Multi-omics approach in Down syndrome resolves new regulators in two regions of the human brain

A dissertation submitted in partial satisfaction

of the requirements for the degree of

Doctor of Philosophy in Neuroscience and Brain Technologies

Cycle XXXIII

By

Mohit Rastogi

Supervisor: Dr. Laura Cancedda

University of Genoa

Genoa, Italy

2021

“What you do makes a difference

and you have to decide what kind of difference you want to make”

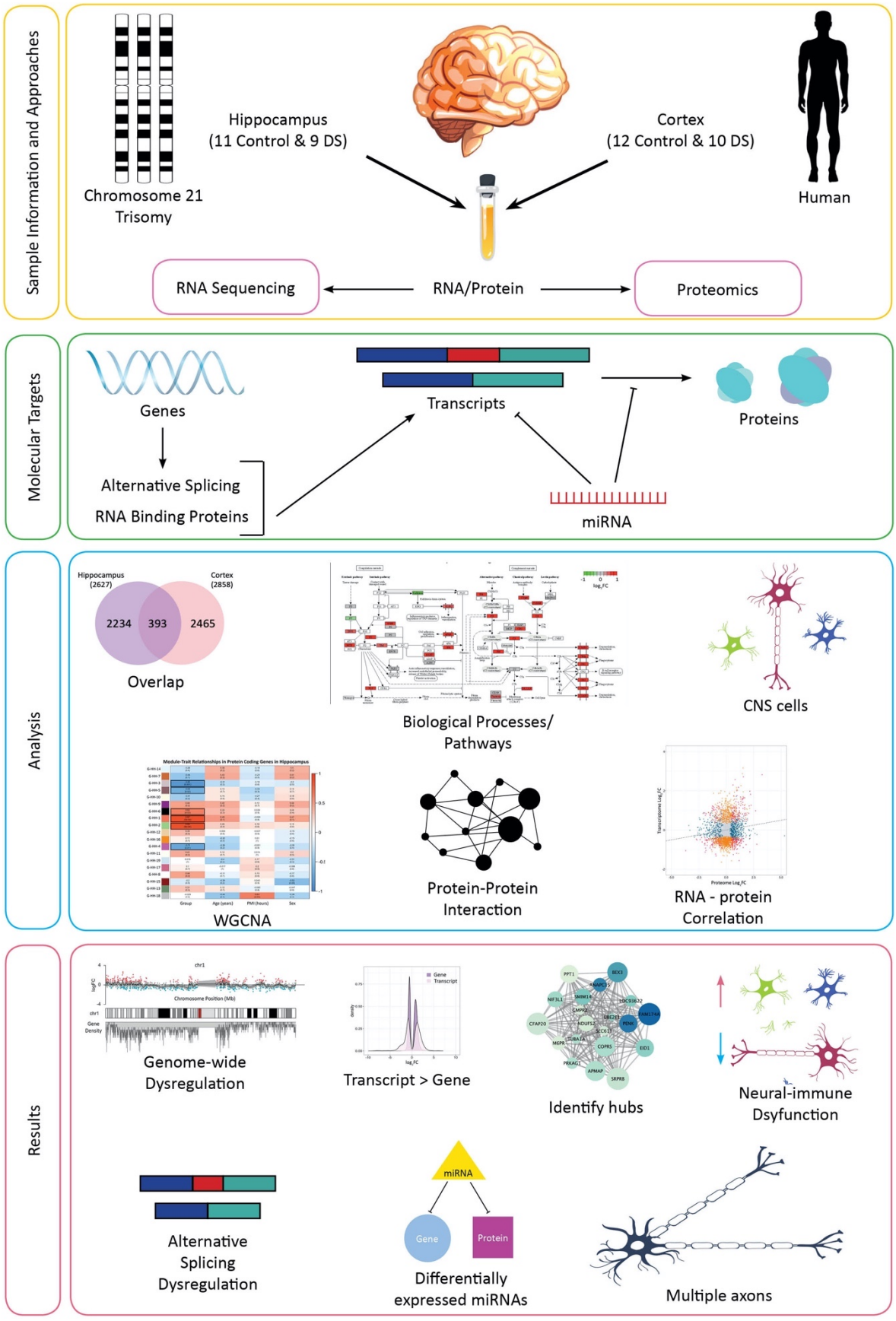
- Jane Goodall

Dedicated to my family, friends, and mentors
for their unwavering belief, support, and teachings.

Abstract

Down syndrome (DS) is the most common genetic form of intellectual disability. DS is a very complex genetic condition and is caused by trisomy of human chromosome 21. Studies mostly on human DS fetuses, plasma, fibroblasts, iPSCs, and mouse models of DS have demonstrated genome-wide dysregulation in genes and proteins involved in brain development and proteostasis. However, these studies were obtained with low-throughput technologies and addressed only either gene or protein level. Thus, it is still unknown how the gene dysregulation relates to protein dysregulation and, in turn, to DS brain deficits. Here, taking advantage of high throughput technologies, we explored in parallel both gene and protein levels on the same postmortem tissue from DS and age/sex-matched control individuals. We have revealed shared patterns of the transcriptome at gene and transcript levels and proteome dysregulation in hippocampus and cortex of DS individuals. We found many non-triplicated genes and proteins differentially expressed along with overexpression in most of the triplicated genes and proteins. We identified many dysregulated biological processes in both brain regions, such as translation, axon development, synaptic signaling, neuron development, and mRNA splicing. We also found differentially expressed genes and proteins involved in extracellular vesicles and cell-substrate junction among all the cellular components identified in our study. In particular, our work highlights downregulation of neuronal genes together with downregulation of neuron-specific pathways such as long term potentiation (LTP) and synaptic vesicle cycle, and upregulation of microglia and astrocytic genes together with the increased response of inflammatory pathways such as complement and coagulation cascades and cytokine-cytokine receptor interaction, as key points in DS in both regions at gene and protein levels. Furthermore, we observed an alteration in RNA splicing in DS brains, which was shared across brain regions and involved many neurodevelopmental genes belonging to axon formation and guidance, dendrite morphogenesis, neurogenesis, and synaptic signaling. We show that genes related to axon formation and one of the RNA binding protein (PTBP2) that regulate these genes are differentially expressed and have differential splicing in DS brains. Interestingly, we collected the first experimental evidence (in murine neurons) of deficit in axonal polarization (multiple axon formation), establishing it as a definite phenotypic defect in DS. Finally, we performed small RNA sequencing on these postmortem human hippocampi and cortex and found several

differentially expressed miRNAs, which provides information about another layer of gene-regulation in DS. Our findings indicate genome-wide dysregulation in adult DS hippocampus and cortex in a comprehensive assessment of both transcriptome and proteome. These data provide a unique and extensive resource for the field of DS. They will likely prove critical in identifying genetic drivers (among triplicated but also non-triplicated genes) to regulate complete biological systems at organelle-, cell-, process- or tissue-level to better understand the impact of trisomy 21 on people with DS, and thus derive future therapeutics.



Study design, methodology overview and results summary

Abbreviations

AD:	Alzheimer's Disease
ADARB1:	Adenosine Deaminase RNA Specific B1
Ago:	Argonaute
ALS:	Amyotrophic Lateral Sclerosis
ALTA:	Alternative Acceptor Site
ALTD:	Alternative Donor Site
Ank3:	Ankyrin 3
APP:	Amyloid Beta Precursor Protein
AS:	Alternative Splicing
ASD:	Autism Spectrum Disorder
BDNF:	Brain Derived Neurotrophic Factor
BP:	Biological Processes
C1Q:	Complement Component 1, Q Subcomponent
CA1:	Cornu Ammonis 1
CC:	Cellular Components
CD14:	Cluster Of Differentiation 14
CD46:	Cluster Of Differentiation 46
CD300A:	Cluster Of Differentiation 300A
CHD:	Congenital Heart Defects
CLIP:	Cross Linking Immunoprecipitation
CNS:	Central Nervous System
CPM:	Counts Per Million
CREB:	Cyclic AMP Response Element Binding
DEmiRNA:	Differentially Expressed miRNA
DGCR8:	DiGeorge Syndrome Critical Region 8
DGE:	Differential Gene Expression
DKC1:	Dyskeratosis Congenita 1
DLPFC:	Dorsolateral Prefrontal Cortex
DNA:	Deoxyribonucleic Acid
DS:	Down syndrome
DSCR1:	Down Syndrome Critical Region Gene 1
DTE:	Differential Transcript Expression
EIF4G3:	Eukaryotic Translation Initiation Factor 4 Gamma 3
EML1:	EMAP Like 1
EWCE:	Expression Weighted Cell-Type Enrichment
FDR:	False Discovery Rate
FC:	Fold Change
FTD:	Frontotemporal Dementia
GAP43:	Growth Associated Protein 43
GEDD:	Gene Expression Dysregulation Domains
G-HC:	Gene Module From Human Cortex
G-HH:	Gene Module From Human Hippocampus
GO:	Gene Ontology
GSEA:	Gene Set Enrichment Analysis
GWAS:	Genome Wide Association Study

HnRNPA1:	Heterogeneous Nuclear Ribonucleoparticle A1
HSA:	Human Chromosome
IFN:	Interferon
IL4R:	Interleukin 4 Receptor
IL-6:	Interleukin 6
iPSC:	Induced Pluripotent Stem Cells
IR:	Intron Retention
KAP:	Keratins Associated Proteins
KEGG:	Kyoto Encyclopedia Of Genes And Genomes
Kb:	Kilobases
LC-MS:	Liquid Chromatography Mass Spectrometry
LIN28A:	Lin-28 Homolog A
LncRNA:	Long Non-Coding RNA
LTD:	Long Term Depression
LTP:	Long Term Potentiation
MAP2:	Microtubule Associated Protein 2
MARK2:	Microtubule Affinity Regulating Kinase 2
Mb:	Megabases
MCP-1:	Monocyte Chemotactic Protein 1
MIC:	Microexon
miRNA:	MicroRNA
MF:	Molecular Functions
MMU:	Mus Musculus
MS:	Multiple Sclerosis
NDD:	Neurodevelopmental Disorders
NF-Kβ:	Nuclear Factor Kappa B Subunit 1
NOVA1:	Neuro-Oncological Ventral Antigen 1
OSMR:	Oncostatin M Receptor
PCA:	Principal Component Analysis
PD:	Parkinson's Disease
P-HC:	Protein Module From Human Cortex
P-HH:	Protein Module From Human Hippocampus
PMI:	Postmortem Interval
Poly(A):	Poly Adenylation
PSI:	Percent Spliced-In Value
PTBP1:	Polypyrimidine Tract Binding Protein 1
PTBP2:	Polypyrimidine Tract Binding Protein 2
qPCR:	Quantitative Polymerase Chain Reaction
RACE:	Rapid Amplification Of cDNA Ends
RBFOX1:	RNA Binding Fox-1 Homolog 1
RBP:	RNA Binding Protein
REEP1:	Receptor Accessory Protein 1
RIN:	RNA Integrity Number
RIP:	RNA Immunoprecipitation
RISC:	RNA Induced Silencing Complex
RNA:	Ribonucleic Acid
RNA-Seq:	Ribonucleic Acid Sequencing

rRNA:	Ribosomal RNA
SAGE:	Serial Analysis Of Gene Expression
SCG5:	Secretogranin V
scRNA:	Single-Cell RNA
SE:	Skipped Exon
SHTN1:	Shootin 1
SMAD1:	SMAD Family Member 1
snRNA:	Small Nuclear RNA
snoRNA:	Small Nucleolar RNA
SOD1:	Superoxide Dismutase 1
SRM:	Selective Reaction Monitoring
SRRM4:	Serine/Arginine Repetitive Matrix 4
TDP43:	TAR DNA Binding Protein 43
TEC:	To Be Experimentally Confirmed
TGF-β:	Transforming Growth Factor Beta
TNF-α:	Tumor Necrosis Factor Alpha
tRNA:	Transfer RNA
UTR:	Untranslated Region
VDAC2:	Voltage Dependent Anion Channel 2
WGCNA:	Weighted Gene Co-Expression Network Analysis
WPCNA:	Weighted Protein Co-Expression Network Analysis
WT:	Wildtype

List of Figures and Tables

Table of Contents

Introduction

1.1. Neurodevelopmental disorders (NDDs):

Neurodevelopmental disorders (NDDs) are diseases where development of brain structure and proper neural circuits are disrupted. Globally, NDDs correspond to >3% of health issues worldwide (Gilissen et al., 2014). NDDs share some common features such as learning and memory abnormalities, disrupted social and emotional aspects, immunological issues, and few comorbidities such as increased susceptibility to seizures (van Bokhoven, 2011). Some examples of NDDs are autism spectrum disorder (ASD), intellectual disability, fragile X syndrome, schizophrenia, and Down syndrome (Abrahams and Geschwind, 2008; van Bokhoven, 2011; Fromer et al., 2014; Geschwind, 2011; Matson and Shoemaker, 2009; Ropers, 2008).

NDDs result from genetic factors (e.g., heterogeneous single-point mutations or polygenic (Niemi et al., 2018)) and environmental factors (e.g., maternal use of alcohol or drugs during pregnancy; preterm birth; exposure to environmental contaminants and socioeconomic status ((Cheroni et al., 2020; Heffernan and Hare, 2018; Tran and Miyake, 2017)). Identifying the causative genes for NDDs is highly important to understand the molecular mechanisms responsible for these disorders' symptomatology. With this respect, the significant challenges are understanding how the genotype relates to the phenotype changes and pinpointing the key players involved in the disruption of physiological neuronal circuit formation.

Despite tremendous technological advances to understand and map the whole brain by high-throughput sequencing technologies, automated microscopy methods, and multi-electrode electrophysiological recordings, the understanding of the formation and function of the neuronal circuits of the brain is far from being completed. In particular, while these technologies have provided a new perspective on molecules and circuits responsible for the brain's functional and structural organization, understanding how dysregulation in the central nervous system (CNS) can lead to various NDDs is still in its infancy. To better understand

genetic disorders complexity, an integrative approach may be needed to assemble information from multiple perspectives. In particular, approaches that generate a high volume of in-depth data can transform the study of neurodevelopmental disorders because the diverse and complex genomic architecture of genetic disorders that begin early in life and through aberrant developmental trajectories have long-lasting consequences in adulthood. The underlying neuropathology, biological mechanisms, and the cells involved in many of the NDDs are largely unknown. A multi-omics approach in adulthood may, in fact, aid in finding the biological processes and pathways that are mostly dysregulated and responsible for NDD overt phenotypes. Then, we might even be able to find out which hub genes or cells or molecular network to target, to be able to regulate the whole system in one go for future therapeutic approaches. Thus, we took this multi-level data acquisition and interpretation approach to integrate transcriptomics and proteomics data from the same set of individuals to better understand a complex and multigenic neurodevelopmental disorder, Down syndrome.

1.2. Down syndrome (DS):

Down syndrome is a neurodevelopmental disorder caused by an extra copy of human chromosome 21 (HSA21). As for all neurodevelopmental disorders, DS is characterized by structural deficiencies in diverse brain regions such as the hippocampus, cerebellum, and cortex. These defects include hippocampal and cortical growth reduction, abnormal lamination, differentiation of neurons, abnormal synaptic plasticity, decreased dendritic branching, and spine density (Kazemi et al., 2016; Rachidi and Lopes, 2011). The reduced brain size and the other brain structural abnormalities are likely key players in DS people's impaired cognition (Lott, 2012; Marin-Padilla, 1972; Suetsugu and Mehraein, 1980). Indeed, DS is the leading cause of intellectual disability worldwide. Cognitive impairment in DS ranges from mild to severe, and the intelligence quotient tends to go down with age. The cognitive defects start to emerge during infancy and keep increasing in early childhood with poor motor skills, speech, language acquisition, adaptive behavior, and cognition. Children with DS have a low ability to encode information and retrieve memories (Carlesimo et al., 1997; Contestabile et al., 2010; Roizen and Patterson, 2003; Sherman et al., 2007; Vicari et al., 2000). In addition to intellectual disability, people with DS suffer from craniofacial defects, congenital heart

defects (CHD), gastrointestinal abnormalities, hypotonia, audiovestibular and visual impairment, thyroid disorders, hematopoietic disorders, and leukemia, together with other dysfunction of immune responses (e.g., increased risk of infections, hematological/autoimmune disorders, and hyperactivation of the interferon (IFN) signaling) (Ana C. Xavier and Jeffrey W. Taub, 2010; Asim et al., 2015; Bull, 2011; Dey et al., 2013; Sherman et al., 2007; Sullivan et al., 2016).

1.2.1. The genetics of DS:

The genetic architecture of DS is extremely complex. The extra chromosome provides people with DS with extra copies of several genes. In humans, approximately 550 genes are located on HSA21, of which 222 encode proteins, and the rest encode microRNAs, long-noncoding RNAs, and other regulatory elements (Gupta et al., 2016). Since the trisomy of chromosome 21 causes DS, theoretically, one would expect that genes on chromosome 21 are present in three copies and have 1.5 fold more expression in people with DS compared to the euploid state. However, indeed, this is not the case.

There are two hypotheses to explain the effect of the extra chromosome in the DS phenotype. The first is known as the gene-dosage effect hypothesis, where one of the triplicated genes can (directly or indirectly through the effect on the downstream signaling pathways) affect the tendency to acquire severe cognitive defects. For example, APP, a triplicated gene, increases susceptibility to early-onset Alzheimer's disease (AD). A second hypothesis is a developmental instability, which implies that the presence of extra-chromosome causes global, non-specific disturbances in gene expression (independently of what the triplicated genes encode for), disrupting biological homeostasis.

Over the last two decades, it has been conclusively demonstrated that the differentially expressed genes are not just located on HSA21 but also on other chromosomes as well, suggesting that both gene dosage hypothesis and a broad dysregulation of gene expression in DS are involved (Araya et al., 2019; Letourneau et al., 2014; Lockstone et al., 2007). In particular, Letourneau et al. have reported that Down syndrome is associated with altered gene expression across every chromosome and not just chromosome 21. They performed transcriptomics of fetal fibroblasts from a pair of monozygotic twins discordant for trisomy

21 (Letourneau et al., 2014). Another study where the authors performed a meta-analysis on various human samples (tissues such as the brain, thymus; cells such as blood cells, fibroblasts, and iPSCs) from DS and healthy individuals found that dysregulated genes (both triplicated and non-triplicated) had a 1.5 times differential expression (Pelleri et al., 2018).

Overall, DS symptoms originate from gene-expression disturbances. Thus, to design new therapeutic approaches, it becomes vital to investigate the molecular mechanisms behind the phenotypic consequences resulting from these genetic alterations.

1.2.2. Mouse models of DS:

The elucidation of human and mouse genomes and their comparative analysis has revealed that the long arm of HSA21 is homologous to portions of three mouse chromosomes (MMU17, MMU16, and MMU10). One of the most extensively studied mouse model of Down syndrome (the Ts65Dn mouse) carries a large segment of mouse chromosome 16 (~110 HSA21 ortholog genes) along with a substantial segment of MMU17 (60 centromeric genes) not syntenic to HSA21 (Duchon et al., 2011). The chromosomal triplication in Ts65Dn mice is carried as an additional freely segregating chromosome generated by cesium irradiation. Thus, Ts65Dn correlates well with aneuploidy observed in individuals with DS and can be used to investigate both the gene-dosage hypothesis and the developmental instability hypothesis. In other models (e.g., Ts1Cje, Dp(16)1/Yey, Dp(17)1Yey, and Dp(10)1Yey), the triplication is not characterized by a freely segregating chromosome. Ts1Cje was generated via reciprocal translocation of the distal portion of mouse chromosome 16 onto chromosome 12, resulting in 71 HSA21 ortholog genes. Dp(16)1/Yey contains ~130 genes, orthologous to HSA21, on the distal portion of mouse chromosome 16 via Cre-mediated recombination and contains the largest HSA21 orthologous genes. At the same time, Dp(17)1Yey and Dp(10)1Yey carry the MMU17 and MMU10 syntenic regions, respectively, with fewer orthologous genes (Li et al., 2007).

1.2.3. Transcriptomics and proteomics studies in DS:

Even though the cause for DS has been known for many decades now, and we also know that both triplicated and non-triplicated genes are largely dysregulated, we still do not know the

precise mechanisms by which triplication of HSA21 genes leads to neuroanatomical and neurobehavioural phenotypes in DS individuals. Moreover, even with the advancements in high throughput technologies, there are very few studies at transcriptome and proteome levels in DS individuals.

At the transcriptome level, there is only a single study by Olmos-Serrano et al. where the authors uncovered defective oligodendrocyte differentiation and myelination after performing transcriptional profiling in postmortem brains from DS individuals, spanning from mid-fetal development to adulthood (Olmos-Serrano et al., 2016). The authors also validated their results via cross-species comparison to the Ts65Dn mouse model. However, this study has been performed using outdated technology, microarray, which requires transcript specific probes and is not very high throughput. There are few other studies on human-derived tissues or cells performed using microarrays, but they are done either on fetal tissues, amniotic fluid, fibroblasts, or iPSC-derived neurons (Gonzales et al., 2018; Guedj et al., 2016; Huo et al., 2018; Stamoulis et al., 2019; Waugh et al., 2019). For example, Gonzales et al. identified significant changes in the transcript expression and changes in alternative splicing and repetitive element transcripts associated with chromosome 21 trisomy. Also, their transcriptome data suggested that trisomy of chromosome 21 may interfere with the maintenance of pluripotency. Another study on iPSCs-derived GABAergic interneurons from DS individuals indicated migration defects from their RNA sequencing data (Huo et al., 2018). Here, the authors transplanted the iPSCs derived GABAergic progenitors from trisomic and euploid control lines into SCID mouse brain in the medial septum. They observed that the neurites from trisomic interneurons were shorter and with fewer branches, and the majority of these cells were immature and had migration defects.

In animal models, we could find only one study where recently Aziz at el. compared brain development, gene expression, and behavior in three different mouse models of Down syndrome and found Ts1Cje and Ts65Dn brains had considerably more differentially expressed genes compared with Dp(16)1/Yey mice. Differentially expressed genes showed little overlap in identity and chromosomal distribution in the three models, leading to dissimilarities in affected functional pathways. The authors concluded with unique limitations of each model of Down syndrome (Aziz et al., 2018).

Interestingly, only a few studies have explored the non-coding RNAs, especially microRNAs. For example, Perez-Villareal et al. have profiled the blood samples from DS people and have in detail checked the miR-155 and let-7c, which are encoded on HSA21 and are overexpressed in DS. They found that these differentially expressed miRNAs regulate genes involved in lipid metabolism and nervous system development (Pérez-Villareal et al., 2020). In another study from 2018, the authors used nanoparticle formulations to analyze a subset of miRNAs in the plasma of young DS people and their siblings. They found many of them to be differentially expressed (Salvi et al., 2019).

Interestingly, in a study done on lymphocytes from children with DS, it was observed that not all triplicated miRNAs were upregulated, and other differentially expressed miRNAs were located outside chr21 (Xu et al., 2013). Five microRNAs (miR-99a, miR-155, miR-802, miR-125b-2, and let-7c) from chromosome 21 were detected to be overexpressed in DS; these miRNAs had an expected ratio of 1.5 *vs.* controls. In a very recent study in 2019, miRNAs' altered expression has been observed in the plasma obtained from mothers with fetal DS. The authors found 13 miRNAs to be differentially expressed, and targets for these differentially expressed miRNAs were involved in CNS development, congenital abnormalities, and heart defects (Zbucka-Kretowska et al., 2019).

At the proteomic level, there are some studies on the fetuses, amniocytes, plasma, and fibroblasts derived from DS individuals (Lanzillotta et al., 2020; Liu et al., 2018, 2017; O'Bryant et al., 2020). For example, in a study on amniocytes and amniotic fluid, the authors found around 904 differentially expressed proteins involved in 25 biological pathways after proteomics analysis. In the top two pathways, they found NF-κB and the other one, APP, which are known DS dysregulated genes. They further validated nine proteins with selected reaction monitoring (SRM) assays to quantitate individual amniocyte samples differential expression. Two proteins (SOD1 and NES) showed consistent differential expression across the samples (Cho et al., 2013). Proteomics studies on DS fetuses started with the work of Opperman et al., in which the authors showed 84% homology between the proteins expressed in DS and control samples, and ones that were differentially expressed were known to have a structural role that was consistent with defects in DS brain development (Oppermann et al., 2000). A more comprehensive study by Cheon et al. led to the

identification and quantitation of ten protein spots that were differentially expressed, and these were also involved in defective brain development (Cheon et al., 2001). Cheon et al. followed up with few more studies in DS fetuses, and they showed many altered protein pathways such as impairment of synaptic plasticity, brain development, and energy metabolism (Sun et al., 2011).

On the other hand, we found one only study at the proteome level on young adult human DS brains. In the cerebral cortex from young DS adults and age-matched controls, the authors found protein carbonylation, which is the oxidation of protein residues by reactive oxygen species in DS individuals. They showed this effect on six proteins that are part of the intracellular quality control system leading them to hypothesize the proteostasis network's impairment (Di Domenico et al., 2013).

In animal models, changes in protein expression level were observed in the hippocampus and cerebellum of the Ts65Dn mouse model at the age of 6 and 12 months after proteomic analysis (Vacano et al., 2018). Notably, the authors observed minimal differences in protein expression between disomic and trisomic groups but identified numerous differences associated with age and brain region. Ahmed et al. also showed similar findings in three brain regions, hippocampus, cortex, and cerebellum, in the Ts65Dn mouse model, suggesting that with age, the protein expression dysregulation is more exacerbated in 12 *vs.* 6 months old animals. They also observed that the dysregulated proteins were involved in mTOR and MAPK pathways (Ahmed et al., 2017).

There is only one study in which both transcriptomics (with RNA-seq) and proteomics (LC-MS/MS) has been done on the same samples. Here, the authors generated iPSCs and induced them towards neuronal lineage to perform differential expression analysis and study the temporal dynamics of dysregulated genes. They compared their gene expression data with data from BrainSpan (Miller et al., 2014), which contains developing human brain transcriptome information from RNA sequencing and microarray experiments. The authors observed that RNA-seq profiles of differentiated neural progenitors exhibited similarly to the developing brain at 20-30 weeks post conception. At both gene and protein levels, the authors found dysregulation in DNA replication, pluripotency, synaptic formation, and neural signaling in neural progenitor cells. When the progenitors were differentiated at later stages

of development, they saw a deficit in TGF-β signaling and SMAD-associated signaling (Sobol et al., 2019).

The efforts mentioned above have generated a large amount of data. They have provided much-needed clues and potential target candidates to go forward with the research in DS to develop novel therapeutics. However, most of the findings are originating from mouse models, and these observations may not translate well in humans. Indeed there may be disparity due to i) failure of the animal models to accurately mimic the human disease condition; ii) the fact that most of the animal models most likely replicate specific processes, but not the whole spectrum of the physiological changes that are present in humans; iii) identified biomarkers between the two species might not be similar and iv) inter-individual genetic variation affects the human gene expression, which cannot be replicated in the mouse models. Also, a large discrepancy remains in how well the transcriptome level's expression profiles match with those at the protein level, posing some questions on what biological processes are the most responsible for the DS phenotype and what should be considered with the highest priority for developing treatments for DS.

Moreover, the knowledge from studies in humans mainly originates from fetal tissues, plasma, fibroblasts, and iPSCs. We do not still understand how well they correspond to the adult human DS condition. Integration of both transcriptome and proteome to study genes, miRNAs and proteins in the same set of samples would provide an excellent opportunity to understand the human DS condition at a substantially higher resolution.

Although I have taken an unbiased approach in addressing the impact of trisomy 21, I briefly introduce one of the aspect of DS: neuroinflammation, as a part of my contribution to a published manuscript (Pinto et al., 2020).

1.2.4. Neuroinflammation:

Neuroinflammation is a multistep phenomenon controlled by resident glial immune cells (mainly microglia and astrocytes) (Norden et al., 2016) and a cascade of pro-inflammatory factors (e.g., cytokines and chemokines).

Microglia are the resident immune cells in CNS. These represent 5-20% of the human brain cells and have a ramified morphology in healthy condition (Butovsky and Weiner, 2018). Microglia plays a significant role in brain development and plasticity (Arcuri et al., 2017; Tremblay et al., 2011). They sculpt the neuronal circuit by synaptic pruning (Schafer and Stevens, 2013), regulate neuronal death, and phagocytose neuronal progenitors (Kaur et al., 2017; Mosher et al., 2012). They also give support to axonal growth and phagocytose debris (Tay et al., 2018). Critical period plasticity and activity-dependent synaptic maturation are also regulated by microglia activity (Miyamoto et al., 2016; Paolicelli et al., 2011; Squarzoni et al., 2015). In disease or during an infection, microglia adopt an amoeboid-like shape and engage in pro-inflammatory immune responses at the site of infection or injury (Pinto et al., 2020; Salter and Stevens, 2017). During inflammation, microglia upregulate activation and phagocytic markers and produce pro-inflammatory cytokines (Arcuri et al., 2017; Hanisch and Kettenmann, 2007; Shobin et al., 2017; Sipe et al., 2016).

Astrocytes represent around 25-50% of CNS cells. They are neuro-supportive, control brain homeostasis (e.g., control pH and oxidative stress) and are responsible for proper neuronal function (e.g., synaptic plasticity) (Liddelow and Barres, 2017; Sochocka et al., 2017; Verkhratsky et al., 2013). Like microglia, astrocytes also become reactive and contribute to death of neurons (Liddelow et al., 2017).

Microglia and astrocytes communicate with each other either through secreted factors or cell-cell contact. Microglia can make the astrocytes reactive upon release of specific cytokines during diseased states, and in turn, astrocytes can control the functions of microglia by sequestering cholesterol (Liddelow et al., 2020).

1.2.4.1. Neuroinflammation in DS and other brain diseases:

Hemostasis between the CNS and peripheral nervous system fails during pathological conditions, leading to prolonged activation of glial cells, production of pro-inflammatory factors, and reactive oxygen species. Multiple sclerosis (MS) is the most prominent brain disorder which results from CNS inflammation. Advances in understanding the inflammatory mechanisms in MS have made it possible to design therapies that could reduce

neuroinflammation. Similarly, autoimmune diseases such as systemic lupus erythematosus, limbic encephalitis, and others have responded to a treatment designed with the help of studies done in MS. Other diseases with neuroinflammatory response include epilepsy (Vezzani et al., 2019), cerebral ischemia resulting in stroke (Stoll and Nieswandt, 2019), migraine (Edvinsson et al., 2019), and neurodegenerative diseases such as Frontotemporal dementia (FTD) (Yoshiyama et al., 2007), Alzheimer's disease (AD) (Calvo-Rodriguez et al., 2020; Sokolova et al., 2009), Parkinson's disease (PD) (Lecours et al., 2018), and Amyotrophic lateral sclerosis (ALS) (Béland et al., 2020). Interestingly, in neurodevelopmental disorders such as autism spectrum disorder (ASD) and schizophrenia, several studies have pointed to neuroinflammation with the result of microglial and astrocytic activation (Laurence and Fatemi, 2005; Li et al., 2009; Morgan et al., 2010; Vargas et al., 2005).

As already highlighted above, neuroinflammation in DS is very prevalent (Wilcock and Griffin, 2013; Wilcock et al., 2015). A Role of microglia in DS was first reported in 1989, where the authors showed elevated expression of IL-1 in microglia and astrocytes in DS individuals. Other studies have been performed in brain samples from people with DS and mouse models (Griffin et al., 1989). Interestingly, in DS fetuses, an increase in microglia number and increased infiltration of macrophages in the CA1 and subiculum region of the hippocampus has been observed (Kanaumi et al., 2013; Wierzba-Bobrowicz et al., 1999).

In recent studies, an increase in the inflammatory response (including interferon-stimulated genes), differential expression of inflammation-related proteins, and activation phenotype of microglia and astrocyte have been reported in DS individuals (Pinto et al., 2020)*; Wilcock et al., 2013; Wilcock et al., 2015). Microglia activation led to dendritic spine defects in the Ts65Dn mouse (Pinto et al., 2020)*. Another study that showed increase induction of interferon-stimulated genes was performed in fibroblasts and blood cells from DS people with the help of transcriptome analysis (Sullivan et al., 2016), which also confirmed the previous results (Tan et al., 1974). Sullivan et al. followed up on their previous study to confirm chronic autoinflammation in DS in plasma by performing proteomics analysis. They found elevated levels of IL-6, TNF- α, and MCP-1, the potent inflammatory cytokines linked to IFN signaling (Sullivan et al., 2017). Mass cytometry also revealed a similar global immune remodeling level with a signature increase in interferon type I response in DS (Waugh et al., 2019). Two recent

studies have also reported therapies to inhibit and block the microglia activation and reduce the neuroinflammation in mouse models of DS (Pinto et al., 2020; Tuttle et al., 2020).

* Pinto et al. measured the effect of microglial activation on cognitive functions in two mouse models of DS (Dp(16)1Yey and Ts65Dn) and observed microglia activation in the human DS hippocampus. The authors observed an increase in the cell soma and decreased branches for the microglia in both the models and the human brain samples. The mouse models showed a decrease in spine density and electrophysiological activity in the hippocampal neurons. They also observed deficits in hippocampus-dependent cognitive tasks in the mouse models. With the help of acetaminophen, an anti-inflammatory drug, the authors could rescue the cognitive deficits, microglial morphology, and spine density defects in young adult animals. For this study, I performed RNA-seq analysis on human hippocampal samples from DS and control individuals. I found that the upregulated genes were involved with immunological pathways and glial cells (astrocytes, microglia, and oligodendrocytes), while the downregulated genes were enriched in neurological pathways and neuronal cells. Please, find the complete manuscript in the appendix at the end of the thesis.

1.2.5. Pharmacological treatments:

There is currently no pharmacological treatment to rescue cognitive impairment in DS people. Nevertheless, early educational interventions and environmental enrichment lead to an improvement in cognitive tasks and intellectual disability scores in DS children (Engevik et al., 2016; Martínez-Cué et al., 2005). However, these strategies are of limited use in adults with DS as the cognitive decline progresses with age and they lose the acquired abilities (Couzens et al., 2011, 2012).

Interestingly, various drug treatments targeting various putative molecular pathways are effective in rescuing cognition in DS animals. While these preclinical studies performed in different mouse models of DS have reported improvement in cognitive deficits in the models, the same drugs have failed to replicate all the good and promising effects when administered to people with DS during clinical trials. This could be due to genetic differences between mouse and human and to safety, efficacy, compensatory mechanisms, the timing of intervention, or the drug's dosage .

These are some of the main reasons that we need approaches that could identify molecular targets from perspectives that take a more global approach to design therapeutics that are highly effective in humans. One possible way of addressing this issue is with high throughput technologies such as transcriptomics and proteomics.

1.3. The rationale and the questions:

DS is a very complex disorder with global gene dysregulation. Thus a fundamental question in DS is to know what is the impact of the triplication of chromosome 21 on this global gene dysregulation and how does these profound changes at the genetic level translate to phenotypic changes in DS individuals. To understand the genotype to phenotype connection, we need to integrate unbiased, high throughput information at gene and protein levels. In the past, most studies have reported their findings using technologies (i.e., microarrays, serial analysis of gene expression (SAGE), selected reaction monitoring assay (SRM), and gel-based proteomics) that are biased and not high throughput. This is because, until the last five years, performing high throughput transcriptomic and proteomic studies were too costly and slow. Moreover, there were considerable limitations in the availability of postmortem brains. Thus, most of these studies were performed on fetal tissues, plasma, fibroblasts, iPSCs derived from DS individuals and mouse models. Recently, with the advent of better technologies, reduction in cost and increase in speed, and the availability of computational methods to analyze the data obtained from high throughput technologies, along with ongoing efforts to create large brain biobanks, it has become possible to perform multiple analysis on the same set of samples. Here, I will take an integrative high throughput approach to perform a comprehensive assessment of adult DS individuals transcriptome and proteome to find hub genes and proteins that could target future therapeutics.

Below, I introduce some of the concepts involved in the study, such as gene expression, RNA-protein correlation, co-expression of genes, alternative splicing, and miRNAs.

1.4. Gene expression:

Gene expression is a complex process that involves the conversion of DNA to RNA (transcription), post-transcriptional regulation, conversion of RNA to protein (translation), and post-translational modifications. Transcription is controlled by transcription factors, chromatin modifications, splicing, polyadenylation, transport, and degradation. Translation involves initiation, elongation, and termination, as well as localization and degradation. The

interaction between multiple gene products (proteins) converts this genotype to phenotypic information (Buccitelli and Selbach, 2020).

1.4.1. Coding and non-coding genes:

The mammalian genome is transcribed into multiple RNAs that can be grouped into coding and non-coding RNAs. Protein coding genes are mature RNA species that are translated to form a protein product. Non-coding RNAs do not code for any protein product. These non-coding RNAs are of multiple types based on their biosynthesis and length, such as lncRNAs, miRNAs, and pseudogenes. lncRNAs are long non-coding RNAs of more than 200 bp. They do not translate, but they fine-tune the gene expression. miRNAs are 21-22 nucleotides in length and bind the 3' untranslated region of the mRNA, regulating gene expression. Pseudogenes are those regions of the genome that are non-functional and contain defective copies of protein-coding genes due to the accumulation of mutations during evolution (Beermann et al., 2016; Dykes and Emanueli, 2017; Li and Liu, 2019).

1.4.2. Transcriptomics:

A gene is transcribed into multiple transcripts, and the entire set of transcripts originating from an organism is known as the transcriptome of that species. The high throughput sequencing of all the transcript isoforms for all the genes is known as RNA sequencing. For convention, the difference between two groups is known as differential gene expression (DGE) when the gene counts are accumulated from all the transcript counts for that gene. Also, for the convention, differential transcript expression (DTE) is instead the difference in transcript expression between two groups when the diverse type of transcript counts are taken into consideration (Figure 1.1). A differentially expressed gene can have transcripts where only a few of them are differentially expressed, as illustrated in Figure 1.1B (i) & (ii), where Gene A is DGE, but only transcript A.2 is differentially expressed.

On the other hand, it is also possible for a gene to be not differentially expressed but has all its transcripts being differentially expressed between the two groups, as shown in the Figure

1.1B (i) & (ii), where Gene B is not DGE, but all its transcripts (Transcript B.1 and B.2) are differentially expressed. RNA sequencing is used to assess both gene and transcript level abundances. In summary, gene-level abundance is the sum of all transcript-level abundance.

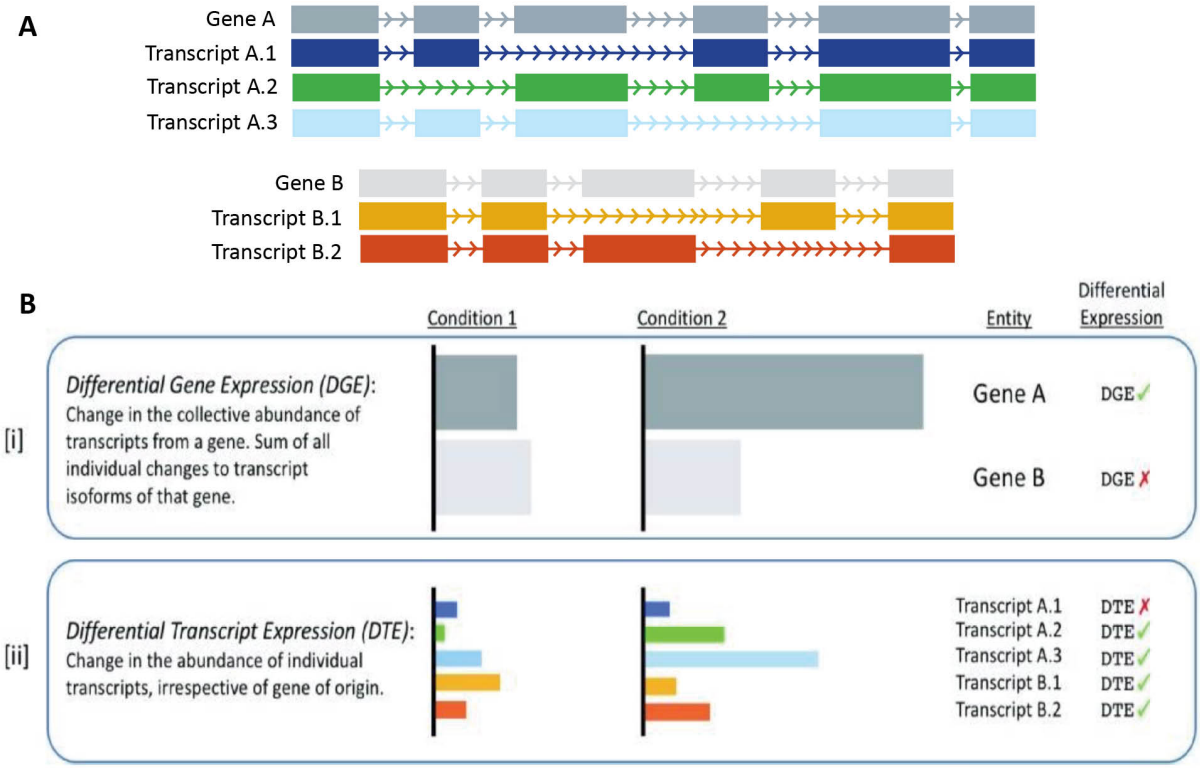


Figure 1.1. Definitions of different types of differential expression analysis. In panel A), two genes (Gene A and Gene B), with 3 and 2 transcripts, are shown. In panel B) the expression of these two genes with their transcripts is compared across two conditions (Condition 1 and Condition 2). The horizontal width of each colored box represents the abundance of the relevant gene or transcript. A negative differential expression result (red cross-mark) for a given entity in any one of the two analysis types does not exclude that same entity from having a positive result (green tick-mark) in one of the other two analysis types. DGE: Differential gene expression, DTE: Differential transcript expression (Froussios et al., 2019; Newman et al., 2018).

In general, the RNA sequencing workflow involves RNA extraction, mRNA enrichment or ribosomal RNA depletion, preparation of cDNA, and addition of unique adapter sequences to each sample to generate sequencing libraries. All the sample libraries are then pooled together to generate a nanomolar pool (this process is called multiplexing) and sequenced at a depth of 10-100 million reads. Following sequencing, the data is demultiplexed, where the sample reads are separated based on unique adapter sequences attached to each sample. Subsequently, they are computationally aligned to the reference genome, and counts for each gene/transcript in each sample is computed. The counts are then filtered and normalized

between samples, and a statistical method is applied to estimate differential gene/transcript expression between sample groups (Stark et al., 2019).

Specifically, mRNA sequencing involves the enrichment of oligo-dT sequences, which contain poly-adenylated (poly(A)) tails and focuses only on protein-coding sequences. For obtaining non-coding RNAs and transcript-level information, ribosomal RNA (rRNA) deletion is required. rRNA-depleted RNA sequencing generates information for both protein-coding genes and some non-coding genes. rRNA-depleted whole transcriptome sequencing is also compatible with degraded samples such as those obtained from humans, as they usually suffer postmortem interval. Small RNAs require their targeted sequencing, where only the diverse class of small RNAs such as microRNA, small nuclear RNAs (snRNAs), transfer RNAs (tRNAs), and small nucleolar RNAs (snoRNAs) are queried (Hafner et al., 2008). For improving the analysis of degraded samples, a higher sequencing depth is required, which reduces the bias between the samples. By taking this approach, it is possible to measure protein-coding genes, their different isoforms, and non-coding transcripts in degraded samples.

1.4.3. Proteomics:

Mass spectrometry (MS)-based proteomics is used to identify and quantify proteins at a large scale, without the use of antibodies (Aebersold and Mann, 2016). MS measures the mass to charge ratio and intensity of charged particles. Proteomics can be used in an unbiased manner to investigate individual proteins and their functions.

Proteins are extracted from the cells or tissues of interest by enzymatic digestion. The resulting peptides are then separated on high-performance liquid chromatography columns, which are packed with micrometer-sized beads and run at nanoliter flow rates. Mass and intensity of peptides are measured in mass spectrometer scans. The peptides are identified by comparing the fragments or MS/MS spectra to a standard database of peptides (Manzoni et al., 2018).

1.4.4. Correlation between mRNA and protein:

Mostly, gene expression is associated with mRNA expression, and it is generally assumed that mRNA expression can be considered a proxy for protein expression. Nevertheless, transcription involves multiple post-transcriptional steps, which result in the generation of not one but many mRNA isoforms. These, in turn, form multiple protein products. Accordingly, at the level of high-throughput technologies such as RNA sequencing (Stark et al., 2019) and proteomics (Aebersold and Mann, 2016), the protein expression does not coincide well with mRNA expression.

The relationship between mRNA and protein expression levels is usually studied at mRNA-protein correlation. This correlation is of two types: a) across-gene correlation explores how well the abundance of mRNAs correlates with their corresponding proteins for different genes in the same condition; b) while within-gene correlation inquires how much the change in mRNA level of one gene can explain the change in corresponding protein level across conditions. Across-gene correlations for mammalian tissues range around 0.6, but these could be higher or lower due to variations in technical biases. Within-gene correlation reveals a significant but modest correlation between mRNA and protein, provided that we look at the genes with similar functional identity to carry out this correlation (Buccitelli and Selbach, 2020).

Both RNA sequencing and MS-based proteomics suffer from noise and biases, which affect both precision and accuracy (Aebersold and Mann, 2016; Stark et al., 2019). Technical and biological noise could be estimated and reduced by incorporating more samples in the study. On the other hand, bias is difficult to quantify precisely but could be estimated approximately by adding reference RNAs or proteins (Schwanhäusser et al., 2011).

On average, expressed genes are present at less than one mRNA copy per cell, whereas the protein abundance is at 10^8 molecules per gene. The difference of several magnitudes in the dynamic range of mRNA and protein expression is mainly due to higher translation efficiencies for more abundant proteins (Li et al., 2014a). Equally important is the speed of transcription and translation. Transcription usually takes about 1 hour to generate ~100 mRNA from the

DNA template, while 10^6 protein molecules are generated after transcription initiation in more than one hour from a single locus (Hausser et al., 2019). So, translation and protein degradation are on-demand and context-specific to reduce the time for protein production or downregulation (Schwanhäusser et al., 2013). Overall, protein abundance is dependent on four critical parameters: rate of transcription, mRNA half-lives, translation rate, and protein half-lives (Baum et al., 2019; Hausser et al., 2019; Kristensen et al., 2013).

1.4.5. RNA sequencing or proteomics: Which one to do?

It depends on the specific question that we would want to address. The gene expression level, mRNA hence RNA-seq, is closer to the genome and would reflect transcription and RNA processing events. In contrast, protein or proteomics looks at the next stage in gene expression, which is related to phenotype. With RNA-seq, it is possible to get information about all the mRNAs as it provides a complete picture of the cell or tissue of interest. Simultaneously, proteomics can inform about less number of proteins as not all the mRNAs have been translated or available as protein products in that particular spatial, temporal context (the time the cells or tissues are collected). Protein is more stable than mRNA, and the post-transcriptional and translational effects can only be studied at the protein level and are not visible at the mRNA level (Buccitelli and Selbach, 2020). On the other hand, proteomics outperforms transcriptomic studies at the level of functional predictions as observed from many co-expression studies (Kustatscher et al., 2019; Lapek et al., 2017; Ori et al., 2016; Romanov et al., 2019; Wang et al., 2017).

Both mRNA and protein expression provide unique information about the biological system in question. Thus, it is essential to integrate RNA sequencing and proteomics technologies to better understand gene expression principles in physiological and pathological conditions. The flow of information from the genome to phenotype regulated at multiple levels will benefit from the acquisition and integration of multi-omics datasets. This would provide insights on more possible targets for designing better therapeutics.

1.5. Multi-omics:

The brain is a complex organ consisting of billions of neurons that make trillions of connections among themselves (Koch and Laurent, 1999). This complex network executes various high function processes and behaviors, including cognitive processes such as learning and memory (Kandel, 2001). Understanding this complexity is one of the most challenging aspects of neuroscience. The perturbations taking place at molecular, cellular, and network-level in developmental disorders make it even more of a priority.

The rapid developments in technologies such as transcriptomics and proteomics have provided essential insights into CNS development and function (McCarroll et al., 2014; Shin et al., 2014). Acquiring data at multiple biological levels has provided a more comprehensive view of gene-expression regulation than any single omics. The development of methods to integrate multi-omics data has provided capabilities to carefully and systematically interpret results and make relationships among the data.

Multi-omics is an integration of data acquired using multiple single omics technologies on the same set of samples to understand the flow of information at multiple levels (Hasin et al., 2017). Each omics provide a different set of information to characterize the normal and diseased condition at a functional level. The integration allows looking for a causative signature rather than a readout at a single omics level. A key challenge in multi-omics is the interpretation of the results, which could benefit from borrowing and applying some of the principles of systems biology. The holistic characterization of a biological network produces a non-linear coherent network from individual multifunctional elements (Hillmer, 2015; Kitano, 2002). Availability of data at multiple levels allows one to look at all the observable entities and their relationships to generate even new relationships and decipher system-level properties such as at organelle, cell, tissue, or organism level. Multi-omics system biology involves discovery, i.e., acquisition of data, analysis, and validation at a single omic level; representation, i.e., integration and visualization in a network; and application, i.e., how these data-driven networks help in clinical investigations (Eckhardt et al., 2020).

1.5.1. Network biology:

Integration and interpretation of multi-omics data lead to a better understanding of the molecular components acting in a system. The brain, being a highly intricate and complex structure, has a dynamic and diverse proteome and these proteins interact with each other to create a stable or transitory protein complex. Moreover, proteins such as transcription factors, RNA binding proteins, among others, interact with RNA and form another layer of regulation, leading to multiple signaling cascades. To unravel these entwined processes and signaling pathways, approaches that can provide a system-wide view of the signaling network are required. These approaches can take the expression profiles of genes and proteins involved in a biological system and the protein-protein interaction knowledge from the literature and put the phenotype in relationship with the data.

Inhibition of single drivers in multi-level signaling cascades can lead to compensatory mechanisms through alternative pathways. For example, a disease such as DS, which leads to genome-wide dysregulation at both RNA and protein levels, requires targeting not just the single drivers but multiple of them and, more specifically, the pathways itself, which could lead to higher efficacy of developed drugs. In this context, the molecular networks generated with data from different technologies may serve as a framework for better biological inferences. Indeed, these biological networks are an intermediate step for studying the genotype to phenotype relationships.

Inside the cell, these molecular interactions often are organized into groups of proteins associated with a similar biological function and form modules (of proteins expression at similar levels) within a vast network (Becker et al.; Costanzo et al., 2010; Furlong, 2013; Ideker and Sharan, 2008; Taylor and Wrana, 2012). The phenomenon of modules is not confined to proteins but is also present for genes, as shown by Goh et al. (Goh et al., 2007). Within an interactome of all the proteins, functional modules are associated with the principle of "guilt by association", where the genes or proteins of similar expression will cluster together to form co-expressed gene clusters. This subnetwork of genes/proteins helps identify the hub molecules within a network and predict functions and novel disease-associated genes (Ideker and Sharan, 2008; Lee et al., 2011). These subnetworks could also help in pointing out never-

before highlighted processes that are disrupted in the disease. Co-expression network analysis is suitable for interpreting and identifying distinct gene regulation levels and across different biological scales from cells to tissue (Hawrylycz et al., 2012; Miller et al., 2014; Oldham et al., 2008; Parikshak et al., 2013).

1.5.2. Cell types from single-cell to bulk RNA-seq:

The cell diversity inside the CNS has been traditionally studied by looking at their morphological and physiological features such as shape, immunohistochemical markers, electrical activity, and the cells' location (Bota and Swanson, 2007; Fishell and Heintz, 2013; Masland, 2004). Transcriptomic studies are required in fact to ascertain the molecular phenotypes of CNS cell types. Nevertheless, most gene-expression studies have analyzed the bulk tissue samples, which reflect an average of what is present in the single-cell types. Ideally, the solution to this problem is to isolate pure populations of single-cell types, which requires fresh extraction of tissue for good quality of expression profile. This is again limited to the model organism, and to take this approach to humans, would require resection of tissue during neurosurgeries. Few efforts have been undertaken along these lines in a limited number of regions, but those are timely, costly, and technically very challenging (Darmanis et al., 2015).

Nevertheless, high-throughput single-cell RNA-seq (scRNA-seq) has started to provide a more comprehensive understanding of the transcriptional programs in an individual cell and identify novel cell types (Macosko et al., 2015; Tasic et al., 2018; Usoskin et al., 2015; Zeisel et al., 2015). Compared to model organism's scRNA-seq studies, where it has been much easier to isolate and sequence the cells, adult human CNS suffers from low throughput and contamination from other cells (Okaty et al., 2011). The approach to isolate single-nuclei from post mortem brains holds excellent promise but currently is limited in scope (Krishnaswami et al., 2016; Lake et al., 2016)

Nevertheless, computation tools exist to obtain cell-specific expression profiles from bulk RNA sequencing by utilizing co-expression of genes. This concept is based on the assumption that different cells from heterogeneous tissues such as the brain express different genes. Thus

in bulk RNA-seq experiments, the genes which are consistently expressed explicitly across the samples will be inferred as a set of marker genes for that particular cell type. This gene expression profile for each cell type helps in deconvoluting the brain RNA-seq expression into individual cell types based on expression levels of genes.

1.6. Alternative splicing in the mammalian brain:

1.6.1. Alternative splicing:

The genetic information is encoded in the DNA, which transcribes to RNA and then translates to protein to partake in various biological functions (CRICK, 1970). However, different cells in different areas of our body have distinct molecular signatures and perform diverse functions even if they carry identical DNA. This diversity is established with post-transcriptional regulation (Franks et al., 2017; Manning and Cooper, 2017). Post-transcriptional regulation consists of alternative splicing, poly(A) tailing, capping, RNA editing, miRNA regulation, transportation, mRNA modifications, and degradation (Bludau and Aebersold, 2020).

Alternative splicing (AS) is the most ubiquitous and crucial phenomenon in generating diverse isoforms from a single gene. During AS of precursor mRNA, combinations of different 5' and 3' splice sites, inclusion or exclusion of specific exons, or intron retention result in many isoforms. These different isoforms get translated to different protein products, which carry out different biological functions. More than 95% of human genes undergo alternative splicing and generate enormous diversity and complexity at the RNA level (Pan et al., 2008; Wang et al., 2008). Across all the tissues, the human brain undergoes the highest AS (Raj and Blencowe, 2015; Yeo et al., 2004). This generates isoforms that affect many aspects of neuronal development and function (Raj and Blencowe, 2015; Zheng and Black, 2013). The number of identified and novel AS events is context-specific, i.e., they are cell-type, tissue-type, and development stage-specific (Calarco et al., 2011; Kalsotra and Cooper, 2011).

Although alternative splicing has progressed a lot from identifying single splicing events to discover global networks, the functional consequences of these events have not been fully understood at the physiological or disease level. Improving the molecular information

regarding splicing networks in physiological conditions would help us understand better the dysregulation in various disorders. This advanced molecular understanding is currently possible thanks to deep RNA sequencing in a cell or tissue of interest in an unbiased manner. (Baralle and Giudice, 2017; Dillman et al., 2013).

Pre-mRNA consists of both introns and exons, and during the splicing process, introns are spliced out, and exons are ligated together to form mature RNA. There are particular sites or sequences at intronic ends that define where the cleavage or ligation will occur. 5' splice site is at 5' end of the intron, and similarly, 3' splice site is at 3' end of the intron. The spliceosome machinery is responsible for splicing mechanisms. There, many factors come together at splice sites to perform the first cleavage and then ligation steps. The most common splicing event is exon skipping or cassette exon, where one exon is either included or excluded in the final mRNA product (Vuong et al., 2016) (Figure 1.2). Mutually exclusive exons are patterns where only one of the exons from a consecutive exon is kept in the mRNA. Alternative 5' and 3' splice sites are donor and acceptor sites that change the length of exon at 3' end and 5' end, respectively. Alternative first exons (promotor) and last exons (poly(A) tail) create different first and last exons, respectively. Intron retention is a splicing event where the intron is retained or excised from the final mRNA. Lastly, microexon are 3-27 nt exons, skipped or retained in the alternative spliced isoform (Raj and Blencowe, 2015).

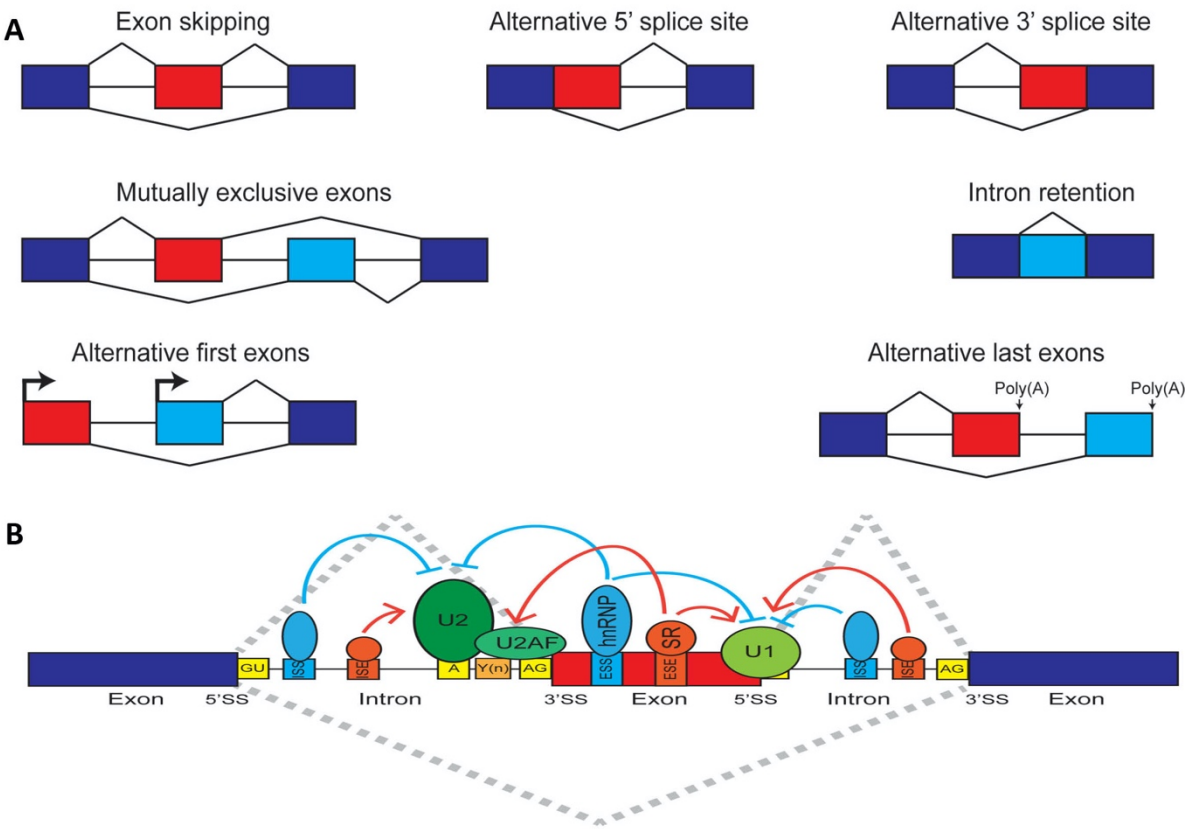


Figure 1.2. Patterns of alternative splicing and the action of spliceosome components. Dark blue boxes represent constitutively expressed exons, and red and light blue boxes represent exons that undergo different kinds of alternative splicing events. (Park et al., 2018).

