative splicing of neuronal pre-mRNAs acting with the same dependence on binding position observed for other RBPs such as TDP-43 (Tollervey et al., 2011), PTB (Llorian et al., 2010) and NOVA (Ule et al., 2006). Moreover, ELAVL3 was found to control glutamate levels, a major neurotransmitter in brain (Ince-Dunn et al., 2012).

NOVA2 is a member of the neuronal RBP collectively known as the NOVA family. The major components of this family are NOVA1 and NOVA2. Compared to NOVA1, which is prevalently expressed in hindbrain and ventral spinal cord, NOVA2 was found in neocortex. The NOVA family was also described to be involved in Paraneoplastic Neurodegeneration (PND). Regarding NOVA2, it has been found to modulate ~ 7% of brain-specific splicing events in neocortex (Ule et al., 2005) and successively it has been demonstrated that it is important in the alternative splicing of genes involved in axon guidance (Leggere et al., 2016; Saito et al., 2016).

Acetylcholinesterase (AChE) is a hydrolase encoded by ACHE gene that is required for the regulation of cholinergic signalling (Soreq and Seidman, 2001). In the brain, two isoforms are expressed: a soluble and predominant variant called AChE-S (synaptic AChE) and an inducible variant called AChE-R (read-through) (Soreq and Seidman, 2001). While AChE-S can exist in both soluble and membrane-bound forms, AChE-R was found only in a soluble form (Perrier et al., 2002). However, the expression of both those variants occurs in cholinergic neurons (Johnston et al., 1979), astrocytes (Sáez-Valero et al., 2003) and
cholinoceptive neurons (Aubert et al., 1994). Several studies have reported an aberrant production of AChE during the pathogenesis of AD (Navaratnam et al., 1991; Fishman et al., 1986; Whitehouse et al., 1982). In particular, it has been demonstrated that AChE could promote the formation of amyloid-β (Aβ) fibrils binding Aβ protein (Alvarez et al., 1997) and determining neurotoxicity (Muñoz and Inestrosa, 1999). Indeed, increased levels of AChE were reported to speed up Aβ plaque formation in brains of transgenic mice suffering from neurodegeneration (Rees et al., 2003). Moreover, N-glycosylated AChE was found to be overexpressed in the cerebrospinal fluid of patients with AD (Sáez-Valero et al., 1997). Recently, it has been demonstrated that increased levels of N-glycosylated AChE can disrupt the interaction between O-glycosylated β-neurexins and neuroligin-1 determining loss of glutamatergic synapses and accumulation of post-synaptic proteins (Xiang et al., 2014).

Sintaxin 3 (STX3) is a member of SNARE (Soluble N-ethylmaleimide Sensitive Factor Attachment Protein Receptor) machinery that in turn is important for the vesicle trafficking in polarized epithelial cells (Jahn and Scheller, 2006). In particular, STX3 is located to apical plasma membranes (Soo Hoo et al., 2016) where it is involved in fusion post-Golgi vesicles with the membrane (Sharma et al., 2006). In 2006, Darios and colleagues demonstrated that the omega-3 and omega-6 polyunsaturated fatty acids (PUFAs) could change the STX3 structure triggering its binding to the syntaxin partner SNAP25 (Synaptosomal-associated Protein of 25 kDa) and therefore strongly stimulating the neurite outgrowth (Darios and Davletov, 2006). In a previous work, our group established an indirect connection between TDP-43 depletion and STX3. In particular, the downregulation of miRNA Let-b induced by loss of TDP-43 could alter the expression of let-7b targets, including STX3 (Buratti et al., 2010).
CELF5 belongs to the Bruno-like (Brunol) proteins and it is one of the mammalian homologues of *Drosophila* Bruno (Good et al., 2000). Successively, the Bruno-like family were collectively referred as CUGBP, Elav-like family (CELF) because of the first two members identified in humans: CUG binding protein (CUG-BP, now called CELF1) and Elav-type RNA Binding Protein 3 (ETR-3, now called CELF2) (Ladd et al., 2001). Overall, this family is composed of six proteins, namely CELF1, CELF2, CELF3, CELF4, CELF5 and CELF6 that are widely expressed in brain even if CELF1 and CELF2 are found in many tissues (Gallo and Spickett, 2010). To date there is little information regarding the function of CELF proteins. However, they seem to regulate the alternative splicing of transcripts expressed in brain and CELF2, CELF3 and CELF5 could also play a role in neuronal development (Ladd, 2013). Regarding CELF5, it was found expressed all over in the brain, albeit low-expression was detected in medulla oblongata and spinal cord (Ladd, 2013). CELF5 was also linked to Myotonic Dystrophy of Type 1 (DM1), since it was observed to modulate the alternative splicing of Tau exon 6 (Leroy et al., 2006).