1.6.2. Regulation of alternative splicing: RNA binding proteins

The outcome of alternative splicing is influenced by the content of consensus sequences at the splice sites. Those are bound by the spliceosome machinery components, cis-regulatory sequences, and the expression levels of trans-acting factors such as RNA binding proteins (RBPs) and other splicing factors.

RBPs constitute a large portion of trans-acting factors. They interact with RNA to control its synthesis and life cycle. Around 1500 proteins have been categorized as RBPs (Gerstberger et al., 2014), and their number continues to grow with the advent of large high throughput technologies. RBPs can shuttle between the nucleus and cytoplasm, affecting the biosynthesis of a large number of transcripts.

Recently, a splicing code has been deciphered (Barash et al., 2010). Thus we are now able to predict splicing in different cells or tissues. The code mainly suggests that different RBPs could bind to the same regulatory sequences on the same or different transcripts; the same RBP could positively or negatively impact the binding motif's location, and RBPs could regulate each other (Figure 1.3). Nevertheless, information regarding the functional consequence of alternative splicing is still missing. However, studies by crosslinking and immunoprecipitation followed by sequencing (CLIP-seq), RNA immunoprecipitation followed by sequencing (RIP-seq), and transgenic models where the RBP of interest is depleted or overexpressed in specific tissues are providing first information. Overall, we now know that multiple RBPs can regulate alternative splicing cooperatively and competitively.

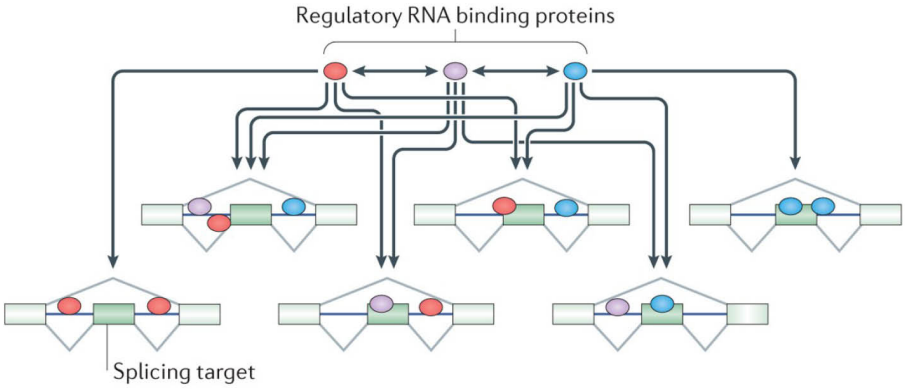


Figure 1.3. RNA binding proteins regulate multiple exons. Splicing regulators control large target exon sets that often overlap with those regulated by other RNA-binding proteins (RBPs) (Vuong et al., 2016).

1.6.3. RBPs and splicing in the brain:

Alternative splicing can impact multiple aspects of the nervous system, such as neuronal development, differentiation, and synaptic transmission (Raj and Blencowe, 2015; Vuong et al., 2016). Diverse RBPs regulate splicing in the brain. For example, expression levels of PTBP1, PTBP2, and SRRM4 play essential roles during neurogenesis. In neuronal progenitors, PTBP1 induces the exclusion of exon 10 of PTBP2, which leads to exon 10 skipping and a transcript with premature termination codon. When progenitors differentiate to neurons, PTBP1 is downregulated, and a positive regulator of PTBP2 splicing, SRRM4, is upregulated. Hence, exon skipping is repressed, leading to the expression of PTBP2 promoting neuronal maturation ((Li et al., 2014b); Figure 1.4).

PTBP2 null mice die shortly after birth, and genes involved in the cytoskeleton and cell proliferation exhibit defects in alternative splicing. This misregulation of AS is also observed in genes regulating neurite outgrowth and synaptic transmission (Li et al., 2014b; Licatalosi et al., 2012).

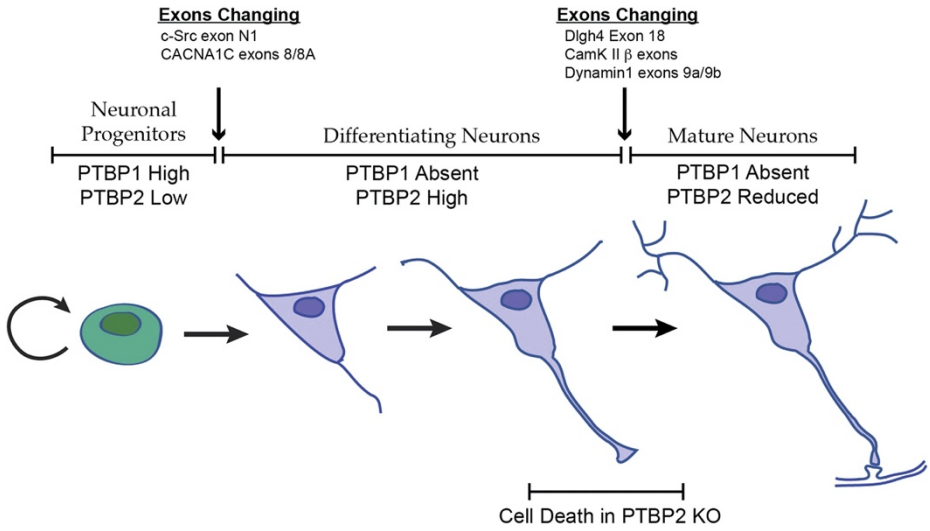


Figure 1.4. Changes in expression of PTBP1 and PTBP2 during neuronal differentiation (Li et al., 2014b).

Another RBP family plays an essential role during neurodevelopment. The RBFOX family consists of three paralogs: RBFOX1, 2 & 3. Deleting RBFOX1 results in increased excitability of neurons in the dentate gyrus and susceptibility to seizures (Gehman et al., 2011) (Figure 1.5). RBFOX1 has also been implicated in neuronal differentiation (Fogel et al., 2012). Notably, RBFOX1 regulates its own alternative splicing for exon 19, resulting in two isoforms. One is nuclear (exon 19 excluded), and the other one is cytoplasmic (exon 19 included). RBFOX2 deletion leads to improper migration, dendritic arborization of Purkinje cells, and increased cell death (Gehman et al., 2012).

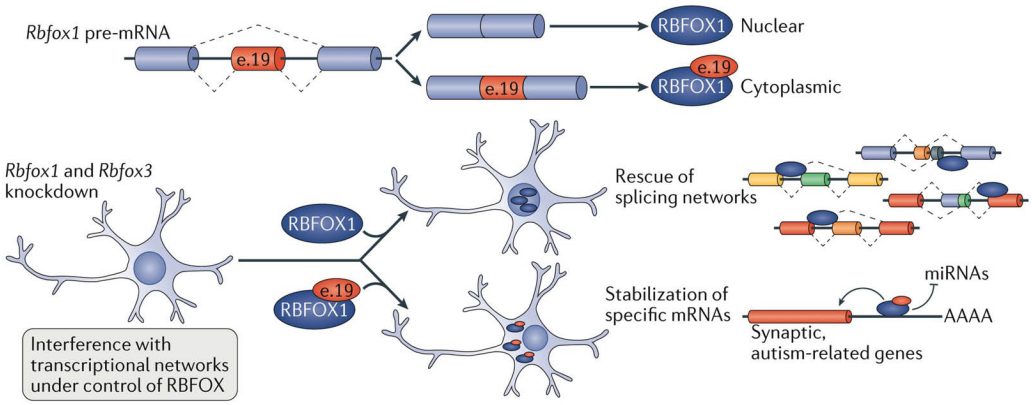


Figure 1.5. Exon 19 (e.19) in RBFOX1 is alternatively spliced, giving rise to nuclear (lacking e.19) and cytoplasmic (containing e.19) protein isoforms. In RBFOX-depleted neurons, the misregulation of splicing networks controlled by RBFOX1 leads to defects in the expression of transcription factors, other splicing factors, and synaptic proteins. The introduction of exogenous nuclear RBFOX1 rescues changes resulting from the misregulation of splicing networks. The introduction of exogenous cytoplasmic RBFOX1 stabilizes multiple mRNAs (for example, by interfering with the binding of some microRNAs (miRNAs)), particularly the mRNAs of synaptic and autism-related genes. This stabilization increases the targeted genes' protein expression, rescuing certain neuronal functions (Baralle and Giudice, 2017).

Splicing regulatory networks within functionally related genes consist of co-regulated exons. NOVA 1 and 2 were the first RBPs found to perform this type of regulation. NOVA1 knockout results in motor failure and postnatal death, while NOVA2 null mice exhibit dysregulation of activity-dependent LTP of slow inhibitory postsynaptic current (Raj and Blencowe, 2015). Moreover, NOVA2 regulates neuronal migration, and possibly, it does so by splicing Dab1, a component of the reelin signaling pathway, which controls neuronal migration and cortical lamination. NOVA2 acts by repressing exon 7b and 7c of Dab1. This results in a product that undergoes ubiquitylation upon Reelin activation (Rice et al., 1998; Yano et al., 2010).

1.6.4. Defects in AS results in neurological disorders:

During neurodevelopment, the gene expression has to be tightly regulated, and AS is one crucial step in this regulation (Vuong et al., 2016). As neurons mature, they undergo dramatic and sophisticated changes in their morphology, such as the formation of synapses to build an intricate circuit. These processes are finely tuned by AS. These exact and subtle changes occur through tuning of gene expression.

Accordingly, abnormalities in AS have been associated with various neurodevelopmental disorders, including autism spectrum disorder (ASD) (Parikshak et al., 2016; Quesnel-Vallières

et al., 2019). For example, microexons are frequently misregulated in ASD and correlate well with reduced levels of SRRM4 (Irimia et al., 2014). Loss of SRRM4 leads to disruption in neurite growth, axon guidance, and neuronal migration in the forebrain (Raj et al., 2011). Interestingly, RBFOX splicing network is also misregulated in ASD individuals (Weyn-Vanhentenryck et al., 2014).

Moreover, DSCR1 gene from chromosome 21 (triplicated in Down syndrome) has alternative first exons as detected by rapid amplification of cDNA ends (RACE) technique (which provides the full length of RNA transcript). Interestingly cDNA library screening in Down syndrome has shown that all the exons of the DSCR1 gene are flanked by consensus sequence for RBPs (Fuentes et al., 1997). However, the functional consequence of these alternative exons has not been established. In DS fetal brains and Dyrk1a overexpression mouse models, a modest amount of splicing associated transcripts are modified due to change in location of spliceosome machinery components, which results in exon inclusion in few essential synaptic genes (Toiber et al., 2010). In an RNA-seq analysis from endothelial progenitor cells originating from one DS individual and one age and sex-matched control, the authors showed large-scale evidence of DS- specific alternative splicing (Costa et al., 2011). Global changes in alternative splicing have also been reported in individuals with schizophrenia. In an RNA-seq study, more than 1000 genes showed differential splicing compared to the superior temporal gyrus of control individuals (Wu et al., 2012). Interestingly in another study on the dorsolateral prefrontal cortex (DLPFC), 798 differentially expressed transcripts were identified from 316 genes involved in the inflammatory response (Fillman et al., 2013).

Finally, AS has also been implicated in neuropsychiatric disorders and neurodegenerative diseases such as Parkinson's (Trabzuni et al., 2012), Alzheimer's (Raj et al., 2018; Rockenstein et al., 1995), Amyotrophic lateral sclerosis (ALS) (Ling et al., 2015), and bipolar disorder (BP), for example, in BP, several risk genes have been associated with splice variants. One of the prominent genes with risk variants is ankyrin-G (Ank3), which has been identified in several GWAS studies in BP (Ferreira et al., 2008).

Although there is a clear association of AS to brain disorders, to what extent the dysregulation of these neuronal exons and splicing networks contributes to various neurological disorders remains an important open question.

1.6.5. Alternative splicing in axon formation:

Neurons consist of a single axonal protrusion responsible for conducting action potential from the cell body to other cells. During axonogenesis, neurons first produce multiple, indistinguishable neurites. As the neuron matures, one of the neurites undergoes molecular and morphological changes to become the axon (Cheng and Poo, 2012) and grows into the final targeting area with the help of axon guidance molecules (Figure 1.6).

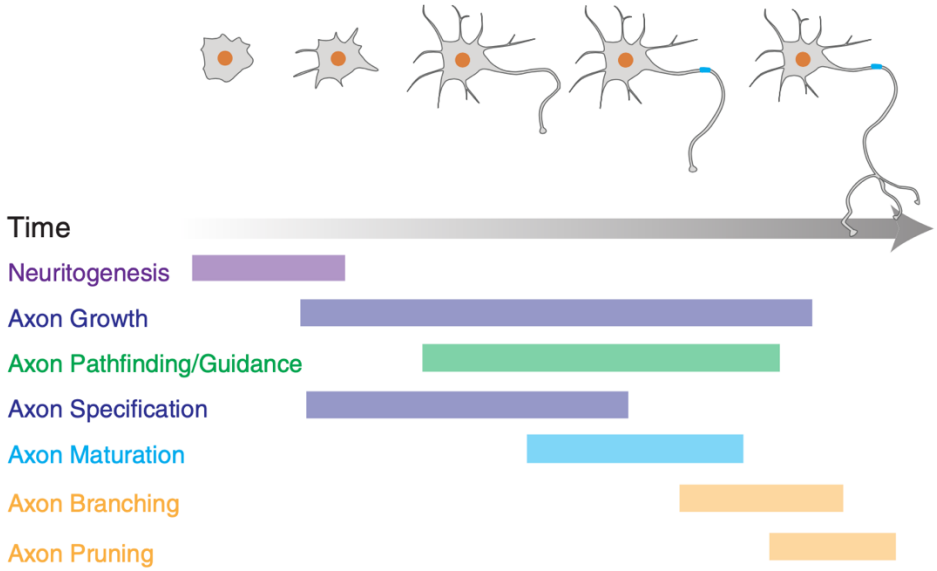


Figure 1.6. Stages of axon formation (Zheng, 2020).

Axon specification is established through morphological and molecular hallmarks of axons (Cáceres et al., 2012). For example, the axon is demarcated by Tau1 at the distal part and depletion of MAP2. After demarcation, microtubule polarity and axon initial segment acquisition occur, leading to axonal maturation (Szu-Yu Ho and Rasband, 2011). Finally, the axon forms branches and gets localized to the final destination. Notably, some of these axonal branches are pruned away during development, following an activity-dependent matching process (Zheng, 2020). Despite numerous studies on molecular players involved in

axonogenesis, we still do not fully understand the regulatory mechanisms that control the formation of a single axon, and thus the maintenance of neuronal polarity (Zheng, 2020).

One of the possible mechanisms could be alternative splicing. This level of regulation has been recently studied in more detail through an unbiased transcriptome profiling of neurons before and after the initial axon formation (Zhang et al., 2019). The authors found many genes to undergo isoform expression changes during this period and exhibit neural-specific splicing patterns. Many genes did not change their gene expression when the neurons were growing out axons but underwent splicing (Zhang et al., 2019).

For example, Shootin1 (SHTN1) changes from the long to the short isoform during axon formation, and excessive short isoforms lead to multiple axons (Kubo et al., 2015). Indeed, the short isoform (SHTN1S) is located at the axonal tip when an axon protrudes from the neuron. The SHTN1L shifts to SHTN1S form gradually, suggesting axonal growth precedes axon specification. This is because the gain and loss of function analysis for SHTN1S have shown that it is not responsible for promoting axonal growth, but SHTN1L is the one playing that role. This was supported by a previous study from Goslin and Banker in 1989, showing that axons and dendrites can switch fates even after initial axon outgrowth (Goslin and Banker, 1989). The authors performed axotomy experiments in rat hippocampal cultures. They reasoned that when the axon is 100-200 μm long, they can still become either axon or dendrite depending on each neurite's rate of growth.

Furthermore, Zhang et al. found out that RNA binding protein PTBP2 is a master regulator of axonogenesis-associated splicing. A significant number of axonogenesis-associated genes are affected by PTBP2 gain and loss (Zhang et al., 2019). Interestingly, the expression of PTBP2 peaks during early axonogenesis in the brain (Zheng, 2016; Zheng et al., 2012) and Ptbp2 $^{-/-}$ axons are shorter, and 25% of neurons can generate two or more axons. Notably, the splicing of SHTN1 is regulated by PTBP2, and the long isoform is decreased, and the short isoform is increased in Ptbp2 $^{-/-}$ neurons (Figure 1.7). Thus, alternative splicing can regulate different aspects of early axonogenesis and ensure the proper expression of isoforms of axonogenesis-related genes.

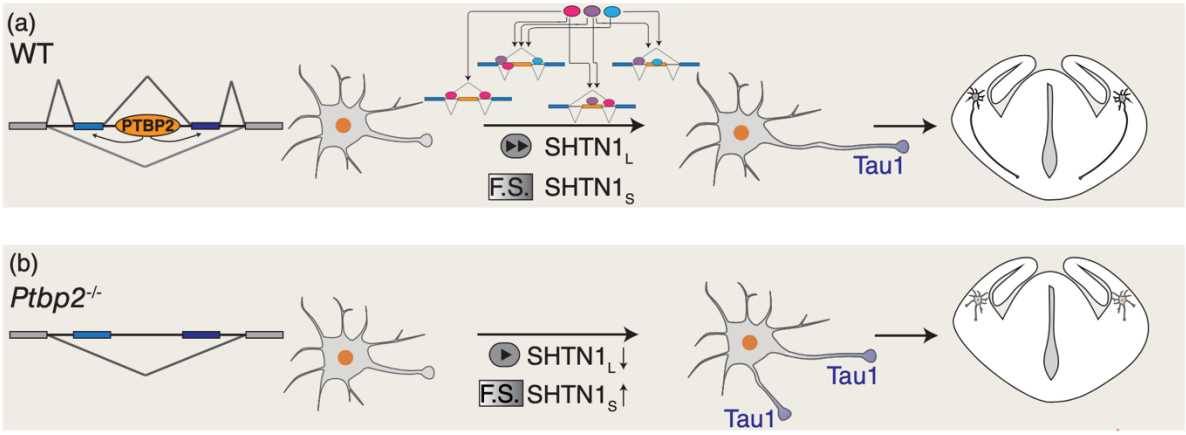


Figure 1.7. Alternative splicing regulation of early axonogenesis. (a) A neuron-specific alternative splicing program coordinates early axonogenesis (axonal growth and specification). A combinatorial and coordinated alternative splicing regulatory network encompassing many RBPs and alternative splicing events accompanies the morphological and molecular changes of axon formation. Alternative splicing of Shtn1 produces two isoforms. SHTN1L promotes axonal growth. SHNT1S promotes axon specification. FS indicates axon fate specification. (b) The axonogenesis-associated splicing program, including Shtn1, is coordinated by PTBP2. PTBP2 loss decreases SHTN1L and reduces axonal elongation. Meanwhile, SHTN1S is increased, and some Ptbp2$^{-/-}$ neurons can extend two TAU1^{+} axons (Zheng, 2020).

1.7. Role of miRNA regulation in the mammalian brain:

1.7.1. miRNAs biology:

miRNAs are 21-22 nucleotides in length, non-coding RNAs, and regulate gene expression (Bartel, 2004). They mainly interact with 3' untranslated regions (UTR) of their target genes through their interaction with seed sequences of 6-8 nucleotides (nt) at their 5' end (Lewis et al., 2005). Some miRNAs often belong to miRNA families where the members share identical seed sequences and perform similar physiological functions. For example, the let-7 family (let-7a/7b) promotes neuronal differentiation (Bian et al., 2013). miRNAs are transcribed from intra- and inter-genic regions of the genome by RNA polymerase II. The primary miRNAs (pri-miRNAs) consists of an imperfectly paired stem of about 33 nt, with a terminal loop and two flanking segments. After the pri-miRNAs are processed to precursor miRNAs (pre-miRNAs) and then to mature miRNAs, they are incorporated in the RNA-induced silencing complex (RISC) (Carthew and Sontheimer, 2009), which leads to either mRNA degradation or translation inhibition.

In particular, the pri-miRNA is first processed by RBPs DGCR8 and Drosha in the nucleus. Those remove the stem-loop to form pre-miRNA. This pre-miRNA is then exported to the cytoplasm with exportin 5 and another RNase III enzyme, RBP Dicer. This processes the pre-miRNA to remove the terminal loop to finally form mature miRNA duplexes. The duplexes associate with the RBP Argonaute (Ago) proteins in the RISC, where one strand of the duplex is degraded, and another one is retained (Sharp, 2009) (Figure 1.8). Every step of miRNA biogenesis is tightly regulated, and any dysregulation has been associated with a variety of developmental defects and diseases (Rajman and Schratt, 2017).

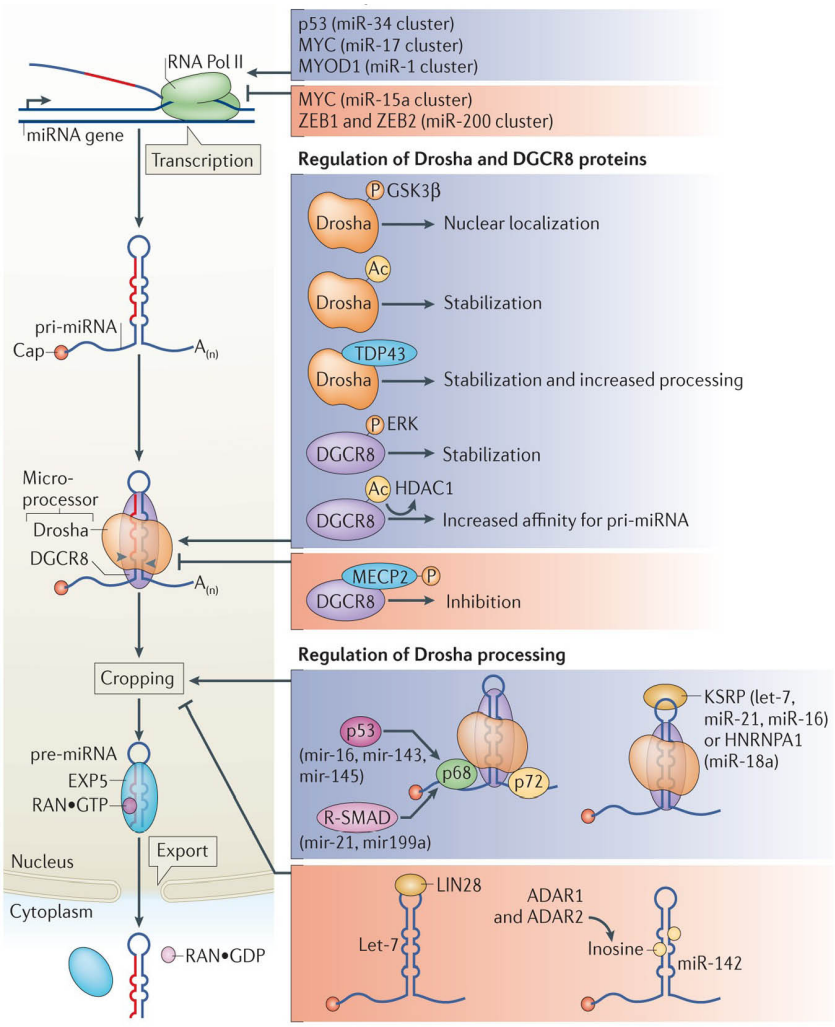


Figure 1.8. Biogenesis of miRNA in animals (Ha and Kim, 2014).

The accurate prediction of miRNA targeting is essential for miRNA function in physiology and pathology. Three types of target sites have been described in terms of conservation and efficacy: 8mer, 7mer-m8, and 7mer-A1 sites (Figure 1.9). More than 60% of human protein-

coding genes are targeted by one or more miRNA, and each miRNA could target multiple mRNAs. Similarly, each mRNA could be targeted by multiple miRNAs resulting in a complex regulatory network (Bartel, 2009).

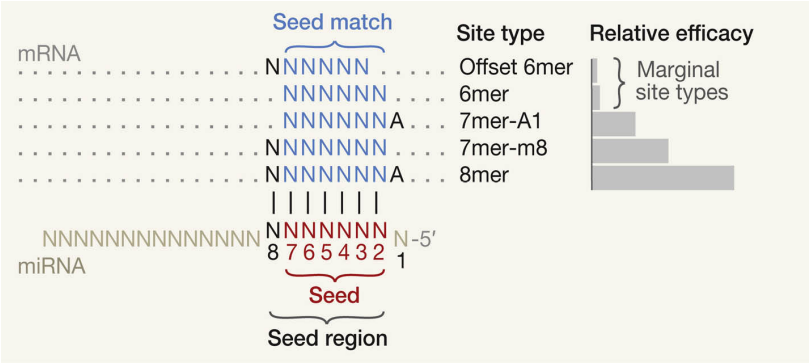


Figure 1.9. microRNA target sites. These sites each have 6-7 contiguous Watson-Crick pairs (vertical lines) to the miRNA's seed region (miRNA positions 2-8). Two of these sites also include an A at position 1. Relative site efficacy in mammalian cells is graphed to the right (log scale). The most effective sites are 7-8 nt sites that include a perfect match to the miRNA seed (positions 2-7, red), whereas the 6 nt sites are the least effective (Bartel, 2018).

1.7.2. RNA binding proteins and microRNAs:

Whereas RBPs such as DGCR8 and Drosha are very well characterized in the processing of pri-miRNAs, other RBPs in miRNA biogenesis has not been completely defined. Few RBPs that play specifically targeted roles include DKC1, which promotes excision of snoRNA-derived miRNAs from snoRNA host genes (Scott et al., 2009), and LIN28A, which inhibits the differentiation-inducing let-7 family expression in stem cells by interacting with pri-miRNA and pre-miRNA (Heo et al., 2009). RBPs can positively and negatively impact miRNA biogenesis and could affect different steps of the processing. This has led to identifying 116 RBPs that could influence miRNA processing (Nussbacher and Yeo, 2018). For instance, hnRNPA1 stimulates microprocessor processing of miR-18a by changing the structure of pri-miR-18a and make it more accessible to Drosha (Guil and Cáceres, 2007). On the other hand, hnRNPA1 negatively regulates let-7a-1 by binding to the conserved terminal loop (Michlewski and Cáceres, 2010). Moreover, Nussbacher and Yeo have recently done a systematic analysis to identify RBP-pre-miRs interaction. RBPs can also regulate miRNA RISC loading. For instance, TDP43, a RBP associated with ALS, disrupts the loading of miR-1 and miR-206 in the RISC complex (King et al., 2014). These findings highlight that many RBPs could interact and

function in intermediary steps of miRNA processing, providing a broad layer of regulation of miRNA activity and abundance.

1.7.3. miRNA in nervous system development:

miRNAs represent an essential layer of post-transcriptional regulation, which fine-tunes expression profiles of genes necessary for neuronal development, maturation, and synaptogenesis. They have also been recognized as critical players in the formation and retrieval of memory and are associated with LTP and LTD (Hu and Li, 2017; Sambandan et al., 2017). During synapse development, some miRNAs induce mRNA degradation, and others are involved in the regulation of mRNA translation (Schratt et al., 2006; Swanger and Bassell, 2011). In polarized neurons, miRNA regulation can occur at the global level or specifically in axons or dendrites. These studies highlight the importance of miRNAs in the strict spatiotemporal regulation of gene expression, which is essential for neural circuit formation (Olde Loohuis et al., 2012; Wang et al., 2012; Ye et al., 2016).

Several miRNAs are differentially expressed across developmental periods, and differential expression across regions increases over developmental time (Ziats and Rennert, 2014). Putative targets of these differentially expressed miRNAs are involved in biological processes such as neurodevelopment and transcriptional regulation. Some miRNAs are dynamically expressed during development resulting in some miRNAs being continuously expressed across all developmental phases (miR-9 and let-7), and few expressed for a short time. This suggests that the latter group of miRNA might be involved in time-specific processes during development, such as proliferation, migration, and network formation (Barca-Mayo and De Pietri Tonelli, 2014). In particular, miRNAs can regulate neurite/synapse formation and function. For example, BDNF is downregulated by miR-375 to inhibit dendritic growth (Abdelmohsen et al., 2010), miR-138 regulates the palmitoylation of several proteins functioning at synapses (Siegel et al., 2009), and miR-124 enhances long-term facilitation by de-repressing activity-dependent transcription factor CREB (Rajasethupathy et al., 2009).

Materials and methods

2.1. Human samples RNA sequencing:

2.1.1. Sample preparation:

Human brain tissues for DS and control individuals were obtained from NIH NeuroBioBank. Brain sample and donor information are available in Table 1 and 2 in the appendix. No statistical tests were done before the acquisition of samples. A piece of the human hippocampus and cortex from both DS and control individuals were cryo-pulverized with the help of pestle and mortar. The crushed tissue was collected in a 1.5 ml eppendorf tube. About 50 mg of crushed tissue was collected in another 1.5 ml Eppendorf and kept on dry ice until use. Briefly, RNA was extracted with QIAzol reagent and purified on miRNeasy spin columns (QIAGEN). RNA samples were quantified at 260 nm with an ND1000 Nanodrop spectrophotometer (Thermo Scientific). RNA purity was also determined by absorbance at 280 and 230 nm. All samples showed A260/280 and A260/230 ratios greater than 1.9. RNA quality and integrity were verified by microfluidic assay with Agilent 2100 Bioanalyzer. The RNA integrity number (RIN) ranged between 2.1 and 7.3.

2.1.2. Library preparation for whole transcriptome sequencing:

500 ng RNA was used to prepare sample libraries. Briefly, Illumina's TruSeq stranded total RNA library preparation kit was used. First, ribosomal RNA (rRNA) was depleted using Ribo-Zero rRNA removal beads. Then the RNA is fragmented, and the first-strand cDNA synthesis is initiated. Following this, second-strand cDNA is synthesized, and 3' ends are adenylated. It is followed by adapter ligation to the cDNA strands, amplification, and purification of DNA fragments. All the sample libraries are then quantitated on Agilent 2100 Bioanalyzer. The size in bp and concentration (ng/μl) of the peak corresponding to the amplified products are taken for normalizing the libraries. The normalized libraries were then pooled together in a process called multiplexing (Figure 2.1). The pooled libraries were then loaded onto the Illumina S2 flow cell. Paired-end sequencing for 100 bp each was performed on NovaSeq 6000 system to

obtain around 100 million reads per sample. The basecall files were then converted to fastq files and are demultiplexed to obtain each sample for further analysis.

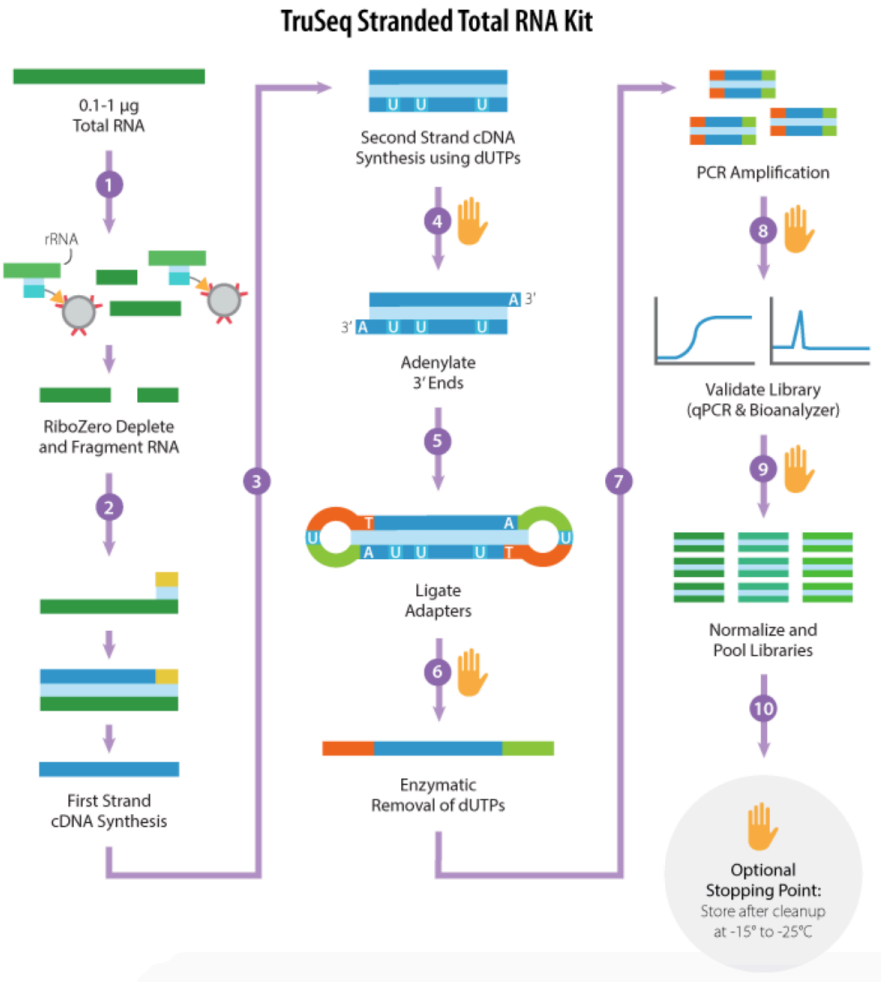


Figure 2.1. Workflow for preparing total RNA sequencing libraries.

2.1.3. RNA-seq read alignment and quality check:

First, all the raw reads were run through FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FastQC output was compiled by use of multiQC and visually inspected to look for "sequence quality", "per tile sequencing quality", "overrepresented sequences" , "adapter content" and other quality parameters. All the samples with good quality parameters were carried forward for the alignment. Fastq files were aligned with STAR v2.7.3a (Dobin et al., 2013) to *Homo sapiens* GRCh38 genome, Ensembl version 99 (Yates et al., 2020). The fastq files were mapped using 2pass mapping parameters. In the first pass mapping, each file was mapped to the reference

genome with 1pass mapping parameters. Then, the mapped spliced junctions for each sample were extracted and concatenated. The junctions originating from mitochondrial genes were filtered out. A new index was created based on the concatenated splice junctions, and then the 2pass mapping was performed on the 1pass mapped files. The output of the 2pass mapping consisted of bam files and reads per gene information for each sample. These were the raw counts to be used for further analysis. The bam files were also checked for quality controls using PicardTools v2.23.6 (http://broadinstitute.github.io/picard). The number of reads obtained for each sample is available in Table 3 in the appendix.

2.1.4. Differential gene expression:

The raw counts from STAR were imported in RStudio (RStudio Team (2020)) for performing differential expression analysis. Differential expression analysis was performed using edgeR v3.32.1, a statistical package based on generalized linear models. The normalization was performed using the Trimmed means of M values (TMM) method (Robinson et al., 2010). The differential expression analysis was performed for those genes which were present in at least "n-1" samples for the group with less number of samples. The counts for filtered genes were subsequently used to calculate $\log_2$ fold changes for DS *vs.* control samples, and the significant genes were taken to be those with a false discovery rate (FDR) of <0.05.

2.1.5. Expression profile visualization:

The differentially expressed genes were visualized onto different chromosomes with the help of rtracklayer (Lawrence et al., 2009) and ggbio (Yin et al., 2012) R packages. Further, the gene expression profile for all the genes was visualized using karyoploteR (Gel and Serra, 2017) and regionR (Gel et al., 2016) R packages to obtain genome-wide dysregulation domains (GEDDs). The identification of GEDDs for each chromosome was validated by performing loess regression with 3% smoothing and 1000 bootstrap tests. The genes located in the GEDDs and their $\log_2$ fold changes and start location were extracted into a table for each chromosome separately. The default parameters from the original study were used to calculate the start and end of each GEDD (Letourneau et al., 2014). The number of GEDDs for each chromosome from hippocampus and cortex gene expression data is available in Table 5 in the appendix.

2.1.6. Identifying biological pathways:

clusterProfiler (Yu et al., 2012) R package was used to perform Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) on the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa et al., 2016). The gene expression profile for each brain region was tested against the KEGG pathways database with 10000 permutations, and p-value cutoff <0.05, and the enrichment was plotted using the ridgeplot function. Pathway maps for enriched pathways were plotted using pathview (Luo and Brouwer, 2013) R package with human pathway IDs.

2.2. Gene Ontology analysis:

GO enrichment analysis was performed using The Database for Annotation, Visualization, and Integrated Discovery (DAVID) (Huang et al., 2009) with a high level of stringency among three different levels (medium, high and highest). The GO analysis was performed for biological processes (BP), cellular components (CC), and molecular functions (MF), and the results were multiple-testing corrected with the Benjamini-Hochberg test (Benjamini and Hochberg, 1995). Heatmaps were created using the ComplexHeatmap R package (Gu et al., 2016).

2.3. Cell-type enrichment analysis:

We integrated the Zeisel et al. (Zeisel et al., 2015) data with cell-type-specific transcriptome signatures from mouse cortex and hippocampus with the bulk RNA-sequencing data for the human hippocampus and cortex that we generated. The Expression Weighted Cell-type Enrichment (EWCE) R package (Skene and Grant, 2016) was used to obtain this integration. Two ranked gene lists, one for significantly upregulated genes for both the regions and another for significantly downregulated genes for both regions separately, were ordered according to their Log_2 Fold Change (FC). In EWCE, random samples were obtained by reordering the ranked list 100,000 times. Transcript length and GC content were controlled during the analysis. Data were represented as Z-scores, i.e., standard deviations from the mean. Values below 0 indicate depletion in expression, which has been assigned the value of 0. Heatmaps were created using superheat R package (Barter and Yu, 2018).

2.4. Weighted Gene / Protein Co-expression Network Analysis:

R package WGCNA (Langfelder and Horvath, 2008) was to construct co-expression networks, as previously done (Oldham et al., 2008). Briefly, the soft power parameter is chosen between 6 and 9 for the signed hybrid networks when the network achieves scale-free topology (R^2>=0.9) and a negative slope. Networks were constructed using the 'blockwiseModules' function. Then the topological overlap matrices (TOM) are constructed and clustered hierarchically using average linkage hierarchical clustering using "1-TOM" as dissimilarity measure (dissTOM). TOM reflects how close the neighbors of a gene are to neighbors of another gene, i.e., network interconnectedness. Modules were defined as branches of dendrogram using hybrid dynamic tree cutting method with biweight midcorrelation (bicor), with a minimum module size of 30, deepsplit of 2. The modules generated initially were merged based on module eigengenes using correlation-based adjacency as dissimilarity matrix. Modules with a distance of <0.25 were merged into a single module.

Each module was summarized by module eigengene (ME), which reflected the characteristic expression profile of a module. For each gene, the module membership measure (kME) was defined as the correlation between gene expression values and the module eigengene (ME). To assess the statistical significance of module membership, a Bonferroni corrected p-value was also calculated. Further derivation of intramodular connectivity measures helped identify the key players in the network, which were the hub genes. Module-Trait associations are computed with biological (Group: DS or control, Sex: Female or male, Age in years) and technical traits such as postmortem interval (PMI) in hours. Modules were considered to be significant with any trait if the Bonferroni corrected p-value <0.05. The Cytoscape edges and nodes parameters were obtained from modules with significant module-trait association with group trait. These nodes and edge parameters for significant modules were imported in Cytoscape (Shannon et al., 2003). The genes present in each module were used for creating interaction networks, cell-type enrichment, and gene ontology analysis. Correlation analysis was performed between the modules using the module membership values and Pearson's correlation test in Rstudio.

2.5. Human samples Quantitative PCR:

Reverse transcription was performed according to the manufacturer's recommendations on 1 μg of RNA with the QuantiTect Reverse Transcription Kit (QIAGEN), including a genomic DNA–removal step. SYBR green qRT- PCR was performed in triplicate with 10 ng of template cDNA using QuantiTect Master Mix (QIAGEN) on a 7900-HT Fast Real-time System (Applied Biosystems) and using the following universal conditions: 5 min at 95°C, 40 cycles of denaturation at 95°C for 15 s, and annealing/extension at 60°C for 30 s. Product specificity and occurrence of primer dimers were verified by melting-curve analysis. Primers were designed with Beacon Designer software (Premier Biosoft) to avoid template secondary structure and significant cross-homology with other genes by BLAST search. For each target gene, primers were designed to target all possible transcript variants annotated in the RefSeq database (https://www.ncbi.nlm.nih.gov/refseq).

In each experiment, no-template controls and RT-minus controls were run in parallel to the experimental samples. The PCR reaction efficiency for each primer pair was calculated via the standard curve method with four serial-dilution points for cDNA (32, 8, 2, and 0.5 ng). The PCR efficiency calculated for each primer set was used for subsequent analysis. All experimental samples were detected within the linear range of the assay. Gene-expression data were normalized via the multiple-internal-control-gene method (Vandesompele et al., 2002). To determine an accurate normalization factor for data analysis, we evaluated the expression stability of different control genes with the GeNorm algorithm available in qBasePlus software (Biogazelle). The tested control genes were GAPDH (glyceraldehyde-3-phosphate dehydrogenase) and ACTB (actin B). Based on the relative expression stability of the control genes calculated via GeNorm analysis, expression data for the different samples were normalized with GAPDH and ACTB. The list of oligonucleotides with primer efficiencies and sequences is available in Table 8 in the appendix.

2.6. Mass-Spectrometry Based Proteomics:

2.6.1. Sample preparation:

Samples of the human DS hippocampus and cortex and their age-/sex- matched healthy controls were lysed in Radio immunoprecipitation assay (RIPA) buffer. The prepared samples were loaded directly into the separation column, and the peptides were eluted. The peptide separations were carried out at 55°C by a 75-µm ID×50cm 2µm, 100 Å C18 column-mounted in the thermostatic column compartment of the machine. Eluting peptides were electrosprayed and analyzed by tandem mass spectrometry.

2.6.2. Proteomic data analysis:

The raw data were processed with MaxQuant software (Cox and Mann, 2008). A false discovery rate (FDR) of 0.01 was requested for the identification of proteins. Quantification in MaxQuant was performed using the built-in label-free quantification algorithm (Luber et al., 2010), enabling the "Match Between Runs" (Nagaraj et al., 2012). All proteins and peptides matching to the reversed database were filtered out. Label-free protein quantitation (LFQ) was performed with a minimum ratio count of 1 (Cox et al., 2014).

2.6.3. Proteomic bioinformatic analysis:

All bioinformatics analyses were performed with the Perseus software of the MaxQuant computational platform (Tyanova et al., 2016). Protein groups were filtered to require 100% valid values in at least one experimental group. The label-free intensities were expressed as base $\log_2$, and empty values were imputed with random numbers from a normal distribution for each column to best simulate low abundance values close to the noise level. A Student's t-test with permutation-based FDR statistics was run for each sample and 250 permutations were performed, with an s0 of 0.1 and required an FDR of 0.05.

2.7. Transcript quantitation:

Transcripts for each sample were quantified using RSEM (Li and Dewey, 2011), and the differential expression for transcripts was performed with the EBSeq R package (Leng et al., 2013).

2.8. Alternative splicing detection:

We performed alternative splicing detection using VAST-TOOLS and vastdb version hs2.23.06.20 (Tapial et al., 2017). First, each sample fastq file was aligned to the human vastdb database with default parameters. These aligned outputs were merged according to the group status (i.e., DS or control). The 'vast-tools combine' command was used to create a final output file. This output file was taken to calculate different splicing for five different kinds of alternative splicing events (ALTD, ATLA, EX, MIC, IR) using the 'tidy' module. The tidy module extracts inclusion levels (events) for each exon from all the genes with their percent spliced in values (PSI). Events were significant when average changes between the two groups with percent spliced in value |deltaPSI|>10 with sufficient read coverage in >80% samples from each group. For the intron retention events, percent intron retention (PIR) values are reported.

2.9. RT-PCR validation of splicing quantification:

1 μg of RNA for each sample was reversed transcribed by Qiagen QuantiTect RT kit in 20ul reaction. A control cDNA was PCR amplified for a different number of cycles (20, 25, 30, 35, and 40) to test for the number of cycles before the product reached saturation and resolved by electrophoresis on 2% agarose gel. After obtaining the desired number of cycles for each gene, all the control and DS human sample cDNAs were PCR amplified and run on agarose gels. Gels were quantified using ImageJ (NIH). The list of primers with sequences and number of cycles used is available in Table 6 in the appendix.

2.10. Agarose Gel Electrophoresis:

All agarose gels for electrophoresis were prepared using 1X TBE supplemented with ethidium bromide. The 2% gels were run with 1X TBE running buffer for 40 minutes at 80V.

2.11. Small RNA sequencing:

2.11.1. Small RNA library preparation:

500 ng RNA was used to prepare small RNA sequencing libraries with the help of NEXTFLEX Small RNA-seq kit v3 (PerkinElmer). This kit utilizes adapters with 4 nucleotides (4N) long randomized ends in order to reduce ligation bias. First, the RNA was denatured at 70^{o}C, and the 3' 4N adapter was ligated. The excess adapter was removed, and then the 5' 4N adapter is ligated onto the sequences. This is followed by reverse transcription to generate cDNA and was purified using bead cleanup. The purified cDNA was PCR amplified with 18 cycles and barcoded primers. Then gel-free cleanup was performed to obtain the final libraries (Figure 2.2). The libraries were quantified with Qubit, and quality checked on Agilent bioanalyzer. Like total RNA sequencing, the libraries were normalized and pooled to run on Illumina S1 flowcells on the Novaseq 6000 instrument.

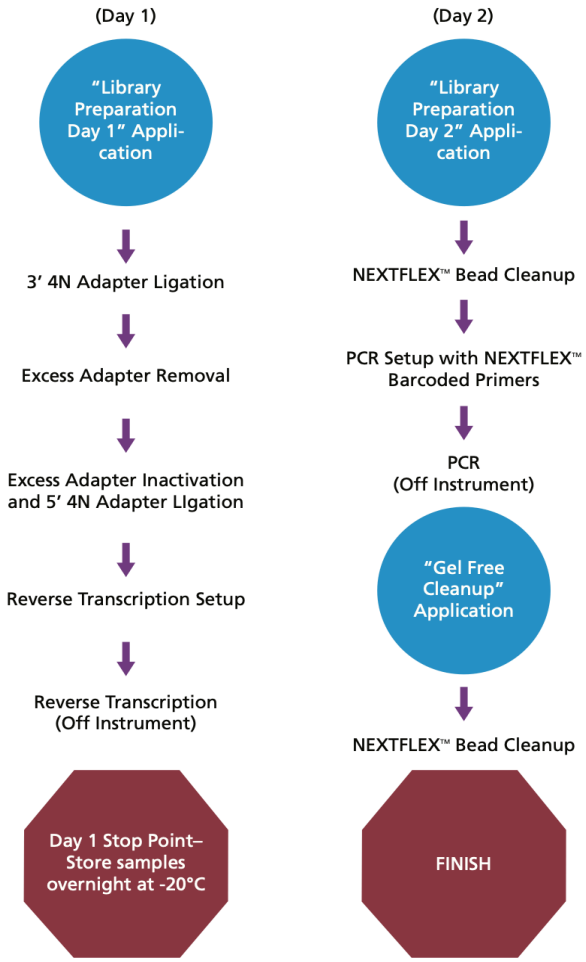


Figure 2.2. Workflow to prepare small RNA libraries

2.11.2. Small RNA alignment, differential expression, and target prediction:

Like total RNA sequencing, the basecalls files were converted to fastq files and checked for quality control using FastQC. The adapter sequences from fastq files for each sample were trimmed using Cutadapt v1.16 (Martin, 2011), followed by trimming 4 nucleotides from each end. The trimmed fastq files were mapped using miRge2.0 (Lu et al., 2018) onto the miRBase (Kozomara et al., 2019). Subsequently, the counts were obtained for each miRNA, and they were used to perform differential expression analysis with edgeR package.

To look for targets of these differentially expressed miRNAs, conserved family predictions were downloaded from TargetScanHuman v7.2 (Agarwal et al., 2015). The targets for each differentially expressed miRNA were filtered out separately from the downloaded file. The

predicted targets were then annotated with log_2FC and FDR from RNA-seq and proteomics experiments to find differentially expressed targets.

2.12. Primary culture:

The primary hippocampal culture was prepared from wildtype and Ts65Dn postnatal day 2 pups as described. Brains were dissected under a stereomicroscope in ice-cold dissection buffer (Hank's Balanced Salt Solution (HBSS) with 6 mg/ml glucose, 3 mg/ml bovine serum albumin (BSA), 5.5 mM $MgSO_4$, 5 ug/ml gentamycin, and 10 mM HEPES, pH 7.4). Hippocampal tissue was minced and enzymatically digested with 0.25% trypsin in HBSS consisting of 0.6 mg/ml DNase for 5 mins at 37°C. Tissue chunks were washed in DB, incubated for 5 min in DB supplemented with 1 mg/mL of Soybean trypsin inhibitor (Sigma), and mechanically dissociated in DB supplemented with 0.6 mg/mL DNase. Cells were passed through a 40 µm cell strainer and then centrifuged (*110 x g* for 7 min at 4°C) to remove cellular debris. Cells were plated on glass coverslips, 6-wells plates, or MEA coated with 0.1 mg/ml poly-L-lysine in 100 mM borate buffer, pH 8.5 at a density of 250-500 cells/mm. Neurons were maintained in a culture medium consisting of Neurobasal-A supplemented with 2% B27, 1% GlutaMax and 5 µg/mL gentamycin (all from Gibco) at 37 °C in humidified atmosphere (95% air, 5% CO_2).

2.13. Immunocytochemistry:

The primary hippocampal culture was grown on sterilized, and poly-L-lysine coated coverslips. The cells were fixed at days in vitro (DIV14) with 4% paraformaldehyde in PBS for 20 minutes at room temperature. Fixed cells were washed 4 times 10 minutes each with PBS to wash away the PFA. The cells were then permeabilized with 0.1% Triton/PBS (PBST) for 3 times 10 minutes each. This was followed by blocking with 5% normal goat serum (NGS) in PBST for 1 hour and then primary antibody incubation overnight at 4°C. The coverslips were washed 4 times 10 minutes each with PBST and followed by 2 hours incubation at room temperature with secondary antibodies. Coverslips were mounted on the glass slides with ProLong Gold with DAPI and imaged on Leica SP5 confocal microscope using 40X oil objectives. The cells were imaged using 0.5 µm z-steps, and the images were projected into a single plane using

maximum intensity projections. The list of antibodies used for immunocytochemistry is available in Table 7 in the appendix.

2.14. Axon initial segment analysis:

The maximum projected images were imported into MATLAB software (Mathworks) for AIS analysis using custom scripts as previously developed by Grubb's lab (Grubb and Burrone, 2010). The script determines the AIS length based on the AnkG intensity profile along a traced path. The AIS start and end is based on the intensity reaching 1/3 of the maximum intensity. For each genotype (wildtype and Ts65Dn), each neuron was analyzed to measure the AIS length. The neurons were also manually counted in ImageJ (NIH) containing no axons, single axons, or with two or more axons.

Results

3.1. Large gene-dysregulation in human hippocampus and cortex in DS *vs.* control individuals.

To start our investigation on new molecular targets that could rescue cognitive impairment in DS, we first performed total RNA sequencing on 11 control and 9 DS hippocampal human samples. Among all the RNA biotype entities identified by the RNA sequencing, we found the majority of them were protein-coding genes followed by long non-coding RNAs (lncRNAs) and then pseudogenes and other RNAs such as unprocessed pseudogenes, miscellaneous RNAs (misc RNA), small nuclear RNAs (snRNAs), microRNA (miRNA), To be experimentally confirmed (TEC) RNAs, small nucleolar RNAs (snoRNAs), and others small RNAs (Figure 3.1). We restricted our analysis to the protein-coding genes. The expression levels of these genes were normalized according to the library size for each sample to obtain counts per million (CPM) reads. After acquiring the normalized counts for each gene within each specimen, we performed a quality control check by performing a Principal Component Analysis (PCA, which indicates variation in the data, i.e., how well individual samples from each group vary among themselves), a hierarchical clustering analysis (which assembles samples that are more similar to each other than the samples from another group) and sample RNA integrity number (RIN, which gives a measure of RNA integrity, i.e., is the RNA stable or degraded based on the ratio of 18S to 28S ribosomal subunits) on all data. So, we discarded samples based on a combination of these three criteria. In particular, we discarded outliers by PCA and hierarchical clustering analysis, and we excluded samples with degraded RNA. We eventually retained RNA-seq data from 6 controls and 6 DS individuals for the hippocampus. Similarly, cortical samples were checked and filtered from initial 12 control and 10 DS samples to retain 8 control and 8 DS samples.

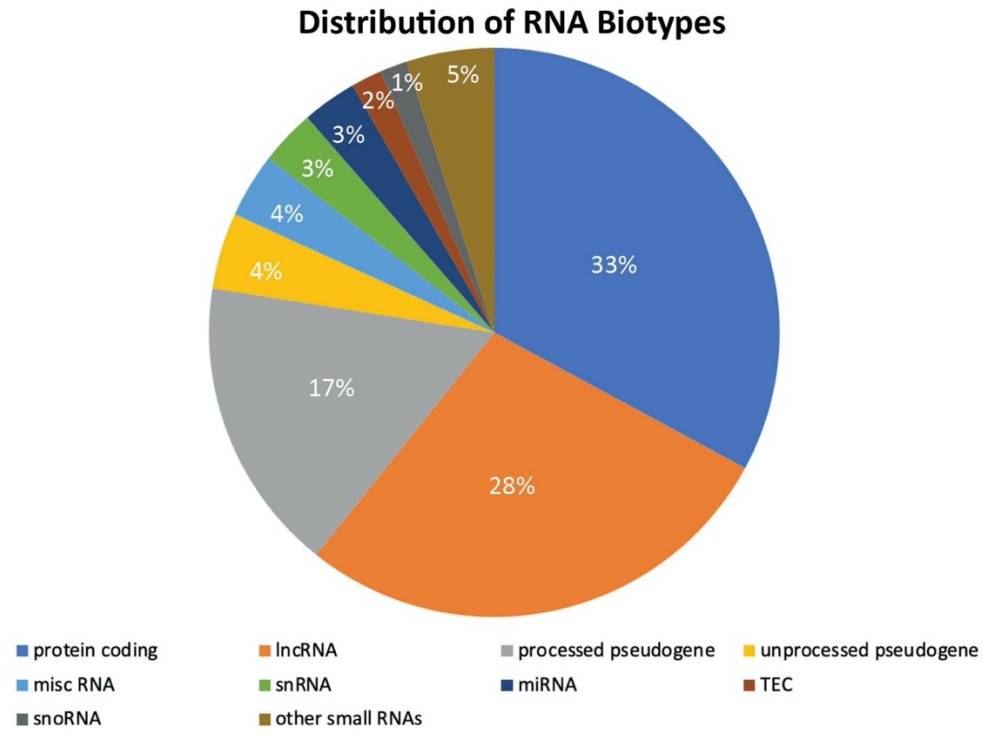


Figure 3.1. Diverse RNA biotypes identified from total RNA-seq experiments. In the human samples, diverse RNA biotypes were identified and are shown with the help of a pie chart. lncRNA: long non-coding RNA, misc RNA: miscellaneous RNA, snRNA: small nuclear RNA, miRNA: microRNA, TEC: To be Experimentally Confirmed RNA, snoRNA: small nucleolar RNA. The numbers inside the pie chart shows the percentages of each biotype.

To investigate gene expression changes in the DS hippocampus and cortex *vs.* control individuals at the level of a single gene, we carried out a differential expression analysis separately for both regions. We defined differentially expressed genes (DEGs) by a false discovery rate (FDR) of < 0.05 (Figure 3.2). Overall, we found 2622 up- and 2205 down-regulated genes in the hippocampus and 2769 up- and 2779 down-regulated genes in the cortex (Figure 3.3A, 3.3B). Of the 180 triplicated genes present on the Chromosome 21, excluding the keratins associated proteins (KAPs), 63 and 88 genes were differentially expressed in DS *vs.* control samples in the hippocampus and cortex, respectively. All differentially expressed triplicated genes were upregulated (Figure 3.3C).

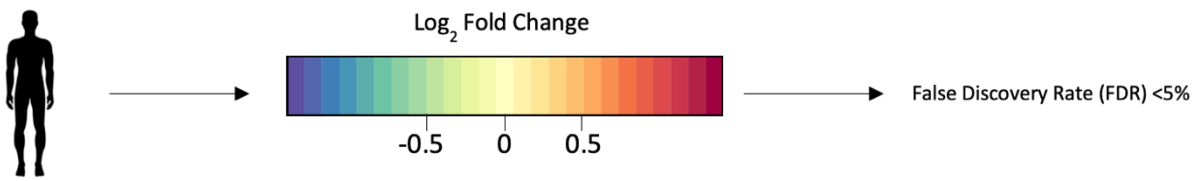


Figure 3.2. Differentially expressed genes defined by FDR <5%. Schematic diagram to represent differentially expressed genes that we defined by FDR <5%.