YPEL4 belongs to the Yippe-like (YPEL) family and contains a Yippe-like domain that is putatively identified as a zing-finger-like, metal-binding domain (Roxström-Lindquist and Faye, 2001). This protein family takes its name from the *Drosophila* counterpart namely *Yippe*, which shares the 76% of similarity with the human cDNA clone (Roxström-Lindquist and Faye, 2001). In human there are five YPEL proteins: YPEL1, YPEL2, YPEL3, YPEL4 and YPEL5 whose functions still remain to be elucidated (Hosono et al., 2004). However, it is known that YPEL4 is expressed in all human tissues and its functions seem to be connected with ETS Domain-containing Protein Elk-1 (Elk-1) /MAPK signalling pathways (Liang et al., 2010) and cell division (Hosono et al., 2004). Moreover, it was identified as a putative gene involved in Schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).
Finally, we also included in our qPCR validation \textit{TNF}, \textit{ICAM1} and \textit{TNFRSF9} genes.

\textit{TNF-\(\alpha\)} is encoded by the \textit{TNF} gene and it is a pro-inflammatory cytokine implicated in the regulation of the inflammatory response in body and brain (Neniskyte et al., 2014). In healthy brains, it is expressed at low levels but during injuries or in the presence of ischemia and neurodegenerative disorders the levels of \textit{TNF-\(\alpha\)} increased promoting neuronal inflammation (Neniskyte et al., 2014; Santello and Volterra, 2012).

\textit{TNF-\(\alpha\)} is also important for release of the neurotransmitter norepinephrine (NE) in the central nervous system (Ignatowski et al., 1997), as well as it contributes to the synapse plasticity (Stellwagen and Malenka, 2006).

\textit{TNFRSF9} (also known as CD137) is a member of the Tumor Necrosis Factor Receptor (TNFR) Superfamily and is found in wide range of tissues, including brain (Reali et al., 2003). In particular, in central nervous system (CNS) is expressed on plasma membrane and cytoplasm of neurons and astrocytes and its upregulation was described to be mediated by Fibroblastic Growth Factor 2 (FGF2) (Reali et al., 2003) and associated with melanoma brain metastasis (Mittelbronn et al., 2014). The reserve signalling between CD137 and its ligand CD137L was reported to inhibit cell survival in mouse neural stem cells (Yun et al., 2013).

\textit{ICAM-1} is an adhesion molecules encoded by \textit{ICAM1} gene and produced by endothelial cells. It is important for mediating the transfer of neutrophil into the brain during brain injuries (Hess et al., 1994). Interestingly, it was described to be upregulated in age-dependent neurodegeneration (Miguel-Hidalgo et al., 2007) and localized in amyloid plaques of Alzheimer’s patients (Verbeek et al., 1994).

\textit{TNF}, \textit{ICAM1} and \textit{TNFRSF9} genes are very interesting because they are involved inflammation pathways. Indeed, the activation of the immune system is commonly accepted as one of the central mechanisms underlying neurodegeneration, and capable to
modulate it in a beneficial or detrimental manner. Different events can trigger neurodegenerative pathways, including the deposition of misfolded proteins in CNS, the release of toxic components associated with neurons or synapses destruction and the dysregulation of immune system (Glass et al., 2010; Block and Hong, 2005; Wyss-Coray and Mucke, 2002). For example, in ALS pathology, infiltration of leucocytes (mostly T cells) and deposition of IgG were found in motor neurons subjected to injuries (Holmøy, 2008). T cells infiltration could results in neuronal damage or neuronal protective action. Indeed it has been found that they could lead to cells death through release of cytokine and TNF-α (Neumann et al., 2002), or promote neurogenesis by secretion of Brain Derived Neurotrophic Factor (BDNF) supporting motor neuron survival in mice (Serpe et al., 2005). Moreover, increasing levels of several cytokines were also observed in the cerebrospinal fluid of ALS patients with advanced disease, including Transforming Growth Factor Beta (TGF-β) (McGeer and McGeer, 2002).

Overall, my results have demonstrated the potential impact of not well-characterized hnRNPs, in particular DAZAP1 in brain functions and I have linked this activity to genes affected by TDP-43 misregulation. Moreover, my experiments have strengthened the importance of inflammation during neurodegeneration, since many of the DEGs identified following siTDP-43 and siDAZAP1 are related to inflammatory response.

5.2 hnRNP Q and hnRNP R have evolved to perform different functions within cells

The second part of my PhD thesis was focused on the establishment of a functional diversity between to closely related hnRNPs, namely hnRNP Q and hnRNP R in mammalian cells. As described in the paragraph above, these two proteins were able to modulate in a different manner the splicing profile and mRNA expression of TDP-43
controlled genes. The peculiarity of hnRNP Q and hnRNP R is the extreme similarity in their protein sequences, reflecting the fact that in lower organisms, such as *Drosophila*, they are present as one single protein (McDermott et al., 2012). Here, we demonstrated that in SH-SY5Y cells, hnRNP Q and hnRNP R are expressed in different splicing isoforms and also that they show different cell distribution (Figure 35-36). Regarding this issue, I found that the majority of hnRNP Q protein is located in the cytoplasm (Figure 35) and this was confirmed by the expression of the Q1 counterpart in the cytosol fractions (Figure 36A). On the other hand, I identified hnRNP R as mainly expressed in the nucleus (Figure 35), as also confirmed in the nuclear cytoplasm analysis of SH-SY5Y cells (Figure 36B). Moreover, the cellular localization of both proteins was not altered by the downregulation of each other or after cell differentiation into more neuronal-like cells (Figure 37-38).

These findings indicate that in mammals hnRNP Q and hnRNP R have differentiated in two distinctive proteins in order to carry out specific functions within cells. In fact, following an RNA-seq analysis of the whole transcriptome status of cells after siRNA depletion of hnRNP Q and hnRNP R, we were able to observe several common but also many unique features associated with the depletion of both these proteins (Figure 43). Regarding hnRNP Q depleted cells, the most enriched in GO terms were related to membrane architecture, while for hnRNP R depleted cells, the most significant GO terms were connected to tissue and system development, as well as cellular signalling. Taken together these data indicate the functionally diversification of this two proteins during evolution. Interestingly, although the amount of DEGs derived from hnRNP Q silencing was similar between up- and downregulated genes (Figure 39A), in cells silenced with siRNA against hnRNP R we observed a prevalently overexpression of target genes (Figure 40A).
Nonetheless, in both the conditions we identified several genes related to neural and inflammatory pathways (Figure 39-40). In particular, our study on hnRNP Q depletion in SH-SY5Y cells suggests a potential role of this factor in the regulation of three immune-related categories of KEGG pathway analysis namely "Rheumatoid arthritis", "Toxoplasmosis" and "Toll-like receptor signalling pathway" that were enriched after its depletion (Figure 44A). Taking a closer look at these three pathways we found that the infection of the parasite Toxoplasma gondii can affect different neuronal cells, including astrocytes and neurons, and trigger a chronic inflammation due to the contemporary action of host immunity and the evasion of the immune response by the parasite (Carruthers and Suzuki, 2007). Interestingly, an epidemiological study indicated that pharmacological therapy with TNF blocking agents used for the treatment of rheumatoid arthritis could be protective against AD (Chou et al., 2016). Finally, the activation of Toll-like receptor (TLR) signalling has been suggested to contribute to neuronal injury in several neurodegenerative disorders, including ALS (Casula et al., 2011), AD (Reed-Geaghan et al., 2009) and multiple sclerosis (Marta et al., 2008; Prinz et al., 2006).

The same KEGG pathway analysis was repeated for hnRNP R DEGs and the results obtained showed not only its engagement in the regulation of inflammatory response but also the participation in brain functions (Figure 44B). Indeed, three main pathways have been identified by this analysis: "Cell adhesion molecules (CAMs)", "Neuroactive ligand-receptor interaction" and ECM-receptor interaction". Intriguingly, in mature neurons, neural CAMs participates to the control of synaptic formation, structure, plasticity, and release of neurotransmitters (Leshchyns’Ka and Sytnyk, 2016). Furthermore, the expression and function of these molecules was found to be changed in neurodegenerative disorders, such as AD (Leshchyns’Ka and Sytnyk, 2016). Other factors involved in the organization of brain architecture are the extracellular matrix (ECM) molecules (Barros et
In fact, ECM molecules are important for preserving the neural stem cell behavior, for providing a scaffold for neuronal migration and for supporting axonal growth and myelination (Barros et al., 2011). Aberrant protein levels of ECM molecules were also described in Parkinson’s disease and ALS (He et al., 2013; Lim et al., 1996).

For both hnRNP Q (Figure 39B-C) and hnRNP R (Figure 40B-C), the qPCR validation of DEGs was carried out through the selection of genes that are ranked in the top 100 of most enriched genes and that correlated with the immune response and brain functions.

Taking a closer look at genes regulated by hnRNP Q depletion, we analysed overexpressed genes, such as PENK, NRG3, KLF4, KLHL4, TNF, ICAM1 and TNFRSF9, as well as downregulated genes, such as RAB26 and ARHGAP36. In particular, proenkephalin A (encoded by PENK gene) is the opioid peptide precursor of enkephalin peptides. The proteolytic cleavage of this precursor is able to generate neuropeptides, including peptide F, peptide E, Met-enkephalin, Leu-enkephalin, BAM20 and BAM22, that seem to be involved in nociception (Lembo et al., 2002; Höllt et al., 1982). Proenkephalin A is expressed in the nervous system and in other tissues such as heart, kidney, intestine and skeletal muscle (Denning et al., 2008) and it can be also released from different immune cells (Plotnikoff et al., 1997). Moreover, it has been found that the expression of proenkephalin A could be altered in patients suffering from dementia or acute neuroinflammatory disorders (Ernst et al., 2010).