We then compared DEGs from the hippocampus and cortex to determine whether there are region-specific changes in DS *vs.* control samples. We found that the number of up-regulated and downregulated genes in the hippocampus was less than in the cortex. Moreover, there were more up-regulated genes than down-regulated genes in the hippocampus, but almost an equal amount of genes were up-regulated and downregulated in the cortex. Notably, many differentially expressed genes were shared between the hippocampus and cortex (i.e., 1852 up-regulated, 1654 downregulated, and 30 triplicated genes; Figure 3.3). Altogether, these results suggest a high level of (common) dysregulation in the hippocampus and cortex in DS individuals.

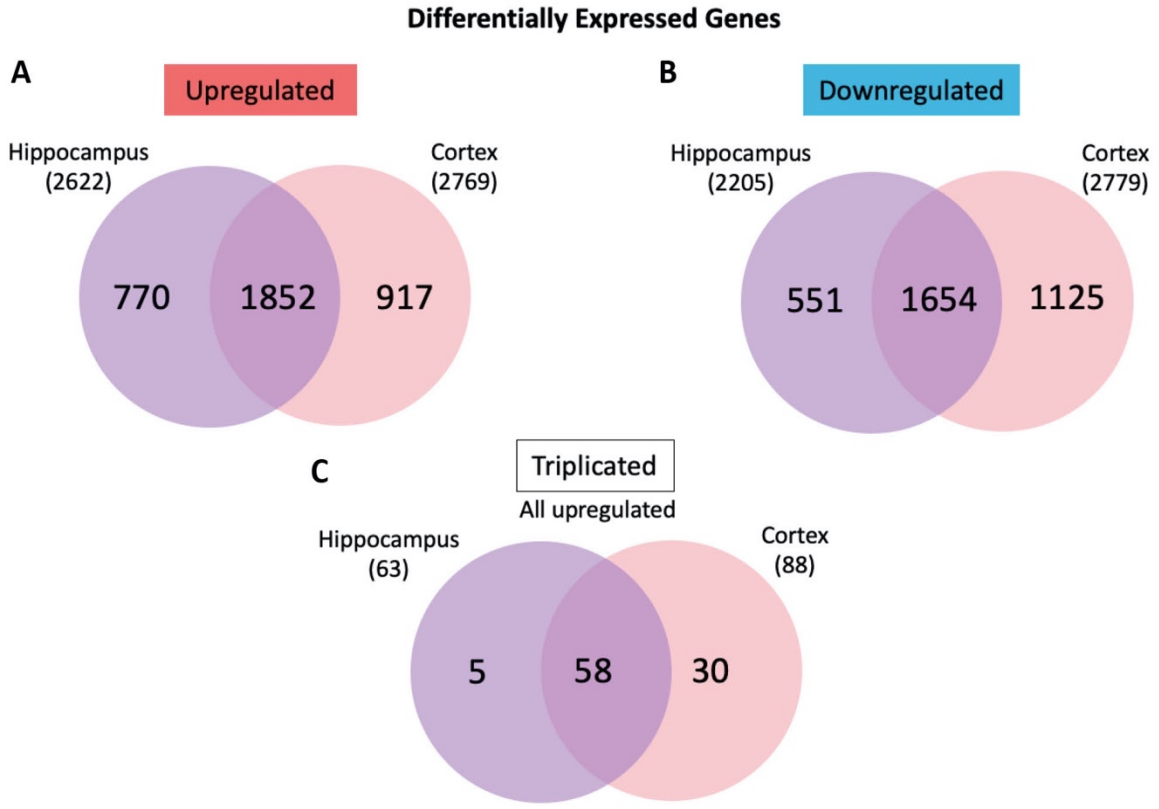


Figure 3.3. Gene expression dysregulation in brain samples from individuals with DS. Venn diagram of differentially expressed genes in human hippocampus and cortex reveals common and non-common upregulated, downregulated and triplicated genes between the hippocampus and the cortex.

3.2. Differentially expressed genes belong to biologically significant processes and pathways.

To understand the biological significance of these differentially expressed genes, we next performed a gene ontology (GO) analysis on the DEGs separately for the hippocampus and cortex. GO analysis is an unbiased enrichment analysis that reveals the representation of a set of genes in different biological processes (BP), cellular components (CC), and molecular functions (MF). In particular, we analyzed the upregulated and downregulated genes for each brain region individually. We found many processes shared between the two regions (Figure 3.4). For example, biological processes such as cell death, collagen metabolic process, interferon-gamma mediated signaling, locomotion, phosphorylation, and cellular components such as extracellular vesicle and cell-substrate junction were shared by the upregulated genes in both regions.

Similarly, many shared GO terms for downregulated genes include cell projection morphogenesis, macromolecular complex assembly, mRNA splicing, protein localization, translation, transport, and respiratory electron transport chain (Figure 3.5). On the other hand, we also found region-specific processes and functions. For example, for upregulated genes, we found cellular response to stimulus, ensheathment of neurons, neurogenesis, endoplasmic reticulum specific to the hippocampus, and terms like activation of immune response, hematopoiesis, histone deacetylation, mononuclear cell proliferation, MAPK cascade, and viral process specific to the cortex. For downregulated genes, axon development, RNA capping, and cellular biosynthetic process were hippocampus specific. Simultaneously, neurotransmitter secretion, synaptic signaling, proton transport, and regulation of mitotic spindle assembly were cortex specific biological processes. These DEGs provide a molecular signature for distinguishing between the DS and control brains with regional level resolution.

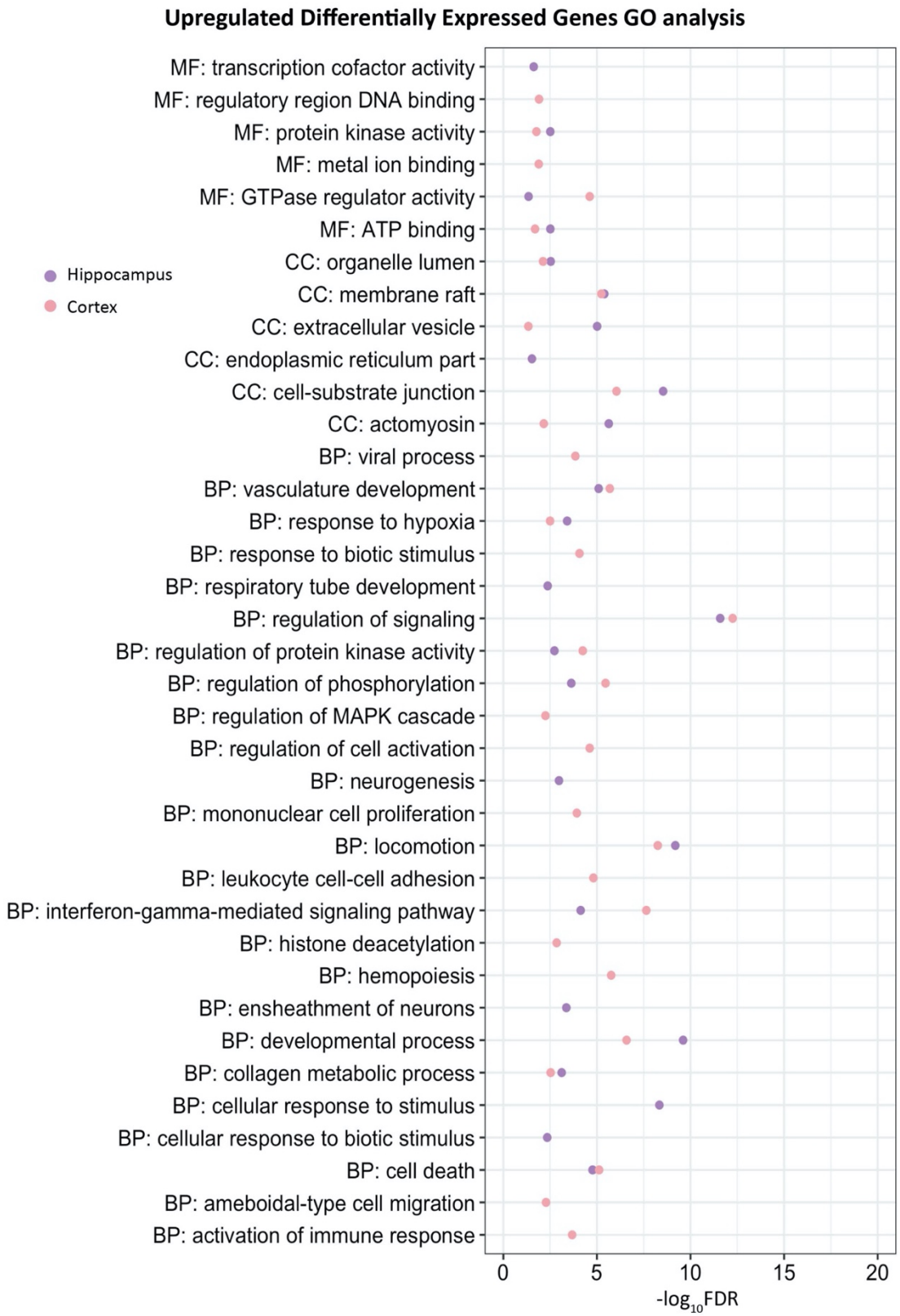


Figure 3.4. Differentially expressed upregulated genes show region-wise conserved and specific GO terms. Gene ontology (GO) analysis for upregulated genes for both hippocampus and cortex shows high levels of overlapping terms, but also specific ones. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate.

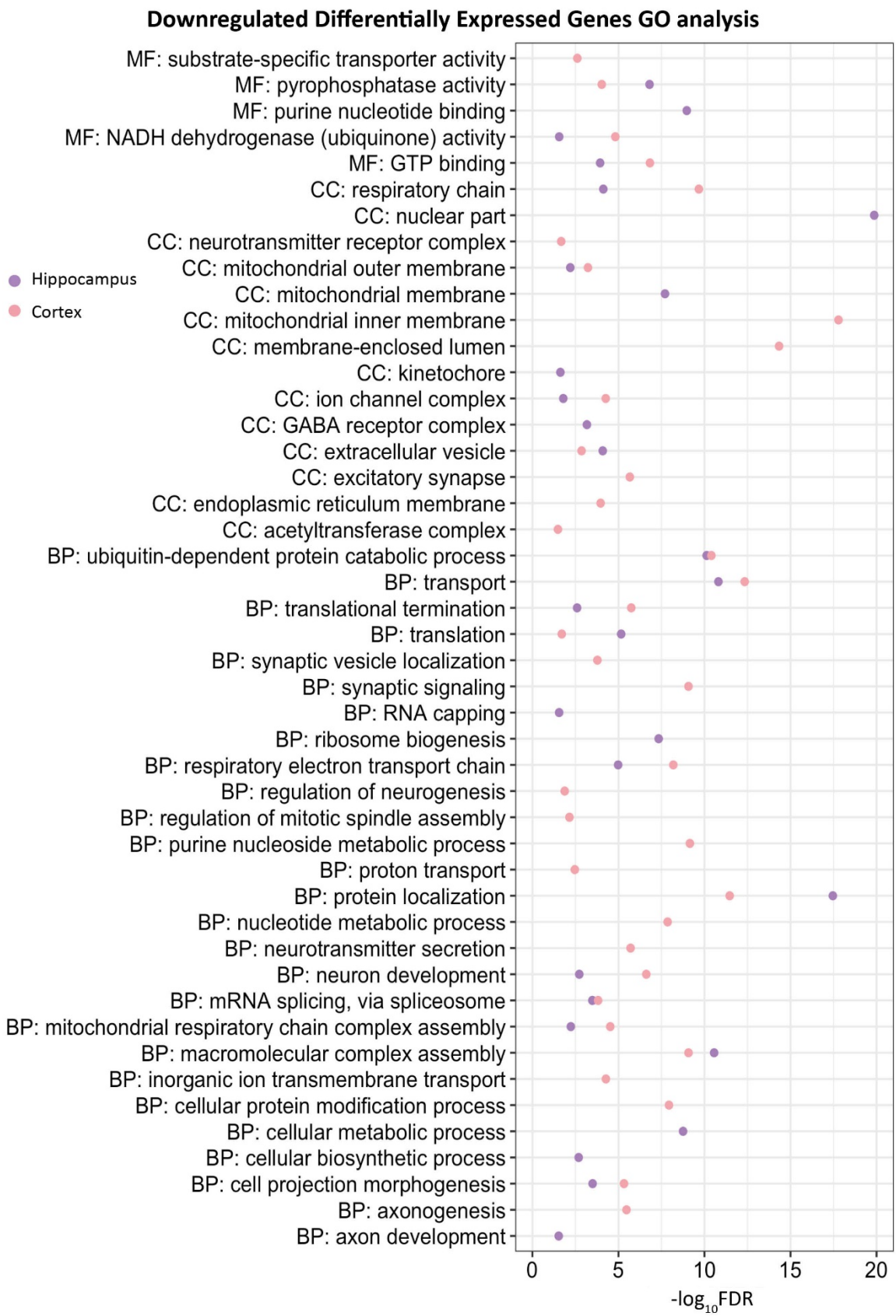


Figure 3.5. Differentially expressed downregulated genes show region-wise conserved and specific GO terms. Gene ontology (GO) analysis for downregulated genes for both hippocampus and cortex shows some overlapping terms and few specific ones. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate.

Many different studies have reported dysregulation in several biological pathways in DS (Cairney et al., 2009; Yu et al., 2020) and have found cell- and tissue-specific pathways to be disturbed. To look for dysregulated pathways in our datasets, we took advantage of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database originating from literature. The KEGG pathways database consists of manually curated pathway maps to find gene interactions and network relationships. To establish the enrichment of genes highlighted by our RNA seq analysis in KEGG pathways, we performed a Gene Set Enrichment Analysis (GSEA). It considers all the genes according to their Log_2 fold change (FC) irrespective of their statistically significant differential expression status. In the hippocampus, we found that the genes with $log_2FC > 0$ were enriched in mostly immunological pathways such as complement and coagulation cascades, staphylococcus aureus infection, and antigen processing and presentation, cytokine-cytokine receptor interaction, and many more. On the other hand, the genes with $log_2FC < 0$ were enriched in ribosomal and proteasomal pathways (Figure 3.6). These results suggest immune dysregulation and protein synthesis, and degradation defects in the hippocampus. In figure 3.7, we showed two KEGG pathways as an example (i.e., complement and coagulation cascades; Figure 3.7A) where most of the genes were upregulated and another example (i.e., RNA transport; Figure 3.7B) where a majority of pathway components were downregulated. In the cortex, the genes with $log_2FC > 0$ were enriched in immunological pathways similarly to the hippocampus (Figure 3.8). On the other hand, the genes with $log_2FC < 0$ were enriched in neural pathways, mostly involving different kinds of synapses such as GABAergic, dopaminergic and glutamatergic synapses, and enriched in proteasomal and ribosomal pathways. They also showed enrichment in LTP, synaptic vesicle cycle, gap junction, circadian entrainment, and few neurodegenerative disorders (Figure 3.8). This result would suggest a dysregulation in neural-immune trajectory. As we did for the hippocampus examples, we similarly map the downregulated genes from the cortex on the synaptic vesicle cycle KEGG pathway as an example (Figure 3.9).

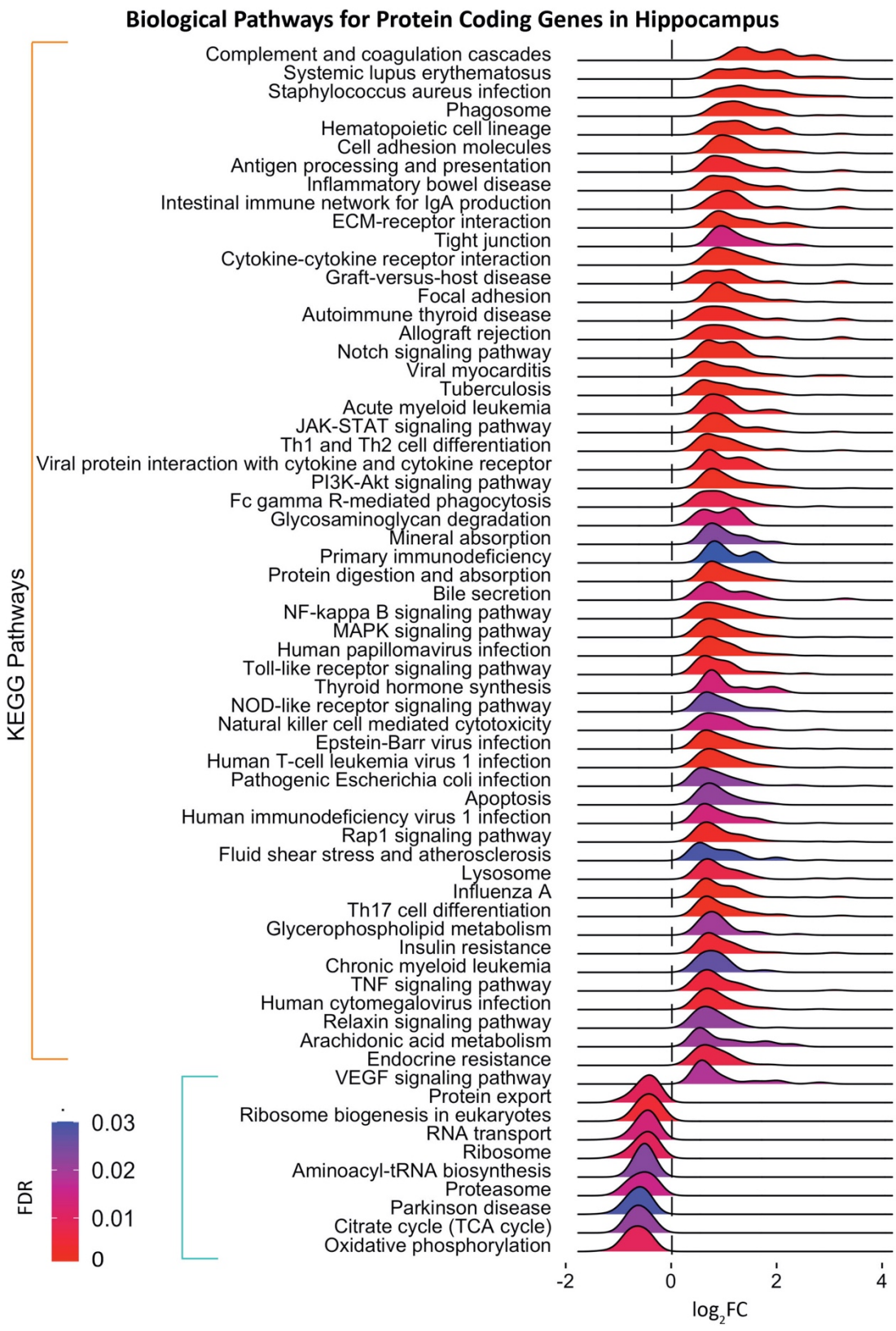


Figure 3.6. Enrichment of immunological, ribosomal and proteasomal pathways in human DS hippocampus. GSEA analysis for all the protein-coding genes identified in the human hippocampus samples reveals KEGG pathways which were found to be enriched in immunological and ribosomal and proteasomal pathways. In red brackets are immunological pathways, in blue brackets are ribosomal and proteasomal pathways. The color bar on the left indicates FDR (obtained after Benjamini-Hochberg multiple testing correction). KEGG: Kyoto Encyclopedia of Genes and Genomes, FC: Fold change , FDR: False Discovery Rate.

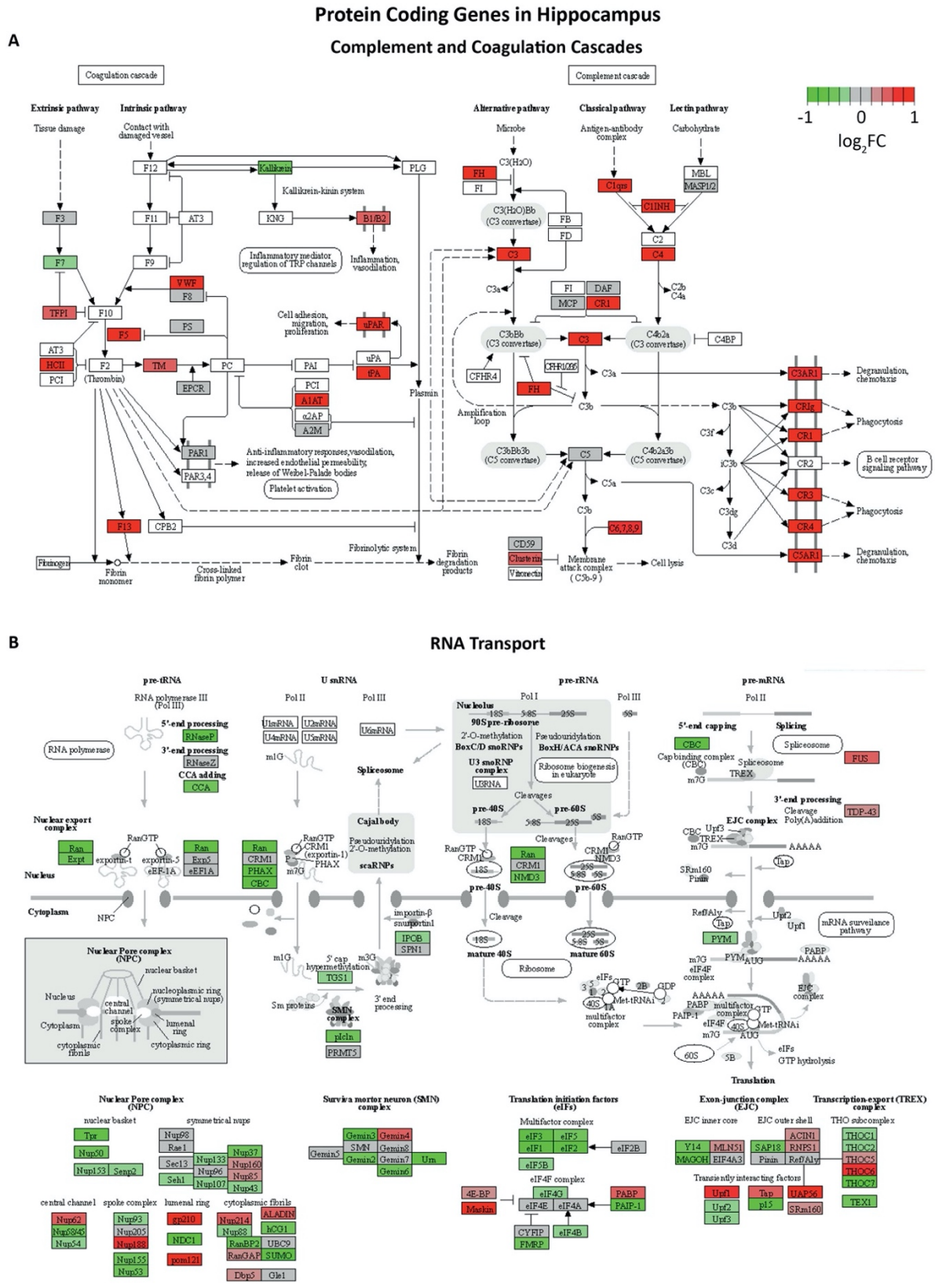


Figure 3.7. Examples of enriched KEGG pathways in human DS hippocampus. Mapping of log_2FC values on the KEGG pathways from hippocampus reveals which genes from these pathways are differentially expressed. A) Complement and coagulation cascades pathway B) RNA transport pathway. Genes are color coded by their log_2FC values where red indicates upregulated, green indicates downregulated, grey indicates no differential expression and white represent those genes not found in our data. Range of log_2FC is shown by the color bar on top. FC: Fold change.

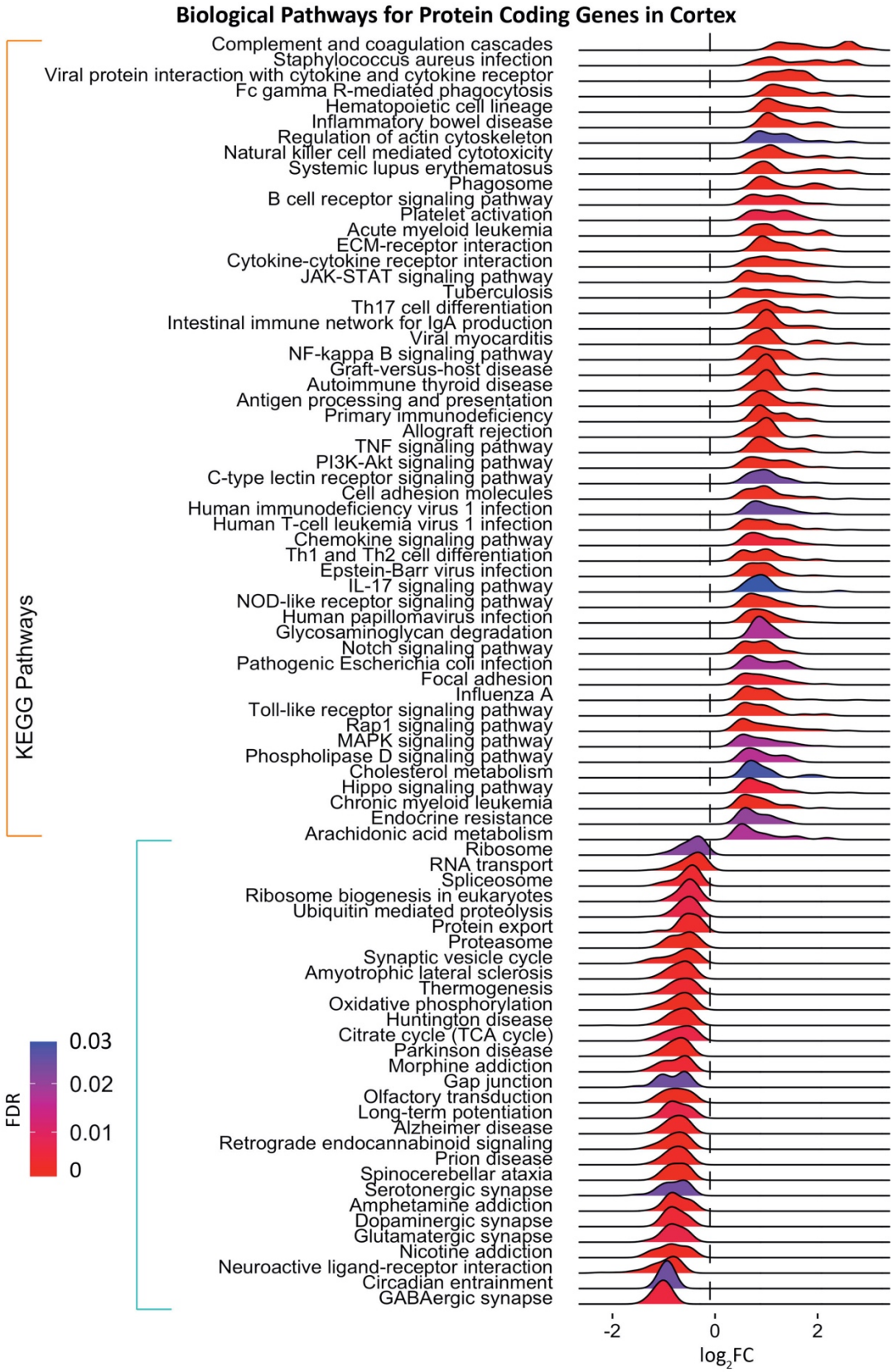


Figure 3.8. Enrichment of immunological and neuronal pathways in human DS cortex. GSEA analysis for all the protein-coding genes identified in the human cortex samples shows KEGG pathways which were found to be enriched in immunological neural pathways. In red brackets are immunological pathways, in blue brackets are ribosomal, proteasomal and neural pathways. The color bar on the left indicates FDR (obtained after Benjamini-Hochberg multiple testing correction). KEGG: Kyoto Encyclopedia of Genes and Genomes, FC: Fold change, FDR: False Discovery Rate.

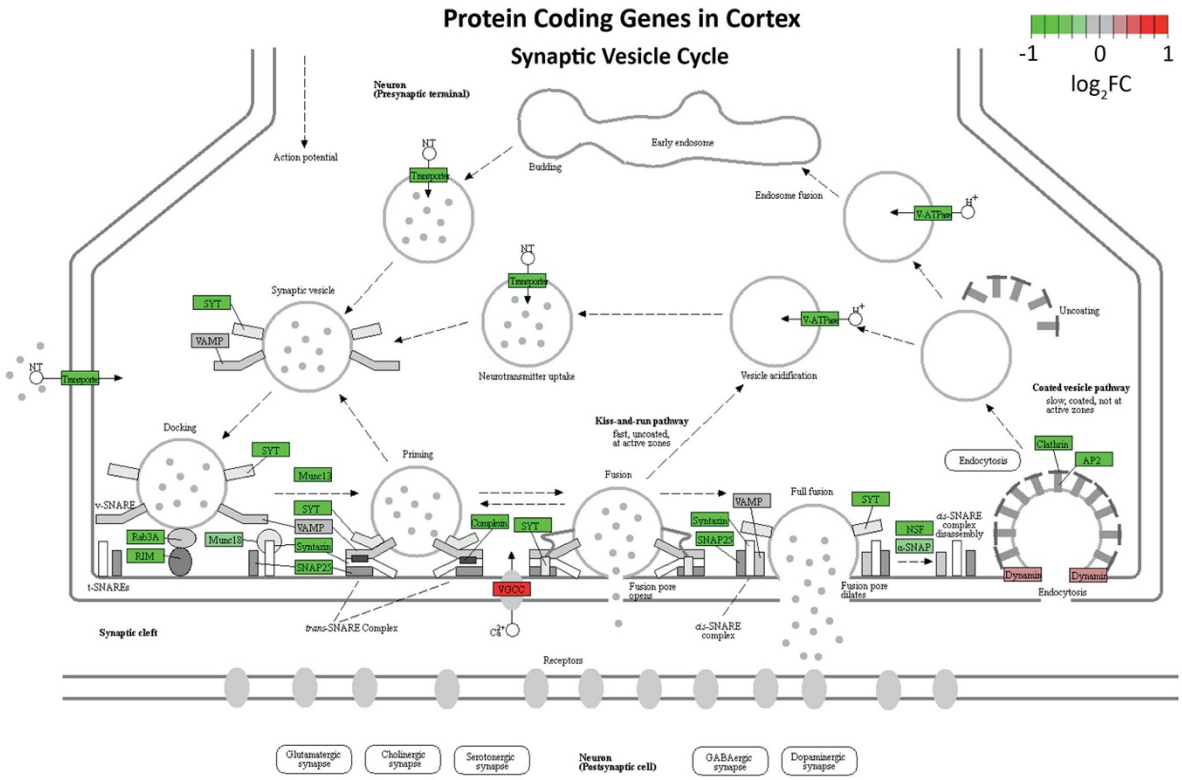


Figure 3.9. Example of enriched KEGG pathway in human DS cortex. Mapping of log_2FC values on the downregulated pathway synaptic vesicle cycle shows differentially expressed genes involved in the pathway. Genes are color-coded by their log_2FC values where red indicates upregulated, green indicates downregulated, grey indicates not differentially expressed and white represent those genes not found in our data. Range of log_2FC is shown by the color bar on top. FC: Fold change.

3.3. Differentially expressed genes are divided into significant brain cell types.

The human brain comprises multiple cell types (e.g., astrocytes, microglia, oligodendrocytes, interneurons, pyramidal neurons). The expression changes we observed at the bulk tissue level could result from the change in proportions or expression levels of these cells among the diverse experimental groups. We cannot calculate the proportion of different cells in a bulk RNA sequencing dataset, but it is possible to look for expression changes originating from different cells. To this aim, we took advantage of single-cell RNA sequencing (scRNA-seq) results on astrocytes, microglia, oligodendrocytes, interneurons, and pyramidal cells from mouse cortex and hippocampus available in the literature (Zeisel et al., 2015). We used a recently developed algorithm known as Expression Weighted Cell-type Enrichment (EWCE) analysis (Skene and Grant, 2016). The EWCE analysis tests for enrichment for a gene list of interest (in our case, the DEGs from the bulk RNA-seq data from hippocampus and cortex)

against the cell-type expression levels and computes whether the enrichment is statistically significant or it just occurs by chance.

After performing cell-type enrichment analysis on the DEGs from hippocampus and cortex, we found that enrichment in different cells was consistently shared between the two regions with high significance (high z-scores, which denotes the number of standard deviations away from mean); Figure 3.10). In particular, both the hippocampal and the cortical upregulated genes showed clear enrichment in glial cells (astrocytes, microglia, and oligodendrocytes), whereas the downregulated genes showed clear enrichment in neural cells such as interneurons and pyramidal cells. Altogether our results indicated that there is an involvement of neural-immune dysregulation in Down syndrome.

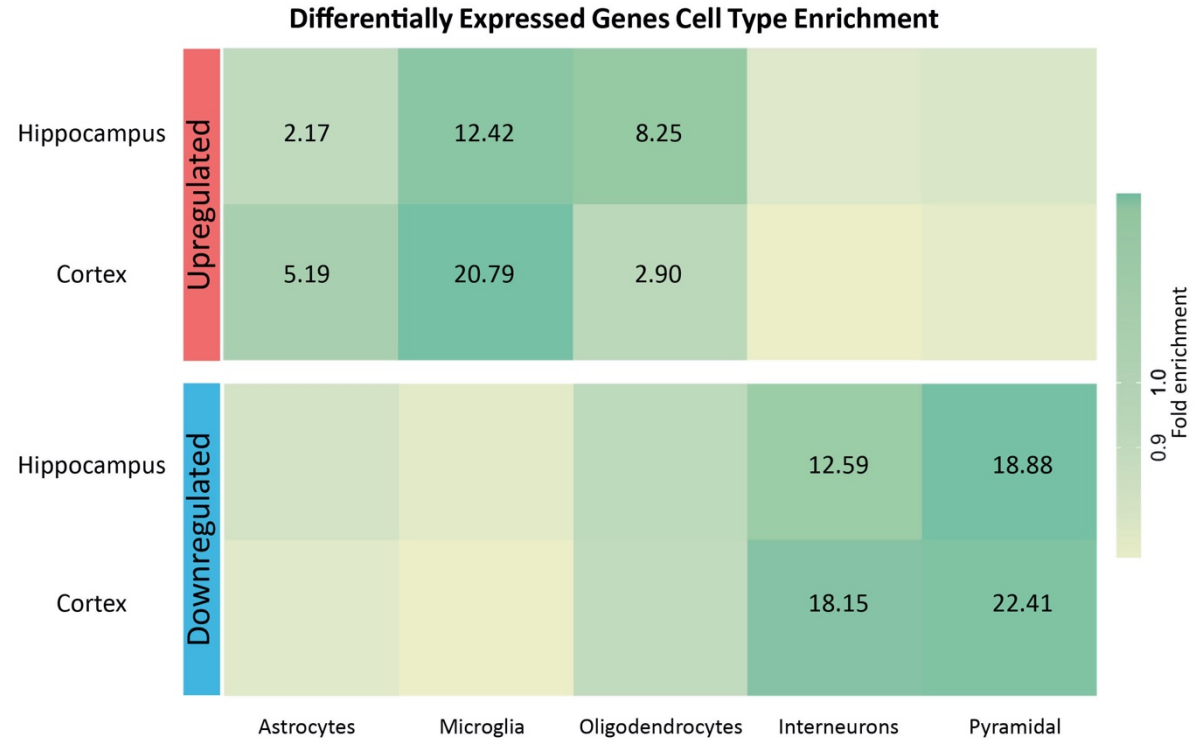


Figure 3.10. Differentially expressed genes show cell-type enrichment in human DS hippocampus and cortex. Upregulated genes in hippocampus and cortex show enrichment in glial cells and downregulated genes reveal neural cells enrichment. The numbers in the heatmap represent the z-score (number of standard deviations away from mean) for the significant enrichment (Bootstrap significance testing with 100,000 repetitions) and the color bar on right represent the level of fold enrichment. Dysregulation of gene expression in DS is present at genome-wide level.

Since we found a high amount of dysregulation among the triplicated and non-triplicated genes and dysregulation in a wide variety of biological processes and pathways, we thought

to map the DEGs that we found in our bulk RNA-seq analysis based on their chromosome location. This analysis allowed us to identify which chromosomes showed the most dysregulation in DS (Figure 3.11). As expected, many upregulated genes were present on chromosome 21 because of the trisomy. Notably, there were other chromosomes, too, showing many dysregulated genes in both the hippocampus and the cortex (Figure 3.12).

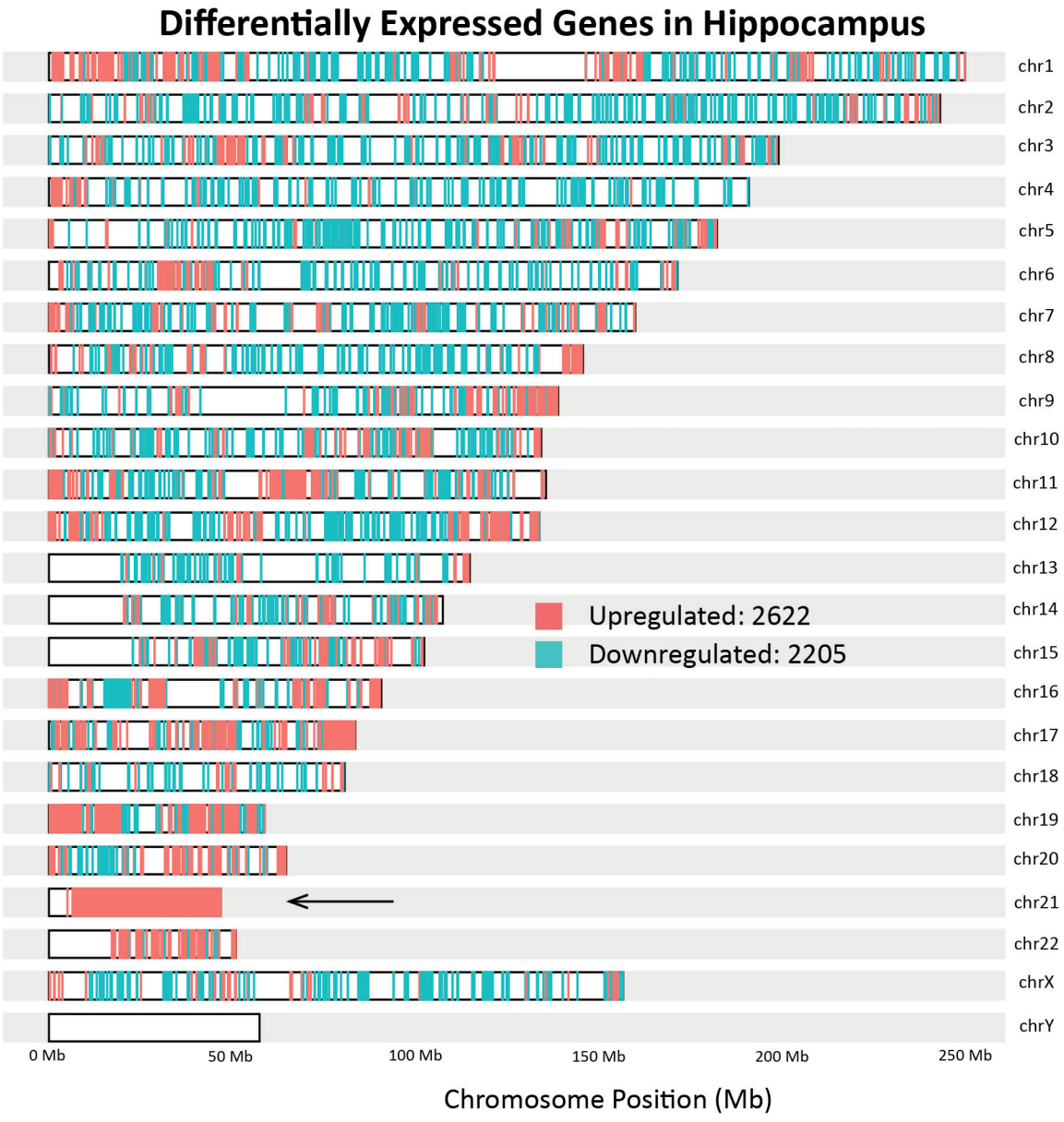


Figure 3.11. Differentially expressed genes show genome-wide dysregulation in human hippocampus. Differentially expressed genes are represented by their location in megabases (Mb) (x-axis) on different chromosomes, and shows upregulated genes in red and downregulated genes in blue. Numbers inside the plot shows the number of up and downregulated genes, and black arrow represents chromosome 21.

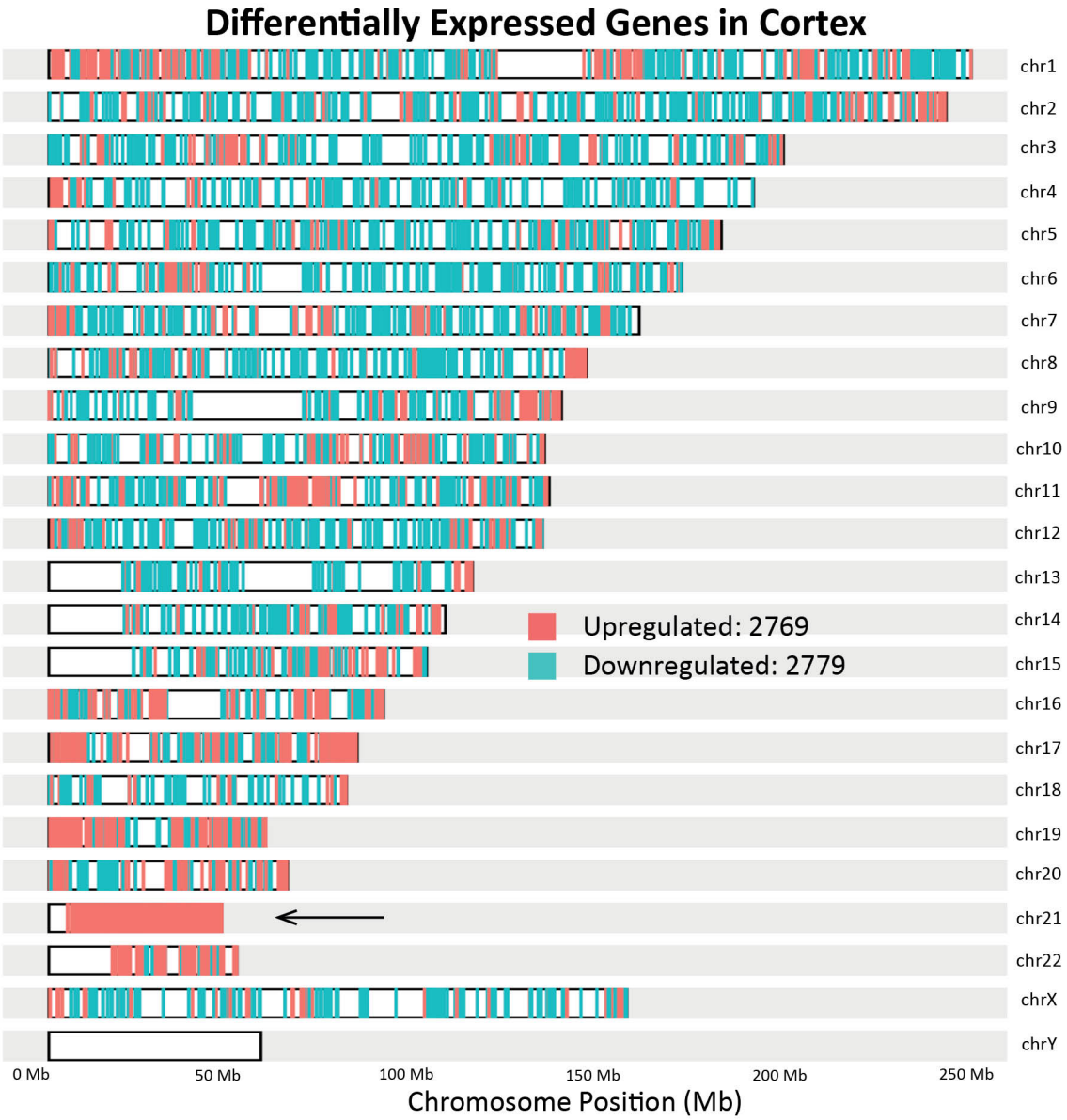


Figure 3.12. Differentially expressed genes show genome-wide dysregulation in human cortex. Differentially expressed genes are represented by their location in megabases (Mb) (x-axis) on different chromosomes, and shows upregulated genes in red and downregulated genes in blue. Numbers inside the plot shows the number of up and downregulated genes, and black arrow represents chromosome 21.

Next, we investigated whether the DEGs are organized in domains termed as gene-expression dysregulation domains (GEDDs) of upregulated and downregulated genes, as previously observed in fetal fibroblasts from DS individuals (Letourneau et al., 2014). The authors observed a consistent pattern across all the chromosomes for an alternating increase and decrease in gene expression (GEDDs) across large segments of chromosomes. To find GEDDs in our gene expression data from the bulk RNA-seq experiments from the hippocampus and the cortex, we performed loess smoothing on the gene expression for all the chromosomes. We found around 1356 and 1676 GEDDs in total for both the hippocampus and the cortex

respectively across all the chromosomes combined. Figure 3.13 shows two examples (chr1 and chr11) from hippocampus gene expression, where we found 148 and 77 GEDDs, respectively.

Similarly, in the cortex, for example, on chr1 and 11 (Figure 3.14), we found 210 and 85 GEDDs, respectively. In our analysis of adult human hippocampal and cortical samples, we similarly found GEDDs on all chromosomes. The size of the domains ranges from 0.2 kilobases (kb) to 42 megabases (Mb) (Figure 3.13).

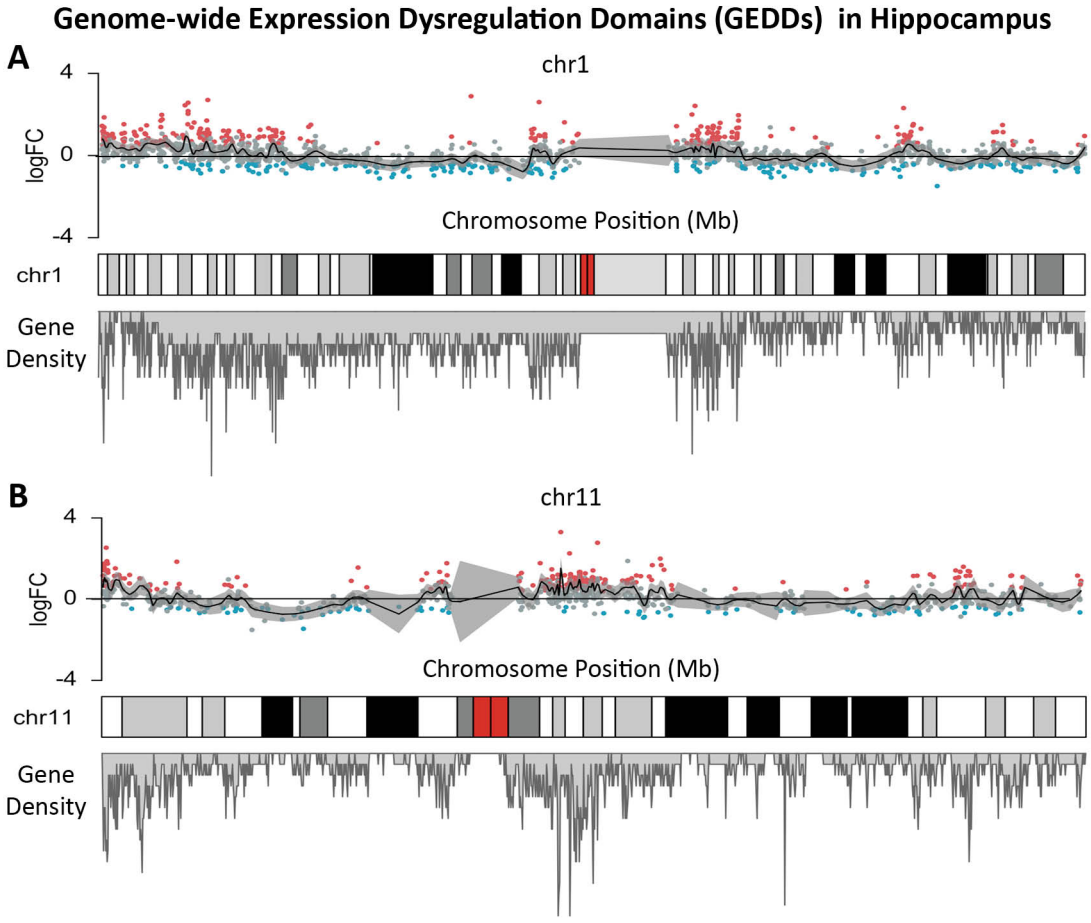


Figure 3.13. Identification of GEDDs in human DS hippocampus. Genome-wide Expression Dysregulation Domains (GEDDs) present on chromosome 1 and 11 are shown on hippocampus gene expression data. Each panel is divided into three subgraphs: the top graph represents GEDDs by the $\log_2$ fold changes (on y-axis) across the chromosome position (x-axis) with red dots showing upregulated genes, blue dots showing downregulated genes and grey dots non differentially expressed genes. This graph shows clear domains of upregulated and downregulated genes across the length of the chromosome. The gray shaded region represents the 95% confidence interval for the loess smoothing performed on the gene expression to define the domains. The middle graph shows the karyotype diagram of the respective chromosome and bottom graph represent the number of genes present on that specific chromosome region (gene density) for all the genes normally present on that chromosome, shown here to indicate regions of chromosome with different densities.

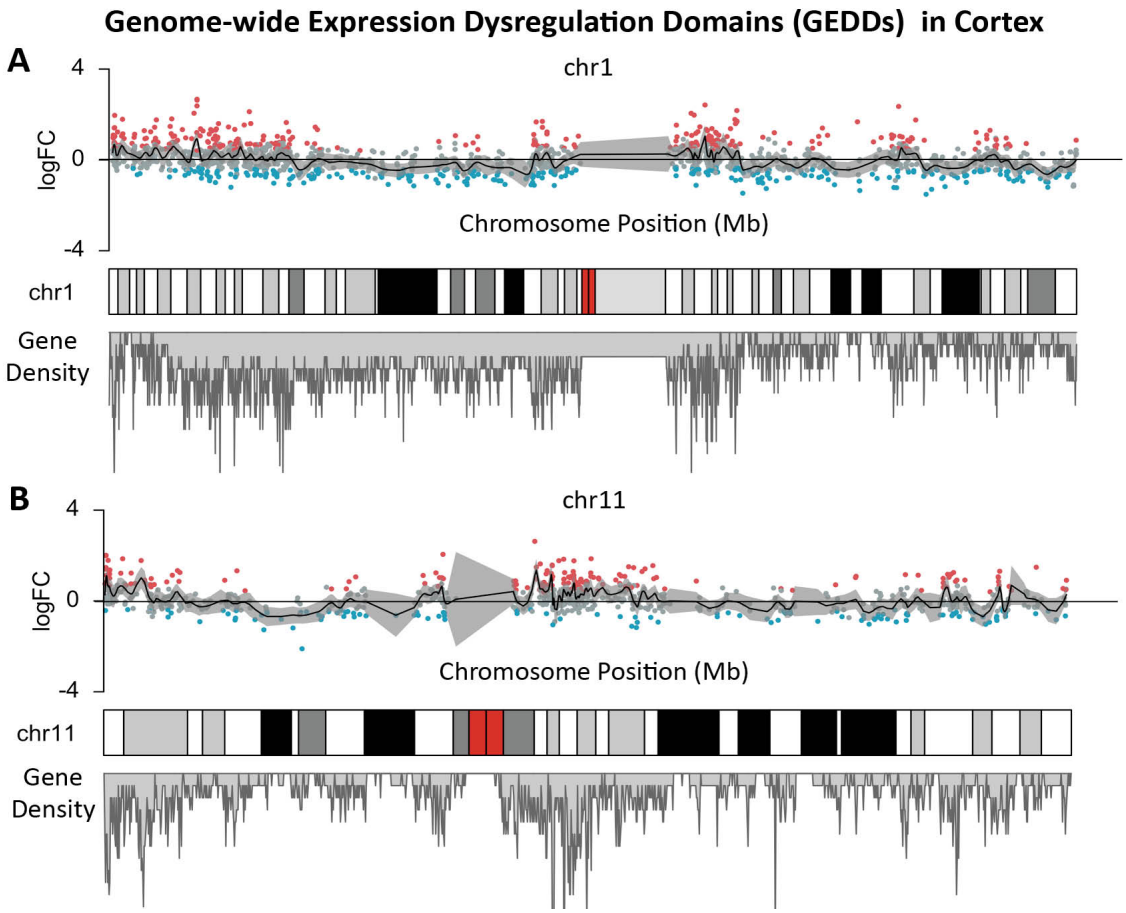


Figure 3.14. Identification of GEDDs in human DS cortex. Cortex gene expression data exhibits Genome-wide Expression Dysregulation Domains (GEDDs) present on chromosome 1 and 11. Each panel is divided into three subgraphs: the top graph represents GEDDs by the $\log_2$ fold changes (on y-axis) across the chromosome position (x-axis) with red dots showing upregulated genes, blue dots showing downregulated genes and grey dots non differentially expressed genes. This graph shows clear domains of upregulated and downregulated genes across the length of the chromosome. The gray shaded region represents the 95% confidence interval for the loess smoothing performed on the gene expression to define the domains. The middle graph shows the karyotype diagram of the respective chromosome and bottom graph represent the number of genes present on that specific chromosome region (gene density) for all the genes normally present on that chromosome, shown here to indicate regions of chromosome with different densities.

3.4. Expression weighted network analysis predicts candidate genes influencing genome dysregulation.

Noting the high level of dysregulation and clustering of genes according to their $\log_2$FC between the DS and control individuals, we next sought to determine the extent to which these gene expression changes impacts DS in a systems-level framework (i.e., construct interaction networks of co-expressed genes). To this aim, we performed a weighted gene co-expression network analysis (WGCNA) on the bulk RNA-seq data from hippocampal and cortical samples (Langfelder and Horvath, 2008; Zhang and Horvath).

The WGCNA is a widely used method to identify clusters (modules) of genes with a similar level of expression among all the genes and under the assumption that they might thus be involved in similar biological processes and molecular pathways. The WGCNA analysis helped us to identify functionally related groups of genes influenced by DS. Overall, we found 19 modules for the human hippocampus (Figure 3.15) and 13 modules for the human cortex (Figure 3.16), including a module for each brain region where the grouped genes did not belong to any other module. We tested for association against some biological traits (i.e., control or DS group, age, and sex) or technical trait (post-mortem interval (PMI)) for these modules. Overall, the modules significantly associated only with groups (control or DS), indicating that the disease status strongly influences them (healthy or DS) and not with any other biological trait that could affect the interpretation of our data. Notably, we found 6 modules in the hippocampus which were significantly correlated with either control (G-HH-3, G-HH-4, G-HH-5) or DS samples (G-HH-1, G-HH-2, G-HH-6) (Figure 3.15) and 4 modules in the cortical control (G-HC-2, G-HC3) or cortical DS (G-HC-1, G-HC-4) samples (Figure 3.16).

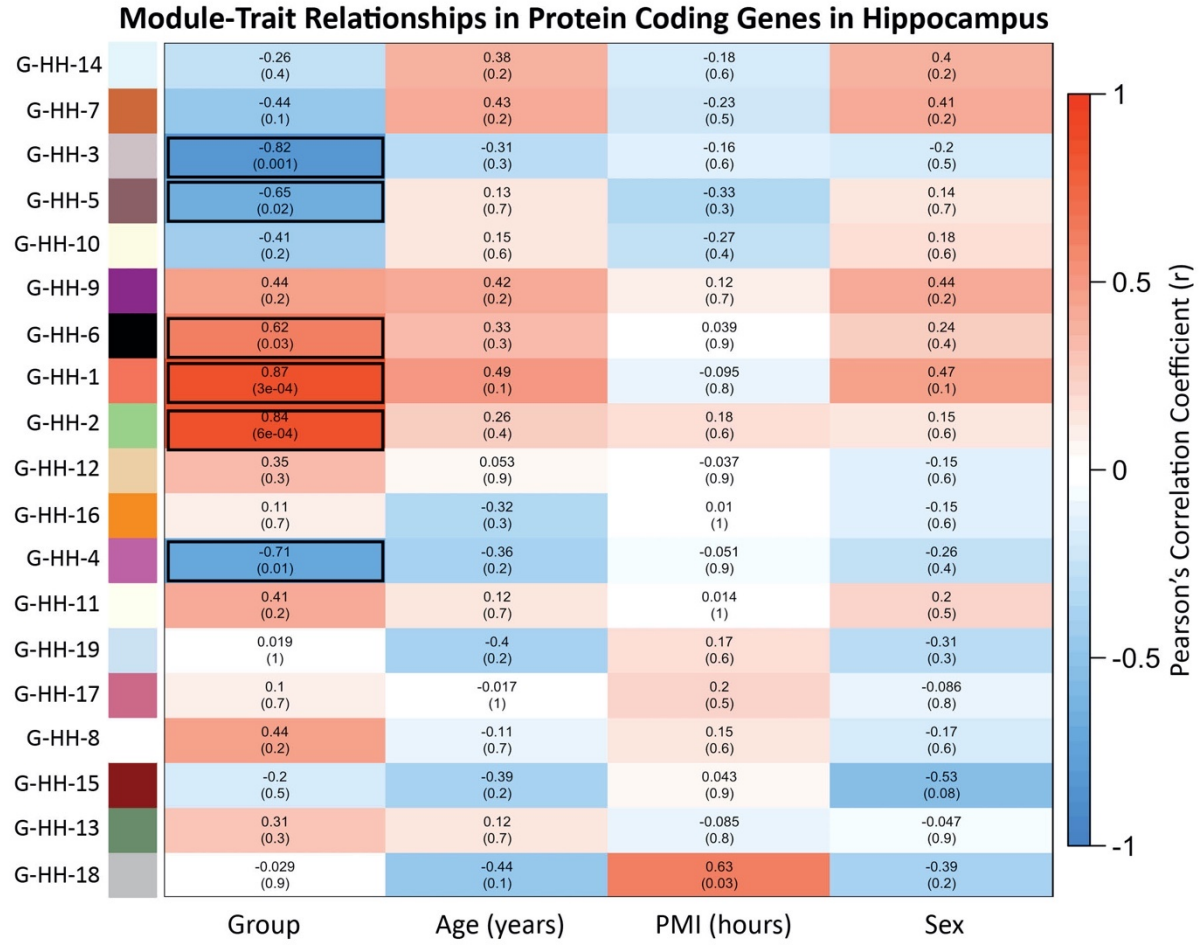


Figure 3.15. WGCNA-derived modules from human hippocampus gene expression data capture disease association. The modules were tested for association against Group (Control, DS), Age in years, post-mortem interval (PMI) in hours and Sex (F,M). The top numbers inside the heatmap show the color-coded (color bar on the right) Pearson's correlation coefficient associated with each trait and the number in parenthesis indicate their statistical significance in terms of p-value. The black boxes represent the modules having statistically significant association with either group category ($p<0.05$, Pearson's correlation test). G-HH represents gene modules in human hippocampus.