Neuregulin 3 (encoded by NRG3 gene) is the neural-enriched member of the Epidermal Growth Factor (EGF) family proteins and it is structurally related to neuregulin 1 (Zhang et al., 1997). In humans, the interaction between neuregulin 3 and the tyrosine kinase Erb4 was associated with synaptogenesis (Müller et al., 2018) and oligodendrocyte survival (Carteron, 2006). In rat, this interaction was also described to promote the development of adult forebrain through the migration and differentiation of precursor progenitors (Anton et
Interestingly, several polymorphisms in the *NRG3* genes were described to be a risk factor for neuropsychiatric disorders (Chen et al., 2009).

Klf4 (encoded by *KLF4* gene) is a transcription factor containing a zing-finger motif that is involved in several cellular functions, including proliferation and differentiation. In particular, together with Oct4, c-Myc and Sox2 is one of the induced pluripotent stem (iPS) genes able to induce cell pluripotency (Takahashi and Yamanaka, 2006), as well as being important for maintaining self-renewal of embryonic stem cells (ESCs) (Hall et al., 2009). Moreover, Klf4 was reported to be a transcription repressor of axon growth of central nervous system neurons (Moore et al., 2009) and the regulation of its production was demonstrated to be critical for the proper formation of mouse cerebral cortex (Qin and Zhang, 2012).

Rab26 (encoded by *RAB26* gene) is a small GTPase belonging to the Rab family. This large Rab family is implicated in the control of vesicle trafficking from their formation to membrane fusion. Although the function of Rab26 is still not fully elucidated, it was reported to localize with presynaptic membrane vesicles in fly neuromuscular junctions and recruit two components of pre-autophagosomal machinery, namely Atg16L1 and Rab33B (Binotti et al., 2015). This observation is particularly intriguing because it suggests a role of hnRNP Q in the recycling of synapse vesicles via autophagy system, thus providing a further link with TDP-43, since the fly orthologous of TDP-43 was demonstrated to be important for the formation of NMJs. Indeed, the generation of TBPH-null allele flies (Feiguin et al., 2009; Godena et al., 2011; Langellotti et al., 2016; Romano et al., 2018) and hTDP43A315T transgenic mouse (Magrané et al., 2013) dramatically induced morphological modifications at synaptic terminals and a strong reduction of synaptic vesicles in the NMJs, respectively.

Furthermore, neuronal expression was also described for both KLHL4 and ARHGAP36...
even if possible neuronal implications, including relationship with neurological disorders, are still unknown (Rack et al., 2014; Braybrook et al., 2001).

On the other hand, looking at DEGs obtained from silencing of hnRNP R we validated six overexpressed genes, including *CARTPT, FOSB, JAG1, DUOX1, HMOX1* and *ICAM5* and four downregulated genes, such as *KCNAB1, SDCBP2, EFEMP1* and *APC5*. Notably, the *CARTPT* gene is responsible for the production of CART mRNA whose expression was found to be increased in the rat brain after stimulation with psychomotor agents (Douglass et al., 1995). The resulting CART neuropeptides was successively described to localize with gamma-aminobutyric acid (GABA) (Smith et al., 1999) and substance P (Hubert and Kuhar, 2005), indicating that these species could act as a neurotransmitter.

The *FOSB* gene encodes both FosB and its alternative splicing variant ΔFosB. Both proteins belong to the Fos family of transcription factors and their production can be promoted especially in the nucleus accumbens and dorsal striatum following administration of many drugs (Nestler et al., 2001). Expression of Fos proteins was also detected in dentate gyrus of hippocampus and cerebral cortex (Ohnishi et al., 2011). The activated form of Fos proteins is achieved by the heterodimerization with Jun proteins resulting in the generation of the Active Activator Protein-1 (AP-1) complex that recognizes AP-1 sites in various promoters (Chinenov and Kerppola, 2001). Fosb expression was also discovered to be linked to neurological disorders, since Fosb-null mice highlighted the adult-onset of spontaneous epilepsy (Yutsudo et al., 2013) and exhibited depression-like behaviors (Ohnishi et al., 2011), as well as to neuroinflammation through the regulation of the receptors of the anaphylatoxin C5a: C5AR1 and/or C5AR2 (Nomaru et al., 2014).

The *DUOX1* gene encodes the maturation factor of DUOX1 (Dual Oxidase 1) that is a member of the NADPH oxidase family (Bedard and Krause, 2007). It has been reported
Discussion

that NADPH oxidase provides the major source of reactive oxygen species (ROS) essential for inducing neuronal differentiation, neuronal signalling and neurite formation (Munnamalai and Suter, 2009; Tejada-Simon et al., 2005; Knapp and Klann, 2002). In mouse embryonal carcinoma p19 cells, it has been also demonstrated that their neuronal-like differentiation resulted from the interaction between p53 and DUOXA1 (Ostrakhovitch and Semenikhin, 2011). Moreover, an excessive production of ROS by NADH oxidase can trigger neuroinflammation during CNS pathology, as described for Alzheimer’s and Parkinson’s disease (Kim et al., 2015).

Jagged1 (encoded by the JAG1 gene) is one of five ligands that participate to the Notch signalling pathway (Grochowski et al., 2016). In mammals, it is widely expressed during development and it is also present in postnatal brain (Irvin et al., 2004; Stump et al., 2002; Jones et al., 2000). In mouse, Jagged1/Notch signalling was found to control oligodendrocyte differentiation and myelination (Wang et al., 1998). Moreover, Jagged1 is essential for the maintenance of neuronal stem cell niche in adult brain (Nyfeler et al., 2005), as well as being important for synaptic signalling (Alberi et al., 2011). Jagged1/Notch signalling is also involved in spatial memory. Indeed, in hippocampus of patients suffering from Alzheimer’s disease their expression was found to be altered (Marathe et al., 2017).

Heme oxygenase 1, also known as HO-1 (encoded by the HMOX1 gene) is the first of three members belonging to the HO family. The production of HO-1 is regulated by stress stimuli, such as UVA irradiation, sodium arsenite and hydrogen peroxide (Keyse and Tyrrell, 1989). Neuronal protection against glutamate-induced oxidative cell death was observed after overexpression of HO-1 (Chen et al., 2001). Notably, upregulation of the HMOX1 gene was also found in neurons and astrocytes derived from hippocampus, cerebral cortex, and subcortical white matter of Alzheimer’s patients, supporting the role of
Discussion

chronic oxidative stress in the establishment of neuropathological disorders (Schipper et al., 1995).

*KCNAB1* (also known as *Kβv1* gene) encodes the subunit β of the shaker-related voltage-gated potassium channel Kv1 (Leicher et al., 1996). Three different splicing isoforms are present for this gene: *Kβv1.1, Kβv1.2* and *Kβv1.3* (Rettig et al., 1994). *Kβv1.2* and *Kβv1.3* are preferentially transcribed in heart, while *Kβv1.1* is expressed in brain, especially in hippocampus and striatum (Rettig et al., 1994). This particular potassium channel is able to inactivate the subunit α, that is the second subunit that constitutes the Kv1 channel. Potassium channels are extremely important for transducing signals in nervous system (Kuang, 2015) and *Kβv1.1* knockout mice showed normal synaptic plasticity but impairment of learning skills (Giese et al., 1998). Furthermore, a specific mutation in *KCNAB1* gene was also found to be the causative agent of epilepsy and developmental disabilities in a Chinese children (Zhang et al., 2015).

Syntenin-2 (encoded by *SDCBP2*) is a novel PDZ (Postsynaptic density protein, Discs-large, Zona occludens) protein able to bind phosphatidylinositol 4,5-bisphosphate (PIP2) (Geeraerts et al., 2013). This protein is encoded by two alternative splicing isoforms, namely 2α and 2β, and it shares ~70% with the most famous syntenin-1 (Koroll et al., 2001), that is a scaffold protein involved in the regulation of the neural membrane architecture (Hirbec et al., 2005). The available literature has not fully elucidated the cellular function of syntenin-2 in depth. However, syntenin-2 isoform 2α was found to interact with the neural cell-adhesion molecules, namely neurexin-1, suggesting a potential role in brain dynamics (Koroll et al., 2001).

Fibulin-3 (encoded by *EFEMP1*) is a calcium-binding ECM protein belonging to the Fibulin protein family (Timpl et al., 2003) consisting of Fibulin-1, Fibulin-2, Fibulin-3, Fibulin-4 and Fibulin-5. These proteins contain three domains, where the central domain
III contains nine epidermal growth factor (EGF)-like motifs able to bind calcium (Timpl et al., 2003). For most fibulin proteins, the interaction with calcium is essential for binding other ECM proteins and cellular receptors, allowing them to carry out their cellular function in joining supramolecular structures such as basement-membrane networks, elastic fibres, several types of microfibrils and proteoglycan aggregates (Timpl et al., 2003). The function of fibulin-3 is still unknown. However, fibulin-3 expression was strongly detected only in the primary visual cortex (V1C) region of prenatal brain (Ayana et al., 2018) while it was found weakly expressed in adult normal brain but upregulated in gliomas where it promotes cell tumor growth and invasion through the regulation of Notch signalling (Hu et al., 2009, 2012). Moreover, a single mutation in EFEMP1 was described to be associated with age-related macular degeneration (Stone et al., 1999).

Finally, we noticed that the silencing of hnRNP Q could lead to the overexpression of three genes related to inflammatory pathways that we previously illustrated to be upregulated in response to siRNA depletion of both TDP-43 and DAZAP1: ICAMI, TNF and TNFRSF9. This finding corroborated our previous results, regarding the ability of DAZAP1 and hnRNP Q to modulate TDP-43 activity in SH-SY5Y cells.

By contrast, the silencing of hnRNP R led to the overexpression of other two immune-related proteins (ICAM5 and ACP5) without inducing a significant alteration of the same inflammatory genes controlled by TDP-43, DAZAP1 and hnRNP Q.