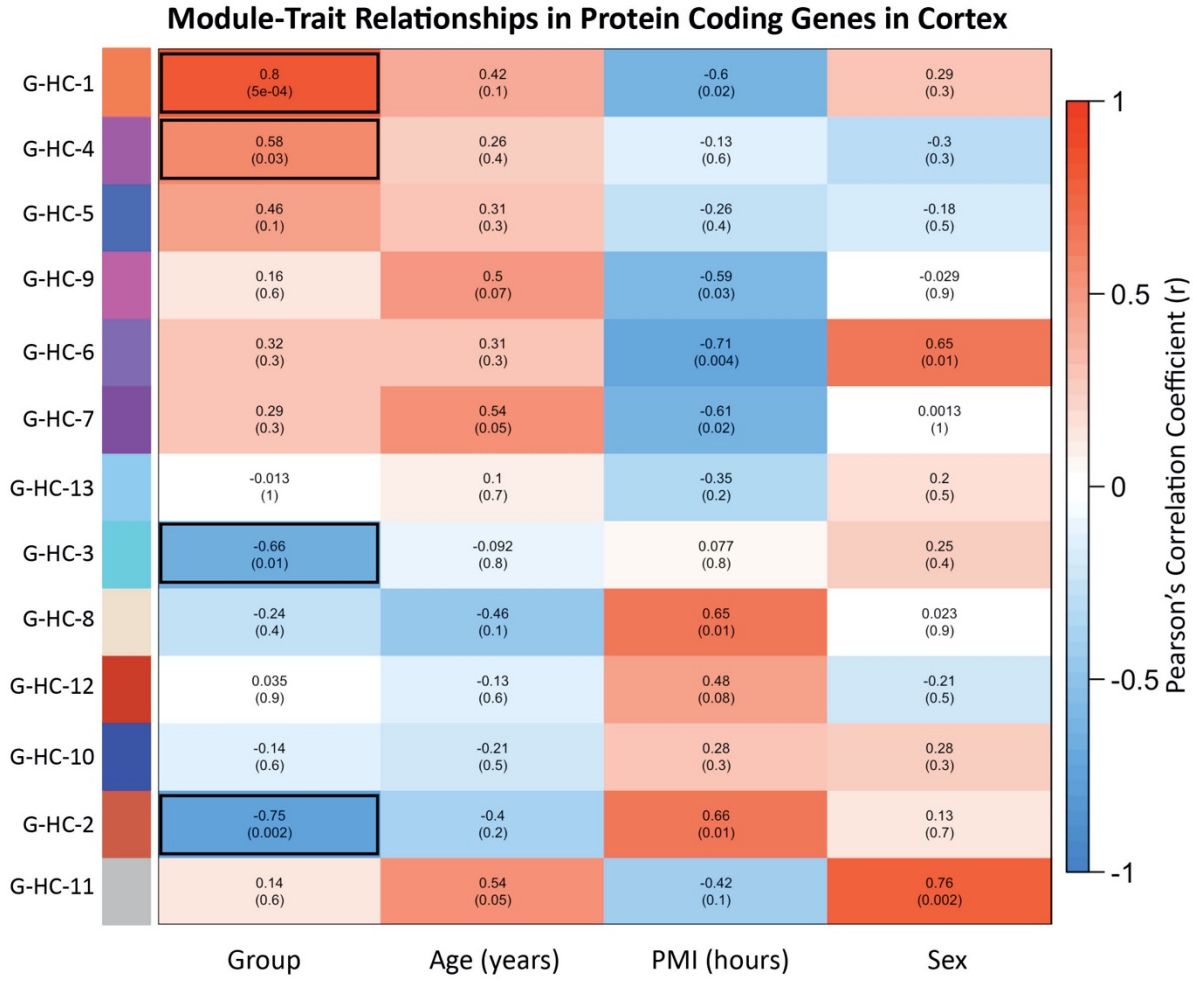


Figure 3.16. WGCNA derived modules for human cortex gene expression data capture disease associations. The modules were tested for association against Group (Control, DS), Age in years, PMI in hours and Sex (F,M). The top numbers inside the heatmap show the color-coded (color bar on the right) Pearson's correlation coefficient associated with each trait and the number in parenthesis indicate their statistical significance in terms of p-value. The black boxes represent the modules having significant association with either group category ($p<0.05$, Pearson's correlation test). G-HC represents gene modules in human cortex.

To understand the biological importance of the significant modules and their relevance to DS, we first performed a gene ontology (GO) analysis on the genes belonging to each of the modules. We found that the most significant upregulated module G-HH-1 ($R=0.8$, $p=5 \times 10^{-4}$) in DS hippocampus was enriched in locomotion, cell-cell adhesion, cell communication and also in a number of immune-regulated processes such as leukocyte-mediated immunity,

response to LPS, response to biotic stimulus and type I interferon signaling pathway (Figure 3.17 left). This highlighted that G-HH-1 has biological importance in terms of immune response and cell communication. Interestingly, many of these latter processes were shared with the cortical module G-HC-1.

On the other hand, the module G-HH-3, which is significantly correlated (R=-0.82, p=0.001) (Figure 3.14) with control samples (i.e. negatively correlated with DS samples), was enriched in biological processes such as ATP metabolic processes and electron transport chain (Figure 3.17, right). This points to an association of this module with mitochondrial dysfunction. G-HH-3 also showed significant enrichment in axonogenesis, neuron-projection development and synaptic signaling. Interestingly, module G-HH-3 shared some of these biological processes with the two downregulated modules from the cortex (G-HC-2 and G-HC-3; Figure 3.17, right). Thus modules of co-expressed genes reflect enrichment in certain biological processes for both regions.

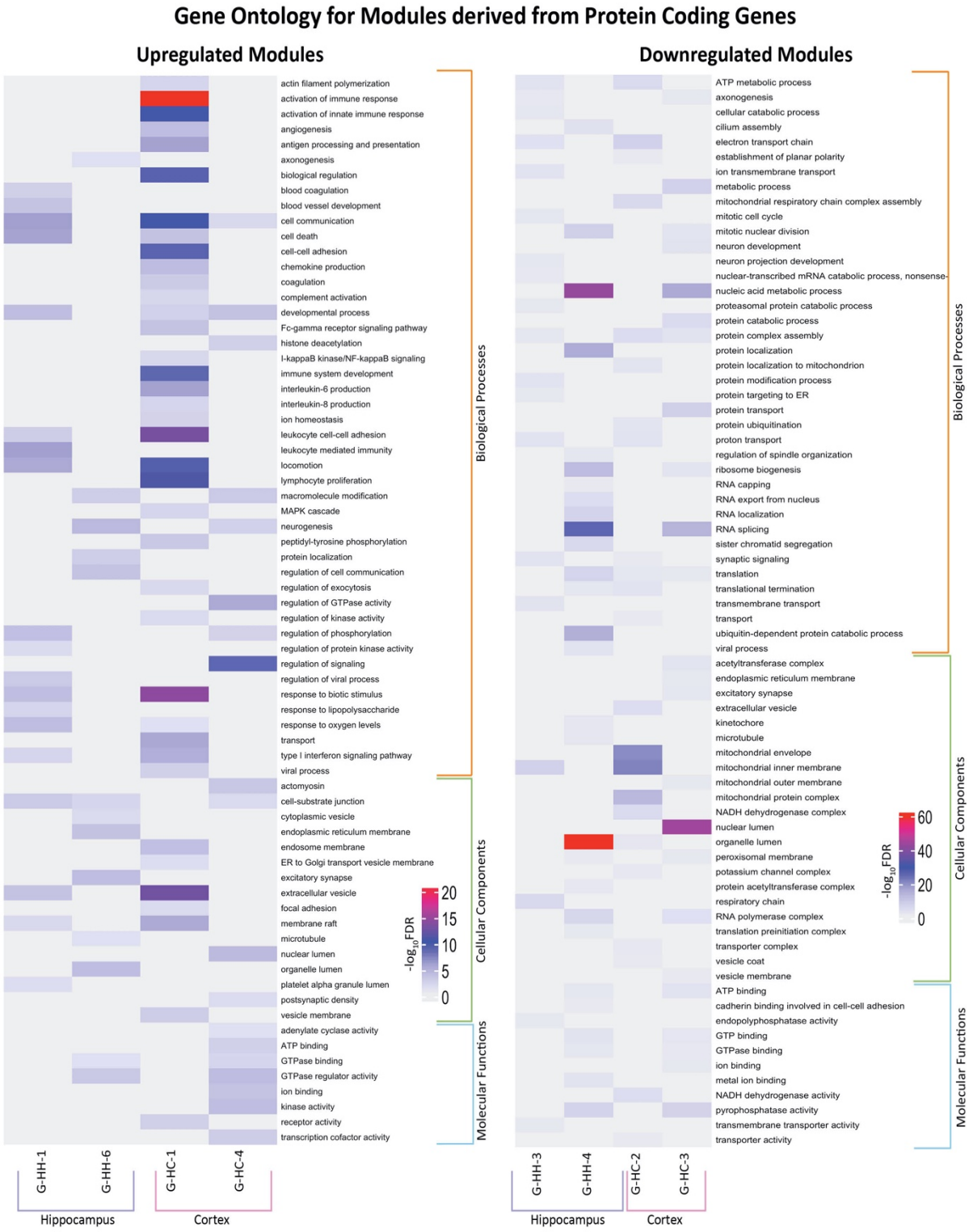


Figure 3.17. WGCNA derived co-expression modules associated with multiple GO terms. Multiple Biological Processes (BP), Cellular Components (CC) and Molecular Functions (MF) terms are enriched for significant modules from both human hippocampus and cortex gene expression data. The GO terms are separated by parenthesis, where orange represents biological processes, green represents cellular components and blue represents molecular functions. The region-wise modules are also separated by parenthesis, where purple shows hippocampus and pink shows cortex. The color bar on the right shows the range of significance defined by -$\log_{10}$FDR (after Benjamini-Hochberg multiple testing correction) for each term and each module. FDR: False Discovery Rate, G-HH: Gene module in Human hippocampus, G-HC: Gene module in Human cortex.

To further resolve the biological impact of these significant modules from hippocampus and cortex, we investigated cell-type enrichment for each of these modules. Strikingly, we found that the upregulated hippocampal module, G-HH-1 (Figure 3.18 top) was specifically enriched in microglia. Another upregulated module, G-HH-2 was enriched in all the glial cell types (astrocytes, microglia and oligodendrocytes). Similarly, the upregulated module in cortex, G-HC-1 (Figure 3.18 top) showed enrichment in only microglial cells. Conversely, the downregulated hippocampal module G-HH-3 (Figure 3.18 bottom) was enriched in neural cell types (pyramidal cells and interneurons), similar to the downregulated modules G-HC-3 and G-HC-2 from the cortex. We also found regional differences, implying that certain enrichments and dysregulation may potentially occur in specific regions in DS. For example, the upregulated module from hippocampus G-HH-6 had cell enrichment in oligodendrocytes and pyramidal cells (Figure 3.18 top) which was not present in any module originating from the cortical samples. Thus, markers and expression profiles of specific brain cell types from scRNA-seq data likely reflects sources of variation for each of the WGCNA-derived modules in DS. The enrichment for most of the upregulated modules in glial cells and downregulated modules in neural cells closely resembles the clear distinction observed at differential gene expression level (Figure 3.10).

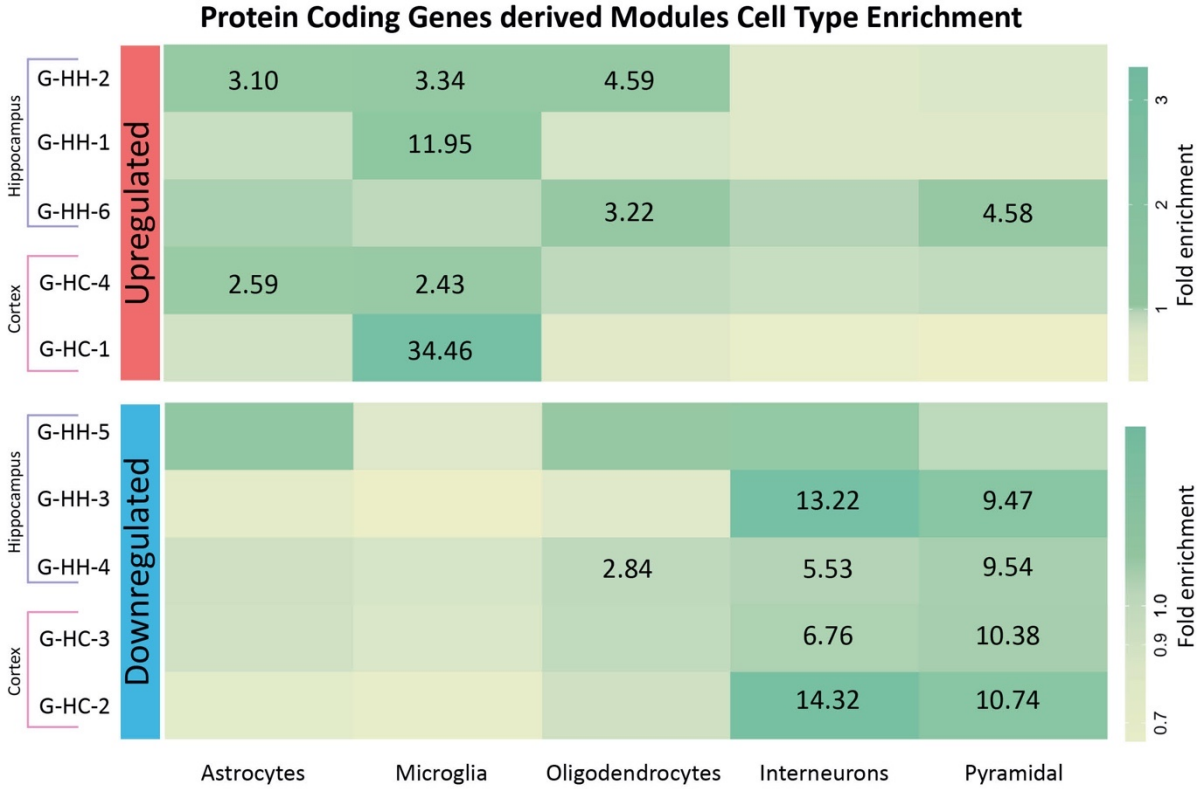


Figure 3.18. WGCNA derived co-expression modules capture differences in cell enrichment from both human hippocampus and cortex gene expression. They are separated by their enrichment in different cell types and divided in up- (top panel) or down-regulated modules (bottom panel). The color bars on the right for upregulated and downregulated modules are different and represent the fold enrichment. The region-wise modules are also separated by parenthesis, where purple shows hippocampus and pink shows cortex. The numbers inside the heatmap squares show the z-score (number of standard deviations away from mean) for significant cell enrichment (Bootstrap significance testing with 100,000 repetitions) G-HH: Gene module in Human hippocampus, G-HC: Gene module in Human cortex.

All the above analysis indicated a high level of correlation between different modules from hippocampus and cortex which would make it easier to look for common pathways and genes involved in both the regions. We substantiated some of this correspondence between the region-wise modules with the help of a correlogram (Figure 3.19). We took the module membership value for each gene which is defined as the correlation of each gene within a module with the gene that most represents that specific module (known as module eigengene) and computed the correlation between the modules. We found that many modules were highly correlated between the hippocampus and the cortex. For example, modules G-HH-1 and G-HH-2 from hippocampus were significantly well correlated with G-HC-1 from the cortex. To further corroborate our findings, we are currently comparing the results obtained by Olmos-Serrano et al. to our results. The authors analyzed DS and euploid control individuals from fetal to adult developmental stages in multiple brain regions with microarray technology (Olmos-Serrano et al., 2016). To compare our results, we need to select the age-matched samples from the study mentioned above, considering that the genes analyzed by microarray are much lesser than total RNA sequencing performed here.

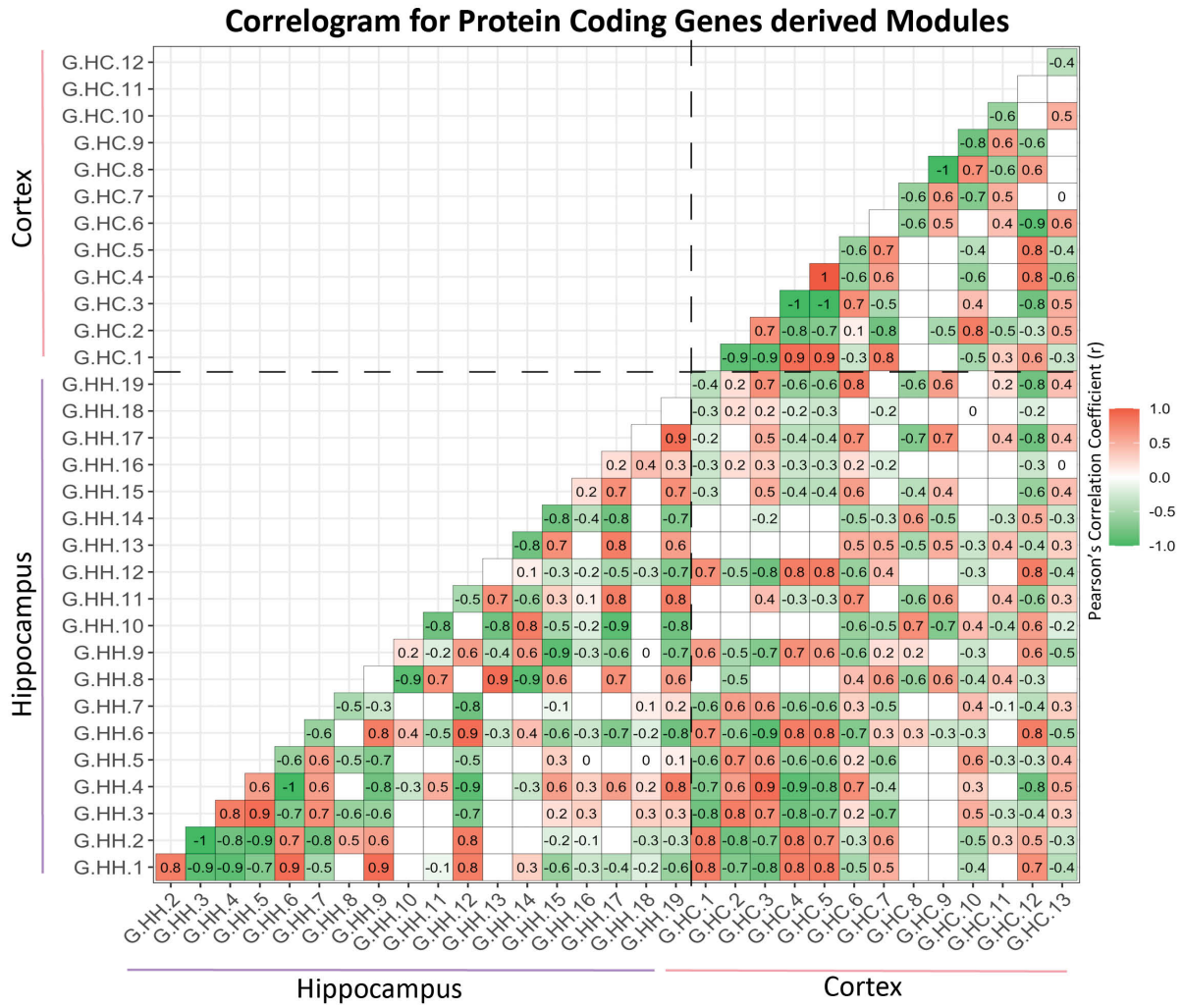


Figure 3.19. Module-based correlation in gene expression from human hippocampus and cortex. Correlation analysis among all the modules derived from the WGCNA analysis for protein coding genes from hippocampus and cortex shows some correlated modules. The red color indicates positive correlation, green indicates anti-correlation and white indicates no correlation. Numbers inside the boxes are the correlation values for only statistically significant correlations ($p<0.05$, Pearson's Correlation test). The entity of the correlation is represented by a color bar on the right. G-HH: Gene module in Human hippocampus, G-HC: Gene module in Human cortex.

Next, we analyzed the specific genes involved in each of the significant modules from the WGCNA analysis by making a gene-gene interaction network using Cytoscape (Shannon et al., 2003). Cytoscape is a tool used to visualize the gene-gene interaction network of co-expressed genes. This interaction network allowed us to investigate hub genes that could regulate or be responsible for some of the defects and disturbances we observed in various biological processes and functions when we compared DS people samples to those of controls. As module sizes had a broad range and, in general, high numbers of genes, we only focused on the top (those with the highest module membership within each module) 20 genes for each module for the Cytoscape analysis. In particular, to identify hub genes in each

module, we created a network from these 20 genes considering their number of connections (degree, denoted by the size of the nodes in the network, Figure 3.20) and the degree of gene expression (denoted by log_2FC). For example, in the hippocampal module G-HH-1 having 607 members (Figure 3.20A), six genes out of the top 20 consisted of C1QC, CD14, CD300A, IL4R, SERPINA1, OSMR, which are involved in the complement system, cytokine signaling, and immune response signaling pathways. This enrichment of hub genes to immune pathways was consistent with the module's high association with microglia (Figure 3.19), the brain's immune cells. This upregulated module G-HH-1 has few triplicated genes in the top hub genes, potentially responsible for some of the dysregulation in the rest of the genes from the same module. On comparing the G-HH-1 module with cortical module G-HC-1 (Figure 3.21A), we found out that none of the network-driving genes (hubs) are the same, and only one of them, ITGB2, is triplicated. On the other hand, in the hub genes from both hippocampal module G-HH-3 (Figure 3.20B) and cortical module G-HC-2 (Figure 3.21B), four genes (VDAC2, SCG5, GAP43, and REEP1) are common. All these results suggest that there could be potentially different drivers or hubs for each region, depending on which group of genes and biological processes they control.

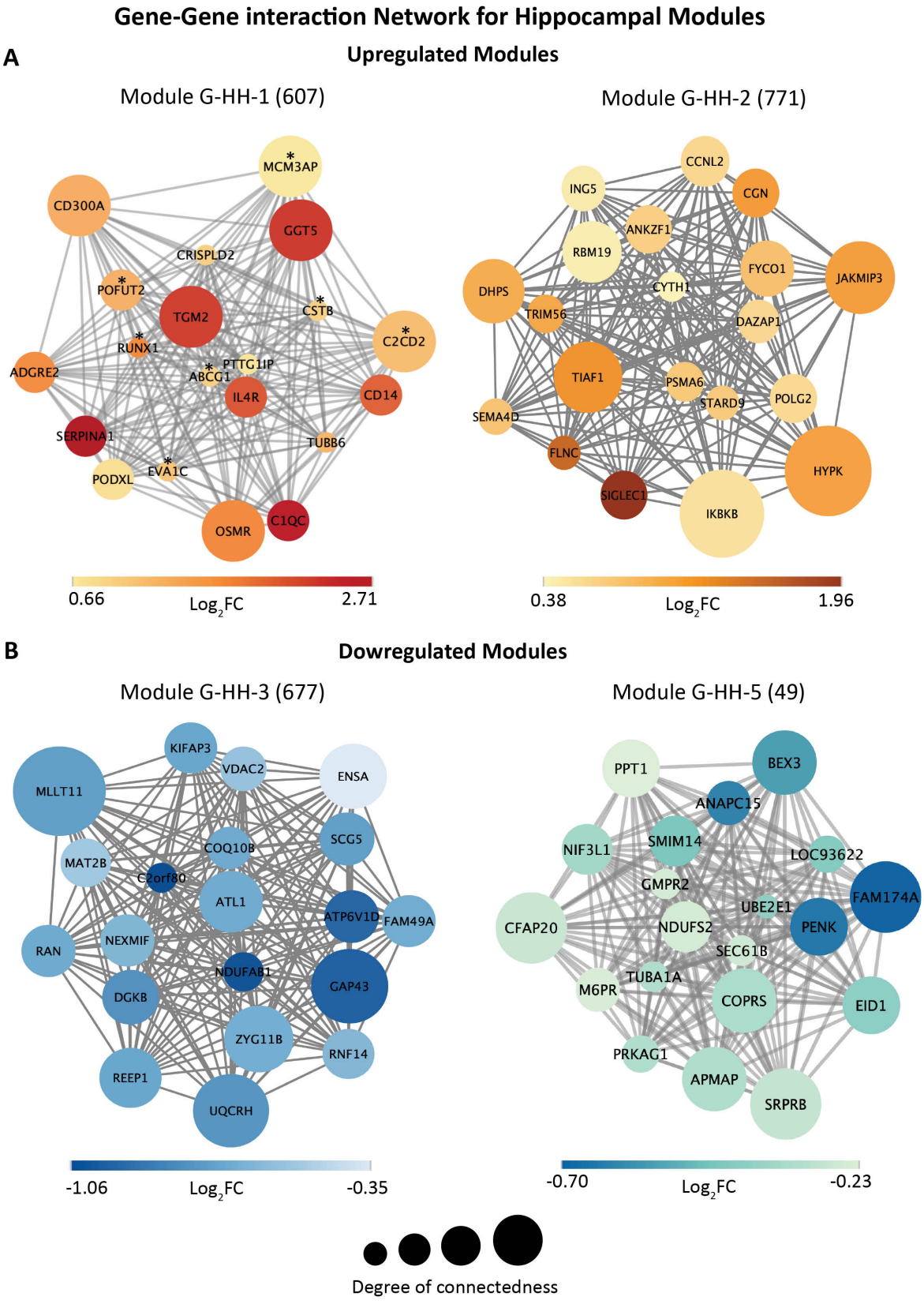


Figure 3.20. Gene-gene interaction network for WGCNA-derived significant modules from hippocampus gene expression data comprising of the top 20 genes. The size of the nodes denotes the degree of connectedness, the color of the node (coded in the color maps below) represents the log_2FC. A) Upregulated modules G-HH-1 and G-HH-2 B) Downregulated modules G-HH-3 and G-HH-5. Number of genes in each module are denoted in the brackets. Asterisks denote triplicated genes. G-HH: Gene module in Human hippocampus, FC: Fold change.

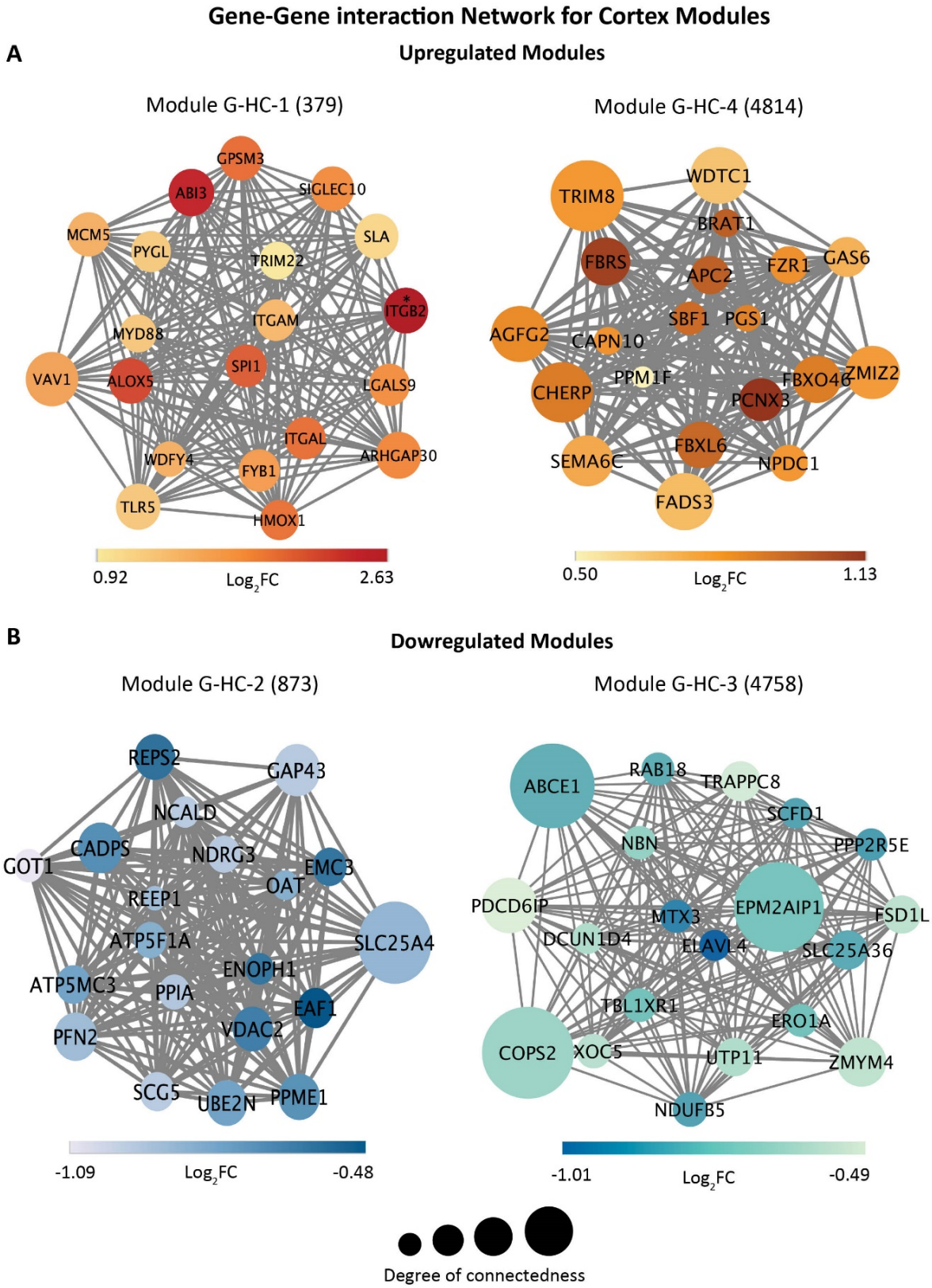


Figure 3.21. Gene-gene interaction network for WGCNA-derived significant modules from cortex gene expression data comprising of the top 20. The size of the nodes denotes the degree of connectedness, the color of the node (coded in the color maps below) represents the log_2FC. A) Upregulated modules G-HH-1 and G-HH-2 B) Downregulated modules G-HH-3 and G-HH-5. Number of genes in each module are denoted in the brackets. Asterisks denote triplicated genes. G-HH: Gene module in Human hippocampus, FC: Fold change.

Next, to validate DGE results, we performed real-time PCR on selected genes from different modules for both hippocampus and cortex and found significant correlation with the RNA-seq data.

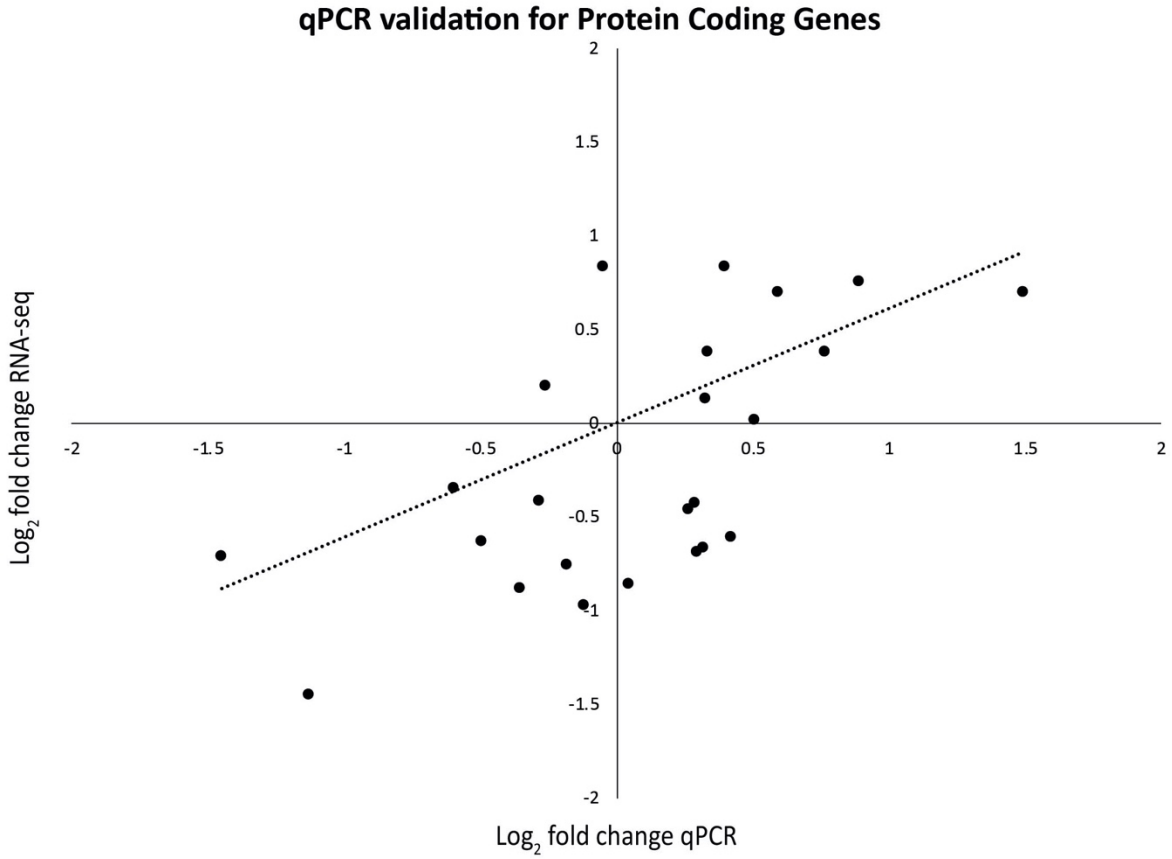


Figure 3.22. Correlation between RNA-seq and qPCR results. Significant correlation is observed between RNA-seq data (y-axis) and qPCR (x-axis) after performing validation for some of the triplicated genes and other non-triplicated, dysregulated genes showing fold-change comparison. The line represents linear regression with intercept at 0 and it shows a significant correlation R = 0.633 (p=0.001, Pearson's Correlation test).

3.5. Dysregulation in DS present also at transcript level.

One of the outstanding questions in transcriptomics is whether the information generated by analysis at the level of total gene expression (commonly referred to as gene) by standard RNA-seq techniques replicates changes at the level of transcript isoforms (commonly referred to as transcripts). With the advent of more advanced technology for RNA sequencing, reduction in cost and increase in the number of reads that could be sequenced in a short time has provided the much-needed evidence that both gene-level and transcript-level changes reflect

both similar and distinct functional characteristics (Gandal et al., 2018). More than 95% of the human genome undergoes alternative splicing, thus often resulting in more than one mature RNA (mRNA) from each gene. These different transcript isoforms code for different protein isoforms when translated. Thus, in order to find out whether there is any dysregulation at the transcript level in DS *vs.* control and whether this disturbance in transcript expression matches the observations at gene-level differential expression, we expanded our differential gene expression analysis to transcript level on the same datasets and followed a similar pattern of analysis as we did for gene-expression. We first considered the differential transcript expression (DTE) and found 2627 and 2463 up-and down-regulated transcripts from 1954 and 2078 genes, respectively, in the hippocampus. In the cortex, we identified 2858 and 1860 up-and down-regulated transcripts corresponding to 2076 and 1610 genes, respectively (Figure 3.23).

There were fewer differentially expressed transcripts common between the hippocampus and the cortex in comparison to DEGs. In particular, we found only 393 and 613 common up and downregulated transcripts, respectively. Nevertheless, we found an increase in the number of differentially expressed triplicated transcripts compared to DEGs with 77 DE transcripts for the hippocampus and 190 in the cortex. These were corresponding to 50 genes for the hippocampus and 88 genes for the cortex. They were thus similar in number to DEGs, but they were not identical genes. In contrast to what we described at the gene level, the cross-region overlap is lower at the DTE level, suggesting that transcripts confer higher regional specificity.

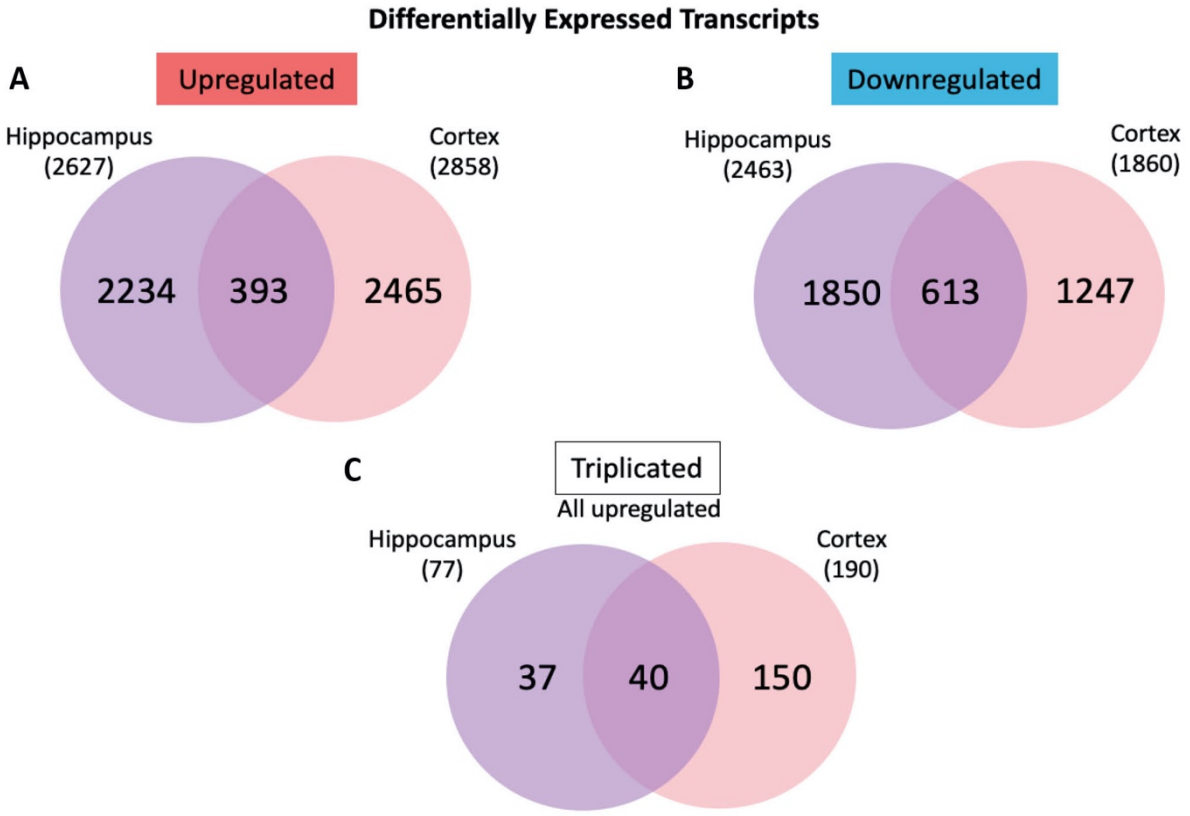


Figure 3.23. Transcript expression dysregulation in brain samples from individuals with DS. Venn diagram of differentially expressed transcripts in human hippocampus and cortex reveals common and non-common upregulated, downregulated and triplicated transcripts between the hippocampus and cortex.

Also, when we compared the log_2FC (calculated between DS and control samples) (Figure 3.24), for the protein-coding genes and transcripts, we found a more extensive range of log_2FC for both upregulated transcripts (right side of the curve in Figure 3.24A & B) and downregulated transcripts (left side of the curve). This reflects more considerable differences (between DS and control samples) when analyzing at the level of transcripts *vs.* genes. This observation, along with the attenuated cross-region overlap (Figure 3.23) at the transcript level, demonstrates the need to study splicing mechanisms in DS, which is underexplored.

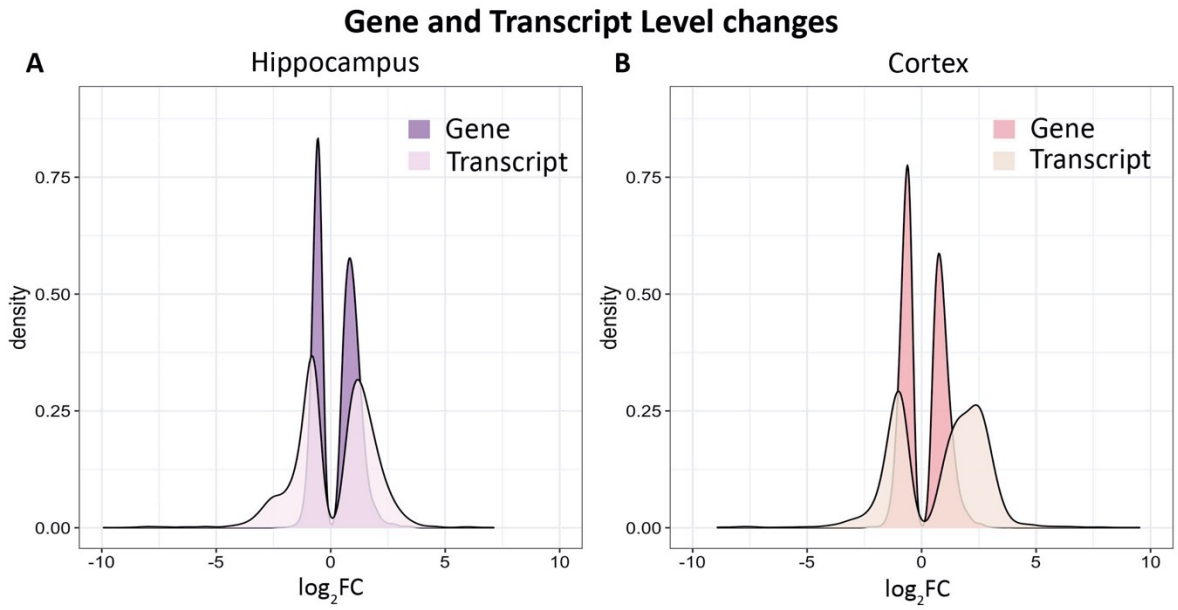


Figure 3.24. Larger range of fold changes for transcripts *vs.* genes. The log_2FC of all the identified genes and transcripts in the human hippocampus and cortex reveals larger range of fold changes for transcripts *vs.* genes. This density plot represents the number of genes with a certain distribution of log_2FC, and shows higher range of fold changes for transcripts (darker colors) in comparison to gene (lighter colors) for both A) hippocampus (violet) and B) cortex (pink). FC: Fold change.

3.6. Differentially expressed transcripts play important roles in many biological processes similar to differentially expressed genes.

As for DEGs, we also performed a Gene Ontology (GO) analysis on DE transcripts to determine the biological importance of these dysregulated transcripts. We found similar biological processes, cellular components, and molecular functions involved in both hippocampus and cortex when looking at the up-or down-regulated transcripts (Figure 3.25), as we had found at the gene-level changes.

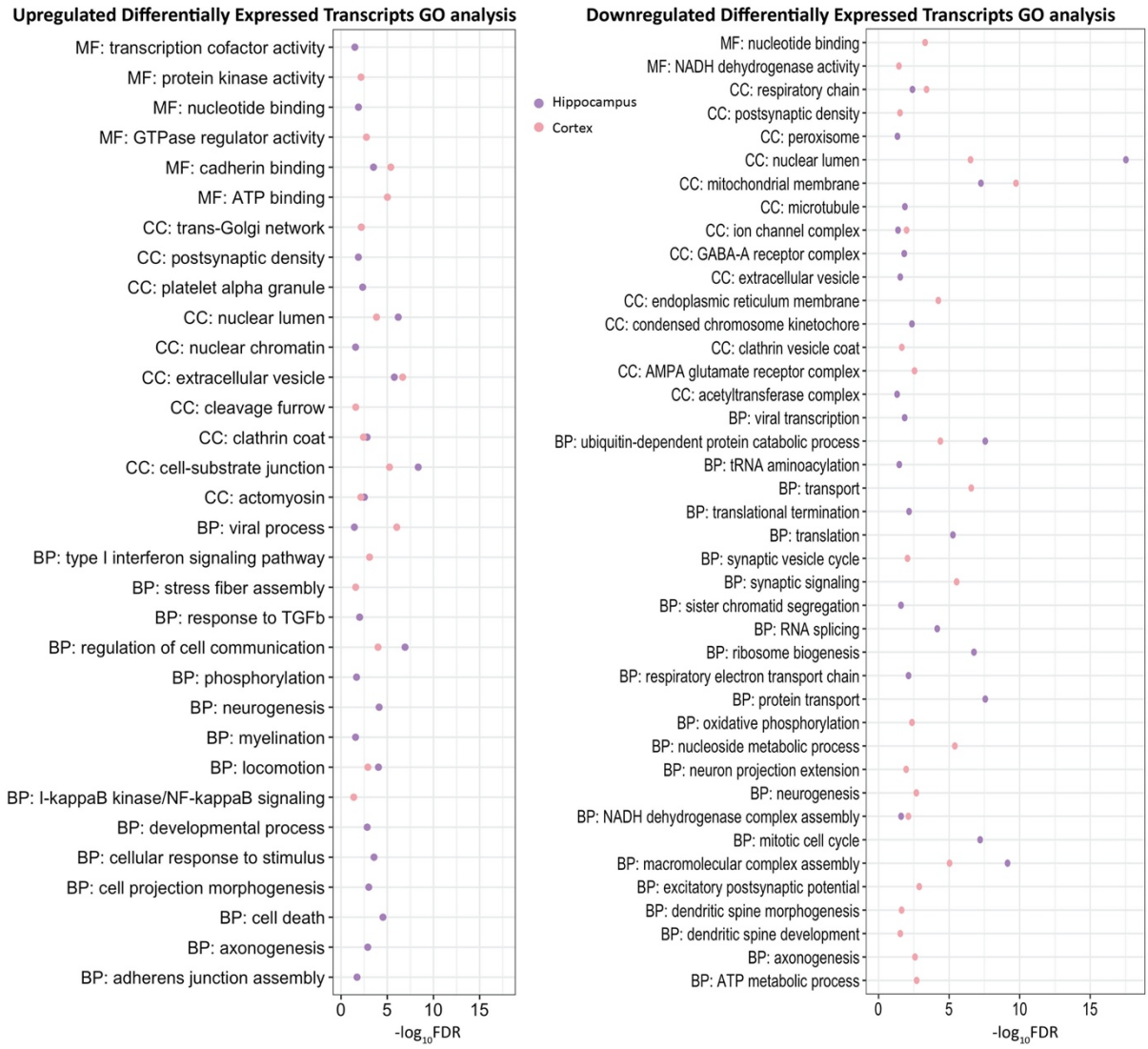


Figure 3.25. Differentially expressed transcripts show region-wise conserved and specific GO terms. Gene ontology analysis for differentially expressed transcripts in hippocampus and cortex shows some overlapping terms and few specific ones. On left, the GO analysis for upregulated transcripts and on right, for downregulated transcripts. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate

Similar to gene-level expression changes, we performed cell-type enrichment analysis for the DE transcripts (Figure 3.26) and found many similarities but few inconsistencies between the gene and transcript levels. In the hippocampus, for both upregulated and downregulated transcripts, we found significant enrichment for all the five cell types with higher values in microglia and oligodendrocytes for upregulated transcripts and higher values in interneurons, pyramidal cells, and oligodendrocytes in downregulated transcripts. On the other hand, the results for the cortex were more consistent with what we found at the gene level, but again with enrichment in oligodendrocytes for both up-and down-regulated transcripts. This observation suggests a dysregulation in all cell types in the hippocampus at the transcript

level and the cortex. There is consistent dysregulation in the neural-immune trajectory at the transcript level as observed at the gene expression. One unique feature from transcript level analysis was the enrichment of oligodendrocytes in both the regions when looking at up-and down-regulation.

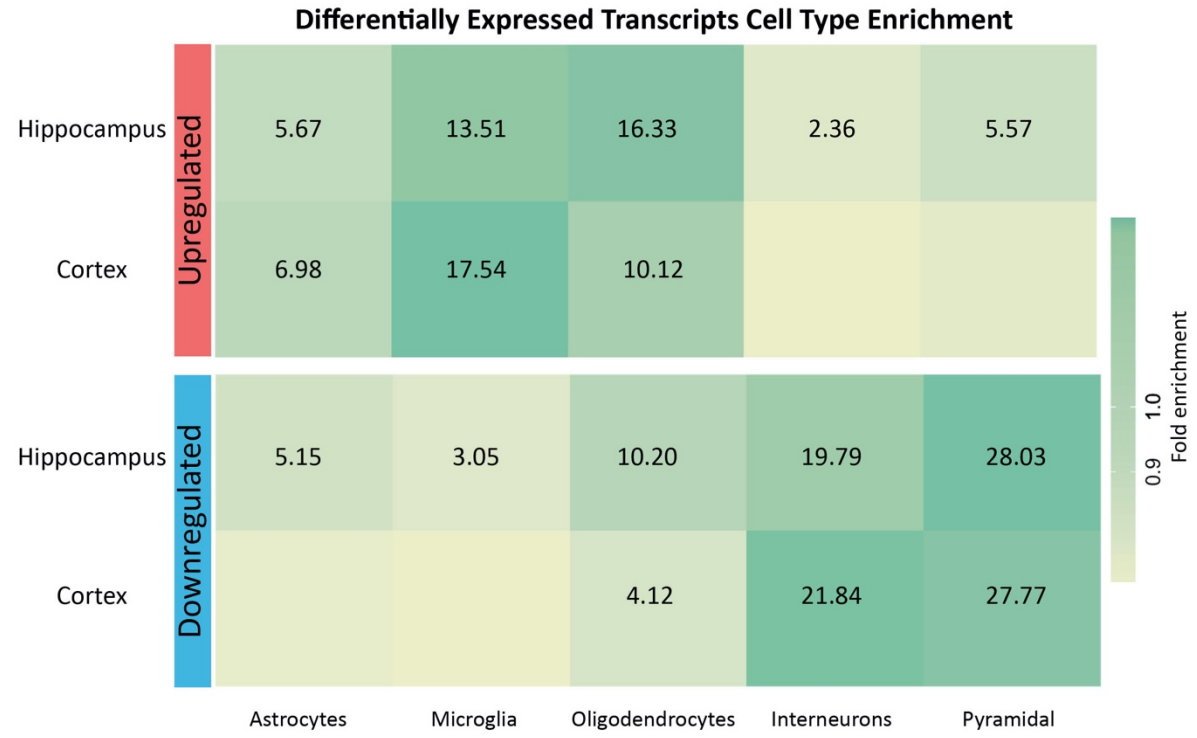


Figure 3.26. Differentially expressed transcripts show cell-type enrichment in human DS hippocampus and cortex. Cell-type enrichment represented by fold enrichment for the upregulated (top panel) and downregulated (bottom panel) transcripts in different cells for hippocampus and cortex. The numbers in the heatmap represent the z-score (number of standard deviations away from mean) for the significant cell enrichment (Bootstrap significance testing with 100,000 repetitions).

Further, we compared (Figure 3.27) the differentially expressed genes and transcripts for both the regions and found many transcripts that did not show DGE. These isoform-only DE genes (1489 in hippocampus and 1350 in the cortex) do not have changes at the gene-expression level where the expression is cumulative of all the transcripts of one particular gene, but rather the increase or decrease in expression is in the different transcript isoforms. This point further highlights the fact there are differences at the transcript level that are not present at the gene level. These differences could potentially be responsible for higher region-specificity.

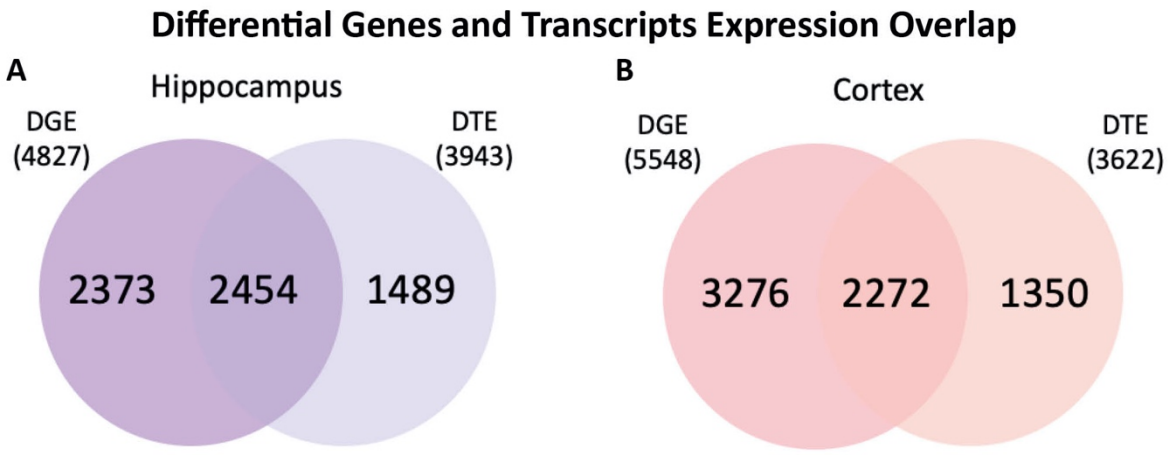


Figure 3.27. Venn diagram of all the differentially expressed genes and transcripts reveals common and non-common genes in A) human hippocampus and B) cortex DGE: Differential Gene Expression, DTE: Differential Transcript Expression.

3.7. Alternative splicing is responsible for generating multiple biologically important isoforms in DS.

To identify the underlying regulatory mechanisms and the differences we observed between the gene and transcript expression profiles in DS, we examined the role of splicing. Alternative splicing can generate diverse mRNA isoforms using a combination of alternative exons. We analyzed six different kinds of splicing events: alternative 5' and 3' splice sites, exon skipping, microexons, mutually exclusive exons, and intron retention. We performed alternative splicing analysis on hippocampus and cortex RNA sequencing data using a recently developed method, "Vast-tools," which calculates all the above mentioned splicing events. Splicing is calculated in terms of percent spliced-in (PSI) value, and the difference in the PSI for the same exon between two groups (in our case DS and controls) is known as deltaPSI. We found a large number of differential splicing events in both hippocampus and cortex. Overall, intron retention (IR) and exon skipping (SE) were the two most abundant events, and we focused our further analysis on these two kinds of events (Figure 3.28).

Exon skipping events consist of microexon events (these are the events where the exon's size is between 3-27 bp). So, we pooled these microexon events (Figure 3.28) together with the rest of the exon skipping events to finally get 888 events originating from 771 genes in the hippocampus (Figure 3.28). We also found few triplicated genes to have exon skipping events associated with them (Figure 3.29, red lines and numbers). We then functionally

characterized by GO analysis the differential exon skipping events. Only in the hippocampus, we found significant enrichment for axon and dendrite formation, protein localization, protein phosphorylation, and synaptic signaling (Figure 3.30) across upregulated events defined by PSI (DS) > PSI (control) (Figure 3.29). Overall, we had similar observations for the cortex, too, with 1086 exon skipping events in 837 genes (Figure 3.28) and enrichment in biological processes such as locomotion and stress fiber assembly. When we observed the downregulated events defined by PSI (DS) < PSI (Control) (Figure 3.29), we found no biological processes which were only specific to the hippocampus (Figure 3.30). However, it shared many processes and cellular components with the cortex, such as neurogenesis, neurotransmitter secretion as biological processes, cell-substrate junction, microtubule, nucleoplasm, and vesicle membrane as cellular components. The downregulated exon skipping events in the cortex was important for axon guidance, cell communication, CNS development, cilium morphogenesis, locomotion, protein transport, and synaptic signaling. Overall, exon skipping events in the cortex were responsible for many biological processes. Together, these results suggest that splicing dysregulation is a shared phenomenon between the two regions and is highly important in DS. The genes that are differentially spliced mostly corresponded to multiple steps involved in brain development.

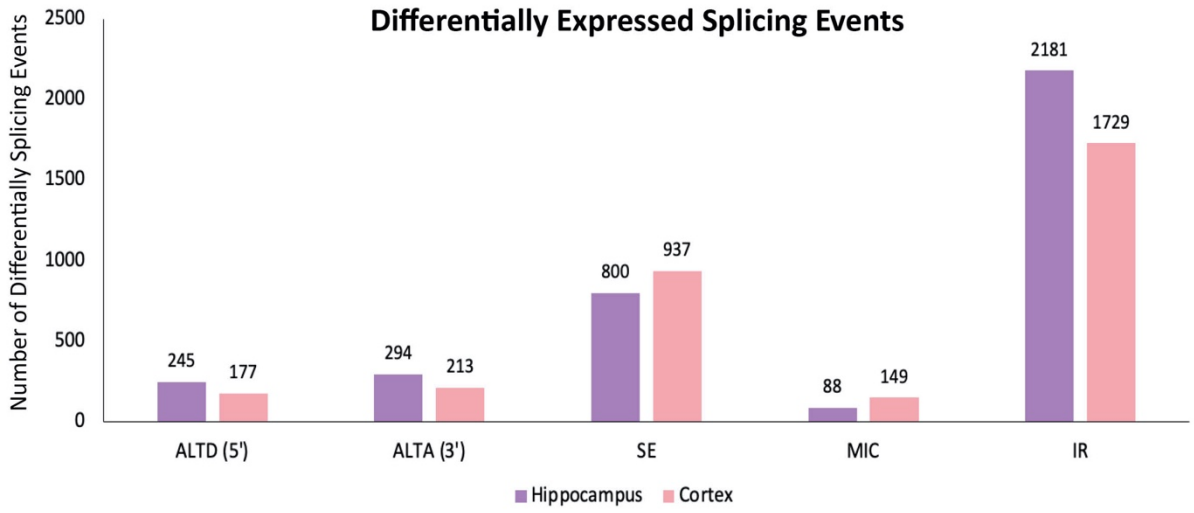


Figure 3.28. Distribution of alternative splicing events which were significant with |dPSI| >10 in both hippocampus and cortex. IR: Intron Retention, MIC: microexon, SE: Skipped exon, ALTA (3'): Alternative acceptor site (3' splice site), ALTD (5'): Alternative donor site (5' splice site).