Overall, this second part of the work allowed a better characterization of the functional diversification of hnRNP Q and hnRNP R, highlighting their importance in controlling genes associated with immune system and neuronal development, differentiation and cell signalling. Most importantly, I was able to relate the majority of them to their aberrant regulation during neurodegenerative diseases.
5.3 TDP-43, DAZAP1 and hnRNP Q potentially regulate same transcripts important for ALS pathogenesis

Over the last decades, perturbations of RBP homeostasis have increasingly acquired significance in the pathogenesis of ALS. Especially TDP-43 was the first hnRNP proteins associated with ALS, since in 2006 it was discovered as a component of ubiquitinated cytoplasmic inclusions in brain of patients (Neumann et al., 2006). Subsequent studies have also identified ALS-causative mutations in the TDP-43 encoding gene, and in other members of hnRNP family, including FUS, hnRNP A1, hnRNP A2/B1, TIA-1 and matrin-3 (Purice and Taylor, 2018). Therefore, the identification of all possible targets regulated by TDP-43 and other hnRNP proteins acting on the same biological pathway could lead to the understanding of the molecular mechanisms behind the onset and progression of neurodegeneration. Thanks to sequencing technologies it has been possible to characterize gene expression changes or splicing variations associated with RBP misregulation (Buratti et al., 2013).

In this study DAZAP1 and hnRNP Q were found to be the most promising tested hnRNPs able to modulate TDP-43 activity. Interestingly, the combination of RNA-seq results obtained after depletion of TDP-43, DAZAP1 and hnRNP Q have highlighted the presence of brain /ALS-related genes commonly affected by these proteins (Figure 45), contributing to expand the list of genes that can potentially affect neurodegeneration.

Table 23 presents some promising genes and summarizes the knowledge already available in literature regarding their relationship with brain function and/or ALS pathology.
Table 23. Genes commonly regulated by depletion of TDP-43, DAZAP1 and hnRNP Q in SH-SY5Y cells which showed an involvement in brain disorders and/or ALS pathology.

<table>
<thead>
<tr>
<th>Gene Symbol/(Description)</th>
<th>Brain disorders /ALS pathology</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IGF2</strong> (<em>Insuline-like Growth Factor 2</em>)</td>
<td>IGF-2 could protect against ALS degeneration of human motor neurons derived from iPSCs through phosphorylation of Akt and glycogen synthase kinase-3β and upregulation of β-catenin.</td>
<td>(Allodi et al., 2016)</td>
</tr>
<tr>
<td><strong>SYT14</strong> (<em>Synaptotagmin 14</em>)</td>
<td>SYT14 mutations are causative of autosomal-recessive spinocerebellar ataxia-11 (SCAR11). t(1;3)(q32.1;q25.1) leads to disruption of SYT14 and it is associated with developmental abnormalities.</td>
<td>(Doi et al., 2011) (Quintero-Rivera et al., 2007)</td>
</tr>
<tr>
<td><strong>ABAT</strong> (<em>4-Aminobutyrate Aminotransferase</em>)</td>
<td>Strong expression in neuronal tissue; responsible for catabolism of GABA (neurotransmitter) into succinic semialdehyde. ABAT plays an essential role in the conversion of dNDPs to dNTPs in the mitochondrial nucleoside salvage pathway and causes a neurometabolic disorder that includes encephalomyopathic MDS. Identification of ABAT as putative gene in autism following genotype-based association analysis. ABAT was found to be associated with neurofibrillar tangles (NFTs) following genome-wide analysis of 5-hydroxymethylcytosine regions in AD brain.</td>
<td>(Besse et al., 2015) (Barnby et al., 2005) (Zhao et al., 2017)</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>References</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>ABCC9 (ATP Binding Cassette Subfamily C Member 9)</strong></td>
<td><em>ABCC9/SUR2</em> gene encoded an ATP-sensitive potassium channel that is strongly implicated in human aged-related neurological disorders, including hippocampal sclerosis of aging (HS-Aging).</td>
<td>(Nelson et al., 2015)</td>
</tr>
<tr>
<td><strong>ADCY1 (Adenylate Cyclase 1)</strong></td>
<td><em>ADCY1</em> gene encoded an adenylyl cyclase specifically expressed in brain; it was found to be upregulated in frontal cortex of ALS patients. Notably, in our results it was found downregulated in all tested lnRNPs.</td>
<td>(Andrés-Benito et al., 2017)</td>
</tr>
<tr>
<td><strong>ADD3 (Adducin 3)</strong></td>
<td><em>ADD3</em> gene encoded γ-adducin subunit of membrane-skeletal protein, namely Adducin. Mutation of ADD3 was found to inhibit the binding of γ-subunit to α-subunit and determinate cerebral palsy. ADD2-knockout mice showed synaptic plasticity, motor coordination and behavioral deficits accompanied by changes in the expression and phosphorylation levels of the α- (ADD1) and γ-adducin (ADD3) subunits. Adducin is highly phosphorylated in spinal cord tissue derived from ALS patients comparing to individuals without neurological disease. Upregulated upon TDP-43 depletion in adult mouse brain.</td>
<td>(Kruer et al., 2013) (Porro et al., 2010) (Hu et al., 2003) (Polymenidou, Lagier-Tourenne, et al., 2011)</td>
</tr>
<tr>
<td><strong>CAPN5 (Calpain 5)</strong></td>
<td>CAPN5 has been shown to be the second most abundantly expressed calpain in the mammalian central nervous system.</td>
<td>(Singh et al., 2014)</td>
</tr>
</tbody>
</table>
### Discussion