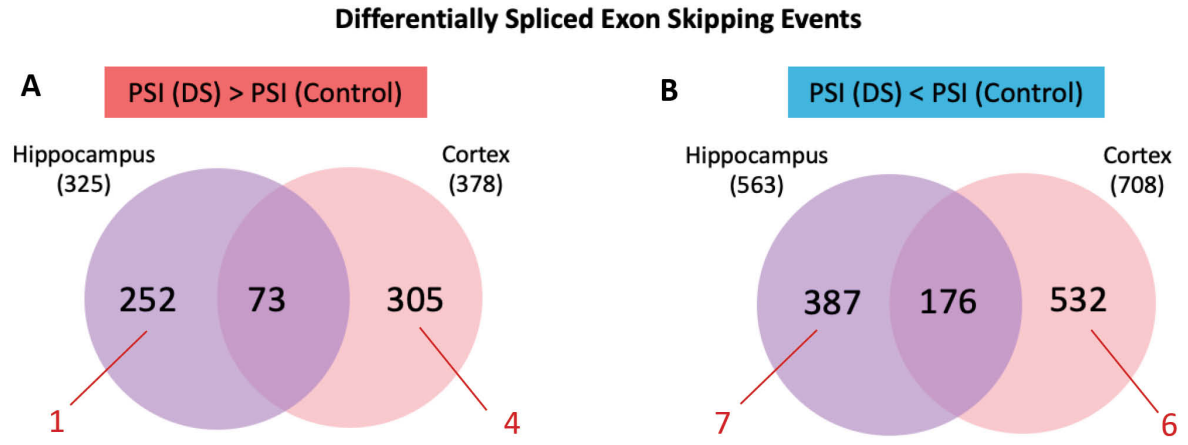


Figure 3.29. Overlap of differentially spliced exon skipping events. Venn diagram shows the overlap between Hippocampus and Cortex for the A) upregulated, B) downregulated and triplicated differential exon skipping events. Red numbers and lines indicate number of triplicated genes found in the respective overlaps. PSI: Percent spliced-in index, DS: Down syndrome.

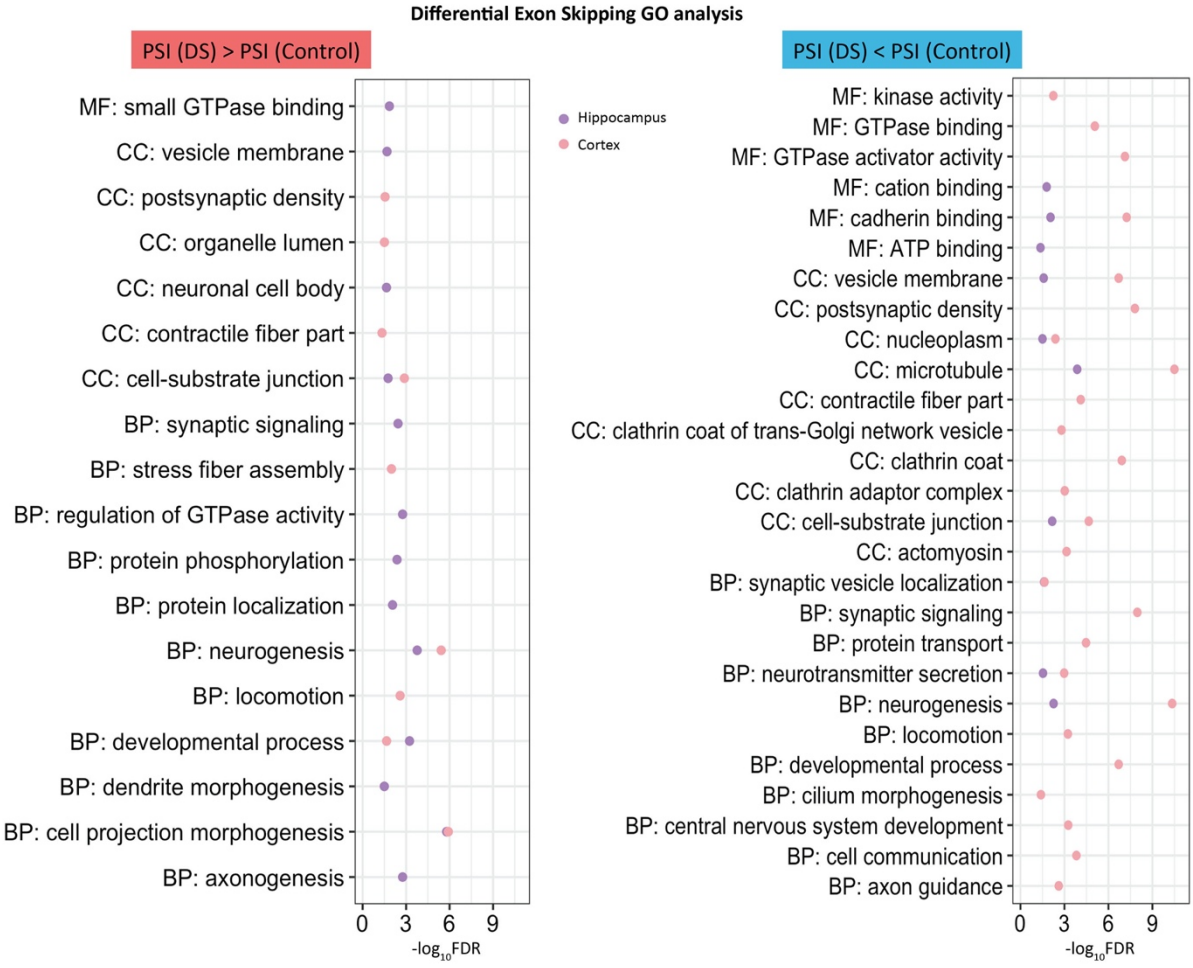


Figure 3.30. Differentially spliced exon skipping events show region-wise conserved and specific GO terms. Gene ontology analysis for significant exon skipping events in hippocampus and cortex shows biological importance for genes having less exon skipping in DS samples (PSI (DS) > PSI (Control)) on left and for genes having more exon skipping in DS samples (PSI (DS) < PSI (Control)) on right. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate

Next, we performed cell-type enrichment analysis on the significant skipped exon events from both regions (Figure 3.31). We observed that there was a broad cell-type enrichment except in microglia. This result suggests that splicing dysregulation might be more reflective for neural lineage, which produces neurons, astrocytes, and oligodendrocytes, rather than the hematopoietic lineage, which is responsible for generating microglia.

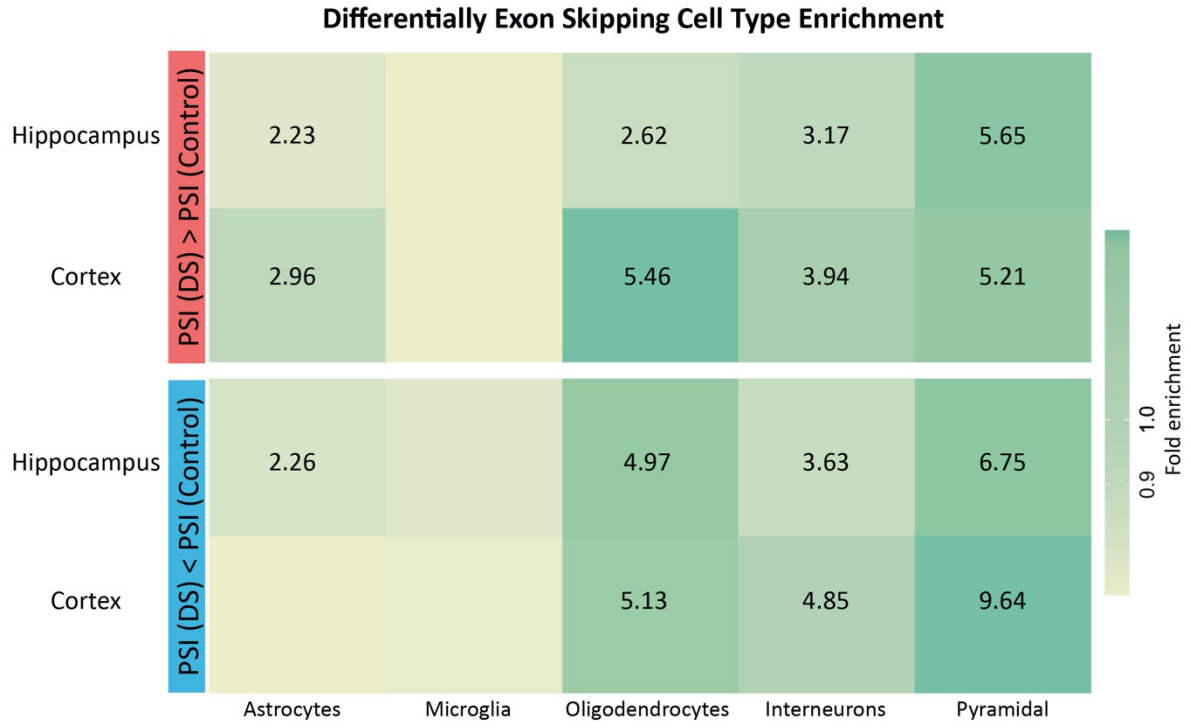


Figure 3.31. Differentially spliced exon skipping events show cell-type enrichment in human DS hippocampus and cortex. Cell-type enrichment analysis reveals enrichment for upregulated and downregulated exon skipping events in different cells for hippocampus and cortex. The numbers in the heatmap represent the z-score (number of standard deviations away from mean) for the significant enrichment (Bootstrap significance testing with 100,000 repetitions). The top panel indicates enrichment for exon skipping events where PSI (DS) > PSI (Control) while bottom panel represents events with PSI (DS) < PSI (Control). PSI: Percent Spliced-In.

When we performed the same analysis on spliced genes characterized by intron retention, we found 2181 events from 1708 genes in the hippocampus, where only 175 events were downregulated in DS samples and 1728 events from 1368 genes in the cortex with 362 events downregulated (Figure 3.28, 3.32). Very few of these events were associated with triplicated genes (Figure 3.32, red lines and numbers). As the number of downregulated events was far less than the upregulated events, and it was impossible to perform GO analysis on fewer downregulated events, we decided to pool all the genes with a dysregulation in intron retention for further biological analysis. Gene ontology (GO) analysis on these differentially

spliced intron retention events revealed that they were involved in cilium organization, macromolecule modification, regulation of GTPase activity processes, and localized in microtubule, nuclear lumen, and postsynaptic density for both hippocampus and cortex. Specifically, in the hippocampus, they are involved in RNA localization, RNA splicing, and ion binding and in the cortex, in macromitophagy, mitosis, neurogenesis, neurotransmitter secretion, and kinase activity (Figure 3.33). Overall, in this RNA study, we provided the first detailed evidence of splicing dysregulation in DS individuals with the help of total RNA sequencing.

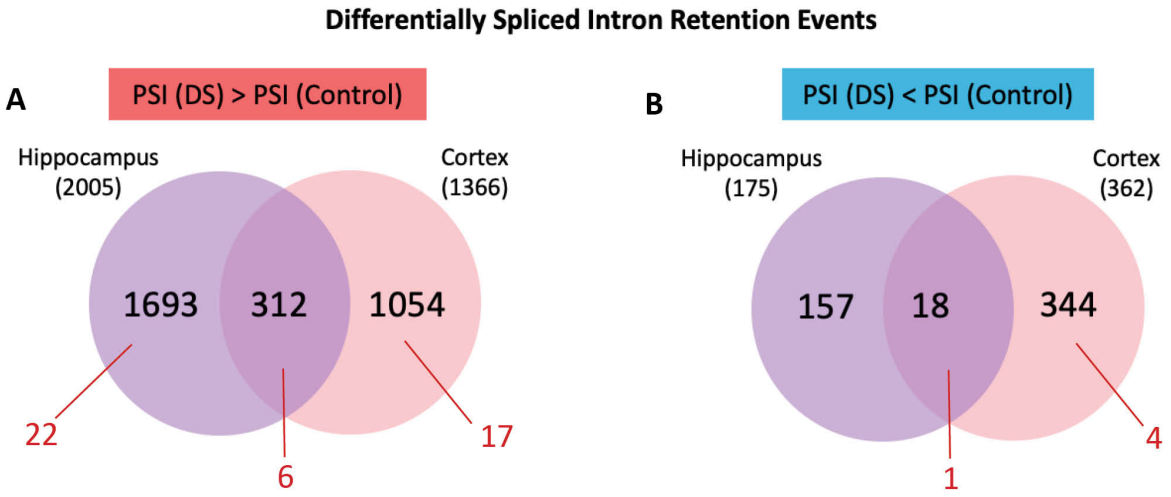


Figure 3.32. Overlap of differentially spliced intron retention events Venn diagram shows the overlap between Hippocampus and Cortex for the upregulated, downregulated and triplicated differential intron retention events. Red numbers and lines indicate number of triplicated genes found in the respective overlaps. PSI: Percent spliced-in index, DS: Down syndrome.

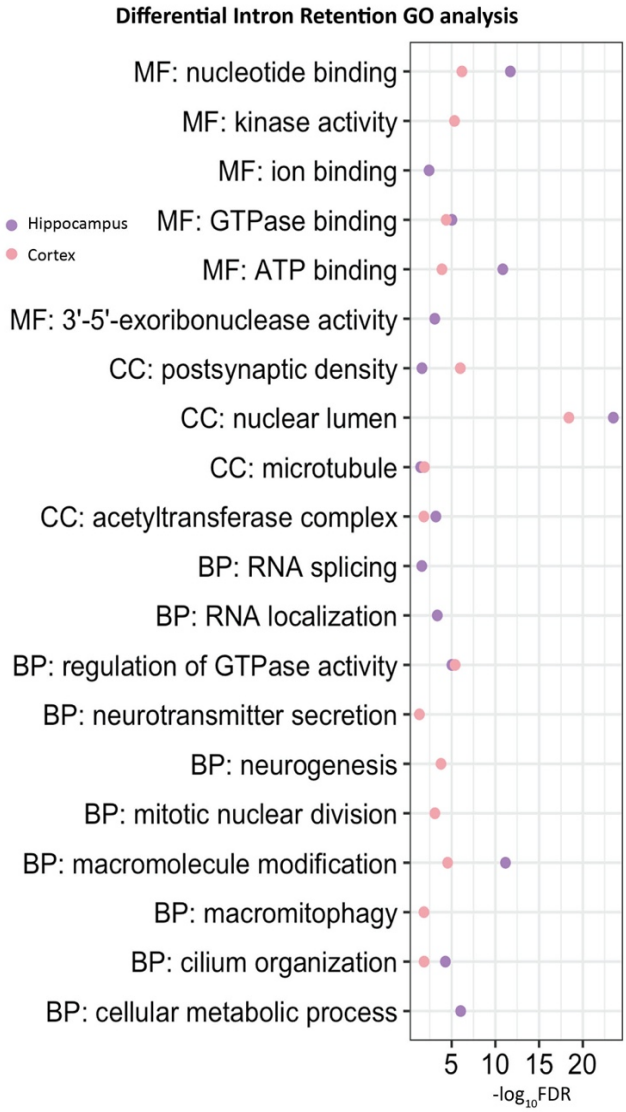


Figure 3.33. Differentially spliced intron retention events show region-wise conserved and specific GO terms. Gene ontology analysis for significant intron retention events in hippocampus and cortex shows specific and common GO terms. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate

3.8. RNA-binding Proteins regulates alternative splicing and are dysregulated in DS.

Extensive alteration of splicing is often controlled by RNA binding proteins (RBPs). Indeed, RBPs can impact alternative splicing by enhancing or repressing exon inclusion to generate multiple RNA transcripts and impact the overall gene expression. Thus, we looked at the RBPs, which were differentially expressed in DS and could influence the alternative splicing of specific genes that we found to be differentially expressed either at gene or transcript level. To this aim, we downloaded a list of RNA-binding proteins from the literature (Gerstberger et al., 2014) obtained by a detailed manual curation of all the known RBPs in the human genome.

First, we checked the expression of the RBPs from the literature list in our results on DGE and DTE from the hippocampus. From this list of 1542 RBPs, which represents ~7.5% of protein-coding genes, we found several RBPs to be either DGE (445) or DTE (420) in our samples (Figure 3.34). Seven RBPs belonged to triplicated genes at the DGE level and three at the DTE level. ADARB1, responsible for A to I RNA editing, is the only RBP that originates from chr21 and present at the differential gene and transcript levels.

Similarly, in the cortex, we found several RBPs to be either DGE (492) or DTE (311) (Figure 3.34). Nine RBPs are triplicated at DGE level and nine at DTE level, and interestingly, all of them were the same at both gene and transcript level. Although, at least for now, we do not follow the effect of these differentially expressed RBPs on the dysregulation we already observed in DS at gene and transcript level, it is clear that the number of RBPs we found to be differentially expressed are many and could potentially impact the expression levels of a large number of genes.

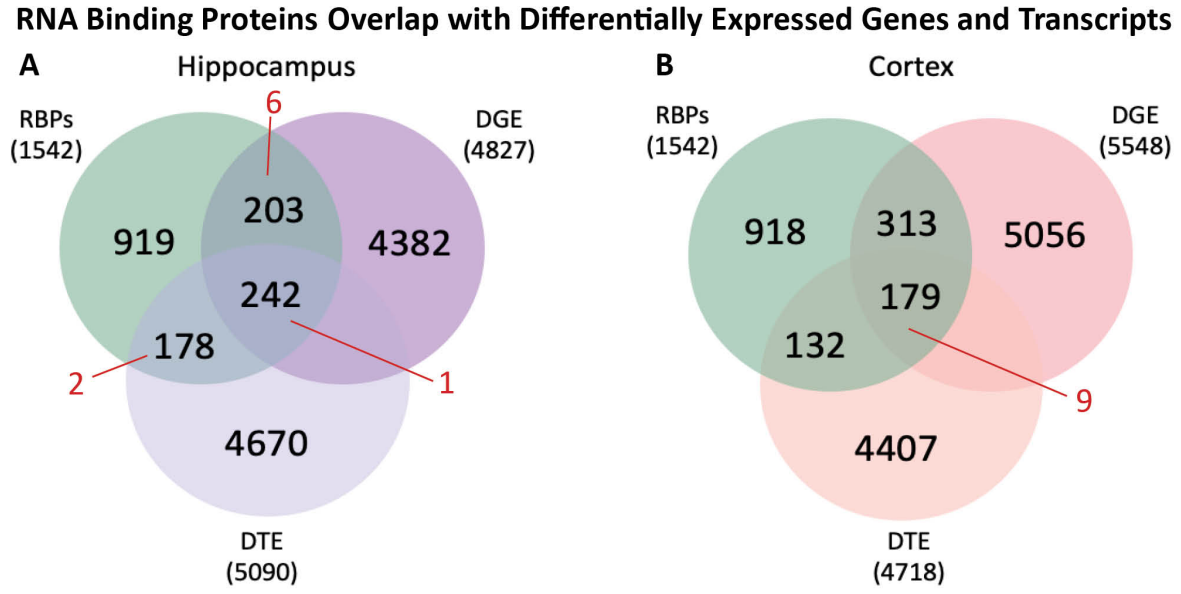


Figure 3.34. Overlap of 1542 RBPs with the differential expressed genes and transcripts in A) hippocampus and B) cortex reveals differentially expressed RBPs at gene and transcript level. DGE: Differential Gene Expression, DTE: Differential Transcript Expression, RBP: RNA Binding Protein. Red numbers and lines indicate number of triplicated genes found in the respective overlaps.

Next, we also looked at whether these RBPs were differentially spliced too in the hippocampus. Here, we focused only on those genes which were skipping exons, excluding other kinds of alternative splicing events for simplicity. In the hippocampus, we found 63

differentially spliced genes encoded by RNA binding proteins, among which 12 were differentially expressed at the gene level and 25 at the transcript level. When we overlapped differentially spliced RBPs to DGE and DTE, we found 9 RBPs (Figure 3.35A), PRPF18, MRPS18C, MRPL47, EIF4G3, ADARB1, PTBP2, TCERG1, DIS3, and PRPF40A. Among these, again, ADARB1 is the only triplicated RBP that is also differentially spliced.

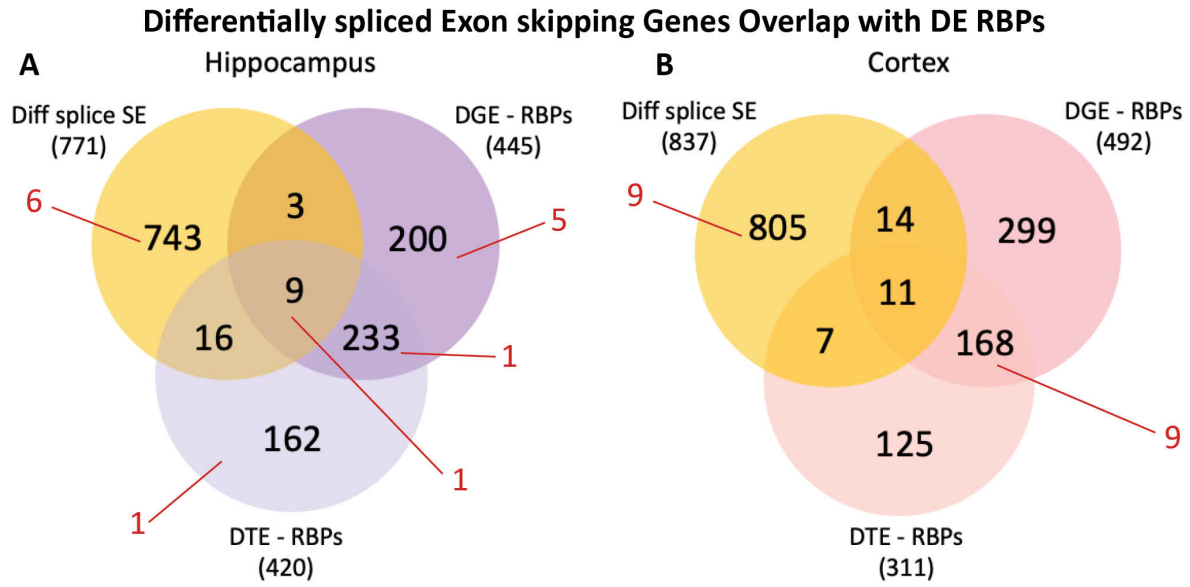


Figure 3.35. Overlap of differentially expressed RBPs at gene and transcript level with differentially spliced genes with exon skipping event in A) Hippocampus and B) Cortex reveals those RNA binding proteins which are differentially spliced by exon skipping event. DGE: Differential Gene Expression, DTE: Differential Transcript Expression, RBP: RNA Binding Protein, SE: Skipped Exon, Diff splice: Differentially spliced.

Among the overlap, we also found polypyrimidine tract binding protein 2 (PTBP2). PTBP2 belongs to a family of multifunctional RBPs, which play a significant role in tissue-specific transcriptional programs. PTBP2 is a crucial regulator of alternative splicing in neural progenitors and immature neurons. Most importantly, PTBP2 is known to be enriched in neurons and involved in axonogenesis (Zhang et al., 2019). This was particularly interesting considering our DTE results associated with cytoskeleton and axon formation to be differentially spliced.

Similarly, in the cortex, we assessed whether the differentially expressed RBPs were also differentially spliced. We found 25 RBPs to be differentially spliced at the gene level and 18 RBPs at the transcript level (Figure 3.35B). In the overlap between differentially spliced RBPs

to DGE and DTE, we found 11 RBPs, which are SRRM2, TCERG1, RBM3, PRPF18, YTHDC2, MRPS18C, MRPL33, EIF4G3, DAZAP1, G3BP2, and CELF2. None of these were triplicated.

With multiple lines of evidence at gene, transcript, and alternative splicing level in hippocampus towards dysregulation in axon formation (Figures 3.5, 3.17, 3.25, and 3.30), we decided to validate some of the exon-skipping events in human hippocampal samples with semi-quantitative PCR (Figure 3.36). Among all 888 exon skipping events from 771 genes, we selected those genes involved in axon formation and one of the RBPs, PTBP2, which is known to act on these genes and is itself alternatively spliced (Figure 3.35A). These results confirmed that alternative splicing is involved in genes responsible for axon formation and is highly dysregulated in DS.

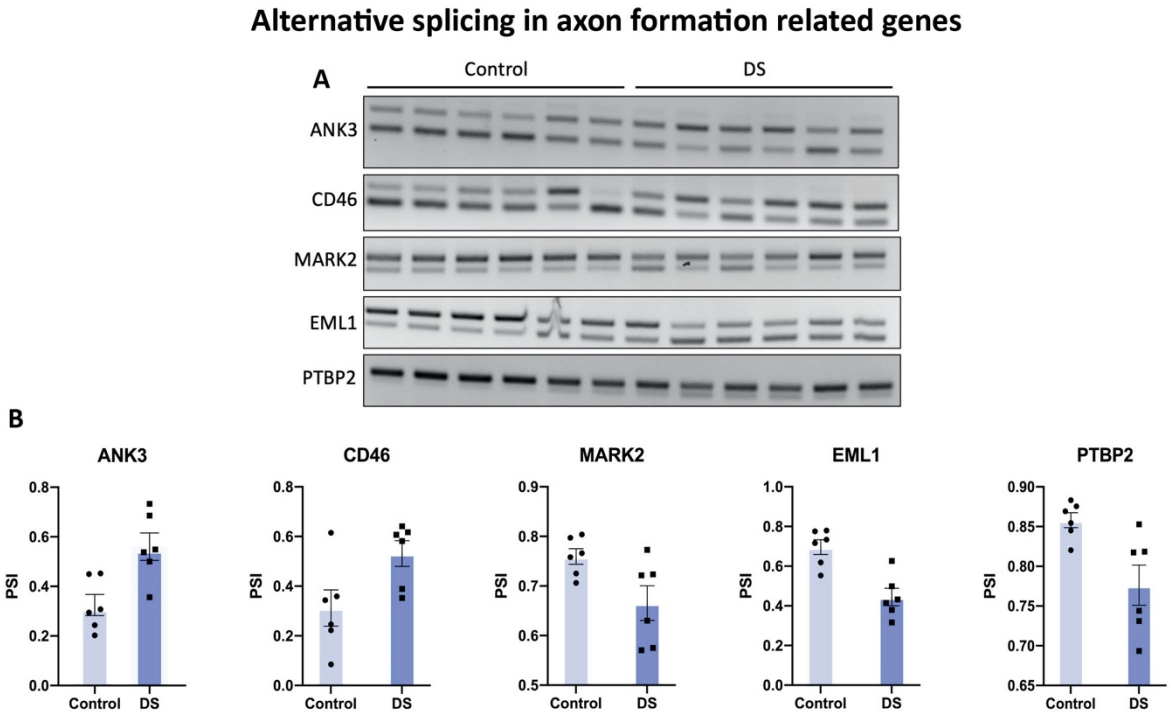


Figure 3.36. Semi-quantitative PCR validation and quantitation for some genes involved in axon formation, as indicated by the experiments performed in Figures 3.5, 3.17, 3.25 and 3.30 confirmed that axonogenesis related genes are differentially spliced and PTBP2 may be responsible for this effect. A) Agarose gel images for each gene with 6 control and 6 DS hippocampal samples. B) Quantitation of the two bands for each gene to calculate amount of splicing denoted by PSI. All the quantitations are significant ($p<0.05$, student's t-Test) PSI: Percent spliced-in index, ANK3: ankyrin 3, CD46: cluster of differentiation 46, MARK2: Microtubule affinity regulating kinase 2, EML1: EMAP like 1, PTBP2: Polypyrimidine Tract Binding Protein 2.

3.9. Protein expression alteration provide another point of view for DS.

Differential gene or transcript expression often results in proteins that are also differentially expressed. This is because alternative splicing results in different transcripts for a gene, producing protein isoforms with different functions and possible changes in the expression levels. So, to complement our results from RNA sequencing, we compared the transcriptome with proteome. First, we performed proteomics on the same set of samples for both hippocampus and cortex. After PCA analysis and hierarchical clustering, we obtained the same samples as we obtained in RNA sequencing, confirming the consistency in the samples we selected (6 samples each for control and DS individuals for hippocampus and 8 samples each for cortex).

Next, we performed differential expression analysis on both the hippocampus and cortex proteomic data (Figure 3.37). We found 741 and 417 up and down-regulated proteins in the hippocampus and 542 and 487 up and down-regulated proteins in the cortex. Overall, we found 285 upregulated proteins common between the two regions and 142 overlapping downregulated proteins. Some triplicated proteins were also differentially expressed, but fewer than what we observed at the gene-expression level.

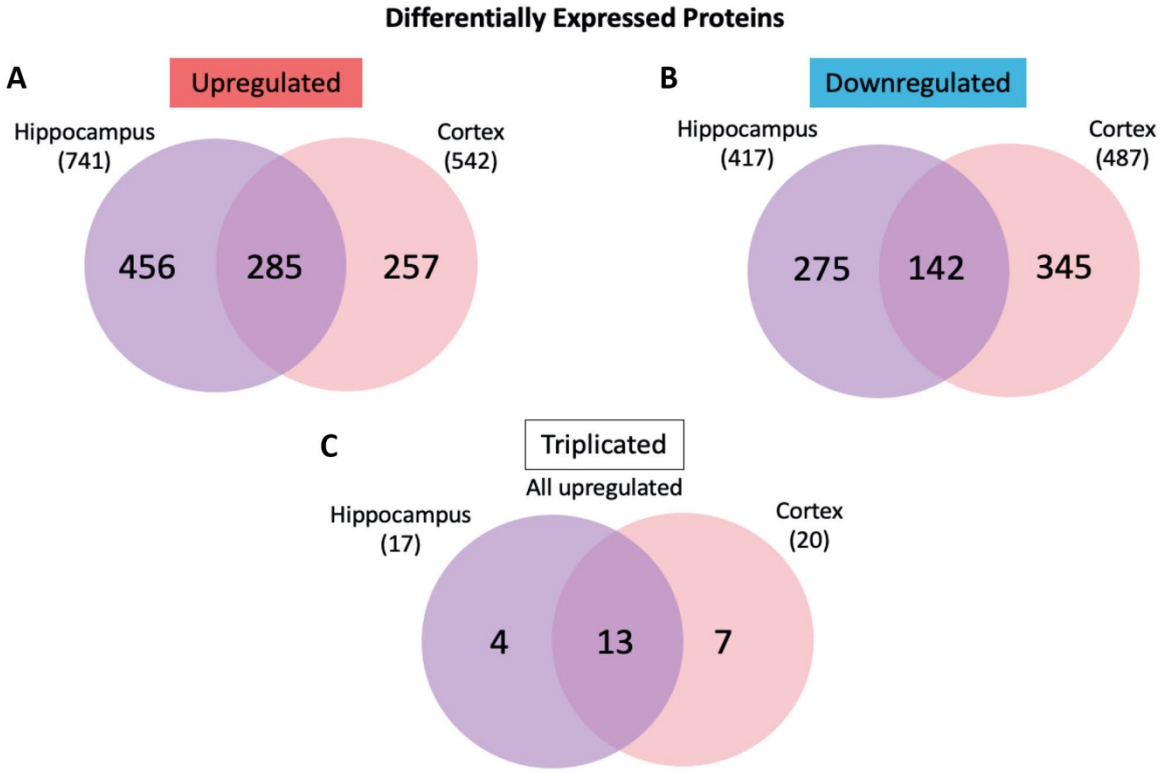


Figure 3.37. Venn diagram of differentially expressed proteins from hippocampus and cortex shows common, non-common upregulated, downregulated and triplicated proteins between hippocampus and cortex.

3.10. Multiple biological processes are dysregulated at protein level.

Next, we performed the Gene Ontology analysis and looked for enrichment in various biological processes, organelle localization, and molecular functions. The upregulated proteins in the hippocampus and cortex indicated enrichment in viral process, transport, translation, response to unfolded protein, protein targeting to ER, protein complex biogenesis, and many other shared processes between the two regions (Figure 3.38). Only cell-substrate junction was shared between the hippocampus and cortex in all the cellular components for upregulated proteins. The downregulated proteins from both regions showed enrichment in neurotransmitter secretion, synaptic signaling, and synaptic vesicle transport (Figure 3.39). In terms of cellular components, the downregulated proteins showed enrichment in the postsynaptic density, mitochondrial membrane, trans-Golgi network, and terminal bouton.

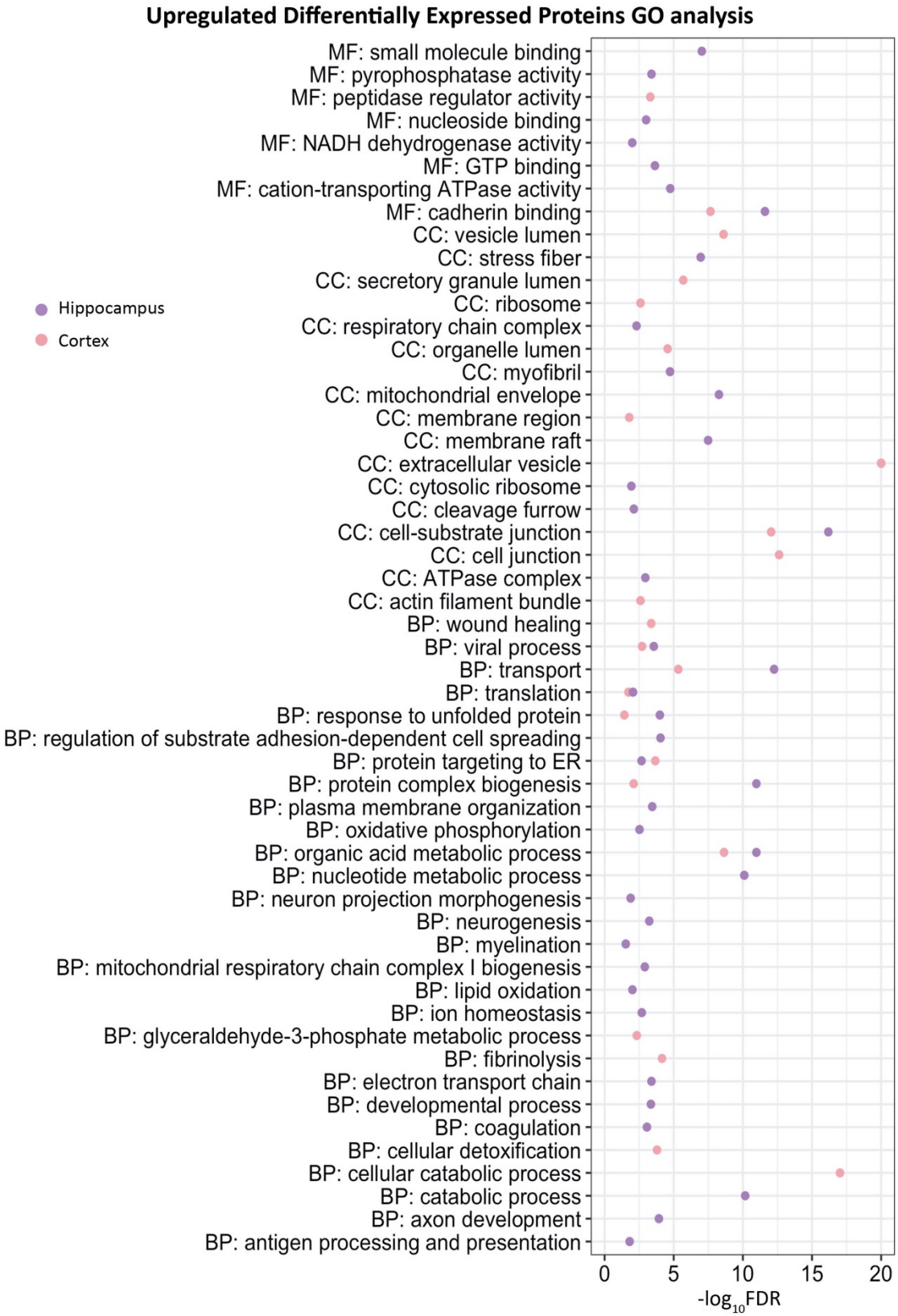


Figure 3.38. Differentially expressed upregulated proteins show region-wise conserved and specific GO terms. Gene ontology (GO) analysis for upregulated proteins for both hippocampus and cortex shows some overlapping terms and few specific ones. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate

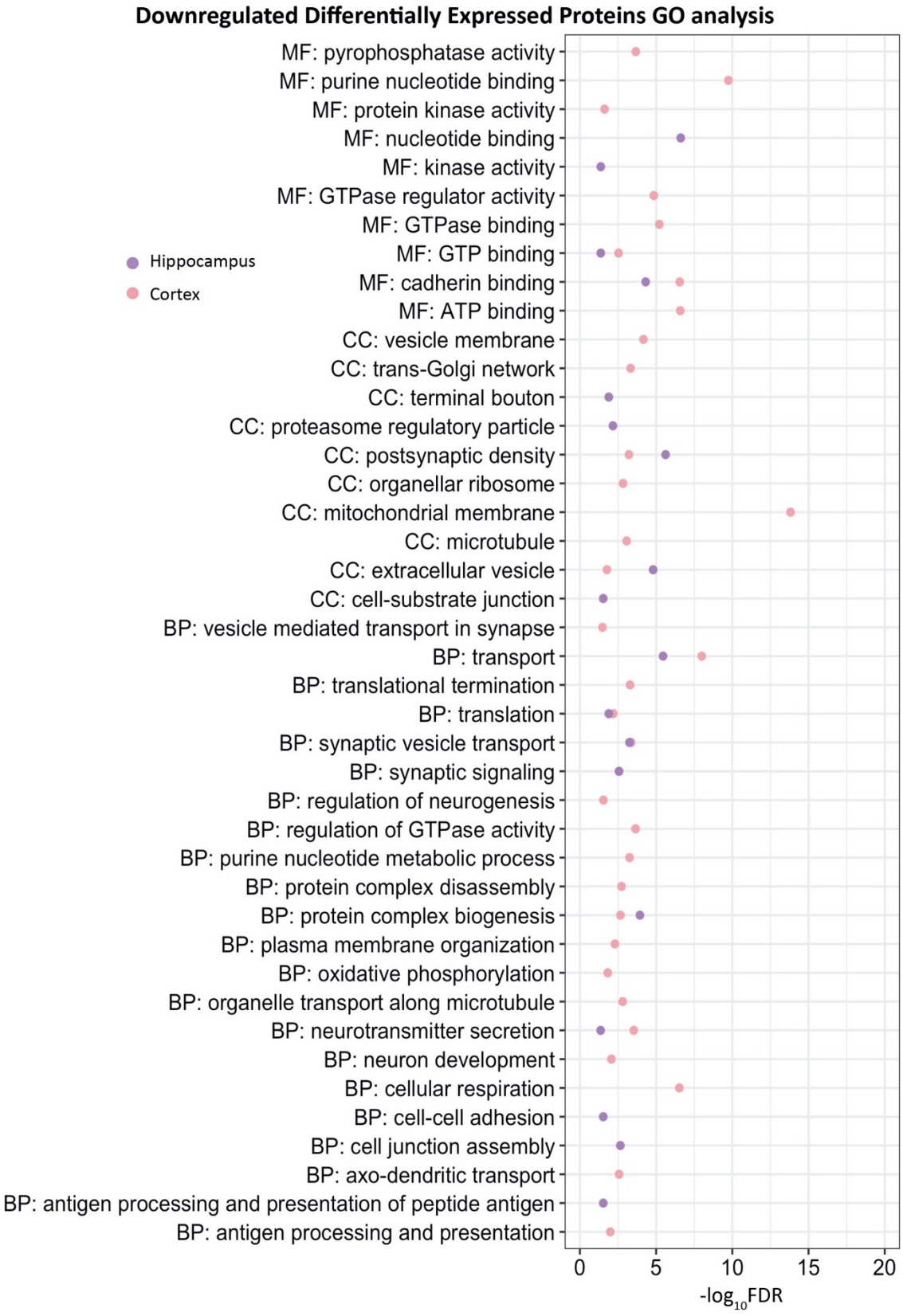


Figure 3.39. Differentially expressed downregulated proteins show region-wise conserved and specific GO terms. Gene ontology (GO) analysis for downregulated proteins for both hippocampus and cortex shows some overlapping terms and few specific ones. Purple dots represent hippocampus and pink dots represent cortex. BP: Biological Processes, CC: Cellular Components, MF: Molecular Function, FDR: False Discovery Rate

We next performed EWCE analysis to look for overrepresentation in different cell types. The upregulated proteins in the hippocampus showed enrichment in all cell types. (Figure 3.40, top). On the other hand, the upregulated proteins from the cortex showed similar enrichment to upregulated genes, i.e., in all the glial cells. The downregulated proteins from the hippocampus and cortex were enriched in interneurons and pyramidal cells (Figure 3.40, bottom), similar to what we observed at the gene level. Nevertheless, this analysis's most striking observation was the enrichment in oligodendrocytes that we found for both the upregulated and downregulated proteins in both regions, suggesting a high dysfunction taking place in these cells. This observation is supported by a previous study highlighting the role of oligodendrocytes dysregulation in DS (Olmos-Serrano et al., 2016).

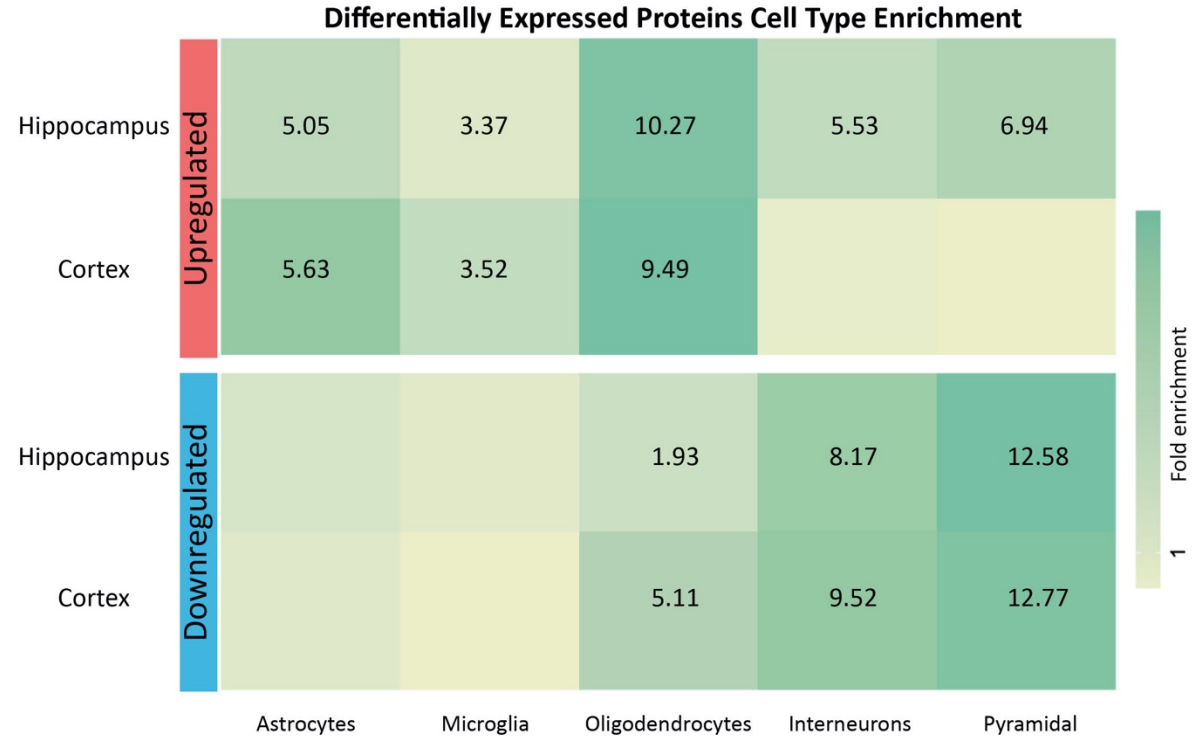


Figure 3.40. Differentially expressed proteins show cell-type enrichment in human DS hippocampus and cortex. Cell-type enrichment reveals differences among the upregulated (top panel) and downregulated (bottom panel) proteins in different cells for hippocampus and cortex. The numbers in the heatmap represent the z-score (number of standard deviations away from mean) for the significant enrichment (Bootstrap significance testing with 100,000 repetitions).

3.11. Co-expression analysis helps identify drivers of proteome dysregulation.

At the gene level, we observed that the co-expression network analysis arranged the genes into clusters related to function, cellular location, and cell types. We wanted to investigate

whether the protein level information provided us with somewhat similar observations. We performed Weighted Protein Co-expression Network Analysis (WPCNA) and found many protein modules associated with DS. We identified 16 modules from the hippocampal samples (Figure 3.41) and 14 from the cortical samples (Figure 3.42). We found three upregulated modules (P-HH-2, P-HH-4, P-HH-5) and three downregulated modules (P-HH-1, P-HH-3, P-HH-6) to be significantly associated with DS and control individuals, respectively in the hippocampus (Figure 3.41). While, there were two modules (P-HC-1, P-HC-3) associated with DS and three modules (P-HC-2, P-HC-4, P-HC-5) associated with control samples in the cortex (Figure 3.42). None of these significant modules were influenced by the other biological (age and sex) or technical (PMI) factors.

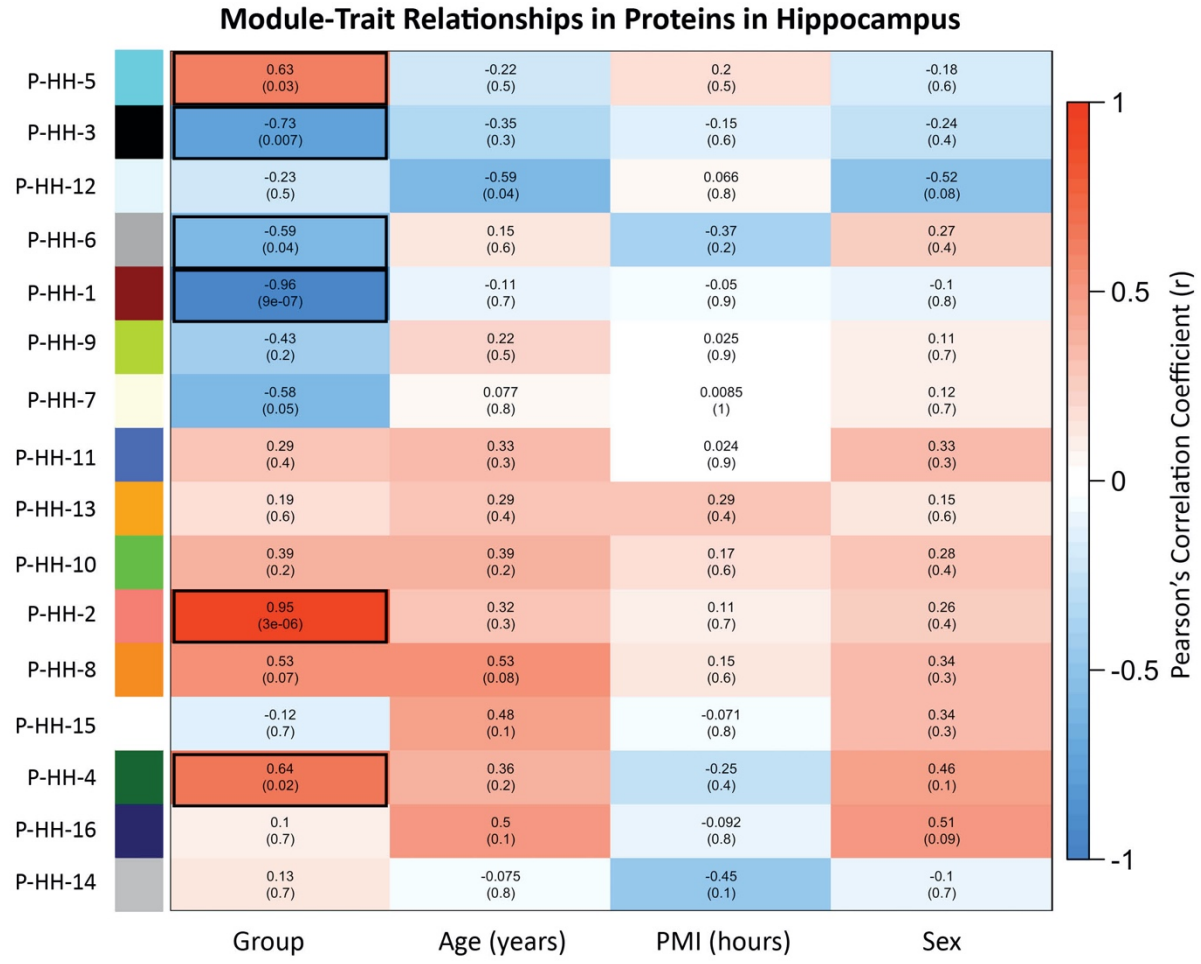


Figure 3.41. WPCNA-derived modules from human hippocampus protein expression data capture disease association. The modules were tested for association against Group (Control, DS), Age in years, post-mortem interval (PMI) in hours and Sex (F,M). The top numbers inside the heatmap show the color-coded (color bar on the right) Pearson's correlation coefficient associated with each trait and the number in parenthesis indicate their statistical significance in terms of p-value. The black boxes represent the modules having significant association with either group category ($p<0.05$, Pearson's correlation test). P-HH represents protein modules in human hippocampus.

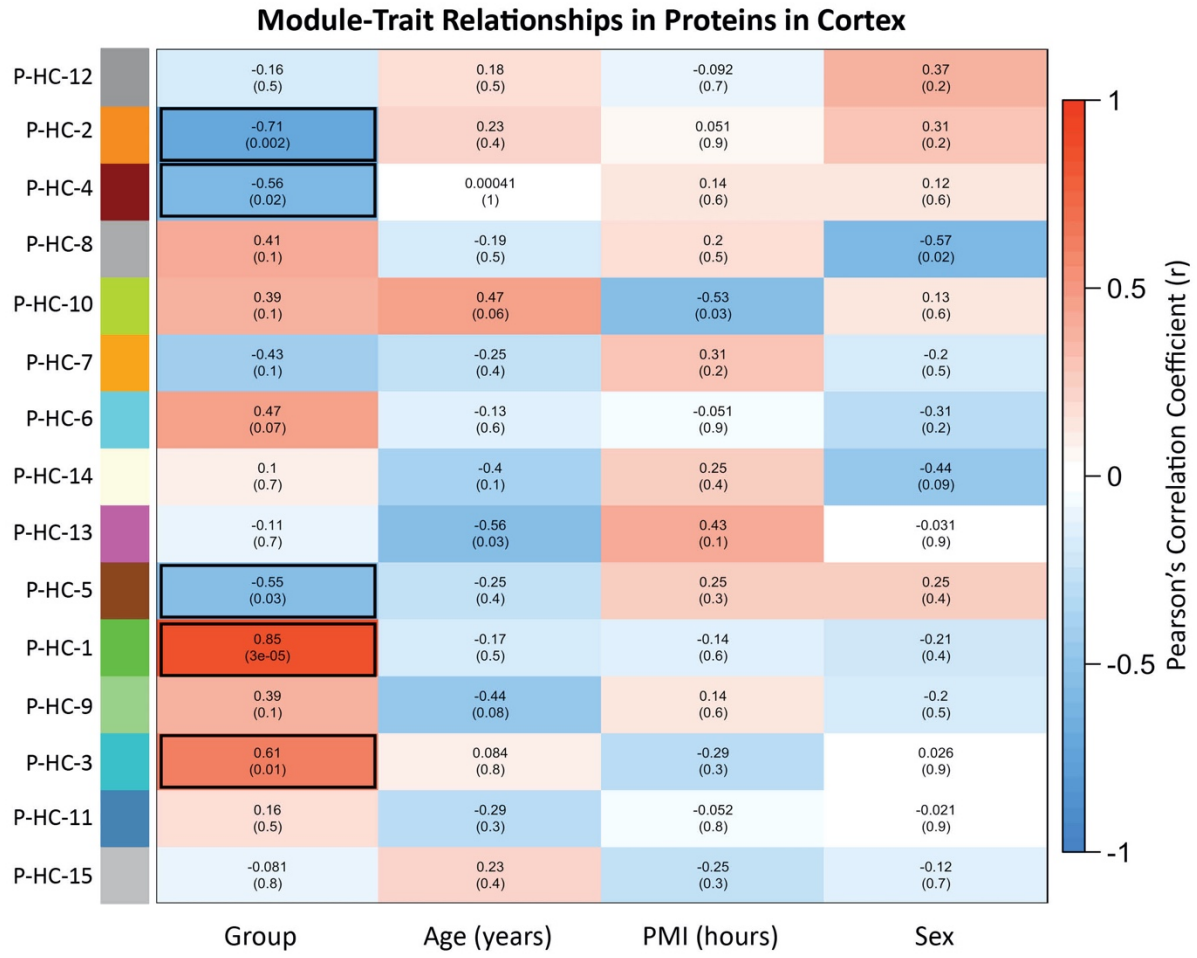


Figure 3.42. WPCNA-derived modules from human cortex protein expression data capture disease association. The modules were tested for association against Group (Control, DS), Age in years, post-mortem interval (PMI) in hours and Sex (F,M). The top numbers inside the heatmap show the color-coded (color bar on the right) Pearson's correlation coefficient associated with each trait and the number in parenthesis indicate their statistical significance in terms of p-value. The black boxes represent the modules having significant association with either group category ($p<0.05$, Pearson's correlation test). P-HH represents protein modules in human hippocampus.

We next sought to determine the biological significance of these significant modules from the hippocampus and cortex. Similar to gene ontology analysis performed on data from modules derived from gene-expression experiments, we performed the same analysis on the modules from proteomics data for both regions. We found that the upregulated modules P-HH-2, P-HH-4, P-HH-5 from hippocampus and P-HC-1, P-HC-3 from cortex showed significant enrichment in extracellular vesicle (Figure 3.43 left). Moreover, four of these upregulated modules from both the regions showed enrichment in cell-substrate junction. This highlighted that the upregulated modules might be involved in cell-cell communication, which usually occurs with the help of vesicles released by the cells or establishing contact between different cell types.

On the other hand, the downregulated protein modules from the hippocampus, P-HH-3, and cortex, P-HC-2 (Figure 3.43 right) shared many biological processes and cellular components (e.g., antigen processing and presentation, membrane fusion, neurogenesis, neurotransmitter secretion, protein transport, synaptic signaling, cell-substrate adhesion). Interestingly, many downregulated modules from both regions showed enrichment in extracellular vesicles similar to upregulated modules. This suggests that the cell communication which results from the release of the extracellular vesicle is highly impaired in DS.

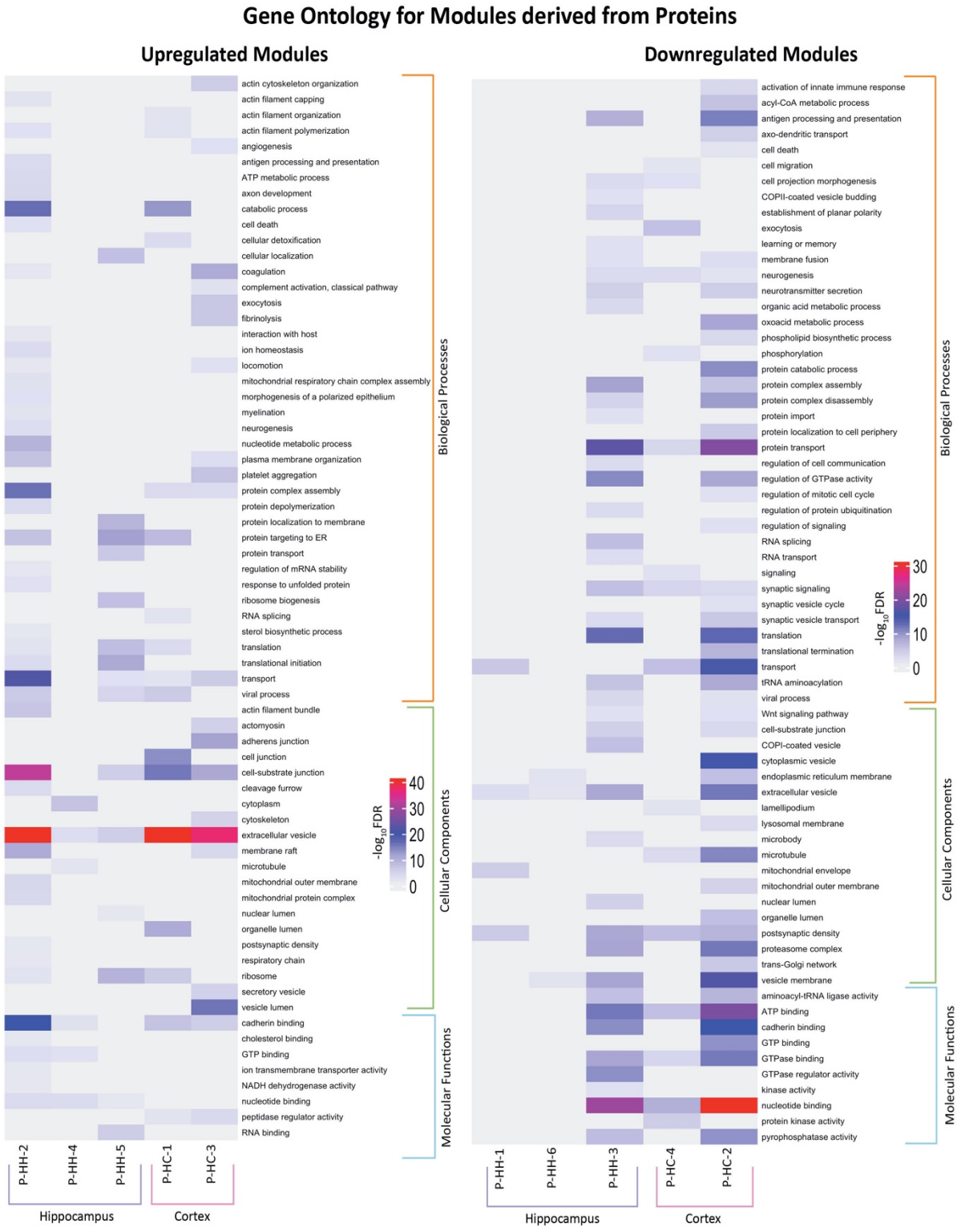


Figure 3.43. WPCNA derived co-expression modules associated with multiple GO terms. Multiple Biological Processes (BP), Cellular Components (CC) and Molecular Functions (MF) terms are enriched for significant modules from both human hippocampus and cortex protein expression data. The GO terms are separated by parenthesis, where orange represents biological processes, green represents cellular components and blue represents molecular functions. The region-wise modules are also separated by parenthesis, where purple shows hippocampus and pink shows cortex. The color bar on the right shows the range of significance defined by -$\log_{10}$FDR (after Benjamini-Hochberg multiple testing correction) for each term and each module. FDR: False Discovery Rate, P-HH: Protein module in Human hippocampus, P-HC: Protein module in Human cortex.

We further resolved the biological significance of these significant modules by determining their enrichment with cell-type-specific profiles in the brain (Figure 3.44). Upregulated protein modules for both hippocampus and cortex were broadly significant in all cell types. This result is consistent with the modules obtained from protein coding-genes at the RNA level. Moreover, the downregulated protein modules in both hippocampus and cortex were significantly associated with neural cell types such as interneurons and pyramidal cells, similar to our observations from modules derived from gene-expression data. Interestingly, one module from hippocampus P-HH-3 and cortex P-HC-2 were also associated with oligodendrocytes.

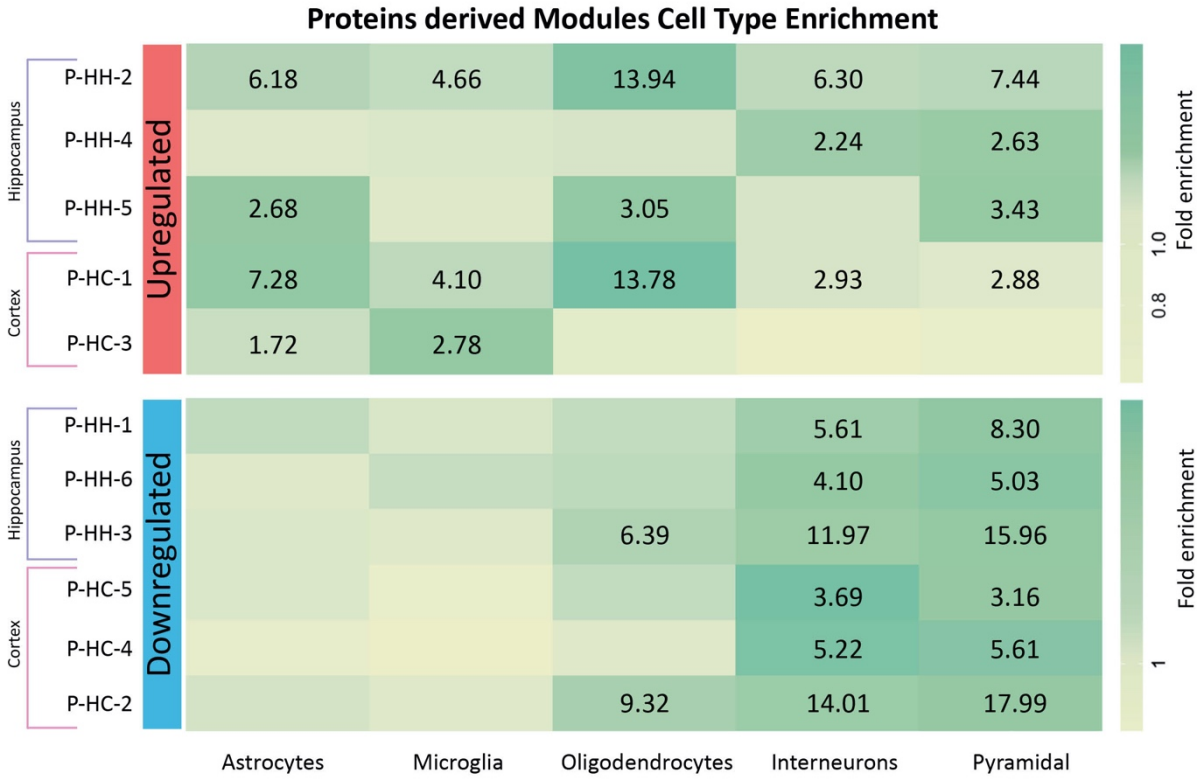


Figure 3.44. WPCNA-derived modules capture differences in cell enrichment from both human hippocampus and cortex protein expression. They are separated by their enrichment in different cell types and divided in up- (top panel) or down-regulated modules (bottom panel). The color bars on the right for upregulated and downregulated modules are different and represent the fold enrichment. The region-wise modules are also separated by parenthesis, where purple shows hippocampus and pink shows cortex. The numbers inside the heatmap squares show the z-score (number of standard deviations away from mean) for only significant cell enrichment (Bootstrap significance testing with 100,000 repetitions). P-HH: Protein module in Human hippocampus, P-HC: Protein module in Human cortex.

On seeing a considerable overlap between biological processes, cellular components, and cell enrichment between the modules from the proteomic experiments, we prepared a correlogram to find out how much correlation was present between each module from both

hippocampus and cortex (Figure 3.45). The correlation significance was tested using hypergeometric Fisher's exact test, and the modules showed a high (red), anti (green), and low (white) correlation in protein membership. We found a high correlation (R=0.7, p=0.01) between P-HH-3 and P-HC-2, consistent with the biological and cell enrichment we observed previously.

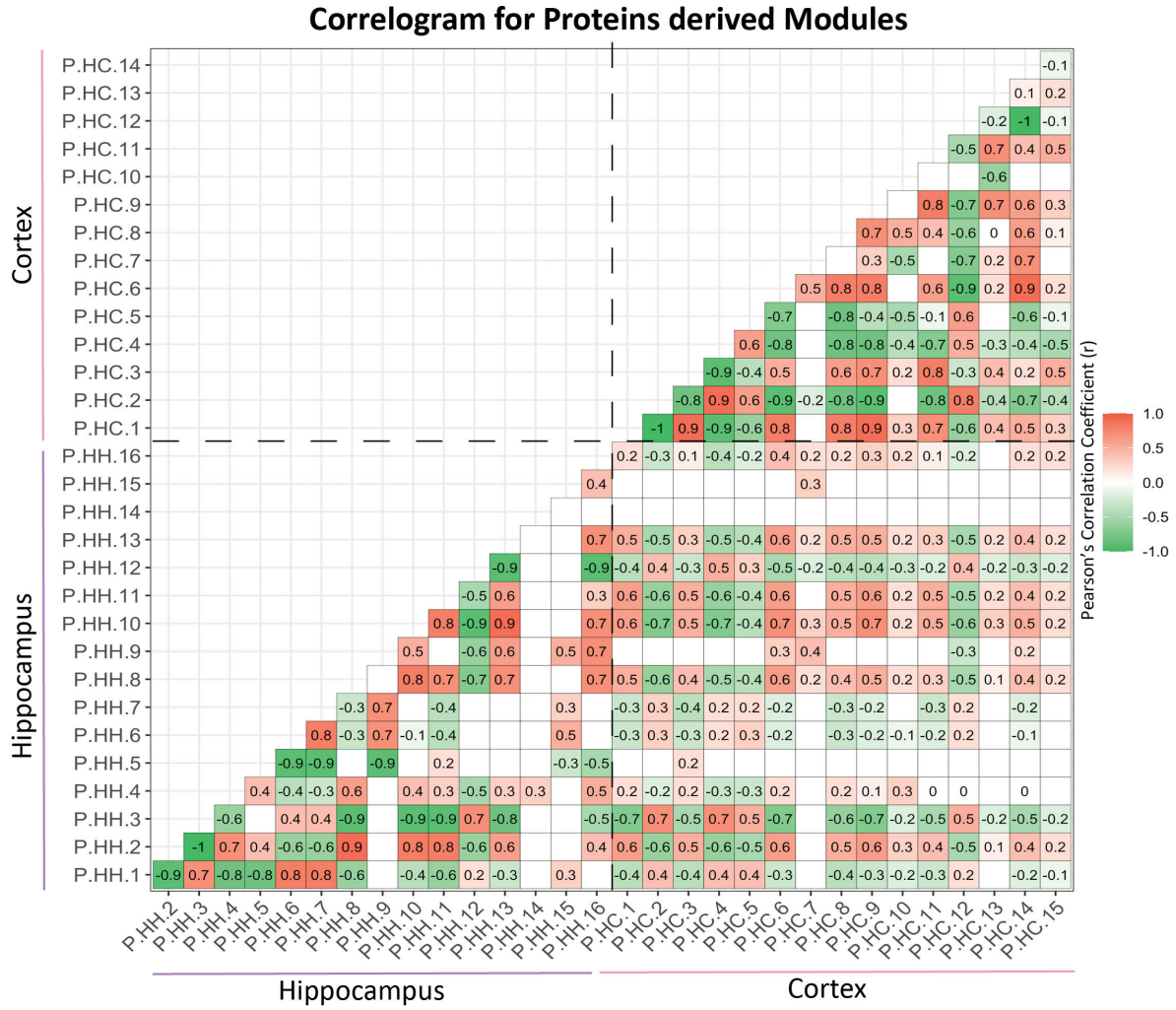


Figure 3.45. Module-based correlation in protein expression from human hippocampus and cortex. Correlation analysis among all the modules derived from the WPCNA analysis for proteins from hippocampus and cortex shows correlation among them. The red color indicates positive correlation, green indicates anti-correlation and white indicates no correlation. Numbers inside the boxes are the correlation values for only statistically significant correlations (p<0.05, Pearson's Correlation test). The correlation is represented by a color bar on the right. P-HH: Protein module in Human hippocampus, P-HC: Protein module in Human cortex.

Similar to the data derived from gene expression experiments, we created protein-protein interaction networks from the top 20 proteins for each significant module to identify the driving proteins (hubs) in the network (Figure 3.45 and 3.46). We did not find many triplicated proteins among the top hubs in many modules. We found only one triplicated protein,

TRAPPC10, in the hippocampal upregulated protein module P-HH-5 with 120 members (Figure 3.46A). In the upregulated modules from the cortex's protein expression data, the P-HC-3 module had three triplicated proteins, COL18A1, COL6A1, and COL6A2, whereas P-HC-1 had two triplicated proteins, APP and S100B (Figure 3.47A). As we observed from Figure 3.45, there were very few well-correlated modules from the two regions. Moreover, we did not observe any common hub genes from the upregulated modules from the hippocampus and cortex.

On the other hand, the downregulated module P-HH-3 (Figure 3.46B) was correlated well with P-HC-2 (Figure 3.47B), as previously observed. These two modules from the hippocampus and cortex respectively shared two proteins among the top hub proteins, which are TNPO2 and AARS1.

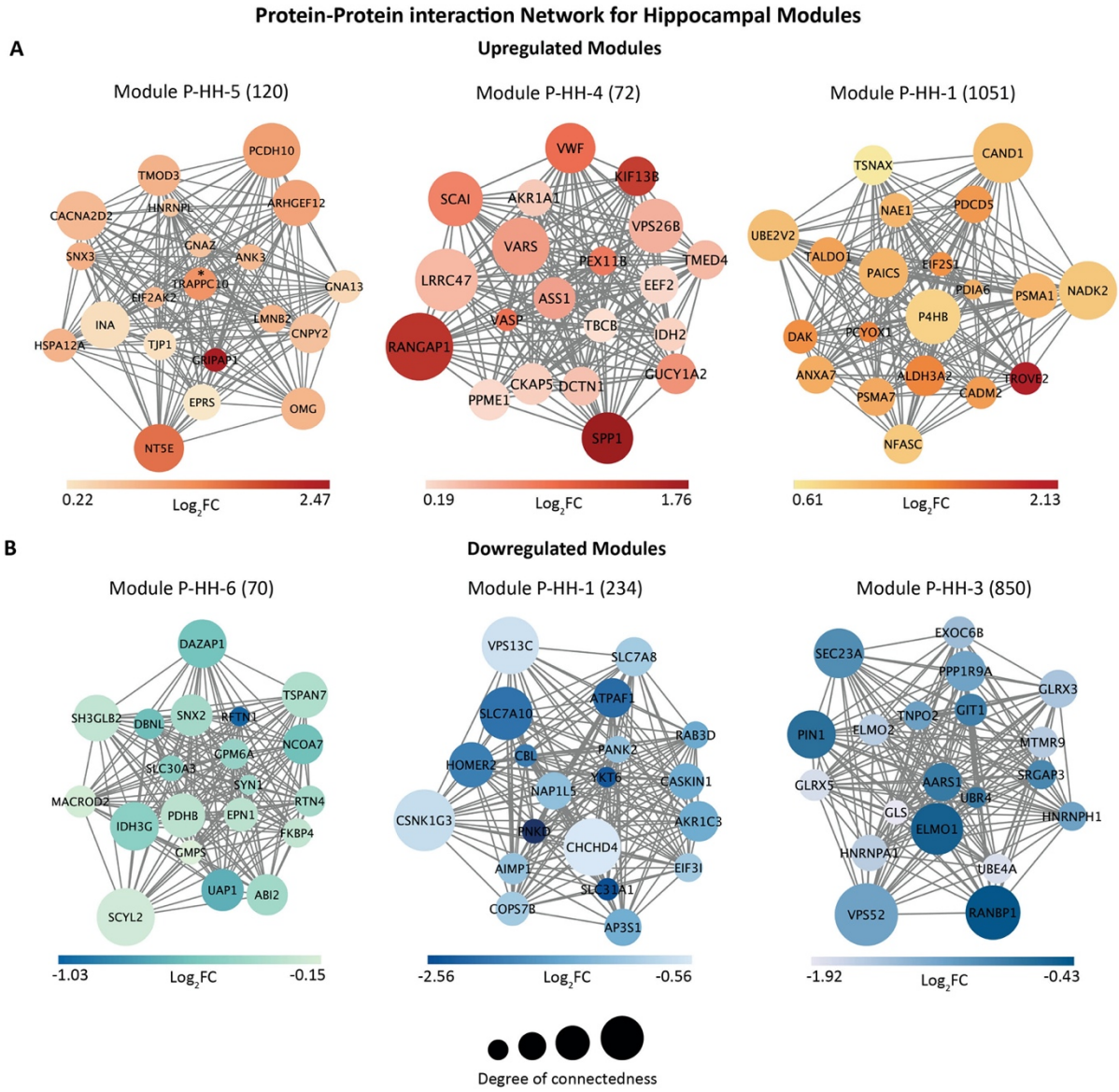


Figure 3.46. Protein-protein interaction for WPCNA-derived significant modules from hippocampus protein expression data shows the top 20 proteins. The size of the modes denotes the degree of connectedness, the color of the node (coded in the color maps below) represents the log_2FC. A) Upregulated modules P-HH-5, P-HH-4 and P-HH-1 B) Downregulated modules P-HH-6, P-HH-1 and P-HH-3. Number of proteins in each module are denoted in the brackets. Asterisks denote triplicated proteins. P-HH: Protein module in human hippocampus, FC: Fold change.

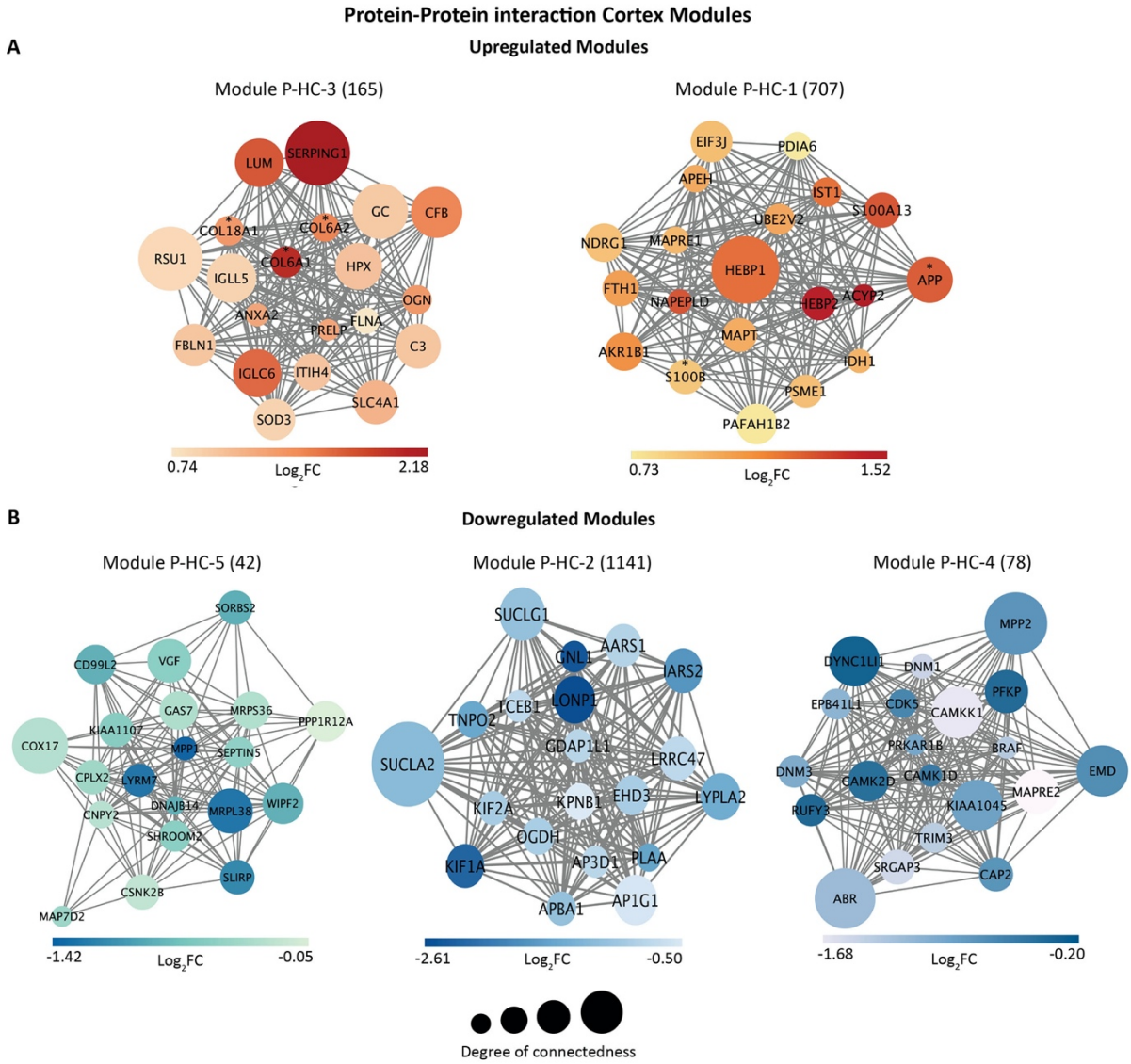


Figure 3.47. Protein-protein interaction for WPCNA-derived significant modules from cortex protein expression data shows the top 20 proteins. The size of the modes denotes the degree of connectedness, the color of the node (coded in the color maps below) represents the log_2FC. A) Upregulated modules P-HC-3 and P-HC-1 B) Downregulated modules P-HC-5, P-HC-2 and P-HC-34. Number of proteins in each module are denoted in the brackets. Asterisks denote triplicated proteins. P-HC: Protein module in human cortex, FC: Fold change.

3.12. Alterations at the levels of RNA and protein expression are largely distinct.

The number of proteins identified in the proteomic analysis was much less than the number of protein-coding genes identified through RNA-seq. Even though fewer proteins are detected in proteomics, there are still around 3800 proteins common between both technologies for which we had expression values. We thus investigated how expression changes at the RNA level correspond with protein level changes in DS *vs.* control brains. We plotted the log_2FC for all genes common between RNA and protein data for both hippocampus and cortex and

found very little correlation between the two (Figure 3.48A, B). On the other hand, there is a high correlation when we compare the two regions either at RNA level (0.9) or protein level (0.5) separately (Figure 3.43C). This result suggests that although overall RNA and protein are poorly correlated, but comparing the two regions at either RNA or protein level individually, there is a high cross-region overlap.

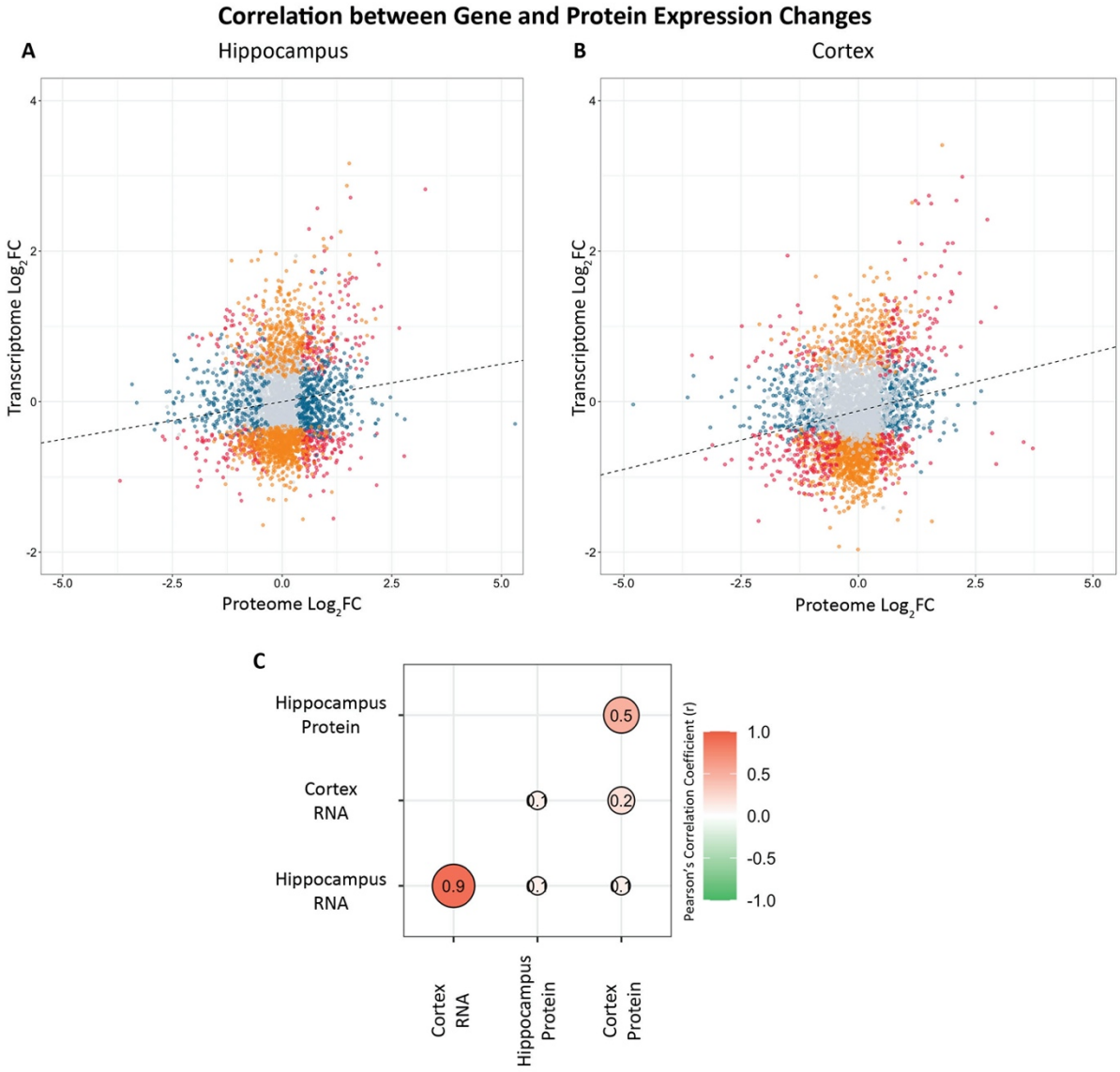


Figure 3.48. Low correlation between RNA and protein for both regions but high cross region correlation at RNA and protein level. Correlation plot shows low correlation between RNA and protein for hippocampus and cortex but high cross-region correlation at both RNA and protein. A) Plot shows correlation between RNA and protein for all common genes between RNA and protein for the hippocampus with r=0.129 (p=0.001, Pearson's correlation test) and B) for the cortex with r=0.198 (p=0.001, Pearson's correlation test). x-axis represents log_2FC from proteomics data while y-axis represents log_2FC from gene expression data. The line represents linear regression with intercept at 0. Red dots represent genes with differential expression at both RNA and protein level, orange dots represent differentially expressed genes only, blue dots represent differentially expressed proteins only and gray dots represent non differentially expressed genes or proteins. C) Correlations at RNA and

protein levels for both hippocampus and cortex. The color bar on right in C represents the correlation from -1 to 1. FC: Fold change.

To further elaborate on these similarities and differences between the two omics (transcriptome and proteome) and the two regions (hippocampus and cortex), we compared their biological processes, cellular locations, and molecular functions (Figure 3.49). We pooled together all the gene ontology analyses previously performed for both the regions at the RNA and protein level to perform this comparison. Interestingly, even with a low level of correlations at expression levels and module levels, we found many processes and functions to be shared between the two regions and two technologies. For example, in terms of biological processes, transport, translation, neurotransmitter secretion, neuron development, and axon development were represented by many datasets (Figure 3.49 left). At the cellular component level, we found overlap for cell-substrate junction and extracellular vesicle (Figure 3.49 right). This was consistent with our observations from modules derived from both gene and protein expression data (Figure 3.17 and 3.43).

Figure 3.49. Comparison of all the enriched gene ontology terms for both RNA and protein and for both hippocampus and cortex reveals a multiple similarities and distinctions. On left, we show the biological processes and on right, the molecular functions and cellular components for the differentially expressed genes and proteins separated by their expression level (up- or downregulated) and region (hippocampus and cortex). The GO terms are separated by parenthesis, where orange represents biological processes, green represents cellular components and blue represents molecular functions. The color bar on the right shows the range of significance defined by -$\log_{10}$FDR (after Benjamini-Hochberg multiple testing correction) for each term and each module. FDR: False Discovery Rate

3.13. microRNA analysis reveals post-transcriptional regulation in DS.

Our data from both hippocampus and cortex at both RNA and protein levels revealed a highly complex and non-trivial scenario, suggesting that many similarities and dissimilarities exist between the two regions and at multiple levels of expression. To target any gene or protein and/or biological process or pathway would require a careful and thorough understanding of the genetic regulation taking place in DS. To further resolve these multiple levels of genetic regulation in DS, we look at another class of gene-expression regulators, the microRNAs.

microRNA (miRNAs) represent one of the fundamental regulatory pathways of RNA and consequent protein expression. miRNAs post-transcriptionally regulate protein-coding genes. miRNAs either degrade the mRNA or inhibit protein translation through complementary binding at 3' untranslated region (3'UTR) of mRNA. As we observed huge dysregulation in gene and protein expression in DS *vs.* control samples in hippocampus and cortex, this encouraged us to look at mRNA-miRNA and miRNA-protein relationships. We first performed small RNA sequencing and obtained around 7 million reads for each sample. Similar to gene and protein expression, we filtered the samples according to PCA, hierarchical clustering, and RIN values and again obtained the same set of hippocampal and cortical samples. The majority (~ 27%) of small RNA reads corresponded to miRNA reads, followed by ~21% transfer RNA (tRNA) reads, 10% small nucleolar RNAs (snoRNA) reads, and the rest were distributed among other small RNAs or remained unmapped.

Here, we focused on miRNA reads. We filtered the miRNAs according to the criteria that in both control and DS brains, at least 5 samples from the hippocampus (total samples were 6 each) and at least 7 samples from the cortex (total samples were 8 each) should express the miRNA to remove the lowly expressed miRNAs. Thus, we obtained ~700 miRNAs for each brain region. We performed differential expression analysis on these miRNAs after normalization of each sample and found 32 differentially expressed miRNAs (DEmiRNAs) in the hippocampal samples (Figure 3.50) and 50 in the cortical samples (Figure 3.51). Among these DEmiRNAs, 6 of them were shared between the two regions, miR-155-5p, miR-99a-3p, miR-504-3p, and miR-181-5p, upregulated and miR-889-3p and miR-412-5p, which were downregulated.

From previous studies, we know that five miRNAs (miR-155, miR-802, miR- 125b-2, let-7c, and miR-99a), which are derived from chromosome 21, are overexpressed in the hippocampus and frontal cortex of DS fetuses (Hill and Lukiw, 2016; Zhao et al., 2015). Here in the hippocampus and cortex from adult DS individuals, we found four of them to be differentially expressed except miR-802 in the hippocampus (Figure 3.50) and two (miR-155 and miR-99a) in the cortex (Figure 3.51).

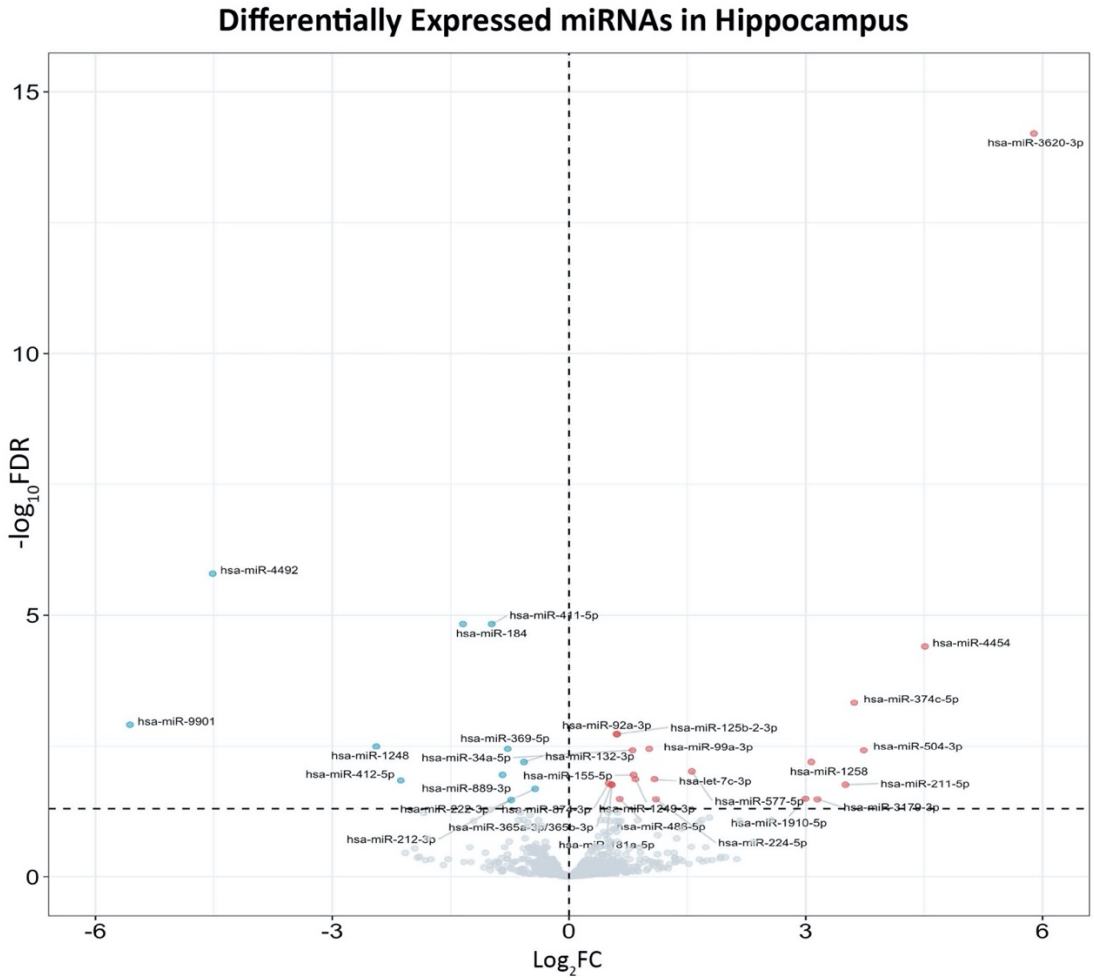


Figure 3.50. Several miRNAs were differentially expressed in the hippocampus. Volcano plot shows the differentially expressed miRNAs identified in the hippocampus upon differential expression analysis. The level of expression changes for each miRNA is shown in Log_2FC and their level of significance in $-log_{10}FDR$. The vertical dotted line represents 0, a point on the x-axis which shows upregulated miRNAs in red on the right side and downregulated miRNAs in blue on the left side. The horizontal line denotes the $-log_{10}FDR=1.3$ which represents the FDR<0.05. The grey dots show non differentially expressed miRNAs. FDR: False discovery rate, hsa-miR: human microRNA.

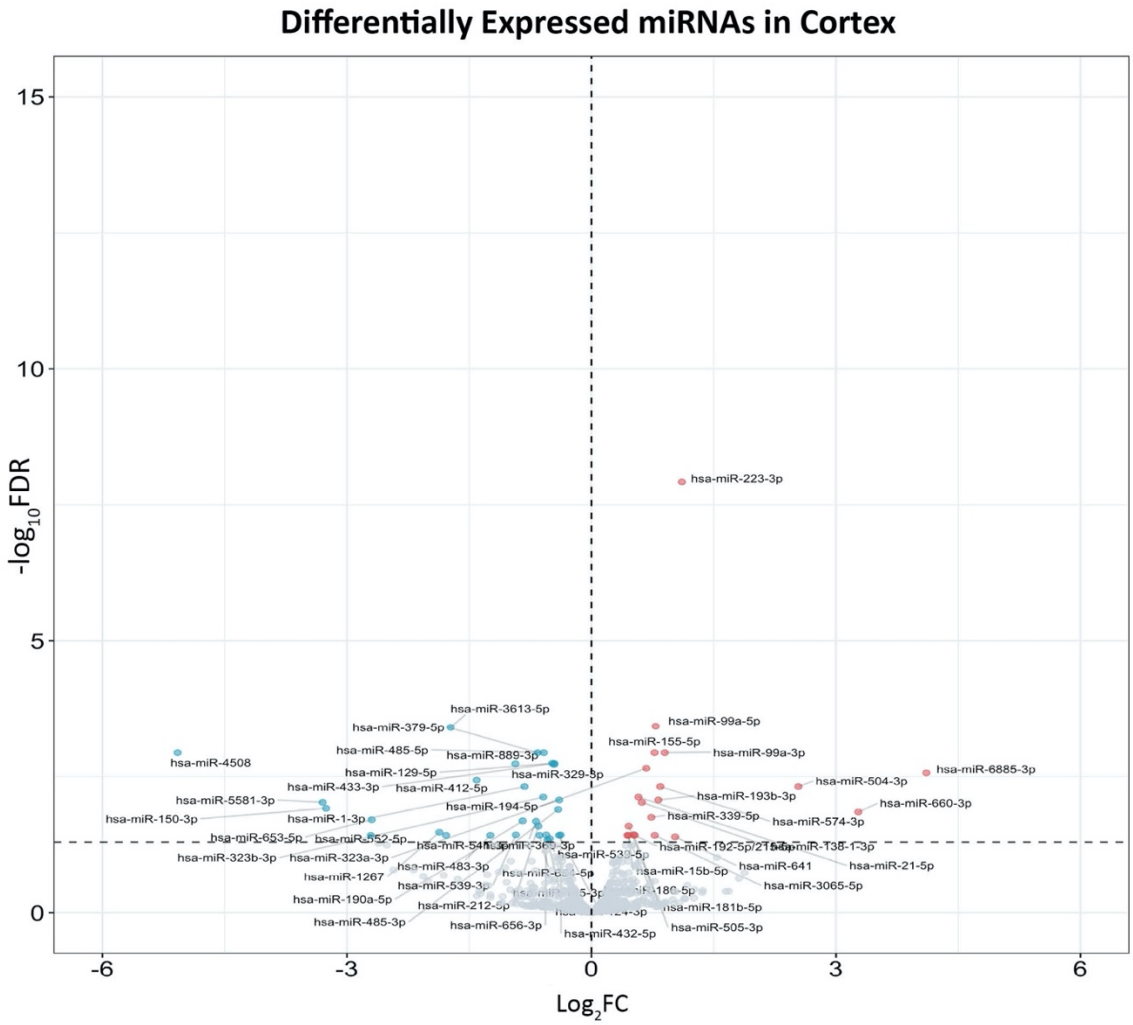


Figure 3.51. Several miRNAs were differentially expressed in the cortex. Volcano plot shows the differentially expressed miRNAs identified in the cortex upon differential expression analysis. The level of expression changes for each miRNA is shown in Log_2FC and their level of significance in $-log_{10}FDR$. The vertical dotted line represents 0, a point on the x-axis which shows upregulated miRNAs in red on the right side and downregulated miRNAs in blue on the left side. The horizontal line denotes the $-log_{10}FDR=1.3$ which represents the FDR<0.05. The grey dots show non differentially expressed miRNAs. FDR: False discovery rate, hsa-miR: human microRNA.

The rest of the other differentially expressed miRNAs in our study were found to be located on other chromosomes. In particular, many downregulated miRNAs were located on chromosome 14, and many upregulated miRNAs were located on chromosome X in both regions.

3.14. Differentially expressed miRNAs have a large repertoire of differentially expressed targets.

As miRNAs are known to regulate multiple genes by binding to their 3'UTRs, we searched for mRNA targets for each DEmiRNAs in the TargetScan v7.2 in the 3'UTRs of about 20,000 genes.

First, we downloaded a list of all conserved miRNA sites in the vertebrates for the conserved miRNA families and filtered the data for our DEmiRNAs. On average, we obtained 4400 mRNA targets for our DEmiRNAs. We then filtered the targets according to their correlation with miRNA expression level by keeping only those inversely-correlated targets to the expression of miRNAs.

We classified the targets into three categories (Figure 3.52): a) differentially expressed at both gene and protein expression levels, which we denoted as "Both" (Figure 3.52A, C); b) differentially expressed at least at gene level, which we denoted as "RNA" (Figure 3.52A, B); and c) differentially expressed at least at protein level, which we denoted as "Protein" (Figure 3.52A, D). Next, we calculated which miRNAs had the most number of differentially expressed targets from each of these three categories. We reasoned that those miRNAs with the most number of targets could act as hub miRNAs.

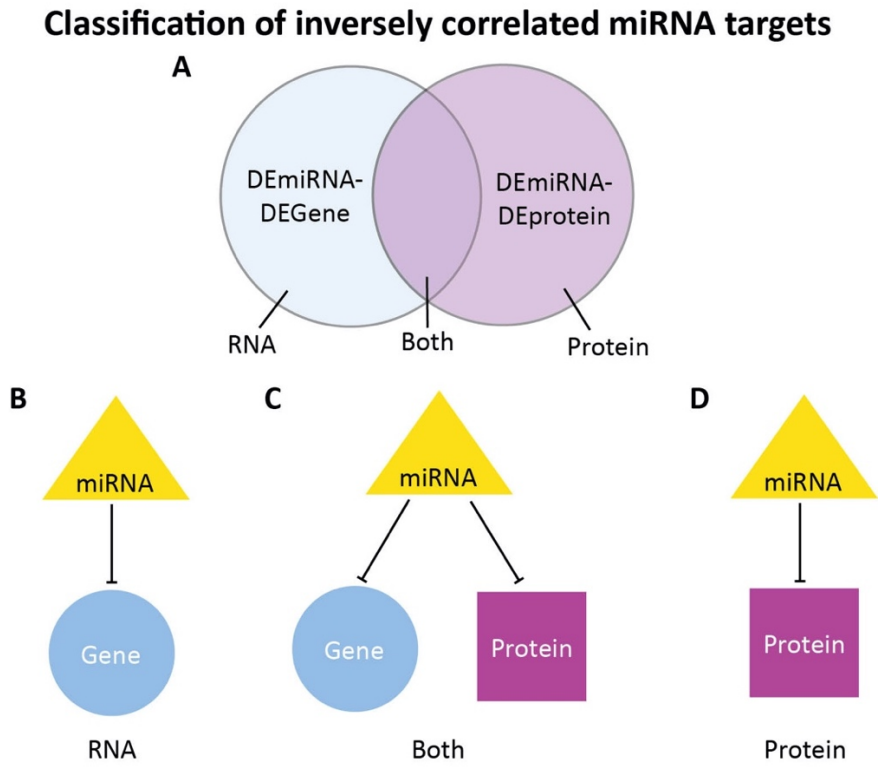


Figure 3.52. Schematic of classification of inversely correlated miRNA targets. A) Venn diagram shows overlap of DEmiRNA-DEgenes with DEmiRNA-DEproteins, where blue circle shows only RNA level inverse correlation, purple circle shows only protein level inverse correlation and the overlap between the two circles shows miRNA inverse correlation with both differentially expressed genes and proteins. B) miRNAs targeting only genes, denoted as "RNA", C) miRNAs targeting both genes and proteins, denoted as "Both" and D) miRNAs targeting proteins only, denoted as "Protein". In B), C) and D) yellow triangles represent differentially expressed miRNAs,

blue circles represent differentially expressed genes and purple circles represent differentially expressed proteins. DE: Differentially expressed.

Among the miRNAs upregulated in the hippocampus, we found miR-181-5p involved with many differentially expressed targets either at RNA level or protein level, followed by miR-92-3p and 374c-5p (Figure 3.53A). On the other hand, in the miRNAs downregulated in the hippocampus, miR-132-3p/212-3p and miR-221-3p/222-3p had the most number of differentially expressed targets (Figure 3.53B). Like upregulated miRNAs in the hippocampus, the targets of upregulated miRNAs in the cortex were mostly targeted by miR-181-5p, followed by the miR-15-5p family miR-186-5p (Figure 3.53C). In the targets obtained from downregulated miRNAs in the cortex, miR-124-3p came out as the hub miRNA (Figure 3.53D). So, overall, we found many hub miRNAs targeting multiple differentially expressed genes or proteins or both.

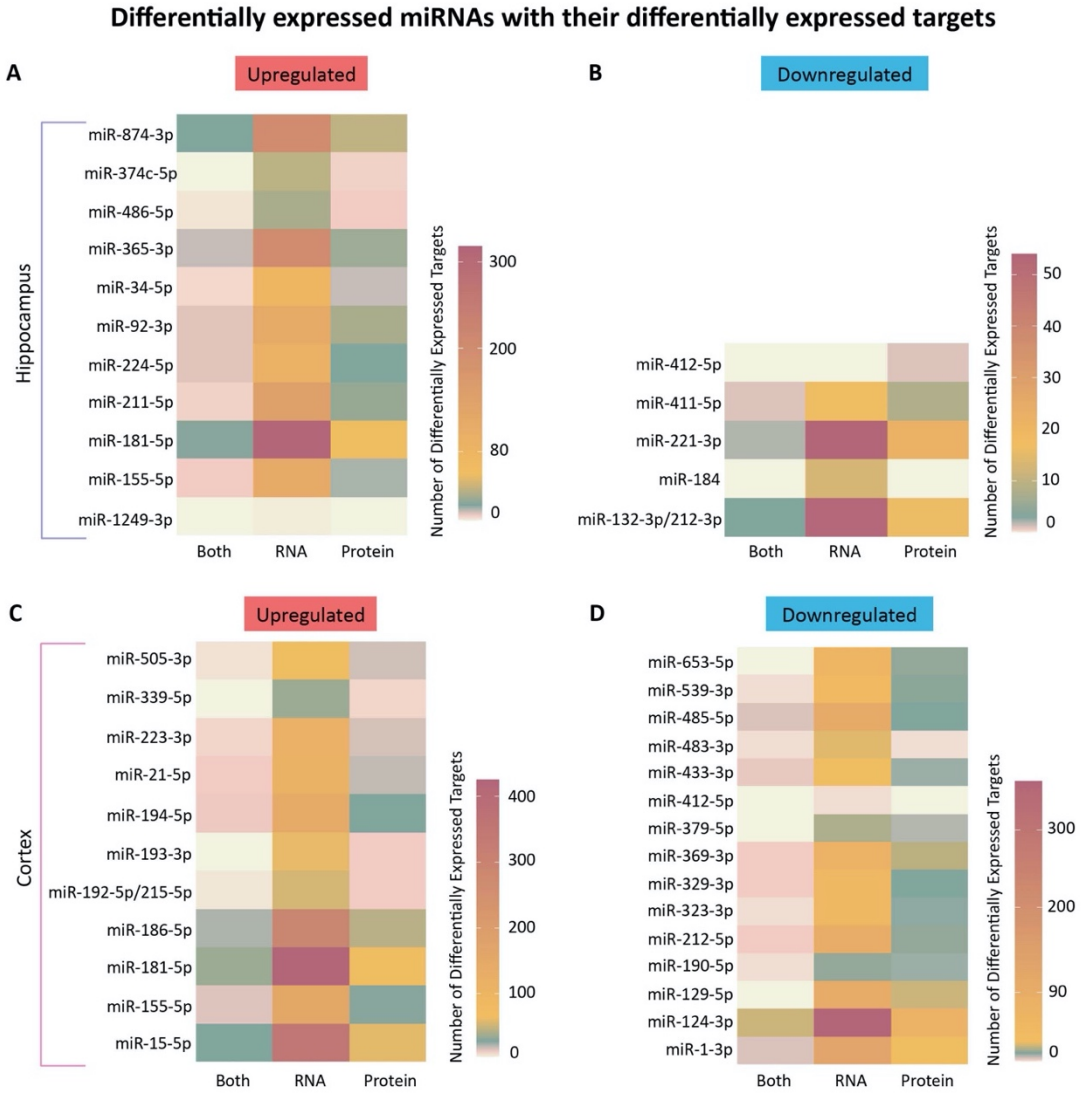


Figure 3.53. miRNA target prediction reveals the top hub miRNAs with the most number of differentially expressed targets. A) miRNAs upregulated in hippocampus B) miRNAs downregulated in hippocampus (C) miRNAs upregulated in cortex and D) miRNAs downregulated in cortex. For all heatmaps, the top miRNAs are those with the most number of differentially expressed targets. The color bar on the right of each heatmap represents the range of the number of differentially expressed targets. Here, miRNAs targeting both genes and proteins, denoted as "Both", miRNAs targeting only genes, denoted as "RNA" and miRNAs targeting proteins only, denoted as "Protein". miR: microRNA.

To investigate the biological significance of the targets of the DEmiRNAs, we performed GO analysis on the targets (Figure 3.54). We found enrichment in biological processes such as axonogenesis, neurogenesis, developmental process, and locomotion. We found high enrichment in nucleoplasm and transcription-factor activity among the cellular components and molecular functions, respectively.

We also found many KEGG pathways enriched in the targets predicted from upregulated miRNAs in both hippocampus and cortex (Figure 3.55). Ras signaling, Rap1 signaling, PI3K-AKT signaling, axon guidance, focal adhesion, and regulation of actin cytoskeleton. On the other hand, there were fewer pathways for downregulated miRNA targets, for example, Wnt signaling for targets of miRNAs downregulated in the hippocampus and axon guidance for targets of miRNAs downregulated in the cortex. Overall, our results suggest that DEmiRNAs could be responsible for some differential expression at gene and protein levels in DS observed in the hippocampus and cortex.

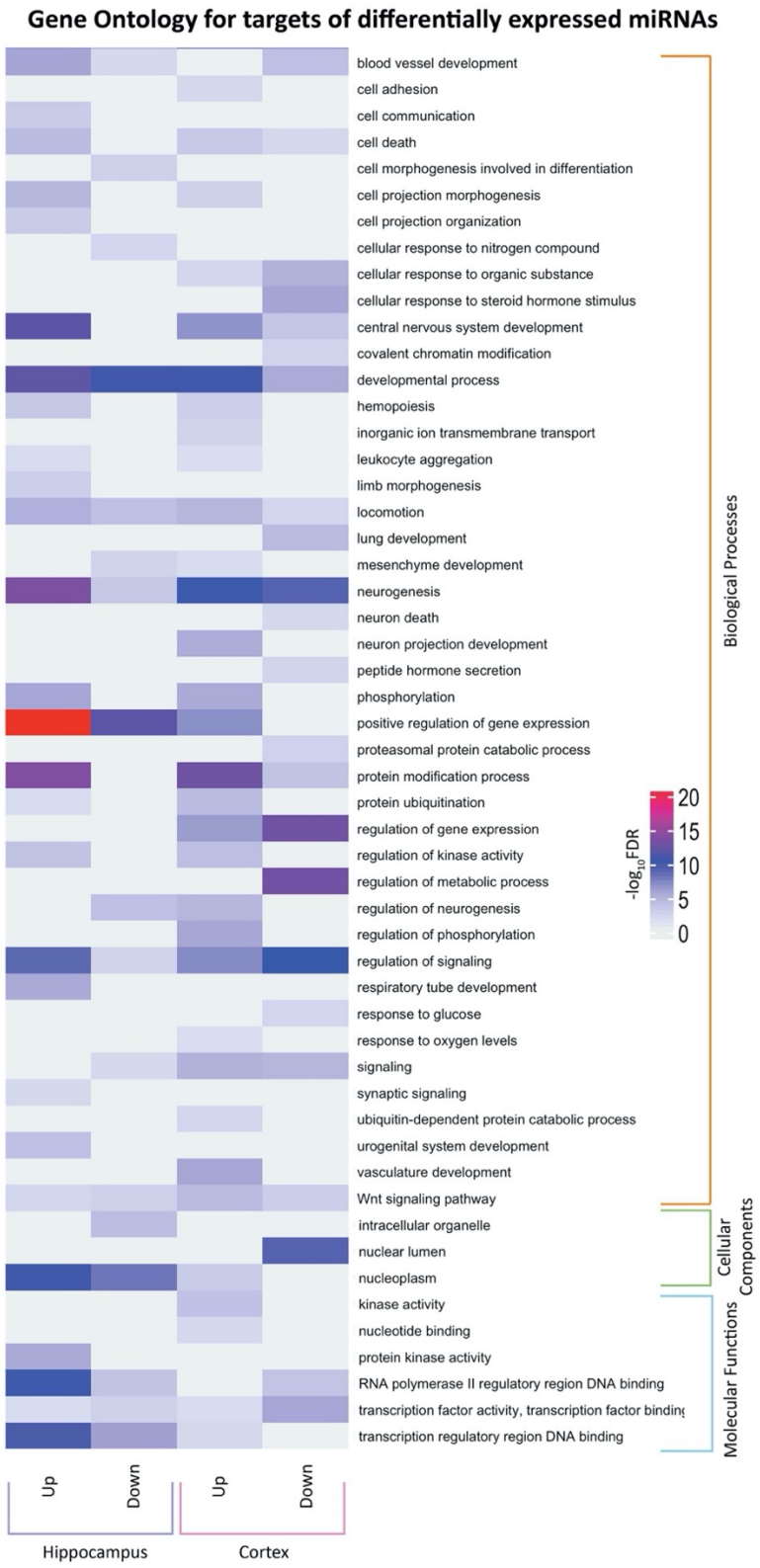


Figure 3.54. Enrichment of targets of differentially expressed miRNAs in multiple Biological Processes (BP), Cellular Components (CC) and Molecular Functions (MF) terms. The GO terms are separated by parenthesis, where orange represents biological processes, green represents cellular components and blue represents molecular functions. The enrichment for targets of up and downregulated miRNAs are separated by hippocampus (purple parenthesis) and cortex (pink parenthesis). The color bar on the right shows the range of significance defined by -$\log_{10}$FDR (after Benjamini-Hochberg multiple testing correction) for each term. FDR: False Discovery Rate

Figure 3.55. Numerous KEGG pathways show enrichment for targets of differentially expressed miRNAs. The enrichment for targets of up and downregulated miRNAs are separated by hippocampus (purple parenthesis) and cortex (pink parenthesis). The color bar on the right shows the range of significance defined by $-\log_{10}$FDR (after Benjamini-Hochberg multiple testing correction) for each term. KEGG: Kyoto Encyclopedia of Genes and Genomes, FDR: False Discovery Rate.

3.15. Multiple axons phenotype in Ts65Dn hippocampal primary culture.

Different analysis throughout our study pointed to dysregulation in genes related to axon formation. To test if there is any defect in axon formation, *in vitro*, we cultured primary hippocampal neurons from Ts65Dn and wildtype (WT) mouse brains. We performed immunocytochemistry analysis by staining with *AnkG* (a marker of axon initial segment), *Pan-Nav* (another protein located in axon initial segment), and *Map2* (dendritic marker). We observed several neurons with three different phenotypes in both cultures: a) neurons with no axons, b) neurons with single axons, and c) neurons with two or more axons (Figure 3.56A). We compared the neurons from these two genotypes and counted the number of neurons with a different number of axons and the length of the axons.

We found that the length of axons with only a single axon in the neuron was similar in both genotypes. In comparison, the neurons with multiple axons were longer in Ts65Dn *vs.* WT (Figure 3.55B). We also observed that Ts65Dn culture had ~30% neurons with two or more neurons compared to 18% neurons from WT. Also, there were 20% neurons from Ts65Dn with no visible axons (Figure 3.55C). These are preliminary results, but they confirm a defective axonal growth and specification in the DS mouse model.

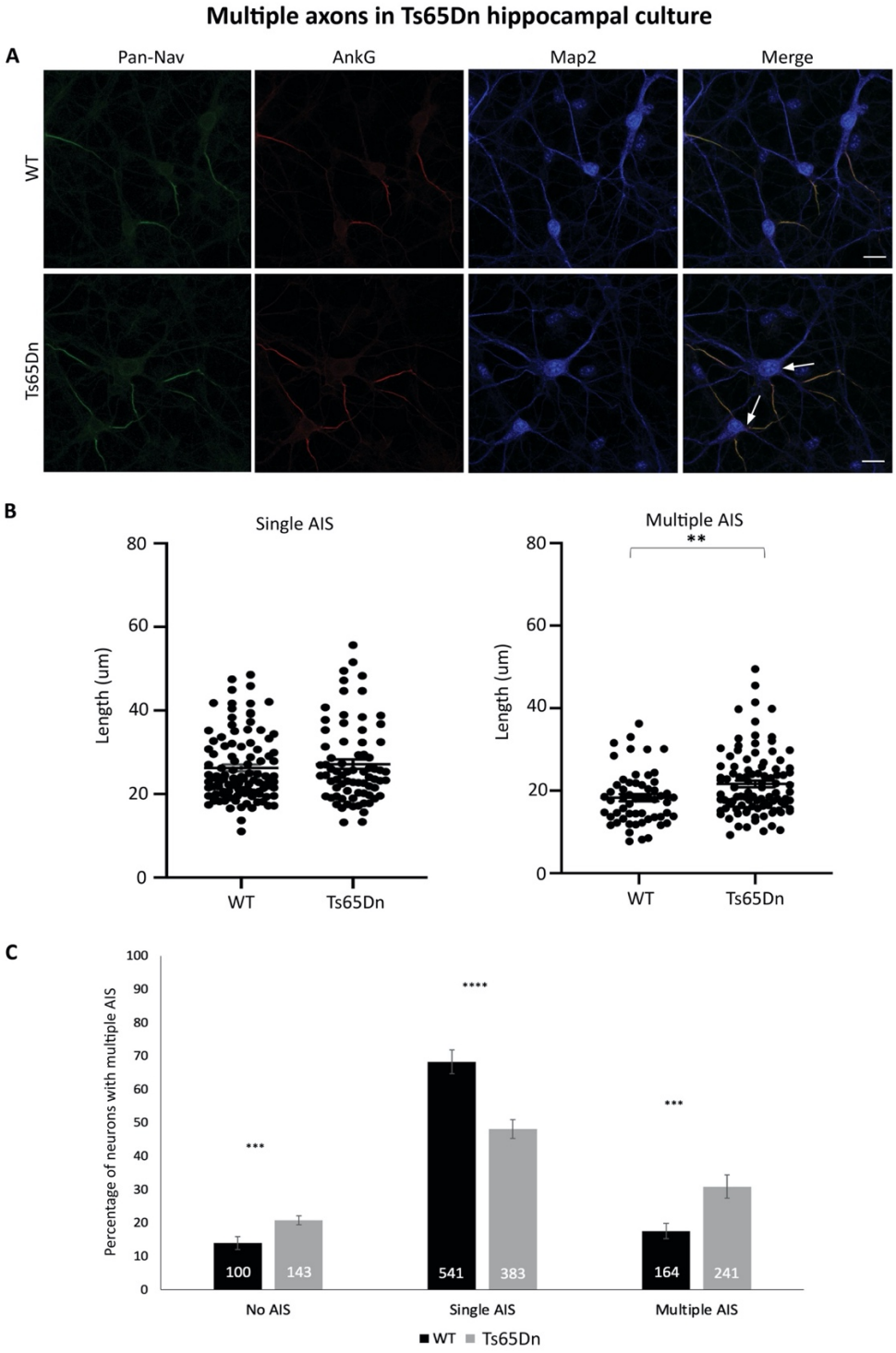


Figure 3.56. Presence of multiple axons in hippocampal primary culture from Ts65Dn mice. A) Representative confocal images of Pan-Nav, AnkG and Map2 stained hippocampal primary culture at days in vitro (DIV) 14 from WT and Ts65Dn mice. Scale bar: 20μm. White arrows in the merged image show the neurons with multiple axons. B) Quantification of the length of the axons from neurons with single or multiple axons as in figure A expressed in mm. Bars represent the average length of each axon ± SEM in all analyzed images and circles represent the single axons from all experiments (12 fields total/genotype, 2 coverslips each, N=2). ** p<0.001, Mann-Whitney U test. C) Quantification of percentage of neurons with multiple axons in all analyzed images, where black bars represent WT and grey bars represent Ts65Dn. The numbers inside the bars show number of neurons having no axons, single axons and multiple axons across all the analyzed experiments (12 fields total/genotype, 2 coverslips each, N=2). ***p<0.0001, ****p<0.00001, multiple t-test followed by Holm-Sidak post-hoc test.