<table>
<thead>
<tr>
<th><strong>Gene</strong></th>
<th><strong>Expression and Function</strong></th>
<th><strong>References</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CASP3 (Caspase 3)</strong></td>
<td>CASP3 is an important mediator of neuronal programmed cell death. It is also associated with chronic neurodegenerative diseases. In transgenic mice expressing mutant SOD1, inhibition of caspase-3 was found to have a protective role in ALS pathology.</td>
<td>(Troy et al., 2011; Ribe et al., 2008; Ekshyyan, O. and Aw, 2004; Li et al., 2000)</td>
</tr>
<tr>
<td><strong>CDK5R1 (Cyclin Dependent Kinase 5 Regulatory Subunit 1)</strong></td>
<td>CDK5R1 gene encoded p35, the main activator of Cyclin-dependent kinase 5 (CDK5). It has been associated with AD onset and progression. It was found to be upregulated in frontal cortex of ALS patients. Notably, in our results it was found to be downregulated in all tested hnRNPs.</td>
<td>(Spreafico et al., 2018; Andrés-Benito et al., 2017)</td>
</tr>
<tr>
<td><strong>CTSS (Cathepsin S)</strong></td>
<td>Significant increased expression of Cathepsin S (encoded by CTSS gene) was found in the anterior lumbar spinal cord in ALS.</td>
<td>(Berjaoui et al., 2015)</td>
</tr>
<tr>
<td><strong>TLR7 (Toll Like Receptor 7)</strong></td>
<td>Significant increased expression of TLR7 was found in the anterior lumbar spinal cord in ALS.</td>
<td>(Berjaoui et al., 2015)</td>
</tr>
<tr>
<td><strong>DDX58 (DExD/H-Box Helicase 58)</strong></td>
<td>RNA helicase upregulated in ALS mouse model: TDP-43 (A315T) mice. In particular, TDP-43 bound to its mRNA.</td>
<td>(MacNair et al., 2016)</td>
</tr>
<tr>
<td><strong>HS3ST2 (Heparan Sulfate-Glucosamine 3-Sulfotransferase 2)</strong></td>
<td>HS3ST2 is an heparan sulfate biosynthetic enzyme showed greatest decreases in expression in FTLD-U human samples.</td>
<td>(Chen-Plotkin et al., 2008)</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>RAB31 (RAB31, Member RAS Oncogene Family)</strong></td>
<td>GTP-binding protein upregulated upon TDP-43 depletion in adult mouse brain.</td>
<td>(Polymenidou, Lagier-Tourenne, et al., 2011)</td>
</tr>
<tr>
<td><strong>PPP1R15A (Protein Phosphatase 1 Regulatory Subunit 15A)</strong></td>
<td>PPP1R15A is also known as GADD34. Mutant SOD1 mice that are haploinsufficient for GADD34 have a remarkably ameliorated disease.</td>
<td>(Wang et al., 2014)</td>
</tr>
</tbody>
</table>
6. CONCLUSION AND FUTURE DIRECTIONS

In conclusion, my work has provided novel insights on the involvement of the hnRNP family in controlling neuronal and inflammatory pathways, suggesting that differential expression of these proteins could play an essential role in modulating the onset and progression of neurodegenerative disorders, especially those related to TDP-43 pathology. The future direction of this work could examine the synaptic plasticity/death-signalling pathways after overexpression/depletion of neuronal genes identified by RNA-seq analysis of TDP-43, DAZAP1, hnRNP Q using a cellular model of TDP-43 aggregation and/or using iPSC cells derived from patients carrying TDP-43 mutations. This data will be useful to identify the most promising common neuronal targets identified by RNA-seq analysis of TDP-43, DAZAP1, hnRNP Q that can affect neurodegeneration. This task will also be made easier by performing RNA-seq analysis for other hnRNP proteins that our group has demonstrated can “worsen” TDP-43 gain-of-function toxicity in the Drosophila eye, namely CG30122 (hnRNP U), Bl/CG13425 (hnRNP K) and Sqd/CG16901 (hnRNP D). This last aim is currently underway and I am also actively involved in taking forward this part of the work.
7. BIBLIOGRAPHY


- Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D.,


ALS models’, *Nature Chemical Biology*, vol. 10, no. 8, pp. 677–685 [Online]. DOI: 10.1038/nchembio.1563.