Discussion

A key issue in designing therapeutic interventions for neurodevelopment disorders such as Down syndrome (DS) is the lack of knowledge and understanding of their underlying mechanisms. Rather than focusing on specific candidate triplicated genes or proteins, we followed a systematic transcriptome and proteome-wide approach. We leveraged the benefits of total RNA-seq, small RNA-seq, and proteomics and performed these analyses on the same set of post-mortem human DS brains and their age-/sex-matched healthy controls. This integrative analysis reduced the potential bias observed when performed isolated single analysis on different samples. While several other single transcriptomic and proteomic studies on DS have provided important clues mostly on the developing brain (Gonzales et al., 2018; Guedj et al., 2016; Huo et al., 2018; Stamoulis et al., 2019; Waugh et al., 2019), our study is the first, to our knowledge, to implement an integrative approach on adult human DS brains.

The intense focus on studying and targeting neurodevelopmental diseases (NDDs) early in life is a reasonable approach. The underlying rationale is that early in life intervention may correct the developmental trajectory and alleviate behavioral phenotype later in life. However, findings from DS mouse models have shown that environmental or pharmacological interventions could reverse molecular, electrophysiological, and behavioral phenotypes even in adulthood (Bianchi et al., 2010; Contestabile et al., 2013; Deidda et al., 2015; Fernandez et al., 2007; Parrini et al., 2017; Savardi et al., 2020). Thus, therapeutic approaches in adults complementary to early interventions should also be investigated.

We also analyzed mRNA transcript (isoform) level expression and identified alternative splicing, RNA binding proteins, and miRNAs as crucial intermediate regulatory steps in DS. Our data provide molecular information signatures that consistently repeat at gene, transcript, and protein levels and may function as critical players in DS pathophysiology.

As DS is such a complex genetic condition, we should tackle this complexity at multiple levels, such as at different ages, brain regions, and at different layers of genetic regulation. Finally,

results from other neurodevelopmental disorders such as autism spectrum disorder, which has been studied more extensively with a large cohort of samples, have provided details about individual genetic variations and suggest following further studies in DS with a larger number of individuals. Thus this may inspire a new layer for therapeutic approaches to aid people with DS in their daily life challenges.

4.1. Gene expression dysregulation in NDDs:

We demonstrated a large number of differentially expressed genes in both hippocampus and cortex of people with DS compared to controls. We mapped these differentially expressed genes onto their chromosomal locations. Along with triplicated genes, we found many non-triplicated dysregulated genes suggesting a genome-wide disturbance in DS. All the chromosomes were found to harbor genome-wide dysregulation domains (GEDDs), as previously observed in fibroblasts and iPSC from DS fetuses (Letourneau et al., 2014). According to their expression, the differentially expressed genes were associated with specific biological processes and could be further partitioned into disease-specific pathways and specific cell types. We found an association of the upregulated genes with immunological pathways and glial cells (astrocytes, microglia, and oligodendrocytes) in the hippocampus and cortex.

Conversely, we observed a predominant association of the downregulated genes with neurological pathways and neural cells (interneurons and pyramidal cells). Interestingly, autism spectrum disorder (ASD), schizophrenia (SCZ), and bipolar disorder (BP) cortices show the same preference of up-and down-regulated genes for immune and neural pathways (Gandal et al., 2018). In particular, upregulation of astrocytes was observed in SCZ and BP and astrocytes and microglia in ASD. Moreover, the interferon response pathway was significantly more upregulated in ASD. Thus, these observations suggest the neural-immune trajectory is dysregulated in different neurodevelopmental and neuropsychiatric disorders and may provide new targets for molecular interrogation.

We also observed region-specific biological processes, cellular components, and molecular functions for the differentially expressed genes among the shared enrichments. These

differences between the two regions were also observed in brains from healthy controls and in individuals with ASD, SCZ, and epilepsy (Collado-Torres et al., 2019; Ramasamy et al., 2014). These observations from our results and literature suggest differences in development and functions of both hippocampal and cortical neural circuits (Sousa et al., 2017) and could coincide with the heterogeneity in the different cell types present in these regions (Bayraktar et al., 2014; Böttcher et al., 2018; Buosi et al., 2018; Matias et al., 2019).

The cell type enrichment analysis on gene expression data from both hippocampus and cortex was consistent with our previous study with fewer human hippocampus samples from DS individuals (Pinto et al., 2020). In our previous study, we observed microglia activation and an increase in the number of microglia upon performing immunohistochemical analysis on tissue obtained from DS hippocampus *vs.* control. When we performed cell-type enrichment on the gene expression results from the previous study, we found enrichment of upregulated genes in the microglia. This led us to reason that cell-type enrichment on other brain regions might also inform us about the molecular pathology and the proportion of different cells. This rationale is consistent with the observation of dynamic alteration of different cell types in DS studies (Griffin et al., 1989; Pinto et al., 2020; Wierzba-Bobrowicz et al., 1999). For example, in an earlier study on DS fetuses, a reduction in the number of neurons in the hippocampus was observed (Guidi et al., 2008). Additionally, an increase in astrocyte number was observed in the DS hippocampus and frontal lobe in early development (Zdaniuk et al., 2011). In the case of Ts65Dn, an increase in GFAP-positive astrocytic cells is observed in the hippocampus during early development, but in adult animals, a decrease in GFAP transcript expression has been reported (Contestabile et al., 2007; Toso et al., 2008).

4.2. Gene expression driven co-expression networks in DS:

The brain is a complex organ, consisting of multitudes of cell types, trillions of connections between them, and performing many functions essential for cognition. So, defining the molecular network involved in this complex structure is highly important to better understand the pathophysiology of the diseases and design new and more efficacious therapeutic approaches. Consequently, we divided gene expression profiles from both regions into co-expression modules. We annotated significant gene-expression driven modules with

biological processes, cellular components, and molecular functions along with individual cell types. The shift from looking only at differentially expressed genes to complete gene expression profile follows our observation that there is a genome-wide dysregulation in DS, influencing the expression of non-triplicated and triplicated genes alike. Multiple studies related to brain diseases have applied this strategy of organizing gene expression into clusters (modules) (Gandal et al., 2018; Parikshak et al., 2015, 2016). In particular, gene clusters have been used to identify genes whose expression differs between different conditions across developmental trajectories, brain cells, and regions (Hawrylycz et al., 2012; Kang et al., 2011; Oldham et al., 2008). The co-expression clusters approach has enabled a much better understanding of diverse diseases than an approach based on looking at only specific genes. For example, Gandal et al., with the help of a co-expression study in ASD at both gene and transcript expression levels, have refined the molecular signatures of glial-immune dysregulation so much so that they were able to identify microglial, astrocytic, and interferon response specific modules along with synaptic vesicle cycle and neural activity-dependent modules.

We used the robustly defined co-expression clusters to identify hub genes that could be responsible for driving the disease in the hippocampus and cortex individually. Different biological processes and cell types defined the clusters for each region. Many modules from the hippocampus and cortex shared biological or cell enrichment, and we observed a high correlation between some of these modules from the two regions. In this regard, we observed that modules derived from the two regions retained the same cell-specific enrichment, even though the top hub (gene) driving these modules were different in both regions. Interestingly, the hub genes in the upregulated modules are not only triplicated but also the non-triplicated ones too. This finding raises an important question that whether targeting non-triplicated genes, which is the downstream effect of triplication, could help rescuing cognitive impairment observed in DS.

Furthermore, we found very few hub genes shared between the two regions. For instance, non-triplicated gene, GAP43, expressed at neuronal growth cones, is one of the hub members of downregulated hippocampal module G-HH-3 and cortical module G-HC-2. Usually, GAP43 is expressed highly in the hippocampus and cortex, two regions implicated in learning and

memory (Holahan and Routtenberg, 2008; Rekart et al., 2005). It is also involved in neurodevelopment as its knockout leads to embryonic lethality, with only 5-10% survival to adulthood (Metz and Schwab, 2004). It plays a vital role in axon guidance (Donovan et al., 2002; Shen et al., 2002) and it is associated with brain diseases (Bogdanovic et al., 2000; Tian et al., 2007). GAP43 also enhances long-term potentiation (Hulo et al., 2002), impaired in DS (Contestabile et al., 2017).

On the other hand, we found a couple of region-specific modules, hippocampal modules G-HH-6 and G-HH-4, which appeared to reflect dysfunction in microtubule organization and protein localization. Indeed, ADGRG1, a hub (non-triplicated) gene in the upregulated module, G-HH-6, is a G-protein coupled receptor and interacts with RhoA to stabilize the oligodendrocyte development (Ackerman et al., 2018). Interestingly, this module enriched in oligodendrocytes. Another example is the TAX1BP1 gene, one of the hub genes of downregulated hippocampal module G-HH-4. This gene is involved in proteotoxicity due to the accumulation of ubiquitin conjugates in the mouse brain upon its loss (Sarraf et al., 2020). Interestingly, disruption in proteostasis and protein quality control in lymphocytes obtained from DS individuals has been observed (Aivazidis et al., 2017). The above finding represents an axis of neurodevelopmental architecture that pertains to protein translation and degradation. Notably this is under the control of integrated stress response in DS (Zhu et al., 2019).

Moreover, we observed that a large proportion of significant modules are enriched for different cell types. This may depend on changes in expression for genes associated with these cells or just simply reflect shifts in the proportions of the numbers of specific cells in DS as discussed previously.

4.3. Transcript expression alterations in DS:

Gene expression is an accumulation of transcript-level expression. Thus, classical RNA-seq gene expression studies lose many changes at the transcript level, which might be highly important. We evaluated differential transcript expression by mapping the RNA-seq reads onto the transcriptome. This analysis provides an opportunity to find molecular interactions

that were not visible at the gene level (Pelechano et al., 2013). One of the most important observations from this comparison at gene and transcript expression levels was that the transcripts had a more extensive range of fold changes than gene expression in DS *vs.* control individuals. In the cell type enrichments for the differentially expressed transcripts, we found high enrichment for oligodendrocytes for all the differentially expressed transcripts for both hippocampus and cortex. The oligodendrocyte enrichment likely relates to the observation made by Olmos-Serrano et al. regarding impairment in oligodendrocyte development and myelination in DS (Olmos-Serrano et al., 2016).

Moreover, the differentially expressed transcripts were less overlapping between the two regions than differentially expressed genes. This regional heterogeneity between hippocampus and cortex has been observed in SCZ (Collado-Torres et al., 2019). Transcript expression also showed many biological processes, cellular components, and molecular functions distinct from the gene expression level. For example, the hippocampus had upregulated transcripts that showed enrichment in axonogenesis, adherens junction assembly, myelination, phosphorylation, trans-Golgi network, and cadherin binding not present at the gene level. Again this phenomenon of enrichment for specific biological processes in different brain regions has been observed previously (Collado-Torres et al., 2019).

Among the downregulated transcripts, mitotic cell cycle, sister chromatid segregation, RNA splicing, microtubule, and peroxisome were few of the transcript-specific enrichments. Few of these signatures for downregulated transcripts were observed partially in ASD, SCZ, and BP (Gandal et al., 2018). Moreover, the neural-immune signature originating from differential transcript expression was more prominent for the cortex, as the hippocampus showed enrichment in all the cell types. This neural-immune heterogeneity between two different regions of the brain has, until now, not been observed.

This observation again highlights the importance of transcript expression, which may unveil the regional heterogeneity much better than gene expression profiles. The difference between gene and transcript expression has been observed in the cortices from ASD, SCZ, and BP. However, there is no evidence of comparison between different regions from

neurodevelopment disorders for differences in gene and transcript expression. Therefore, further studies such as the one discussed here are needed to make a systematic comparison between different brain regions for neurodevelopmental disorders at both gene and transcript level, to understand the regional heterogeneity and find molecular targets that could have local (region-specific) effect or a global effect (brain-wide).

4.4. Alternative splicing and RNA binding proteins:

Alternative splicing is a main regulatory pathway which generates transcript isoforms. Robust evidence from literature suggests the importance of alternative splicing and RNA binding proteins in neurological disorders (Da Cruz and Cleveland, 2011; Fromer et al., 2016; Irimia et al., 2014; Jaffe et al., 2018; Parikshak et al., 2016). Given the functional importance of alternative splicing, recent studies have taken tremendous efforts both experimentally and computationally to annotate and predict different splicing events in the analyzed tissue (Braunschweig et al., 2014; Irimia et al., 2014; Tapial et al., 2017; Yeo et al., 2004). We observed several differentially expressed transcripts in DS *vs.* controls for both the regions. In particular, by analyzing alternative splicing events, we provided a potential of uncovering splicing defects and the master regulators (e.g., RNA-binding proteins and components of the spliceosome machinery) that may play essential roles in DS. For example, we found evidence of dysregulation in axonogenesis and dendritic-spine development-related genes in both hippocampus and cortex. Broadly, for genes undergoing alternative splicing in the form of exon skipping, we found enrichment in cell projection morphogenesis, axonogenesis, dendrite morphogenesis, synaptic signaling, postsynaptic density, microtubule, and contractile fiber. This result underscores previous findings linking alternative splicing to cytoskeleton-related genes (Zhang et al., 2016). For example, the authors found evidence of splicing dysregulation in genes responsible for microtubule dynamics (Mast2, Mark4), focal adhesion (Macf1) and postsynaptic density (Dlg4, Homer). Moreover, converging evidence from other neuropsychiatric and neurodevelopmental disorders suggests impairment in synapses maintenance due to dendritic spine morphogenesis dysregulation (Engle, 2010; Lo and Lai, 2020; Nakai et al., 2018) including DS people and mouse models (Benavides-Piccione et al., 2004; Cramer and Galdzicki, 2012; Haas et al., 2013; Pinto et al., 2020).

Another kind of alternative splicing event we found in our study is intron retention. Intron retention (IR) leads to multiple distinct fates of the transcript. It could lead to degradation via nonsense-mediated decay or via interaction with miRNA-RISC complex as 3' untranslated region (UTR) of mature transcripts with intron retention will provide more miRNA binding sites. IR transcripts can also lead to different protein isoforms or be detained in the nucleus preventing translation (Monteuuis et al., 2019). We observed a large percentage of genes with intron retention in DS. We believe that some of these genes with high IR might explain the great variety of downregulated genes, transcripts, or proteins in DS. Thus this hypothesis warrants further analysis and experimental proof.

RNA binding proteins (RBPs) are essential for multiple neurodevelopmental functions. Supporting this notion, we found PTBP2, an RBP, to be both differentially expressed and have exon 10 skipping in the DS hippocampus and cortex. PTBP2 regulates axonogenesis-related genes and it is a crucial regulator of alternative splicing in neural progenitors and immature neurons (Zhang et al., 2019). Consequently, we validated the accuracy of our results from alternative splicing analysis data by semi-quantitative PCR for PTBP2 and other axon formation-related genes (ANK3, MARK2, EML1 and CD46) that we found differentially expressed and differentially spliced in our gene expression and alternative splicing analysis. We are currently investigating the importance of PTBP2 and the other axon formation-related genes in axon formation in iPSCs derived neurons from DS and their isogenic control lines.

Another exciting question worth exploring is the presence of differentially expressed RNA binding proteins, which play an essential role in alternative splicing. Also, the possibility that RNA binding proteins may be themselves differentially spliced is worth the investigation. Indeed in the neuronal progenitors, PTBP1, another RBP, induces the exclusion of exon 10 of PTBP2, leading to exon 10 skipping and a transcript with premature termination codon (Li et al., 2014b). By integrating a manually curated list of RBPs (Gerstberger et al., 2014) with the differentially expressed genes and transcripts, we found a large number of dysregulated RBPs in both hippocampus and cortex. Among the RBPs that we found differentially expressed at gene and transcript levels and triplicated in DS is ADARB1. This RBP is responsible for base changes from adenosine (A) to inosine (I) RNA editing, a co- or post-transcriptional mechanism of gene expression regulation (Eisenberg and Levanon, 2018). A-to-I editing

occurs within glutamate receptor 2 (GLUR2), which is an AMPA receptor mediating fast excitatory synaptic transmission (Sommer et al., 1991). Interestingly, Gonzales et al. have recently quantified the A-to-I editing in trisomic and euploid cortical neurons derived from induced pluripotent stem cell (iPSC) lines. However, they could not observe any differences in the amount of RNA editing between the two experimental groups in the cortical neurons. They concluded that even though the ADARB1 is triplicated in the iPSC derived trisomic neurons, it does not lead to increase in global levels of RNA editing (Gonzales et al., 2018). In addition, analysis of specific sites (the site of action of ADARB1) in glutamate receptor subunits (GLUR2, GLUR5 and GLUR6) in brains (cortex, cerebellum and cerebral white matter) of both DS and healthy control individuals from fetuses, neonates and adults, there was no increase in expression of ADARB1 and no alteration in editing (Kawahara et al., 2004). These observations could suggest that there is no role of RNA editing in DS, but further experimental validations are needed to thoroughly conclude it in our study.

Furthermore, we extended our analysis by identifying a few RBPs that were differentially expressed at both gene and transcript levels and had an exon skipping event associated with them (i.e. one of their exon was differentially skipped in either DS or control individuals). Among those proteins, we found EIF4G3 in both hippocampus and cortex. EIF4G3 is a critical scaffold protein implicated in protein complexes implied in translation initiation. In DS brain samples, we found a microexon excluded in EIF4G3. Interestingly, EIF4G3 microexon controls the expression of synaptic receptors linked to neuronal activity and cognitive functioning regulation. Also, depletion of EIF4G3 microexon can lead to social behavior, learning, and memory deficits (Gonatopoulos-Pournatzis et al., 2020). Understanding more thoroughly the regulation of alternative splicing in human brain tissue will be essential to add another layer of comprehension because the RBPs, components of spliceosome machinery, and various transcription factors regulate gene expression and can drive genotype to phenotype changes (Hsiao et al., 2016; Yang et al., 2019).

4.5. Translating genotype to phenotype in DS:

Proteins constitute the building blocks and functional machinery of the cell. The polygenic nature of neurodevelopmental disorders such as DS would benefit from an holistic approach

as the brain's core functions require an extensive network of protein-protein interactions. Measuring genes and RNA transcript isoforms is not sufficient and adequate to capture multi-protein functional assemblies. Quantification of changes of protein levels in a neurodevelopmental disorder provides a complementary and direct approach to understanding the phenotype and underlying molecular pathology. Due to technological differences intrinsic to RNA-seq and proteomics, we identified fewer proteins *vs.* genes in our samples. First, any changes in gene expression may be either a cause or consequence of the disorder. Second, biological differences observed between gene and protein may be a limitation of proteomics data. Apart from a low number of proteins detected in MS, the less soluble and large membranous proteins are challenging to detect. These issues highlight why it is crucial to perform integrative multi-level analysis on the same set of samples to obtain complementary information and that critical aspects in both gene and protein expression analysis exist.

Anyhow, even though we found fewer differentially expressed proteins in the hippocampus and cortex, they provided essential insights. Indeed, overall, there were more than 1000 differentially expressed proteins and they were enriched in different biological processes, cellular components and molecular functions. Although many gene ontology terms were shared between the gene and protein expression some of them were distinct from gene expression data. For example, biological processes such as cell death, cellular response to stimulus, interferon gamma mediated signaling pathway, locomotion and mRNA splicing were specific to gene expression data. And antigen processing and presentation, cell junction assembly, cell-cell adhesion, coagulation and ion homeostasis were only present in the protein expression data. Interestingly, few of these enrichments from protein expression data were region-specific. For example, among upregulated proteins, cellular detoxification, fibrinolysis, wound healing, actin filament bundle and vesicle lumen in the cortex and axon development, coagulation, lipid oxidation, myelination, oxidative phosphorylation, cleavage furrow and respiratory chain complex in the hippocampus. Similarly, there were many terms from downregulated proteins specific to each region.

Moreover, we observed the cell-type enrichment for the upregulated proteins in the hippocampus in all cell types, similar to our observation from transcript levels, but the

downregulated proteins in hippocampus were specific to neural cells. Conversely, the cortex showed specific enrichment i.e. the upregulated proteins were enriched in glial cells (astrocyte, microglia and oligodendrocytes) and the downregulated proteins were enriched in neural cells (interneurons and pyramidal cells). Notably, the oligodendrocytes showed the enrichment for all the differentially expressed proteins in both regions. This observation could again be related to the oligodendrocyte dysfunction observed in DS brains throughout development (Olmos-Serrano et al., 2016). Similar observation at both gene and protein expression levels for the oligodendrocytes has also been described in schizophrenia (Martins-de-Souza, 2010). There are multiple studies who took transcriptomics or proteomics approaches on different brain regions of adult individuals with schizophrenia and concluded with the same observation of oligodendrocyte dysfunction (Arion et al., 2007; Hakak et al., 2001; Katsel et al., 2005; Martins-de-Souza et al., 2010; Pennington et al., 2008; Prabakaran et al., 2004; Tkachev et al., 2003).

Further, comparing the expression level changes (log_2FC for DS *vs.* control) of RNA and protein within one region resulted in a very low correlation for both the regions. This observation is in agreement with the multiple studies performed on multiple human tissues including different brain regions and cell types, suggesting a low correlation between RNA and protein (Liu and Aebersold, 2016; Mangleburg et al., 2020; Poulopoulos et al., 2019; Seyfried et al., 2017; Sharma et al., 2015; Wang et al., 2019). In contrast, the expression level changes between DS *vs.* controls (log_2FC) in only gene or only protein comparison between the two regions showed a high correlation. This result suggests a common molecular (gene or protein) signature exists for both regions in DS *vs.* controls that undergo different post-transcriptional and post-translational regulation within a region. This observation requires thorough validation from qPCR and western blot experiments in our study, and we are currently in the process of performing these experiments.

Moreover, by weighted protein co-expression network analysis (WPCNA), we detected robust protein modules associated with DS, which were not conserved with gene expression derived modules. For example, in the cortex, proteomics modules did not include cell death, cell communication, leukocyte cell-cell adhesion, and many immune-related processes highly enriched in gene expression derived modules. Moreover, protein modules had hub proteins

different from hubs of the gene-expression modules. The hubs represented the biological enrichment for the whole module. For example, upregulated hippocampal module P-HH-2 consisted of 6 hub (CAND1, NAE1, PSMA1, PSMA7, P4HB, PDIA6) proteins among the top 20 which were involved in post-translational protein modification, which was not represented by the hubs from modules derived from gene expression data. This difference in hubs and the biological enrichments between gene and protein expression-derived modules suggests that we could obtain complementary information from both RNA-seq and proteomics approaches. One such study exist for DS, where the authors have profiled the neural progenitors and the differentiated neurons from iPSCs derived from trisomic and euploid control individuals at both transcriptome and proteome level. Although, the authors observed a high correlation between the RNA-seq and proteomics results with converging molecular pathways, unlike our own results. Nevertheless, they did not carry out a level of robust network analysis to identify the hub genes as we did. Still, they found 23 hub genes among which, APP (triplicated in DS) (Sobol et al., 2019), one of the hub proteins from our own analysis in the cortical module from protein expression data, P-HC-1. These results suggests that although there are few differences and similarities between the RNA and protein, taking an approach with multiple levels of analysis would lead to wider identification of hub genes or proteins. Moreover, these results indicated that the study of transduction regulatory events and mediators may hold the potential for a better understanding of the mechanisms underlying DS.

4.6. miRNA regulatory relationships:

Multiple factors could contribute to the differences between gene and protein expression changes of DS *vs.* control tissues. One could be differences in expression of non-coding RNAs, especially miRNAs known to degrade the transcript or inhibit translation upon binding at 3'UTR of mRNAs (Bartel, 2004). By differential expression analysis, we found several differentially expressed miRNAs in DS *vs.* controls samples including few encoded on HSA21 (let-7c-3p, miR-125-2-3p, miR-155-5p, miR-99a-3p). When we extended our analysis to look for miRNA binding sites in the differentially expressed genes and proteins, we obtained a few (32 in hippocampus and 50 in the cortex) differentially expressed miRNAs with many (1459-3000 at gene level and 286-543 at protein level) number of differentially expressed targets.

Some of these DEmiRNAs with many differentially expressed putative targets play extensive roles and regulate the nervous system development, plasticity and function. For example miR-34-5p regulates synaptogenesis in the neuromuscular junction of Drosophila larva (McNeill et al., 2020). miR-132 acts as an activity-induced miRNA that can regulate dendritic spine formation (Impey et al., 2010), and deletion of miR-132 and miR-212 in mice impairs synaptic plasticity (Remenyi et al., 2013). miR-181 controls neurites growth and promotes synaptogenesis in cortical neurons (Kos et al., 2016) and is involved in hippocampus-dependent memory formation (Zhang et al., 2017). As potential hub regulators of transduction of multiple mRNAs into proteins, some miRNAs that we found commonly dysregulated in the cortex and hippocampus may become in the future novel therapeutic targets, as suggested by multiple studies having taken antisense oligo-based approaches to treat diseases (Bajan and Hutvagner, 2020; Roberts et al., 2020; Stenvang and Kauppinen, 2008).

We also observed targets for the differentially expressed miRNAs were involved in axon guidance and cytoskeleton regulation. When we investigated whether miRNAs could establish a regulatory network involving the RNA binding proteins implicated in alternative splicing and axon formation, we found one of the topmost influencing miRNAs in the hippocampus and cortex being miR-181-5p. miR-181-5p targets the RNA binding protein PTBP2. At both RNA and protein levels, we found downregulation of this RBP in DS *vs.* control. miR-181-5p has the complementarity in the 3'UTR of PTBP2 to possibly regulate the expression at both RNA and protein levels.

4.7. Physiological consequences of defective neuronal polarization:

Neuronal polarization involves the specification of a single axon and multiple dendrites in undifferentiated neurites (Arimura et al., 2009). Multiple studies in primary culture derived from mouse brains have shown that any perturbation in a specific set of genes responsible for neuronal polarization might lead to axons characterized by no axons or multiple axons.

For example changes in expression of genes such as CRMP-2, PAR3, PAR6, CDC42, GSK3β, MARK2, SHTN1, LKB1 and p75 leads to differences in the percentage of neurons with no axons

or multiple axons (Gomis-Rüth et al., 2008; Shelly et al., 2010, 2011; Takano et al., 2017; Tortosa et al., 2017; Zuccaro et al., 2014). Interestingly, a recent study in primary culture from rat hippocampus has shown that the neurons with multiple axons are electrophysiologically functional with a functional axon initial segments (AIS). AIS is responsible for generating action potential and is enriched with sodium voltage channels. The authors found that sodium channels were closer to the soma in neurons with multiple axons. The depolarizing phase of the action potential was shorter compared to the single axon neurons. Although there were no differences in the whole-cell sodium currents and membrane properties, the change in the distribution of sodium voltage channels was enough to shorten the rising phase of the action potential leading to faster recruitment of somatic sodium channels.

In our differential gene expression, alternative splicing, differential expression of RBPs, differential protein expression and targets for the differentially expressed miRNAs, there were genes involved in axon guidance and cytoskeleton regulation. Our first set of experiments in Ts65Dn murine hippocampal cultures revealed an increased number of neurons with multiple axons and an increased number of neurons with no axons in comparison to wild type culture. While the number of neurons with a single axon were decreased in Ts65Dn culture in comparison to wild-type culture. This supports the results by bioinformatics analysis on DS and control human brains and warrants further experiments in brain slices from Ts65Dn animals and in neurons derived from iPS cells from DS subjects. We are currently analyzing the RNA sequencing and proteomics results from the hippocampi and cortices derived from Ts65Dn and wild-type animals to know and understand if we observe similar expression changes in the mouse model.

These findings of functional neurons with multiple AIS and defects in neuronal polarization due to perturbation in expression levels of essential polarizing molecules at both gene and protein levels suggest that the phenotype of multiple axons we observe in our hippocampal culture derived from Ts65Dn mice might result from changes in expression of polarization essential genes resulting from trisomy.

Conclusion

Our study demonstrates that multi-omics systems biology approach can identify molecular drivers at multiple gene-regulation levels for specific biological processes and cell types. We demonstrate the ability to integrate gene and protein expression data with differential transcript expression and identification of alternative splicing, further increasing the ability of whole transcriptome analysis in discovering gene expression patterns within a context of a neurodevelopmental disorder. Our study represents a proof of concept of a new approach that may lead to a significant advance towards identifying better therapeutics by providing clues to how this data from DS and control individuals could be used as a resource to understand both disease and normal brain function. In our case, it indicated unexpected deficits in neuronal polarization and a general state of neuroinflammation.

In our study, although within each group (Control and DS), the samples differ in terms of sex, race, age and PMI, it was not possible to analyse the results keeping all these parameters separate within a group because of the limited number of samples. So, future experiments may include a larger number of samples with more focus towards these biological and technical traits and performing both transcriptomics and proteomics across different ages, from fetuses to adulthood, to systematically obtain further biological insights into the development of the complex phenotype that characterize DS people. Another step forward would be to isolate other non-coding RNAs such as long non-coding RNAs and circular RNAs, providing further information at a diverse level of investigation and thus possibly new targets for future therapeutic approaches. Given the complexity of DS, a multilayer approach at diverse stage of development leading to time-specific multitarget approach may be advisable.

Appendix

Table 1: Sample and donor information for hippocampal samples (The samples in bold and red were removed based on PCA, hierarchical clustering and RIN values).

BioBank	Code	Sample	RIN	Tissue	Group	Sex	Age (years)	PMI (hours)	Race	M.O.D.
Maryland	4925	C1	6.5	Hp	Control	M	13	16	Black or African-American	Natural
Maryland	1841	C2	7.3	Hp	Control	M	19	14	White	Natural
Maryland	605	C3	6.9	Hp	Control	M	25	19	Black or African-American	Natural
Maryland	5606	C4	6.7	Hp	Control	F	35	2	White	Accidental
Maryland	4782	C5	6.7	Hp	Control	M	18	17	White	Accidental
Miami	**HCT17HFW_06**	**C6**	**2.7**	**Hp**	**Control**	**F**	**45**	**27.31**	**White**	**Natural**
Miami	**HCT17HEL_18_08**	**C7**	**5.4**	**Hp**	**Control**	**M**	**45**	**19.31**	**Black or African-American**	**Natural**
Pitt	**13149**	**C8**	**5.8**	**Hp**	**Control**	**F**	**46**	**7.6**	**White**	**Natural**
Pitt	**13250**	**C9**	**3.0**	**Hp**	**Control**	**F**	**49**	**11.25**	**White**	**Natural**
Maryland	**5762**	**C10**	**3.5**	**Hp**	**Control**	**F**	**39**	**19**	**White**	**Natural**
Maryland	5981	C11	5.5	Hp	Control	F	44	19	White	Unknown
Maryland	1276	Ds1	6.0	Hp	DS	M	13	25	Black or African-American	Natural
Maryland	**5277**	**Ds2**	**3.3**	**Hp**	**DS**	**M**	**19**	**26**	**White**	**Natural**
Maryland	5341	Ds3	5.4	Hp	DS	M	25	24	Black or African-American	Natural
Maryland	5005	Ds4	6.3	Hp	DS	F	39	12	White	Natural
Maryland	M1960M	Ds5	5.4	Hp	DS	M	19	14	Asian	Natural
Miami	**HBNY_18_02**	**Ds6**	**2.7**	**Hp**	**DS**	**F**	**51**	**8.26**	**White**	**Natural**
Miami	**HBNU_18_04**	**Ds7**	**2.6**	**Hp**	**DS**	**M**	**57**	**11.2**	**White**	**Unknown**
Pitt	13202	Ds8	6.5	Hp	DS	F	43	12.42	White	Accidental
Pitt	13235	Ds9	6.0	Hp	DS	F	44	6.53	White	Natural

Table 2: Sample and donor information for cortical samples (The samples in bold and red were removed based on PCA, hierarchical clustering and RIN values).

BioBank	Code	Sample	RIN	Tissue	Group	Sex	Age (years)	PMI (hours)	Race	M.O.D.
Miami	**HCT17HFW**	**C12**	**2.1**	**Cortex**	**Control**	**F**	**45**	**27.31**	**White**	**Natural**
Miami	HCT17HEL	C13	4.8	Cortex	Control	M	45	19.31	Black or African-American	Natural
Pitt	13149	C14	4.7	Cortex	Control	F	46	7.6	White	Natural
Pitt	**13250**	**C15**	**3.1**	**Cortex**	**Control**	**F**	**49**	**11.25**	**White**	**Natural**
Maryland	**5889**	**C16**	**5.1**	**Cortex**	**Control**	**M**	**27**	**12**	**White**	**Natural**
Maryland	5986	C17	5.9	Cortex	Control	M	41	20	Black or African-American	Natural
Maryland	**5762**	**C18**	**4.4**	**Cortex**	**Control**	**F**	**39**	**19**	**White**	**Natural**
Maryland	1668	C19	5.6	Cortex	Control	M	19	24	White	Accidental
Maryland	5563	C20	5.8	Cortex	Control	M	29	22	Black or African-American	Natural
Maryland	5235	C21	6.2	Cortex	Control	M	28	24	White	Accidental
Maryland	6137	C22	6.5	Cortex	Control	F	44	27	Black or African-American	Natural
Maryland	5981	C23	6.5	Cortex	Control	F	44	19	White	Unknown
Miami	**HBNY_18_01**	**Ds10**	**3.9**	**Cortex**	**DS**	**F**	**51**	**8.26**	**White**	**Natural**
Miami	HBNU_18_03	Ds11	2.9	Cortex	DS	M	57	11.2	White	Unknown
Pitt	13202	Ds12	6.0	Cortex	DS	F	43	12.42	White	Accidental
Pitt	13235	Ds13	5.8	Cortex	DS	F	44	6.53	White	Natural
Maryland	**753**	**Ds14**	**5.7**	**Cortex**	**DS**	**M**	**23**	**24**	**White**	**Natural**
Maryland	4904	Ds15	4.4	Cortex	DS	M	40	10	Black or African-American	Natural
Maryland	5005	Ds16	6.5	Cortex	DS	F	39	12	White	Natural
Maryland	5277	Ds17	2.9	Cortex	DS	M	19	26	White	Natural
Maryland	5341	Ds18	3.4	Cortex	DS	M	25	24		Natural
Maryland	5713	Ds19	4.2	Cortex	DS	M	25	22	White	Natural

Table 3. Number of reads mapped in total RNA sequencing.

Sample	Number of input reads	Uniquely mapped reads (number)	Uniquely mapped reads (%)	Number of splices: Total
C1	117867083	111763397	94.82%	17584647
C2	106845629	101329716	94.84%	19857006
C3	105500435	93610678	88.73%	20793072
C4	113429363	103597616	91.33%	23391231
C5	96886236	91673407	94.62%	16091717
C6	108042553	95039843	87.97%	9784247
C7	113299122	103245807	91.13%	19179718
C8	127503211	116900932	91.68%	20780909
C9	106622384	98461971	92.35%	7892502
C10	99430536	89048100	89.56%	9935738
C11	99464380	90748808	91.24%	15150297
C12	90588269	81505761	89.97%	5578697
C13	91598455	84245097	91.97%	14793848
C14	91932859	83519541	90.85%	9305769
C15	97839416	84284628	86.15%	8033066
C16	102011726	89918963	88.15%	17433260
C17	107088974	93125289	86.96%	23489526
C18	106746420	95927148	89.86%	16573042
C19	114899095	105163970	91.53%	19726597
C20	115589683	104597983	90.49%	22425066
C21	106259871	95176864	89.57%	20131540
C22	134184901	122167226	91.04%	25294663
C23	117883109	108388722	91.95%	21877059
Ds1	96625379	86039522	89.04%	17101657
Ds2	91151812	83984223	92.14%	8331928
Ds3	72786555	67434677	92.65%	11108586
Ds4	101787323	92669101	91.04%	18047760
Ds5	105137325	96270115	91.57%	16087522
Ds6	109233617	100252668	91.78%	13772600
Ds7	98043649	89330308	91.11%	11486163
Ds8	104842912	97785977	93.27%	17695414
Ds9	100993469	87129144	86.27%	14571227
Ds10	106544929	93915984	88.15%	14922084
Ds11	104445945	82609678	79.09%	10803106
Ds12	105180273	93772158	89.15%	20480754
Ds13	107449618	96692469	89.99%	19243169
Ds14	113196633	100335858	88.64%	20489965
Ds15	114013839	104013280	91.23%	12971063
Ds16	101111732	86206962	85.26%	16547695
Ds17	113654447	99401525	87.46%	8146059
Ds18	94963002	81718054	86.05%	10747696
Ds19	104288488	91201640	87.45%	14949972

Table 4. Distribution of counts among different RNA biotypes.

Biotype	Count
protein_coding	19968
lncRNA	16880
processed_pseudogene	10168
unprocessed_pseudogene	2627
misc_RNA	2220
snRNA	1910
miRNA	1879
TEC	1060
snoRNA	942
transcribed_unprocessed_pseudogene	922
rRNA_pseudogene	499
transcribed_processed_pseudogene	497
IG_V_pseudogene	188
IG_V_gene	144
transcribed_unitary_pseudogene	135
TR_V_gene	106
unitary_pseudogene	97
TR_J_gene	79
rRNA	58
polymorphic_pseudogene	55
scaRNA	49
IG_D_gene	37
TR_V_pseudogene	33
pseudogene	22
Mt_tRNA	22
IG_J_gene	18
IG_C_gene	14
IG_C_pseudogene	9
ribozyme	8
TR_C_gene	6
sRNA	5
TR_J_pseudogene	4
TR_D_gene	4
IG_J_pseudogene	3
translated_processed_pseudogene	2
Mt_rRNA	2
vaultRNA	1
translated_unprocessed_pseudogene	1
scRNA	1
IG_pseudogene	1
Total	**60676**

Table 5. Number and size of GEDDs for each chromosome from hippocampus and cortex gene expression data.

	Hippocampus		Cortex	
Chromosome	**Total GEDDs**	**Size range**	**Total GEDDs**	**Size range**
Chr1	148	2.2 Kb-25.67 Mb	210	2.73 Kb-25.64 Mb
Chr2	114	0.47 Kb-17.73 Mb	120	7.48 Kb-13.62 Mb
Chr3	82	11.33 Kb-23.68 Mb	93	5.38 Kb-32.37 Mb
Chr4	41	16.21 Kb-35.06 Mb	57	8.25 Kb-16.74 Mb
Chr5	55	5.15 Kb-42.44 Mb	77	0.09 Kb-27.63 Mb
Chr6	79	1.71 Kb-23.43 Mb	74	3.34 Kb-19.95 Mb
Chr7	65	3.93 Kb-15.33 Mb	84	8.05 Kb-14.44 Mb
Chr8	41	14.01 Kb-23.84 Mb	40	23.4 Kb-29.15 Mb
Chr9	45	19.51 Kb-27.13 Mb	68	0.99 Kb-27.9 Mb
Chr10	65	11.17 Kb-19.2 Mb	83	3.12 Kb-13.1 Mb
Chr11	77	3.05 Kb-13.69 Mb	85	0.22 Kb-16.26 Mb
Chr12	83	3.9 Kb-26.92 Mb	104	3.9 Kb-17.91 Mb
Chr13	23	85.64 Kb-23.17 Mb	29	25.61 Kb-42.19 Mb
Chr14	55	13.76 Kb-15.18 Mb	60	10.84 Kb-14.79 Mb
Chr15	67	11.05 Kb-6.01 Mb	73	9.71 Kb-6.3 Mb
Chr16	49	20.3 Kb-16.83 Mb	54	6.6 Kb-21.99 Mb
Chr17	81	3.25 Kb-6.56 Mb	99	2.4 Kb-6.38 Mb
Chr18	24	12.67 Kb-11.65 Mb	23	10.03 Kb-13.45 Mb
Chr19	43	4.08 Kb-10.25 Mb	85	1.86 Kb-6.61 Mb
Chr20	37	0.3 Kb-12.98 Mb	53	5.8 Kb-9.96 Mb
Chr21	7	127.66 Kb-9.73 Mb	5	195.32 Kb-20.17 Mb
Chr22	19	8.98 Kb-13.01 Mb	34	3.59 Kb-6.1 Mb
ChrX	56	43.11 Kb-34.88 Mb	66	8.27 Kb-28.45 Mb
Total	**1356**		**1676**	

Table 6. Oligonucleotide Table for exon skipping genes.

Gene	Event	Product sizes (bp)	Skipped exon size (bp)	No. of PCR cycles	Forward primer	Reverse primer
ANK3	HsaEX0004153	132-195	63	40	ACTGTCACAGAGAAGCACAAA	TTCCTTAAGGTCCTGTGGCCC
MARK2	HsaEX0037917	147-192	45	40	AATTTGCCCTACGGTGTGACC	CTTTTCTTTGTCGTTGCCGCC
EML1	HsaEX0022329	146-203	57	40	CCCTGCCTTTAAGAACCACGG	TCGTTCAGAAGAGCTGGTCCT
CD46	HsaEX0013872	132-177	45	40	ACAGTTATGTTTGAATGCGATAAGGG	TGGAGGCTTGTAAGTAGGCCT
PTBP2	HsaEX0050732	127-161	34	35	GCTGGTGGCAATACAGTCCTG	TGGTTTCCATCAGCCATCTGT

Table 7. List of antibodies used in the work.

Name	Host	Source	RRID Identifier	Concentration
AnkG	Guinea pig	Synaptic Systems	AB_2737033	1:300
Pan-Nav	Mouse	Sigma	AB_477552	1:100
Map2	Rabbit	Covance	AB_291679	1:250

Table 8. List of oligonucleotides used for qRT-PCR analysis

Gene	Sense Primer	Anti-sense Primer	slope	Efficency %	Efficency
hACTB	CAGCAAGCAGGAGTATGAC	GAAAGGGTGTAACGCAACT	-3.242	103.45	2.03
hAGAP3	GTGAGTCCCATGTAAACTTTGT	ACGCACACAGTCTGGTTT	-3.4	96.84	1.97
hAKAP9	ACAACACATGGCACAGAT	ATTAGTCCTAGTTCTTCCTTGAG	-3.18	106.28	2.06
hAPP	CCGCCACAGCAGCCTCTG	AAATGGACACCGATGGGTAGTGAA	-3.507	92.82	1.93
hBDNF	CGAGACCAAGTGCAATCC	TTATGAATCGCCAGCCAAT	-3.282	101.69	2.02
hDGKB	GATTCCAATTATTACGAAGCACTA	CAAGTCATAGGTCACTGATACA	-3	115.08	2.15
hDOCK4	GGCAGTGAGCAGTTGAAT	GTGGGCTAACAGAATCTCTTAA	-3.41	96.45	1.96
hDyrk1a	TGGCAGTTCTTGTCAGTTGGG	TGGCAAGGTCATAAGGCATTCC	-3.554	91.15	1.91
hENSA	AGTATGAATTAGGGCTTGGA	GCTACTCCACTTCCTTCC	-3.12	109.18	2.09
hFBXW7	ACACATTCCTTGGAACAGA	CTAAGACTGACCAGCAACT	-3.6	89.57	1.90
hGAP43	GCAATGTTCCGTTCATCTGAG	GCCTTAGAGCCGCAAGTT	-3.37	98.06	1.98
hGAPDH	AATGAAGGGGTCATTGATGG	AAGGTGAAGGTCGGAGTCAA	-3.493	93.32	1.93
hHOMER1	CATCCACAAGTCATAATTAGGTT	TCGCTTCAAGTCCACATG	-3.179	106.33	2.06
hKCC2	TCCTTCAGTAGACCTCCCT	CACAGCCCATCACATCAG	-3.446	95.07	1.95
hNAP1L1	GAGTGTTAATGGATTATTGTGTT	GATGCTTAGTTAATGGAGGTTA	-3.49	93.43	1.93
hNKCC1	GCTCTATCTAAGGACCTACCACCA	AGGCACTGAAGTACCATTCTGGAG	-3.453	94.81	1.95
hPTTGIP	TCTGTAAGGTCGGTCTTC	AATTAAGGCACTCCAAGC	-2.849	124.39	2.24
hRABEP2	GCTTGCTGTTGCCATCTG	TGTGCTCTCCTGAGTCCAT	-3.19	105.82	2.06
hRAN	TGCCACCTCATTATTATCT	TTAAACTGCCACATTCAC	-2.95	118.26	2.18
hRERE	ACACTCGGATTTGCTACG	AGAACACAGAAGTCACGATT	-3.82	82.71	1.83
hRLF	GAGCCTTCAGAGCACTTA	GTTCACTGTCATTCACCATAT	-3.52	92.34	1.92
hSEMA4D	TTGCTCATCTTCAACTTGT	TTTGTTCTTAACCCTCTCC	-3.18	106.28	2.06
hSTAT3	AATCTCCTACTTCTGCTATC	CTCAGAGAACACATCCTTA	-3	115.44	2.15
hTAXIBP1	ATCAGGGTCAGTCTTTGG	AATCTAGCACACTAATACATACAC	-3.4	96.84	1.97
hTOP2B	CTGTCTGATTGGCTTGTA	GCAGGTCTGTAGTTTGTAA	-3.18	106.28	2.06

References

Abdelmohsen, K., Hutchison, E.R., Lee, E.K., Kuwano, Y., Kim, M.M., Masuda, K., Srikantan, S., Subaran, S.S., Marasa, B.S., Mattson, M.P., et al. (2010). miR-375 Inhibits Differentiation of Neurites by Lowering HuD Levels. Mol. Cell. Biol. *30*, 4197 LP – 4210.

Abrahams, B.S., and Geschwind, D.H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. Nat. Rev. Genet. *9*, 341–355.

Ackerman, S.D., Luo, R., Poitelon, Y., Mogha, A., Harty, B.L., D'Rozario, M., Sanchez, N.E., Lakkaraju, A.K.K., Gamble, P., Li, J., et al. (2018). GPR56/ADGRG1 regulates development and maintenance of peripheral myelin. J. Exp. Med. *215*, 941–961.

Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. Nature *537*, 347–355.

Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. Elife *4*, e05005.

Ahmed, M.M., Block, A., Tong, S., Davisson, M.T., and Gardiner, K.J. (2017). Age exacerbates abnormal protein expression in a mouse model of Down syndrome. Neurobiol. Aging *57*, 120–132.

Aivazidis, S., Coughlan, C.M., Rauniyar, A.K., Jiang, H., Liggett, L.A., Maclean, K.N., and Roede, J.R. (2017). The burden of trisomy 21 disrupts the proteostasis network in Down syndrome. PLoS One *12*, e0176307.

Ana C. Xavier, and Jeffrey W. Taub (2010). Acute leukemia in children with Down syndrome. Haematologica *95*, 1043–1045.

Araya, P., Waugh, K.A., Sullivan, K.D., Núñez, N.G., Roselli, E., Smith, K.P., Granrath, R.E., Rachubinski, A.L., Enriquez Estrada, B., Butcher, E.T., et al. (2019). Trisomy 21 dysregulates T cell lineages toward an autoimmunity-prone state associated with interferon hyperactivity. Proc. Natl. Acad. Sci. *116*, 24231 LP – 24241.

Arcuri, C., Mecca, C., Bianchi, R., Giambanco, I., and Donato, R. (2017). The Pathophysiological Role of Microglia in Dynamic Surveillance, Phagocytosis and Structural Remodeling of the Developing CNS. Front. Mol. Neurosci. *10*, 191.

Arimura, N., Kimura, T., Nakamuta, S., Taya, S., Funahashi, Y., Hattori, A., Shimada, A., Ménager, C., Kawabata, S., Fujii, K., et al. (2009). Anterograde Transport of TrkB in Axons Is Mediated by Direct Interaction with Slp1 and Rab27. Dev. Cell *16*, 675–686.

Arion, D., Unger, T., Lewis, D.A., Levitt, P., and Mirnics, K. (2007). Molecular Evidence for Increased Expression of Genes Related to Immune and Chaperone Function in the Prefrontal Cortex in Schizophrenia. Biol. Psychiatry *62*, 711–721.

Asim, A., Kumar, A., Muthuswamy, S., Jain, S., and Agarwal, S. (2015). "Down syndrome: an

insight of the disease." J. Biomed. Sci. *22*, 41.

Aziz, N.M., Guedj, F., Pennings, J.L.A., Olmos-Serrano, J.L., Siegel, A., Haydar, T.F., and Bianchi, D.W. (2018). Lifespan analysis of brain development, gene expression and behavioral phenotypes in the Ts1Cje, Ts65Dn and Dp(16)1/Yey mouse models of Down syndrome. Dis. Model. Mech. dmm.031013.

Bajan, S., and Hutvagner, G. (2020). RNA-Based Therapeutics: From Antisense Oligonucleotides to miRNAs. Cells *9*, 137.

Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. Nat. Rev. Mol. Cell Biol. *18*, 437–451.

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature *465*, 53–59.

Barca-Mayo, O., and De Pietri Tonelli, D. (2014). Convergent microRNA actions coordinate neocortical development. Cell. Mol. Life Sci. *71*, 2975–2995.

Bartel, D.P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell *116*, 281–297.

Bartel, D.P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. Cell *136*, 215–233.

Bartel, D.P. (2018). Metazoan MicroRNAs. Cell *173*, 20–51.

Barter, R.L., and Yu, B. (2018). Superheat: An R Package for Creating Beautiful and Extendable Heatmaps for Visualizing Complex Data. J. Comput. Graph. Stat. *27*, 910–922.

Baum, K., Schuchhardt, J., Wolf, J., and Busse, D. (2019). Of Gene Expression and Cell Division Time: A Mathematical Framework for Advanced Differential Gene Expression and Data Analysis. Cell Syst. *9*, 569-579.e7.

Bayraktar, O.A., Fuentealba, L.C., Alvarez-Buylla, A., and Rowitch, D.H. (2014). Astrocyte development and heterogeneity. Cold Spring Harb. Perspect. Biol. *7*, a020362–a020362.

Becker, W., Soppa, U., and J. Tejedor, F. DYRK1A: A Potential Drug Target for Multiple Down Syndrome Neuropathologies. CNS Neurol. Disord. - Drug Targets- CNS Neurol. Disord. *13*, 26–33.

Beermann, J., Piccoli, M.-T., Viereck, J., and Thum, T. (2016). Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiol. Rev. *96*, 1297–1325.

Béland, L.-C., Markovinovic, A., Jakovac, H., De Marchi, F., Bilic, E., Mazzini, L., Kriz, J., and Munitic, I. (2020). Immunity in amyotrophic lateral sclerosis: blurred lines between excessive inflammation and inefficient immune responses. Brain Commun. *2*.

Benavides-Piccione, R., Ballesteros-Yáñez, I., Martínez de Lagrán, M., Elston, G., Estivill, X.,

Fillat, C., DeFelipe, J., and Dierssen, M. (2004). On dendrites in Down syndrome and DS murine models: a spiny way to learn. Prog. Neurobiol. *74*, 111–126.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B *57*, 289–300.

Bian, S., Xu, T., and Sun, T. (2013). Tuning the cell fate of neurons and glia by microRNAs. Curr. Opin. Neurobiol. *23*, 928–934.

Bianchi, P., Ciani, E., Guidi, S., Trazzi, S., Felice, D., Grossi, G., Fernandez, M., Giuliani, A., Calzà, L., and Bartesaghi, R. (2010). Early Pharmacotherapy Restores Neurogenesis and Cognitive Performance in the Ts65Dn Mouse Model for Down Syndrome. J. Neurosci. *30*, 8769 LP – 8779.

Bludau, I., and Aebersold, R. (2020). Proteomic and interactomic insights into the molecular basis of cell functional diversity. Nat. Rev. Mol. Cell Biol. *21*, 327–340.

Bogdanovic, N., Davidsson, P., Volkmann, I., Winblad, B., and Blennow, K. (2000). Growth-associated protein GAP-43 in the frontal cortex and in the hippocampus in Alzheimer's disease: an immunohistochemical and quantitative study. J. Neural Transm. *107*, 463–478.

van Bokhoven, H. (2011). Genetic and Epigenetic Networks in Intellectual Disabilities. Annu. Rev. Genet. *45*, 81–104.

Bota, M., and Swanson, L.W. (2007). The neuron classification problem. Brain Res. Rev. *56*, 79–88.

Böttcher, J.P., Bonavita, E., Chakravarty, P., Blees, H., Cabeza-Cabrerizo, M., Sammicheli, S., Rogers, N.C., Sahai, E., Zelenay, S., and Reis e Sousa, C. (2018). NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control. Cell *172*, 1022-1037.e14.

Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res. *24*, 1774–1786.

Buccitelli, C., and Selbach, M. (2020). mRNAs, proteins and the emerging principles of gene expression control. Nat. Rev. Genet. *21*, 630–644.

Bull, M.J. (2011). Health Supervision for Children With Down Syndrome. Pediatrics *128*, 393 LP – 406.

Buosi, A.S., Matias, I., Araujo, A.P.B., Batista, C., and Gomes, F.C.A. (2018). Heterogeneity in Synaptogenic Profile of Astrocytes from Different Brain Regions. Mol. Neurobiol. *55*, 751–762.

Butovsky, O., and Weiner, H.L. (2018). Microglial signatures and their role in health and disease. Nat. Rev. Neurosci. *19*, 622–635.

Cáceres, A., Ye, B., and Dotti, C.G. (2012). Neuronal polarity: demarcation, growth and commitment. Curr. Opin. Cell Biol. *24*, 547–553.

Cairney, C.J., Sanguinetti, G., Ranghini, E., Chantry, A.D., Nostro, M.C., Bhattacharyya, A., Svendsen, C.N., Keith, W.N., and Bellantuono, I. (2009). A systems biology approach to Down syndrome: Identification of Notch/Wnt dysregulation in a model of stem cells aging. Biochim. Biophys. Acta - Mol. Basis Dis. *1792*, 353–363.

Calarco, J.A., Zhen, M., and Blencowe, B.J. (2011). Networking in a global world: Establishing functional connections between neural splicing regulators and their target transcripts. RNA *17*, 775–791.

Calvo-Rodriguez, M., García-Rodríguez, C., Villalobos, C., and Núñez, L. (2020). Role of Toll Like Receptor 4 in Alzheimer's Disease . Front. Immunol. *11*, 1588.

Carlesimo, G.A., Marotta, L., and Vicari, S. (1997). Long-term memory in mental retardation: Evidence for a specific impairment in subjects with Down's syndrome. Neuropsychologia *35*, 71–79.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Cheng, P., and Poo, M. (2012). Early Events in Axon/Dendrite Polarization. Annu. Rev. Neurosci. *35*, 181–201.

Cheon, M.S., Fountoulakis, M., Dierssen, M., Ferreres, J.C., and Lubec, G. (2001). Expression profiles of proteins in fetal brain with Down syndrome BT - Protein Expression in Down Syndrome Brain. G. Lubec, ed. (Vienna: Springer Vienna), pp. 311–319.

Cheroni, C., Caporale, N., and Testa, G. (2020). Autism spectrum disorder at the crossroad between genes and environment: contributions, convergences, and interactions in ASD developmental pathophysiology. Mol. Autism *11*, 69.

Cho, C.-K.J., Drabovich, A.P., Karagiannis, G.S., Martínez-Morillo, E., Dason, S., Dimitromanolakis, A., and Diamandis, E.P. (2013). Quantitative proteomic analysis of amniocytes reveals potentially dysregulated molecular networks in Down syndrome. Clin. Proteomics *10*, 2.

Collado-Torres, L., Burke, E.E., Peterson, A., Shin, J., Straub, R.E., Rajpurohit, A., Semick, S.A., Ulrich, W.S., Price, A.J., Valencia, C., et al. (2019). Regional Heterogeneity in Gene Expression, Regulation, and Coherence in the Frontal Cortex and Hippocampus across Development and Schizophrenia. Neuron *103*, 203-216.e8.

Contestabile, A., Fila, T., Ceccarelli, C., Bonasoni, P., Bonapace, L., Santini, D., Bartesaghi, R., and Ciani, E. (2007). Cell cycle alteration and decreased cell proliferation in the hippocampal dentate gyrus and in the neocortical germinal matrix of fetuses with down syndrome and in Ts65Dn mice. Hippocampus *17*, 665–678.

Contestabile, A., Benfenati, F., and Gasparini, L. (2010). Communication breaks-Down: From

neurodevelopment defects to cognitive disabilities in Down syndrome. Prog. Neurobiol. *91*, 1–22.

Contestabile, A., Greco, B., Ghezzi, D., Tucci, V., Benfenati, F., and Gasparini, L. (2013). Lithium rescues synaptic plasticity and memory in Down syndrome mice. J. Clin. Invest. *123*, 348–361.

Contestabile, A., Magara, S., and Cancedda, L. (2017). The GABAergic Hypothesis for Cognitive Disabilities in Down Syndrome . Front. Cell. Neurosci. *11*, 54.

Costa, V., Angelini, C., D'Apice, L., Mutarelli, M., Casamassimi, A., Sommese, L., Gallo, M.A., Aprile, M., Esposito, R., Leone, L., et al. (2011). Massive-Scale RNA-Seq Analysis of Non Ribosomal Transcriptome in Human Trisomy 21. PLoS One *6*, e18493.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., et al. (2010). The Genetic Landscape of a Cell. Science (80-.). *327*, 425 LP – 431.

Couzens, D., Cuskelly, M., and Haynes, M. (2011). Cognitive Development and Down Syndrome: Age-Related Change on the Stanford-Binet Test (Fourth Edition). Am. J. Intellect. Dev. Disabil. *116*, 181–204.

Couzens, D., Haynes, M., and Cuskelly, M. (2012). Individual and Environmental Characteristics Associated with Cognitive Development in Down Syndrome: A Longitudinal Study. J. Appl. Res. Intellect. Disabil. *25*, 396–413.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526.

Cramer, N., and Galdzicki, Z. (2012). From Abnormal Hippocampal Synaptic Plasticity in Down Syndrome Mouse Models to Cognitive Disability in Down Syndrome. Neural Plast. *2012*, 101542.

CRICK, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

Da Cruz, S., and Cleveland, D.W. (2011). Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. Curr. Opin. Neurobiol. *21*, 904–919.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. *112*, 7285 LP – 7290.

Deidda, G., Parrini, M., Naskar, S., Bozarth, I.F., Contestabile, A., and Cancedda, L. (2015). Reversing excitatory GABAAR signaling restores synaptic plasticity and memory in a mouse model of Down syndrome. Nat. Med. *21*, 318–326.

Dey, A., Bhowmik, K., Chatterjee, A., Chakrabarty, P.B., Sinha, S., and Mukhopadhyay, K. (2013). Down Syndrome Related Muscle Hypotonia: Association with COL6A3 Functional SNP rs2270669. Front. Genet. *4*, 57.

Dillman, A.A., Hauser, D.N., Gibbs, J.R., Nalls, M.A., McCoy, M.K., Rudenko, I.N., Galter, D., and Cookson, M.R. (2013). mRNA expression, splicing and editing in the embryonic and adult mouse cerebral cortex. Nat. Neurosci. *16*, 499–506.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Di Domenico, F., Coccia, R., Cocciolo, A., Murphy, M.P., Cenini, G., Head, E., Butterfield, D.A., Giorgi, A., Schinina, M.E., Mancuso, C., et al. (2013). Impairment of proteostasis network in Down syndrome prior to the development of Alzheimer's disease neuropathology: redox proteomics analysis of human brain. Biochim. Biophys. Acta *1832*, 1249–1259.

Donovan, S.L., Mamounas, L.A., Andrews, A.M., Blue, M.E., and McCasland, J.S. (2002). GAP-43 Is Critical for Normal Development of the Serotonergic Innervation in Forebrain. J. Neurosci. *22*, 3543 LP – 3552.

Duchon, A., Raveau, M., Chevalier, C., Nalesso, V., Sharp, A.J., and Herault, Y. (2011). Identification of the translocation breakpoints in the Ts65Dn and Ts1Cje mouse lines: relevance for modeling Down syndrome. Mamm. Genome *22*, 674–684.

Dykes, I.M., and Emanueli, C. (2017). Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. Genomics. Proteomics Bioinformatics *15*, 177–186.

Eckhardt, M., Hultquist, J.F., Kaake, R.M., Hüttenhain, R., and Krogan, N.J. (2020). A systems approach to infectious disease. Nat. Rev. Genet. *21*, 339–354.

Edvinsson, L., Haanes, K.A., and Warfvinge, K. (2019). Does inflammation have a role in migraine? Nat. Rev. Neurol. *15*, 483–490.

Eisenberg, E., and Levanon, E.Y. (2018). A-to-I RNA editing — immune protector and transcriptome diversifier. Nat. Rev. Genet. *19*, 473–490.

Engevik, L.I., Næss, K.-A.B., and Hagtvet, B.E. (2016). Cognitive stimulation of pupils with Down syndrome: A study of inferential talk during book-sharing. Res. Dev. Disabil. *55*, 287–300.

Engle, E.C. (2010). Human genetic disorders of axon guidance. Cold Spring Harb. Perspect. Biol. *2*, a001784–a001784.

Fernandez, F., Morishita, W., Zuniga, E., Nguyen, J., Blank, M., Malenka, R.C., and Garner, C.C. (2007). Pharmacotherapy for cognitive impairment in a mouse model of Down syndrome. Nat. Neurosci. *10*, 411–413.

Ferreira, M.A.R., O'Donovan, M.C., Meng, Y.A., Jones, I.R., Ruderfer, D.M., Jones, L., Fan, J., Kirov, G., Perlis, R.H., Green, E.K., et al. (2008). Collaborative genome-wide association

analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat. Genet. *40*, 1056–1058.

Fillman, S.G., Cloonan, N., Catts, V.S., Miller, L.C., Wong, J., McCrossin, T., Cairns, M., and Weickert, C.S. (2013). Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol. Psychiatry *18*, 206–214.

Fishell, G., and Heintz, N. (2013). The Neuron Identity Problem: Form Meets Function. Neuron *80*, 602–612.

Fogel, B.L., Wexler, E., Wahnich, A., Friedrich, T., Vijayendran, C., Gao, F., Parikshak, N., Konopka, G., and Geschwind, D.H. (2012). RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. Hum. Mol. Genet. *21*, 4171–4186.

Franks, A., Airoldi, E., and Slavov, N. (2017). Post-transcriptional regulation across human tissues. PLOS Comput. Biol. *13*, e1005535.

Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature *506*, 179–184.

Fromer, M., Roussos, P., Sieberts, S.K., Johnson, J.S., Kavanagh, D.H., Perumal, T.M., Ruderfer, D.M., Oh, E.C., Topol, A., Shah, H.R., et al. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. *19*, 1442–1453.

Froussios, K., Mourão, K., Simpson, G., Barton, G., and Schurch, N. (2019). Relative Abundance of Transcripts (RATs): Identifying differential isoform abundance from RNA-seq. F1000Research *8*, 213.

Fuentes, J.J., Pritchard, M.A., and Estivill, X. (1997). Genomic Organization, Alternative Splicing, and Expression Patterns of theDSCR1(Down Syndrome Candidate Region 1) Gene. Genomics *44*, 358–361.

Furlong, L.I. (2013). Human diseases through the lens of network biology. Trends Genet. *29*, 150–159.

Gandal, M.J., Zhang, P., Hadjimichael, E., Walker, R.L., Chen, C., Liu, S., Won, H., van Bakel, H., Varghese, M., Wang, Y., et al. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science (80-.). *362*, eaat8127.

Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., Ares, M., Mody, I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706–711.

Gehman, L.T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J.E., Meisler, M.H., Ares, M., Otis, T.S., and Black, D.L. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. Genes Dev. *26*, 445–460.

Gel, B., and Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable

genomes displaying arbitrary data. Bioinformatics *33*, 3088–3090.

Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics *32*, 289–291.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Geschwind, D.H. (2011). Genetics of autism spectrum disorders. Trends Cogn. Sci. *15*, 409–416.

Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W.M., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature *511*, 344–347.

Goh, K.-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., and Barabási, A.-L. (2007). The human disease network. Proc. Natl. Acad. Sci. *104*, 8685 LP – 8690.

Gomis-Rüth, S., Wierenga, C.J., and Bradke, F. (2008). Plasticity of Polarization: Changing Dendrites into Axons in Neurons Integrated in Neuronal Circuits. Curr. Biol. *18*, 992–1000.

Gonatopoulos-Pournatzis, T., Niibori, R., Salter, E.W., Weatheritt, R.J., Tsang, B., Farhangmehr, S., Liang, X., Braunschweig, U., Roth, J., Zhang, S., et al. (2020). Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions. Mol. Cell *77*, 1176-1192.e16.

Gonzales, P.K., Roberts, C.M., Fonte, V., Jacobsen, C., Stein, G.H., and Link, C.D. (2018). Transcriptome analysis of genetically matched human induced pluripotent stem cells disomic or trisomic for chromosome 21. PLoS One *13*, e0194581–e0194581.

Goslin, K., and Banker, G. (1989). Experimental observations on the development of polarity by hippocampal neurons in culture. J. Cell Biol. *108*, 1507–1516.

Griffin, W.S., Stanley, L.C., Ling, C., White, L., MacLeod, V., Perrot, L.J., White 3rd, C.L., and Araoz, C. (1989). Brain interleukin 1 and S-100 immunoreactivity are elevated in Down syndrome and Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A. *86*, 7611–7615.

Grubb, M.S., and Burrone, J. (2010). Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability. Nature *465*, 1070–1074.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849.

Guedj, F., Pennings, J.L.A., Massingham, L.J., Wick, H.C., Siegel, A.E., Tantravahi, U., and Bianchi, D.W. (2016). An Integrated Human/Murine Transcriptome and Pathway Approach To Identify Prenatal Treatments For Down Syndrome. Sci. Rep. *6*, 32353.

Guidi, S., Bonasoni, P., Ceccarelli, C., Santini, D., Gualtieri, F., Ciani, E., and Bartesaghi, R. (2008). RESEARCH ARTICLE: Neurogenesis Impairment and Increased Cell Death Reduce

Total Neuron Number in the Hippocampal Region of Fetuses with Down Syndrome. Brain Pathol. *18*, 180–197.

Guil, S., and Cáceres, J.F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat. Struct. Mol. Biol. *14*, 591–596.

Gupta, M., Dhanasekaran, A.R., and Gardiner, K.J. (2016). Mouse models of Down syndrome: gene content and consequences. Mamm. Genome *27*, 538–555.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat. Rev. Mol. Cell Biol. *15*, 509–524.

Haas, M.A., Bell, D., Slender, A., Lana-Elola, E., Watson-Scales, S., Fisher, E.M.C., Tybulewicz, V.L.J., and Guillemot, F. (2013). Alterations to dendritic spine morphology, but not dendrite patterning, of cortical projection neurons in Tc1 and Ts1Rhr mouse models of Down syndrome. PLoS One *8*, e78561–e78561.

Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., Holoch, D., Lim, C., and Tuschl, T. (2008). Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods *44*, 3–12.

Hakak, Y., Walker, J.R., Li, C., Wong, W.H., Davis, K.L., Buxbaum, J.D., Haroutunian, V., and Fienberg, A.A. (2001). Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc. Natl. Acad. Sci. *98*, 4746 LP – 4751.

Hanisch, U.-K., and Kettenmann, H. (2007). Microglia: active sensor and versatile effector cells in the normal and pathologic brain. Nat. Neurosci. *10*, 1387–1394.

Hasin, Y., Seldin, M., and Lusis, A. (2017). Multi-omics approaches to disease. Genome Biol. *18*, 83.

Hausser, J., Mayo, A., Keren, L., and Alon, U. (2019). Central dogma rates and the trade-off between precision and economy in gene expression. Nat. Commun. *10*, 68.

Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391–399.

Heffernan, A.L., and Hare, D.J. (2018). Tracing Environmental Exposure from Neurodevelopment to Neurodegeneration. Trends Neurosci. *41*, 496–501.

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., Yeom, K.-H., Han, J., and Kim, V.N. (2009). TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. Cell *138*, 696–708.

Hill, J.M., and Lukiw, W.J. (2016). microRNA (miRNA)-Mediated Pathogenetic Signaling in Alzheimer's Disease (AD). Neurochem. Res. *41*, 96–100.

Hillmer, R.A. (2015). Systems Biology for Biologists. PLOS Pathog. *11*, e1004786.

Holahan, M., and Routtenberg, A. (2008). The protein kinase C phosphorylation site on GAP-43 differentially regulates information storage. Hippocampus *18*, 1099–1102.

Hsiao, Y.-H.E., Bahn, J.H., Lin, X., Chan, T.-M., Wang, R., and Xiao, X. (2016). Alternative splicing modulated by genetic variants demonstrates accelerated evolution regulated by highly conserved proteins. Genome Res. *26*, 440–450.

Hu, Z., and Li, Z. (2017). miRNAs in synapse development and synaptic plasticity. Curr. Opin. Neurobiol. *45*, 24–31.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Hulo, S., Alberi, S., Laux, T., Muller, D., and Caroni, P. (2002). A point mutant of GAP-43 induces enhanced short-term and long-term hippocampal plasticity. Eur. J. Neurosci. *15*, 1976–1982.

Huo, H.-Q., Qu, Z.-Y., Yuan, F., Ma, L., Yao, L., Xu, M., Hu, Y., Ji, J., Bhattacharyya, A., Zhang, S.-C., et al. (2018). Modeling Down Syndrome with Patient iPSCs Reveals Cellular and Migration Deficits of GABAergic Neurons. Stem Cell Reports *10*, 1251–1266.

Ideker, T., and Sharan, R. (2008). Protein networks in disease. Genome Res. *18*, 644–652.

Impey, S., Davare, M., Lasiek, A., Fortin, D., Ando, H., Varlamova, O., Obrietan, K., Soderling, T.R., Goodman, R.H., and Wayman, G.A. (2010). An activity-induced microRNA controls dendritic spine formation by regulating Rac1-PAK signaling. Mol. Cell. Neurosci. *43*, 146–156.

Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B., O'Hanlon, D., et al. (2014). A Highly Conserved Program of Neuronal Microexons Is Misregulated in Autistic Brains. Cell *159*, 1511–1523.

Jaffe, A.E., Straub, R.E., Shin, J.H., Tao, R., Gao, Y., Collado-Torres, L., Kam-Thong, T., Xi, H.S., Quan, J., Chen, Q., et al. (2018). Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat. Neurosci. *21*, 1117–1125.

Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat. Rev. Genet. *12*, 715–729.

Kanaumi, T., Milenkovic, I., Adle-Biassette, H., Aronica, E., and Kovacs, G.G. (2013). Non-neuronal cell responses differ between normal and Down syndrome developing brains. Int. J. Dev. Neurosci. *31*, 796–803.

Kandel, E.R. (2001). The Molecular Biology of Memory Storage: A Dialogue Between Genes and Synapses. Science (80-.). *294*, 1030 LP – 1038.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44*, D457–D462.

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483.

Katsel, P., Davis, K.L., and Haroutunian, V. (2005). Variations in myelin and oligodendrocyte-related gene expression across multiple brain regions in schizophrenia: A gene ontology study. Schizophr. Res. *79*, 157–173.

Kaur, C., Rathnasamy, G., and Ling, E.-A. (2017). Biology of Microglia in the Developing Brain. J. Neuropathol. Exp. Neurol. *76*, 736–753.

Kawahara, Y., Ito, K., Sun, H., Aizawa, H., Kanazawa, I., and Kwak, S. (2004). RNA editing and death of motor neurons. Nature *427*, 801.

Kazemi, M., Salehi, M., and Kheirollahi, M. (2016). Down Syndrome: Current Status, Challenges and Future Perspectives. Int. J. Mol. Cell. Med. *5*, 125–133.

King, I.N., Yartseva, V., Salas, D., Kumar, A., Heidersbach, A., Ando, D.M., Stallings, N.R., Elliott, J.L., Srivastava, D., and Ivey, K.N. (2014). The RNA-binding Protein TDP-43 Selectively Disrupts MicroRNA-1/206 Incorporation into the RNA-induced Silencing Complex. J. Biol. Chem. *289*, 14263–14271.

Kitano, H. (2002). Systems Biology: A Brief Overview. Science (80-.). *295*, 1662 LP – 1664.

Koch, C., and Laurent, G. (1999). Complexity and the Nervous System. Science (80-.). *284*, 96 LP – 98.

Kos, A., Olde Loohuis, N., Meinhardt, J., van Bokhoven, H., Kaplan, B.B., Martens, G.J., and Aschrafi, A. (2016). MicroRNA-181 promotes synaptogenesis and attenuates axonal outgrowth in cortical neurons. Cell. Mol. Life Sci. *73*, 3555–3567.

Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. Nucleic Acids Res. *47*, D155–D162.

Krishnaswami, S.R., Grindberg, R. V, Novotny, M., Venepally, P., Lacar, B., Bhutani, K., Linker, S.B., Pham, S., Erwin, J.A., Miller, J.A., et al. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. Nat. Protoc. *11*, 499–524.

Kristensen, A.R., Gsponer, J., and Foster, L.J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. Mol. Syst. Biol. *9*, 689.

Kubo, Y., Baba, K., Toriyama, M., Minegishi, T., Sugiura, T., Kozawa, S., Ikeda, K., and Inagaki, N. (2015). Shootin1-cortactin interaction mediates signal-force transduction for axon outgrowth. J. Cell Biol. *210*, 663–676.

Kustatscher, G., Grabowski, P., Schrader, T.A., Passmore, J.B., Schrader, M., and Rappsilber, J. (2019). Co-regulation map of the human proteome enables identification of protein functions. Nat. Biotechnol. *37*, 1361–1371.

Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung,

H.-L., Chen, S., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science (80-.). *352*, 1586 LP – 1590.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Lanzillotta, C., Greco, V., Valentini, D., Villani, A., Folgiero, V., Caforio, M., Locatelli, F., Pagnotta, S., Barone, E., Urbani, A., et al. (2020). Proteomics Study of Peripheral Blood Mononuclear Cells in Down Syndrome Children. Antioxidants *9*, 1112.

Lapek, J.D., Greninger, P., Morris, R., Amzallag, A., Pruteanu-Malinici, I., Benes, C.H., and Haas, W. (2017). Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat. Biotechnol. *35*, 983–989.

Laurence, J.A., and Fatemi, S.H. (2005). Glial fibrillary acidic protein is elevated in superior frontal, parietal and cerebellar cortices of autistic subjects. The Cerebellum *4*, 206–210.

Lawrence, M., Gentleman, R., and Carey, V. (2009). rtracklayer: an R package for interfacing with genome browsers. Bioinformatics *25*, 1841–1842.

Lecours, C., Bordeleau, M., Cantin, L., Parent, M., Paolo, T. Di, and Tremblay, M.-È. (2018). Microglial Implication in Parkinson's Disease: Loss of Beneficial Physiological Roles or Gain of Inflammatory Functions? Front. Cell. Neurosci. *12*, 282.

Lee, I., Blom, U.M., Wang, P.I., Shim, J.E., and Marcotte, E.M. (2011). Prioritizing candidate disease genes by network-based boosting of genome-wide association data. Genome Res. *21*, 1109–1121.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M.G., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035–1043.

Letourneau, A., Santoni, F.A., Bonilla, X., Sailani, M.R., Gonzalez, D., Kind, J., Chevalier, C., Thurman, R., Sandstrom, R.S., Hibaoui, Y., et al. (2014). Domains of genome-wide gene expression dysregulation in Down's syndrome. Nature *508*, 345–350.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. Cell *120*, 15–20.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Li, J., and Liu, C. (2019). Coding or Noncoding, the Converging Concepts of RNAs . Front. Genet. *10*, 496.

Li, J.J., Bickel, P.J., and Biggin, M.D. (2014a). System wide analyses have underestimated protein abundances and the importance of transcription in mammals. PeerJ *2*, e270.

Li, Q., Zheng, S., Han, A., Lin, C.-H., Stoilov, P., Fu, X.-D., and Black, D.L. (2014b). The splicing

regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. Elife *3*, e01201.

Li, X., Chauhan, A., Sheikh, A.M., Patil, S., Chauhan, V., Li, X.-M., Ji, L., Brown, T., and Malik, M. (2009). Elevated immune response in the brain of autistic patients. J. Neuroimmunol. *207*, 111–116.

Li, Z., Yu, T., Morishima, M., Pao, A., LaDuca, J., Conroy, J., Nowak, N., Matsui, S.-I., Shiraishi, I., and Yu, Y.E. (2007). Duplication of the entire 22.9 Mb human chromosome 21 syntenic region on mouse chromosome 16 causes cardiovascular and gastrointestinal abnormalities. Hum. Mol. Genet. *16*, 1359–1366.

Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., and Darnell, R.B. (2012). Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Genes Dev. *26*, 1626–1642.

Liddelow, S.A., and Barres, B.A. (2017). Reactive Astrocytes: Production, Function, and Therapeutic Potential. Immunity *46*, 957–967.

Liddelow, S.A., Guttenplan, K.A., Clarke, L.E., Bennett, F.C., Bohlen, C.J., Schirmer, L., Bennett, M.L., Münch, A.E., Chung, W.-S., Peterson, T.C., et al. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. Nature *541*, 481–487.

Liddelow, S.A., Marsh, S.E., and Stevens, B. (2020). Microglia and Astrocytes in Disease: Dynamic Duo or Partners in Crime? Trends Immunol. *41*, 820–835.

Ling, J.P., Pletnikova, O., Troncoso, J.C., and Wong, P.C. (2015). TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science (80-.). *349*, 650 LP – 655.

Liu, Y., and Aebersold, R. (2016). The interdependence of transcript and protein abundance: new data–new complexities. Mol. Syst. Biol. *12*, 856.

Liu, H., Wang, H., Zhu, H., Zhang, H., and Liu, S. (2018). Preliminary study of protein changes in trisomy 21 fetus by proteomics analysis in amniocyte. Prenat. Diagn. *38*, 435–444.

Liu, Y., Borel, C., Li, L., Müller, T., Williams, E.G., Germain, P.-L., Buljan, M., Sajic, T., Boersema, P.J., Shao, W., et al. (2017). Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells. Nat. Commun. *8*, 1212.

Lo, L.H.-Y., and Lai, K.-O. (2020). Dysregulation of protein synthesis and dendritic spine morphogenesis in ASD: studies in human pluripotent stem cells. Mol. Autism *11*, 40.

Lockstone, H.E., Harris, L.W., Swatton, J.E., Wayland, M.T., Holland, A.J., and Bahn, S. (2007). Gene expression profiling in the adult Down syndrome brain. Genomics *90*, 647–660.

Lott, I.T. (2012). Chapter 6 - Neurological phenotypes for Down syndrome across the life span. In Down Syndrome: From Understanding the Neurobiology to Therapy, M. Dierssen, and R.B.T.-P. in B.R. De La Torre, eds. (Elsevier), pp. 101–121.

Lu, Y., Baras, A.S., and Halushka, M.K. (2018). miRge 2.0 for comprehensive analysis of

microRNA sequencing data. BMC Bioinformatics *19*, 275.

Luber, C.A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., et al. (2010). Quantitative Proteomics Reveals Subset-Specific Viral Recognition in Dendritic Cells. Immunity *32*, 279–289.

Luo, W., and Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics *29*, 1830–1831.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Mangleburg, C.G., Wu, T., Yalamanchili, H.K., Guo, C., Hsieh, Y.-C., Duong, D.M., Dammer, E.B., De Jager, P.L., Seyfried, N.T., Liu, Z., et al. (2020). Integrated analysis of the aging brain transcriptome and proteome in tauopathy. Mol. Neurodegener. *15*, 56.

Manning, K.S., and Cooper, T.A. (2017). The roles of RNA processing in translating genotype to phenotype. Nat. Rev. Mol. Cell Biol. *18*, 102–114.

Manzoni, C., Kia, D.A., Vandrovcova, J., Hardy, J., Wood, N.W., Lewis, P.A., and Ferrari, R. (2018). Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. Brief. Bioinform. *19*, 286–302.

Marin-Padilla, M. (1972). Structural abnormalities of the cerebral cortex in human chromosomal aberrations: a Golgi study. Brain Res. *44*, 625–629.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal; Vol 17, No 1 Next Gener. Seq. Data Anal. - 10.14806/Ej.17.1.200 .

Martínez-Cué, C., Rueda, N., García, E., Davisson, M.T., Schmidt, C., and Flórez, J. (2005). Behavioral, cognitive and biochemical responses to different environmental conditions in male Ts65Dn mice, a model of Down syndrome. Behav. Brain Res. *163*, 174–185.

Martins-de-Souza, D. (2010). Proteome and transcriptome analysis suggests oligodendrocyte dysfunction in schizophrenia. J. Psychiatr. Res. *44*, 149–156.

Martins-de-Souza, D., Maccarrone, G., Wobrock, T., Zerr, I., Gormanns, P., Reckow, S., Falkai, P., Schmitt, A., and Turck, C.W. (2010). Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. J. Psychiatr. Res. *44*, 1176–1189.

Masland, R.H. (2004). Neuronal cell types. Curr. Biol. *14*, R497–R500.

Matias, I., Morgado, J., and Gomes, F.C.A. (2019). Astrocyte Heterogeneity: Impact to Brain Aging and Disease . Front. Aging Neurosci. *11*, 59.

Matson, J.L., and Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. Res. Dev. Disabil. *30*, 1107–1114.

McCarroll, S.A., Feng, G., and Hyman, S.E. (2014). Genome-scale neurogenetics: methodology and meaning. Nat. Neurosci. *17*, 756–763.

McNeill, E.M., Warinner, C., Alkins, S., Taylor, A., Heggeness, H., DeLuca, T.F., Fulga, T.A., Wall, D.P., Griffith, L.C., and Van Vactor, D. (2020). The conserved microRNA miR-34 regulates synaptogenesis via coordination of distinct mechanisms in presynaptic and postsynaptic cells. Nat. Commun. *11*, 1092.

Metz, G.A., and Schwab, M.E. (2004). Behavioral characterization in a comprehensive mouse test battery reveals motor and sensory impairments in growth-associated protein-43 null mutant mice. Neuroscience *129*, 563–574.

Michlewski, G., and Cáceres, J.F. (2010). Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. Nat. Struct. Mol. Biol. *17*, 1011–1018.

Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K., et al. (2014). Transcriptional landscape of the prenatal human brain. Nature *508*, 199–206.

Miyamoto, A., Wake, H., Ishikawa, A.W., Eto, K., Shibata, K., Murakoshi, H., Koizumi, S., Moorhouse, A.J., Yoshimura, Y., and Nabekura, J. (2016). Microglia contact induces synapse formation in developing somatosensory cortex. Nat. Commun. *7*, 12540.

Monteuuis, G., Wong, J.J.L., Bailey, C.G., Schmitz, U., and Rasko, J.E.J. (2019). The changing paradigm of intron retention: regulation, ramifications and recipes. Nucleic Acids Res. *47*, 11497–11513.

Morgan, J.T., Chana, G., Pardo, C.A., Achim, C., Semendeferi, K., Buckwalter, J., Courchesne, E., and Everall, I.P. (2010). Microglial Activation and Increased Microglial Density Observed in the Dorsolateral Prefrontal Cortex in Autism. Biol. Psychiatry *68*, 368–376.

Mosher, K.I., Andres, R.H., Fukuhara, T., Bieri, G., Hasegawa-Moriyama, M., He, Y., Guzman, R., and Wyss-Coray, T. (2012). Neural progenitor cells regulate microglia functions and activity. Nat. Neurosci. *15*, 1485–1487.

Nagaraj, N., Alexander Kulak, N., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. (2012). System-wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-shot Ultra HPLC Runs on a Bench Top Orbitrap*. Mol. Cell. Proteomics *11*, M111.013722.

Nakai, N., Takumi, T., Nakai, J., and Sato, M. (2018). Common Defects of Spine Dynamics and Circuit Function in Neurodevelopmental Disorders: A Systematic Review of Findings From in Vivo Optical Imaging of Mouse Models . Front. Neurosci. *12*, 412.

Newman, J.R.B., Concannon, P., Tardaguila, M., Conesa, A., and McIntyre, L.M. (2018). Event Analysis: Using Transcript Events To Improve Estimates of Abundance in RNA-seq Data. G3 Genes|Genomes|Genetics *8*, 2923 LP – 2940.

Niemi, M.E.K., Martin, H.C., Rice, D.L., Gallone, G., Gordon, S., Kelemen, M., McAloney, K.,

McRae, J., Radford, E.J., Yu, S., et al. (2018). Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. Nature *562*, 268–271.

Norden, D.M., Trojanowski, P.J., Villanueva, E., Navarro, E., and Godbout, J.P. (2016). Sequential activation of microglia and astrocyte cytokine expression precedes increased Iba-1 or GFAP immunoreactivity following systemic immune challenge. Glia *64*, 300–316.

Nussbacher, J.K., and Yeo, G.W. (2018). Systematic Discovery of RNA Binding Proteins that Regulate MicroRNA Levels. Mol. Cell *69*, 1005-1016.e7.

O'Bryant, S.E., Zhang, F., Silverman, W., Lee, J.H., Krinsky-McHale, S.J., Pang, D., Hall, J., and Schupf, N. (2020). Proteomic profiles of incident mild cognitive impairment and Alzheimer's disease among adults with Down syndrome. Alzheimer's Dement. (Amsterdam, Netherlands) *12*, e12033–e12033.

Okaty, B.W., Sugino, K., and Nelson, S.B. (2011). Cell Type-Specific Transcriptomics in the Brain. J. Neurosci. *31*, 6939 LP – 6943.

Olde Loohuis, N.F.M., Kos, A., Martens, G.J.M., Van Bokhoven, H., Nadif Kasri, N., and Aschrafi, A. (2012). MicroRNA networks direct neuronal development and plasticity. Cell. Mol. Life Sci. *69*, 89–102.

Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S., and Geschwind, D.H. (2008). Functional organization of the transcriptome in human brain. Nat. Neurosci. *11*, 1271–1282.

Olmos-Serrano, J.L., Kang, H.J., Tyler, W.A., Silbereis, J.C., Cheng, F., Zhu, Y., Pletikos, M., Jankovic-Rapan, L., Cramer, N.P., Galdzicki, Z., et al. (2016). Down Syndrome Developmental Brain Transcriptome Reveals Defective Oligodendrocyte Differentiation and Myelination. Neuron *89*, 1208–1222.

Oppermann, M., Cols, N., Nyman, T., Helin, J., Saarinen, J., Byman, I., Toran, N., Alaiya, A.A., Bergman, T., Kalkkinen, N., et al. (2000). Identification of foetal brain proteins by two-dimensional gel electrophoresis and mass spectrometry. Eur. J. Biochem. *267*, 4713–4719.

Ori, A., Iskar, M., Buczak, K., Kastritis, P., Parca, L., Andrés-Pons, A., Singer, S., Bork, P., and Beck, M. (2016). Spatiotemporal variation of mammalian protein complex stoichiometries. Genome Biol. *17*, 47.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413–1415.

Paolicelli, R.C., Bolasco, G., Pagani, F., Maggi, L., Scianni, M., Panzanelli, P., Giustetto, M., Ferreira, T.A., Guiducci, E., Dumas, L., et al. (2011). Synaptic Pruning by Microglia Is Necessary for Normal Brain Development. Science (80-.). *333*, 1456 LP – 1458.

Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative Functional Genomic Analyses Implicate Specific

Molecular Pathways and Circuits in Autism. Cell *155*, 1008–1021.

Parikshak, N.N., Gandal, M.J., and Geschwind, D.H. (2015). Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. Nat. Rev. Genet. *16*, 441–458.

Parikshak, N.N., Swarup, V., Belgard, T.G., Irimia, M., Ramaswami, G., Gandal, M.J., Hartl, C., Leppa, V., Ubieta, L. de la T., Huang, J., et al. (2016). Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. Nature *540*, 423–427.

Park, E., Pan, Z., Zhang, Z., Lin, L., and Xing, Y. (2018). The Expanding Landscape of Alternative Splicing Variation in Human Populations. Am. J. Hum. Genet. *102*, 11–26.

Parrini, M., Ghezzi, D., Deidda, G., Medrihan, L., Castroflorio, E., Alberti, M., Baldelli, P., Cancedda, L., and Contestabile, A. (2017). Aerobic exercise and a BDNF-mimetic therapy rescue learning and memory in a mouse model of Down syndrome. Sci. Rep. *7*, 16825.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2013). Extensive transcriptional heterogeneity revealed by isoform profiling. Nature *497*, 127–131.

Pelleri, M.C., Cattani, C., Vitale, L., Antonaros, F., Strippoli, P., Locatelli, C., Cocchi, G., Piovesan, A., and Caracausi, M. (2018). Integrated Quantitative Transcriptome Maps of Human Trisomy 21 Tissues and Cells. Front. Genet. *9*, 125.

Pennington, K., Beasley, C.L., Dicker, P., Fagan, A., English, J., Pariante, C.M., Wait, R., Dunn, M.J., and Cotter, D.R. (2008). Prominent synaptic and metabolic abnormalities revealed by proteomic analysis of the dorsolateral prefrontal cortex in schizophrenia and bipolar disorder. Mol. Psychiatry *13*, 1102–1117.

Pérez-Villareal, J.M., Aviña-Padilla, K., López, E.B., Guadrón-Llanos, A.M., López-Bayghen, E., Magaña-Gómez, J., Meraz-Ríos, M.A., Varela-Echavarría, A., and Angulo-Rojo, C. (2020). Profiling of circulating chromosome 21-encoded microRNAs, miR-155 and Let-7c, in Down Syndrome People. MedRxiv 2020.10.24.20218677.

Pinto, B., Morelli, G., Rastogi, M., Savardi, A., Fumagalli, A., Petretto, A., Bartolucci, M., Varea, E., Catelani, T., Contestabile, A., et al. (2020). Rescuing Over-activated Microglia Restores Cognitive Performance in Juvenile Animals of the Dp(16) Mouse Model of Down Syndrome. Neuron *108*, 887-904.e12.

Poulopoulos, A., Murphy, A.J., Ozkan, A., Davis, P., Hatch, J., Kirchner, R., and Macklis, J.D. (2019). Subcellular transcriptomes and proteomes of developing axon projections in the cerebral cortex. Nature *565*, 356–360.

Prabakaran, S., Swatton, J.E., Ryan, M.M., Huffaker, S.J., Huang, J.-J., Griffin, J.L., Wayland, M., Freeman, T., Dudbridge, F., Lilley, K.S., et al. (2004). Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. Mol. Psychiatry *9*, 684–697.

Quesnel-Vallières, M., Weatheritt, R.J., Cordes, S.P., and Blencowe, B.J. (2019). Autism

spectrum disorder: insights into convergent mechanisms from transcriptomics. Nat. Rev. Genet. *20*, 51–63.

Rachidi, M., and Lopes, C. (2011). Mental Retardation and Human Chromosome 21 Gene Overdosage: From Functional Genomics and Molecular Mechanisms Towards Prevention and Treatment of the Neuropathogenesis of Down Syndrome BT - Genomics, Proteomics, and the Nervous System. J.D. Clelland, ed. (New York, NY: Springer New York), pp. 21–86.

Raj, B., and Blencowe, B.J. (2015). Alternative Splicing in the Mammalian Nervous System: Recent Insights into Mechanisms and Functional Roles. Neuron *87*, 14–27.

Raj, B., O'Hanlon, D., Vessey, J.P., Pan, Q., Ray, D., Buckley, N.J., Miller, F.D., and Blencowe, B.J. (2011). Cross-Regulation between an Alternative Splicing Activator and a Transcription Repressor Controls Neurogenesis. Mol. Cell *43*, 843–850.

Raj, T., Li, Y.I., Wong, G., Humphrey, J., Wang, M., Ramdhani, S., Wang, Y.-C., Ng, B., Gupta, I., Haroutunian, V., et al. (2018). Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. Nat. Genet. *50*, 1584–1592.

Rajasethupathy, P., Fiumara, F., Sheridan, R., Betel, D., Puthanveettil, S. V, Russo, J.J., Sander, C., Tuschl, T., and Kandel, E. (2009). Characterization of small RNAs in Aplysia reveals a role for miR-124 in constraining synaptic plasticity through CREB. Neuron *63*, 803–817.

Rajman, M., and Schratt, G. (2017). MicroRNAs in neural development: from master regulators to fine-tuners. Development *144*, 2310 LP – 2322.

Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., Consortium, U.K.B.E., Consortium, N.A.B.E., Coin, L., et al. (2014). Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat. Neurosci. *17*, 1418–1428.

Rekart, J.L., Meiri, K., and Routtenberg, A. (2005). Hippocampal-dependent memory is impaired in heterozygous GAP-43 knockout mice. Hippocampus *15*, 1–7.

Remenyi, J., van den Bosch, M.W.M., Palygin, O., Mistry, R.B., McKenzie, C., Macdonald, A., Hutvagner, G., Arthur, J.S.C., Frenguelli, B.G., and Pankratov, Y. (2013). miR-132/212 Knockout Mice Reveal Roles for These miRNAs in Regulating Cortical Synaptic Transmission and Plasticity. PLoS One *8*, e62509.

Rice, D.S., Sheldon, M., D Arcangelo, G., Nakajima, K., Goldowitz, D., and Curran, T. (1998). Disabled-1 acts downstream of Reelin in a signaling pathway that controls laminar organization in the mammalian brain. Development *125*, 3719 LP – 3729.

Roberts, T.C., Langer, R., and Wood, M.J.A. (2020). Advances in oligonucleotide drug delivery. Nat. Rev. Drug Discov. *19*, 673–694.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Rockenstein, E.M., McConlogue, L., Tan, H., Power, M., Masliah, E., and Mucke, L. (1995). Levels and Alternative Splicing of Amyloid β Protein Precursor (APP) Transcripts in Brains of APP Transgenic Mice and Humans with Alzheimer's Disease. J. Biol. Chem. *270*, 28257–28267.

Roizen, N.J., and Patterson, D. (2003). Down's syndrome. Lancet *361*, 1281–1289.

Romanov, N., Kuhn, M., Aebersold, R., Ori, A., Beck, M., and Bork, P. (2019). Disentangling Genetic and Environmental Effects on the Proteotypes of Individuals. Cell *177*, 1308-1318.e10.

Ropers, H.H. (2008). Genetics of intellectual disability. Curr. Opin. Genet. Dev. *18*, 241–250.

Salter, M.W., and Stevens, B. (2017). Microglia emerge as central players in brain disease. Nat. Med. *23*, 1018–1027.

Salvi, A., Vezzoli, M., Busatto, S., Paolini, L., Faranda, T., Abeni, E., Caracausi, M., Antonaros, F., Piovesan, A., Locatelli, C., et al. (2019). Analysis of a nanoparticle-enriched fraction of plasma reveals miRNA candidates for Down syndrome pathogenesis. Int. J. Mol. Med. *43*, 2303–2318.

Sambandan, S., Akbalik, G., Kochen, L., Rinne, J., Kahlstatt, J., Glock, C., Tushev, G., Alvarez-Castelao, B., Heckel, A., and Schuman, E.M. (2017). Activity-dependent spatially localized miRNA maturation in neuronal dendrites. Science (80-.). *355*, 634 LP – 637.

Sarraf, S.A., Shah, H. V, Kanfer, G., Pickrell, A.M., Holtzclaw, L.A., Ward, M.E., and Youle, R.J. (2020). Loss of TAX1BP1-Directed Autophagy Results in Protein Aggregate Accumulation in the Brain. Mol. Cell *80*, 779-795.e10.

Savardi, A., Borgogno, M., Narducci, R., La Sala, G., Ortega, J.A., Summa, M., Armirotti, A., Bertorelli, R., Contestabile, A., De Vivo, M., et al. (2020). Discovery of a Small Molecule Drug Candidate for Selective NKCC1 Inhibition in Brain Disorders. Chem *6*, 2073–2096.

Schafer, D.P., and Stevens, B. (2013). Phagocytic glial cells: sculpting synaptic circuits in the developing nervous system. Curr. Opin. Neurobiol. *23*, 1034–1040.

Schratt, G.M., Tuebing, F., Nigh, E.A., Kane, C.G., Sabatini, M.E., Kiebler, M., and Greenberg, M.E. (2006). A brain-specific microRNA regulates dendritic spine development. Nature *439*, 283–289.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Schwanhäusser, B., Wolf, J., Selbach, M., and Busse, D. (2013). Synthesis and degradation jointly determine the responsiveness of the cellular proteome. BioEssays *35*, 597–601.

Scott, M.S., Avolio, F., Ono, M., Lamond, A.I., and Barton, G.J. (2009). Human miRNA Precursors with Box H/ACA snoRNA Features. PLOS Comput. Biol. *5*, e1000507.

Seyfried, N.T., Dammer, E.B., Swarup, V., Nandakumar, D., Duong, D.M., Yin, L., Deng, Q., Nguyen, T., Hales, C.M., Wingo, T., et al. (2017). A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer's Disease. Cell Syst. *4*, 60-72.e4.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. *13*, 2498–2504.

Sharma, K., Schmitt, S., Bergner, C.G., Tyanova, S., Kannaiyan, N., Manrique-Hoyos, N., Kongi, K., Cantuti, L., Hanisch, U.-K., Philips, M.-A., et al. (2015). Cell type– and brain region–resolved mouse brain proteome. Nat. Neurosci. *18*, 1819–1831.

Sharp, P.A. (2009). The Centrality of RNA. Cell *136*, 577–580.

Shelly, M., Lim, B.K., Cancedda, L., Heilshorn, S.C., Gao, H., and Poo, M. (2010). Local and Long-Range Reciprocal Regulation of cAMP and cGMP in Axon/Dendrite Formation. Science (80-.). *327*, 547 LP – 552.

Shelly, M., Cancedda, L., Lim, B.K., Popescu, A.T., Cheng, P., Gao, H., and Poo, M. (2011). Semaphorin3A regulates neuronal polarization by suppressing axon formation and promoting dendrite growth. Neuron *71*, 433–446.

Shen, Y., Mani, S., Donovan, S.L., Schwob, J.E., and Meiri, K.F. (2002). Growth-Associated Protein-43 Is Required for Commissural Axon Guidance in the Developing Vertebrate Nervous System. J. Neurosci. *22*, 239 LP – 247.

Sherman, S.L., Allen, E.G., Bean, L.H., and Freeman, S.B. (2007). Epidemiology of Down syndrome. Ment. Retard. Dev. Disabil. Res. Rev. *13*, 221–227.

Shin, J., Ming, G., and Song, H. (2014). Decoding neural transcriptomes and epigenomes via high-throughput sequencing. Nat. Neurosci. *17*, 1463–1475.

Shobin, E., Bowley, M.P., Estrada, L.I., Heyworth, N.C., Orczykowski, M.E., Eldridge, S.A., Calderazzo, S.M., Mortazavi, F., Moore, T.L., and Rosene, D.L. (2017). Microglia activation and phagocytosis: relationship with aging and cognitive impairment in the rhesus monkey. GeroScience *39*, 199–220.

Siegel, G., Obernosterer, G., Fiore, R., Oehmen, M., Bicker, S., Christensen, M., Khudayberdiev, S., Leuschner, P.F., Busch, C.J.L., Kane, C., et al. (2009). A functional screen implicates microRNA-138-dependent regulation of the depalmitoylation enzyme APT1 in dendritic spine morphogenesis. Nat. Cell Biol. *11*, 705–716.

Sipe, G.O., Lowery R. L., Tremblay, M.-È., Kelly, E.A., Lamantia, C.E., and Majewska, A.K. (2016). Microglial P2Y12 is necessary for synaptic plasticity in mouse visual cortex. Nat. Commun. *7*, 10905.

Skene, N.G., and Grant, S.G.N. (2016). Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment .

Front. Neurosci. *10*, 16.

Sobol, M., Klar, J., Laan, L., Shahsavani, M., Schuster, J., Annerén, G., Konzer, A., Mi, J., Bergquist, J., Nordlund, J., et al. (2019). Transcriptome and Proteome Profiling of Neural Induced Pluripotent Stem Cells from Individuals with Down Syndrome Disclose Dynamic Dysregulations of Key Pathways and Cellular Functions. Mol. Neurobiol. *56*, 7113–7127.

Sochocka, M., Diniz, B.S., and Leszek, J. (2017). Inflammatory Response in the CNS: Friend or Foe? Mol. Neurobiol. *54*, 8071–8089.

Sokolova, A., Hill, M.D., Rahimi, F., Warden, L.A., Halliday, G.M., and Shepherd, C.E. (2009). Monocyte Chemoattractant Protein-1 Plays a Dominant Role in the Chronic Inflammation Observed in Alzheimer's Disease. Brain Pathol. *19*, 392–398.

Sommer, B., Köhler, M., Sprengel, R., and Seeburg, P.H. (1991). RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell *67*, 11–19.

Sousa, A.M.M., Zhu, Y., Raghanti, M.A., Kitchen, R.R., Onorati, M., Tebbenkamp, A.T.N., Stutz, B., Meyer, K.A., Li, M., Kawasawa, Y.I., et al. (2017). Molecular and cellular reorganization of neural circuits in the human lineage. Science (80-.). *358*, 1027 LP – 1032.

Squarzoni, P., Thion, M., and Garel, S. (2015). Neuronal and microglial regulators of cortical wiring: usual and novel guideposts . Front. Neurosci. *9*, 248.

Stamoulis, G., Garieri, M., Makrythanasis, P., Letourneau, A., Guipponi, M., Panousis, N., Sloan-Béna, F., Falconnet, E., Ribaux, P., Borel, C., et al. (2019). Single cell transcriptome in aneuploidies reveals mechanisms of gene dosage imbalance. Nat. Commun. *10*, 4495.

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. Nat. Rev. Genet. *20*, 631–656.

Stenvang, J., and Kauppinen, S. (2008). MicroRNAs as targets for antisense-based therapeutics. Expert Opin. Biol. Ther. *8*, 59–81.

Stoll, G., and Nieswandt, B. (2019). Thrombo-inflammation in acute ischaemic stroke — implications for treatment. Nat. Rev. Neurol. *15*, 473–481.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545 LP – 15550.

Suetsugu, M., and Mehraein, P. (1980). Spine distribution along the apical dendrites of the pyramidal neurons in Down's syndrome. Acta Neuropathol. *50*, 207–210.

Sullivan, K.D., Lewis, H.C., Hill, A.A., Pandey, A., Jackson, L.P., Cabral, J.M., Smith, K.P., Liggett, L.A., Gomez, E.B., Galbraith, M.D., et al. (2016). Trisomy 21 consistently activates the interferon response. Elife *5*, e16220.

Sullivan, K.D., Evans, D., Pandey, A., Hraha, T.H., Smith, K.P., Markham, N., Rachubinski, A.L.,

Wolter-Warmerdam, K., Hickey, F., Espinosa, J.M., et al. (2017). Trisomy 21 causes changes in the circulating proteome indicative of chronic autoinflammation. Sci. Rep. *7*, 14818.

Sun, Y., Dierssen, M., Toran, N., Pollak, D.D., Chen, W.-Q., and Lubec, G. (2011). A gel-based proteomic method reveals several protein pathway abnormalities in fetal Down syndrome brain. J. Proteomics *74*, 547–557.

Swanger, S.A., and Bassell, G.J. (2011). Making and breaking synapses through local mRNA regulation. Curr. Opin. Genet. Dev. *21*, 414–421.

Szu-Yu Ho, T., and Rasband, M.N. (2011). Maintenance of neuronal polarity. Dev. Neurobiol. *71*, 474–482.

Takano, T., Wu, M., Nakamuta, S., Naoki, H., Ishizawa, N., Namba, T., Watanabe, T., Xu, C., Hamaguchi, T., Yura, Y., et al. (2017). Discovery of long-range inhibitory signaling to ensure single axon formation. Nat. Commun. *8*, 33.

Tan, Y.H., Schneider, E.L., Tischfield, J., Epstein, C.J., and Ruddle, F.H. (1974). Human Chromosome 21 Dosage: Effect on the Expression of the Interferon Induced Antiviral State. Science (80-.). *186*, 61 LP – 63.

Tapial, J., Ha, K.C.H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y., et al. (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res. *27*, 1759–1768.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72–78.

Tay, T.L., Béchade, C., D'Andrea, I., St-Pierre, M.-K., Henry, M.S., Roumier, A., and Tremblay, M.-E. (2018). Microglia Gone Rogue: Impacts on Psychiatric Disorders across the Lifespan. Front. Mol. Neurosci. *10*, 421.

Taylor, I.W., and Wrana, J.L. (2012). Protein interaction networks in medicine and disease. Proteomics *12*, 1706–1716.

Tian, S.Y., Wang, J.-F., Bezchlibnyk, Y.B., and Young, L.T. (2007). Immunoreactivity of 43kDa growth-associated protein is decreased in post mortem hippocampus of bipolar disorder and schizophrenia. Neurosci. Lett. *411*, 123–127.

Tkachev, D., Mimmack, M.L., Ryan, M.M., Wayland, M., Freeman, T., Jones, P.B., Starkey, M., Webster, M.J., Yolken, R.H., and Bahn, S. (2003). Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet *362*, 798–805.

Toiber, D., Azkona, G., Ben-Ari, S., Torán, N., Soreq, H., and Dierssen, M. (2010). Engineering DYRK1A overdosage yields Down syndrome-characteristic cortical splicing aberrations. Neurobiol. Dis. *40*, 348–359.

Tortosa, E., Adolfs, Y., Fukata, M., Pasterkamp, R.J., Kapitein, L.C., and Hoogenraad, C.C.

(2017). Dynamic Palmitoylation Targets MAP6 to the Axon to Promote Microtubule Stabilization during Neuronal Polarization. Neuron *94*, 809-825.e7.

Toso, L., Cameroni, I., Roberson, R., Abebe, D., Bissell, S., and Spong, C.Y. (2008). Prevention of developmental delays in a Down syndrome mouse model. Obstet. Gynecol. *112*, 1242–1251.

Trabzuni, D., Wray, S., Vandrovcova, J., Ramasamy, A., Walker, R., Smith, C., Luk, C., Gibbs, J.R., Dillman, A., Hernandez, D.G., et al. (2012). MAPT expression and splicing is differentially regulated by brain region: relation to genotype and implication for tauopathies. Hum. Mol. Genet. *21*, 4094–4103.

Tran, N.Q.V., and Miyake, K. (2017). Neurodevelopmental Disorders and Environmental Toxicants: Epigenetics as an Underlying Mechanism. Int. J. Genomics *2017*, 7526592.

Tremblay, M.-È., Stevens, B., Sierra, A., Wake, H., Bessis, A., and Nimmerjahn, A. (2011). The Role of Microglia in the Healthy Brain. J. Neurosci. *31*, 16064 LP – 16069.

Tuttle, K.D., Waugh, K.A., Araya, P., Minter, R., Orlicky, D.J., Ludwig, M., Andrysik, Z., Burchill, M.A., Tamburini, B.A.J., Sempeck, C., et al. (2020). JAK1 Inhibition Blocks Lethal Immune Hypersensitivity in a Mouse Model of Down Syndrome. Cell Rep. *33*, 108407.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. *11*, 2301–2319.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P. V, et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat. Neurosci. *18*, 145–153.

Vacano, G.N., Gibson, D.S., Turjoman, A.A., Gawryluk, J.W., Geiger, J.D., Duncan, M., and Patterson, D. (2018). Proteomic analysis of six- and twelve-month hippocampus and cerebellum in a murine Down syndrome model. Neurobiol. Aging *63*, 96–109.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. *3*, research0034.1.

Vargas, D.L., Nascimbene, C., Krishnan, C., Zimmerman, A.W., and Pardo, C.A. (2005). Neuroglial activation and neuroinflammation in the brain of patients with autism. Ann. Neurol. *57*, 67–81.

Verkhratsky, A., Rodríguez, J.J., and Parpura, V. (2013). Astroglia in neurological diseases. Future Neurol. *8*, 149–158.

Vezzani, A., Balosso, S., and Ravizza, T. (2019). Neuroinflammatory pathways as treatment targets and biomarkers in epilepsy. Nat. Rev. Neurol. *15*, 459–472.

Vicari, S., Bellucci, S., and Carlesimo, G.A. (2000). Implicit and explicit memory: a functional dissociation in persons with Down syndrome. Neuropsychologia *38*, 240–251.

Vuong, C.K., Black, D.L., and Zheng, S. (2016). The neurogenetics of alternative splicing. Nat. Rev. Neurosci. *17*, 265–281.

Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.-H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. *15*, e8503–e8503.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wang, J., Ma, Z., Carr, S.A., Mertins, P., Zhang, H., Zhang, Z., Chan, D.W., Ellis, M.J.C., Townsend, R.R., Smith, R.D., et al. (2017). Proteome Profiling Outperforms Transcriptome Profiling for Coexpression Based Gene Function Prediction. Mol. &Amp; Cell. Proteomics *16*, 121 LP – 134.

Wang, W., Kwon, E.J., and Tsai, L.-H. (2012). MicroRNAs in learning, memory, and neurological diseases. Learn. Mem. *19*, 359–368.

Waugh, K.A., Araya, P., Pandey, A., Jordan, K.R., Smith, K.P., Granrath, R.E., Khanal, S., Butcher, E.T., Estrada, B.E., Rachubinski, A.L., et al. (2019). Mass Cytometry Reveals Global Immune Remodeling with Multi-lineage Hypersensitivity to Type I Interferon in Down Syndrome. Cell Rep. *29*, 1893-1908.e4.

Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., et al. (2014). HITS-CLIP and Integrative Modeling Define the Rbfox Splicing-Regulatory Network Linked to Brain Development and Autism. Cell Rep. *6*, 1139–1152.

Wierzba-Bobrowicz, T., Lewandowska, E., Schmidt-Sidor, B., and Gwiazda, E. (1999). The comparison of microglia maturation in CNS of normal human fetuses and fetuses with Down's syndrome. Folia Neuropathol. *37*, 227—234.

Wilcock, D.M., and Griffin, W.S.T. (2013). Down's syndrome, neuroinflammation, and Alzheimer neuropathogenesis. J. Neuroinflammation *10*, 864.

Wilcock, D.M., Hurban, J., Helman, A.M., Sudduth, T.L., McCarty, K.L., Beckett, T.L., Ferrell, J.C., Murphy, M.P., Abner, E.L., Schmitt, F.A., et al. (2015). Down syndrome individuals with Alzheimer's disease have a distinct neuroinflammatory phenotype compared to sporadic Alzheimer's disease. Neurobiol. Aging *36*, 2468–2474.

Wu, J.Q., Wang, X., Beveridge, N.J., Tooney, P.A., Scott, R.J., Carr, V.J., and Cairns, M.J. (2012). Transcriptome Sequencing Revealed Significant Alteration of Cortical Promoter Usage and Splicing in Schizophrenia. PLoS One *7*, e36351.

Xu, Y., Li, W., Liu, X., Chen, H., Tan, K., Chen, Y., Tu, Z., and Dai, Y. (2013). Identification of dysregulated microRNAs in lymphocytes from children with Down syndrome. Gene *530*, 278–286.

Yang, E.-W., Bahn, J.H., Hsiao, E.Y.-H., Tan, B.X., Sun, Y., Fu, T., Zhou, B., Van Nostrand, E.L., Pratt, G.A., Freese, P., et al. (2019). Allele-specific binding of RNA-binding proteins reveals functional genetic variants in the RNA. Nat. Commun. *10*, 1338.

Yano, M., Hayakawa-Yano, Y., Mele, A., and Darnell, R.B. (2010). Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. Neuron *66*, 848–858.

Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., et al. (2020). Ensembl 2020. Nucleic Acids Res. *48*, D682–D688.

Ye, Y., Xu, H., Su, X., and He, X. (2016). Role of MicroRNA in Governing Synaptic Plasticity. Neural Plast. *2016*, 4959523.

Yeo, G., Holste, D., Kreiman, G., and Burge, C.B. (2004). Variation in alternative splicing across human tissues. Genome Biol. *5*, R74.

Yin, T., Cook, D., and Lawrence, M. (2012). ggbio: an R package for extending the grammar of graphics for genomic data. Genome Biol. *13*, R77.

Yoshiyama, Y., Higuchi, M., Zhang, B., Huang, S.-M., Iwata, N., Saido, T.C., Maeda, J., Suhara, T., Trojanowski, J.Q., and Lee, V.M.-Y. (2007). Synapse Loss and Microglial Activation Precede Tangles in a P301S Tauopathy Mouse Model. Neuron *53*, 337–351.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omi. A J. Integr. Biol. *16*, 284–287.

Yu, Y.E., Xing, Z., Do, C., Pao, A., Lee, E.J., Krinsky-McHale, S., Silverman, W., Schupf, N., and Tycko, B. (2020). Chapter 1 - Genetic and epigenetic pathways in Down syndrome: Insights to the brain and immune system from humans and mouse models. In Preclinical Research in Down Syndrome: Insights for Pathophysiology and Treatments, M.B.T.-P. in B.R. Dierssen, ed. (Elsevier), pp. 1–28.

Zbucka-Kretowska, M., Niemira, M., Paczkowska-Abdulsalam, M., Bielska, A., Szalkowska, A., Parfieniuk, E., Ciborowski, M., Wolczynski, S., and Kretowski, A. (2019). Prenatal circulating microRNA signatures of foetal Down syndrome. Sci. Rep. *9*, 2394.

Zdaniuk, G., Wierzba-Bobrowicz, T., Szpak, G.M., and Stępień, T. (2011). Original article Astroglia disturbances during development of the central nervous system in fetuses with Down's syndrome. Folia Neuropathol. *49*, 109–114.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science (80-.). *347*, 1138 LP – 1142.

Zhang, B., and Horvath, S. A General Framework for Weighted Gene Co-Expression Network Analysis. Stat. Appl. Genet. Mol. Biol. *4*.

Zhang, M., Ergin, V., Lin, L., Stork, C., Chen, L., and Zheng, S. (2019). Axonogenesis Is Coordinated by Neuron-Specific Alternative Splicing Programming and Splicing Regulator

PTBP2. Neuron *101*, 690-706.e10.

Zhang, S., Chen, J., Zhang, J., and Xu, J. (2017). miR-181a involves in the hippocampus-dependent memory formation via targeting PRKAA1. Sci. Rep. *7*, 8480.

Zhang, X., Chen, M.H., Wu, X., Kodani, A., Fan, J., Doan, R., Ozawa, M., Ma, J., Yoshida, N., Reiter, J.F., et al. (2016). Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. Cell *166*, 1147-1162.e15.

Zhao, C., Tynan, J., Ehrich, M., Hannum, G., McCullough, R., Saldivar, J.-S., Oeth, P., van den Boom, D., and Deciu, C. (2015). Detection of Fetal Subchromosomal Abnormalities by Sequencing Circulating Cell-Free DNA from Maternal Plasma. Clin. Chem. *61*, 608–616.

Zheng, S. (2016). Alternative splicing and nonsense-mediated mRNA decay enforce neural specific gene expression. Int. J. Dev. Neurosci. *55*, 102–108.

Zheng, S. (2020). Alternative splicing programming of axon formation. WIREs RNA *11*, e1585.

Zheng, S., and Black, D.L. (2013). Alternative pre-mRNA splicing in neurons: growing up and extending its reach. Trends Genet. *29*, 442–448.

Zheng, S., Gray, E.E., Chawla, G., Porse, B.T., O'Dell, T.J., and Black, D.L. (2012). PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. Nat. Neurosci. *15*, 381–388.

Zhu, P.J., Khatiwada, S., Cui, Y., Reineke, L.C., Dooling, S.W., Kim, J.J., Li, W., Walter, P., and Costa-Mattioli, M. (2019). Activation of the ISR mediates the behavioral and neurophysiological abnormalities in Down syndrome. Science (80-.). *366*, 843 LP – 849.

Ziats, M.N., and Rennert, O.M. (2014). Identification of differentially expressed microRNAs across the developing human brain. Mol. Psychiatry *19*, 848–852.

Zuccaro, E., Bergami, M., Vignoli, B., Bony, G., Pierchala, B.A., Santi, S., Cancedda, L., and Canossa, M. (2014). Polarized Expression of $p75^{NTR}$ Specifies Axons during Development and Adult Neurogenesis. Cell Rep. *7*, 138–152.