Bibliography


loss of cellular function through aggregation requires additional structural
determinants beyond its C terminal Q/N prion-like domain’, *Human molecular

of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator

splicing regulation, and human disease.’, *Frontiers in bioscience : a journal and

protein-protein and protein-RNA interactions’, *Trends in Biochemical Sciences*,

- Buratti, E., Brindisi, A., Giombi, M., Tisminetzky, S., Ayala, Y. M. and Baralle, F.
Its C-terminal Tail AN IMPORTANT REGION FOR THE INHIBITION OF
CYSTIC FIBROSIS TRANSMEMBRANE’, vol. 280, no. 45, pp. 37572–37584
[Online]. DOI: 10.1074/jbc.M505557200.

- Buratti, E., Conti, L. De, Stuani, C., Romano, M., Baralle, M. and Baralle, F.

‘Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9
skipping’, vol. 20, no. 7.


– D’Ambrogio, A., Buratti, E., Stuani, C., Guarnaccia, C., Romano, M., Ayala,


- Gregory, R. I., Yan, K., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N.


- Hammer, N. A., Hansen, T. v O., Byskov, A. G., Rajpert-De Meyts, E., Grøndahl,


Bibliography


Bibliography


- Lim, G. P., Backstrom, J. R., Cullen, M. J., Miller, C. A., Atkinson, R. D. and


- Magrané, J., Cortez, C., Gan, W., Manfredi, G. and York, N. (2013) ‘Abnormal Mitochondrial Transport and Morphology are Common Pathological Denominators in SOD1 and TDP43 ALS Mouse Models Brain and Mind Research Institute, Weill Medical College of Cornell University, Skirball Institute, Department of Physiology and Neu’, vol. 8174, pp. 1–43.


Bibliography


Prima, V. and Hunger, S. P. (2007) ‘Cooperative transformation by MEF2D/DAZAP1 and DAZAP1/MEF2D fusion proteins generated by the variant


- Quintero-Rivera, F., Chan, A., Donovan, D. J., Gusella, J. F. and Ligon, A. H.
Bibliography


Bibliography


Bibliography


Bibliography

Genovese, G., Georgieva, L., Giegling, I., Giusti-Rodríguez, P., Godard, S.,
Goldstein, J. I., Golimbet, V., Gopal, S., Gratten, J., Haan, L. de, Hammer, C.,
Hamshere, M. L., Hansen, M., Hansen, T., Haroutunian, V., Hartmann, A. M.,
Henskens, F. A., Herms, S., Hirschhorn, J. N., Hoffmann, P., Hofman, A.,
Holtegaard, M. V, Hougaard, D. M., Ikeda, M., Joa, I., Julià, A., Kahn, R. S.,
Kalaydjieva, L., Karachanak-Yankova, S., Karjalainen, J., Kavanagh, D., Keller,
M. C., Kennedy, J. L., Khrunin, A., Kim, Y., Klovins, J., Knowles, J. A., Konte, B.,
Kucinskas, V., Kucinskiene, Z. A., Kuzelova-Ptackova, H., Kähler, A. K., Laurent,
C., Lee, J., Lee, S. H., Legge, S. E., Lerer, B., Li, M., Li, T., Liang, K.-Y.,
Lieberman, J., Limborska, S., Loughland, C. M., Lubinski, J., Lönnqvist, J., Macek,
M., Magnusson, P. K., Maher, B. S., Maier, W., Mallet, J., Marsal, S., Mattheisen,
M., Mattingsdal, M., McCarley, R. W., McDonald, C., McIntosh, A. M., Meier, S.,
Meijer, C. J., Melegh, B., Melle, I., Mesholam-Gately, R. I., Metspalu, A., Michie,
P. T., Milani, L., Milanova, V., Mokrab, Y., Morris, D. W., Mors, O., Murphy, K.
C., Murray, R. M., Myin-Germeys, I., Müller-Myhsok, B., Nelis, M., Nenadic, I.,
Nertney, D. A., Nestadt, G., Nicodemus, K. K., Nikitina-Zake, L., Nisenbaum, L.,
Nordin, A., O’Callaghan, E., O’Dushlaine, C., O’Neill, F. A., Oh, S.-Y., Olincy,
A., Olsen, L., Os, J. Van, Consortium, P. E. I., Pantelis, C., Papadimitriou, G. N.,
Papiol, S., Parkhomenko, E., Pato, M. T., Paunio, T., Pejovic-Milovancevic, M.,
Pulver, A. E., Purcell, S. M., Quested, D., Rasmussen, H. B., Reichenberg, A.,
Reimers, M. A., Richards, A. L., Roffman, J. L., Roussos, P., Rudfer, D. M.,
G., Scolnick, E. M., Scott, R. J., Seidman, L. J., Shi, J., Sigurdsson, E., Silagadze,
T., Silverman, J. M., Sim, K., Slominsky, P., Smoller, J. W., So, H.-C., Spencer, C.
Bibliography


– Troy, C. M., Akpan, N. and Jean, Y. Y. (2011) ‘Regulation of Caspases in the


July 2017).


– Yang, H., Peggie, M., Cohen, P. and Rousseau, S. (2009) ‘Biochemical and Biophysical Research Communications DAZAP1 interacts via its RNA-recognition motifs with the C-termini of other RNA-binding proteins’, Biochemical and


