PhD degree in Systems Medicine

Curriculum in Human Genetics

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

Settore disciplinare: BIO/12

The genetic overlap between neuropsychiatric disorders: a meta-analysis of next generation sequencing data

Iman Sadeghi Dehcheshmeh

Ceinge, Naples

Matricola: R11744

Supervisor: Prof. Lucio Pastore, Ceinge, Napoli

Added Supervisor: Prof. Valeria D'Argenio, Ceinge, Napoli

Internal Supervisor: Prof. Francesco Salvatore, Ceinge, Napoli

External Supervisor: Prof. Roderic Guigo', CRG, Barcellona

Anno accademico 2019-2020

This Predoctoral Research has been performed with the support from the following scientific centers:

- School of European Molecular Medicine (SEMM)

- University of Milan
- University of Naples Federico II
- CEINGE - Biotecnologie Avanzate, Naples, Italy

and:

Centre de regulacio genomica (Centre for genomic regulation), Barcelona, Spain.

"The world is meaningless and it's residents are nothing, you're nothing do not get engaged in pointless things

Do you know what is laid before you after life, love and affection and the rest is nothing"

*- **Rumi***

Acknowledgements

I would like to thank first Dr. Natalia Vilor-Tejedor who has openly accepted to directly supervise my project. Natalia, as an astonishing person, both scientifically and personally, deeply tried to help, guide and teach me through my PhD path. I learned from her how to approach a scientific issue, collaborate with others, and be positive about what I am doing. Thanks a lot for supporting, cheering, and giving me confidence to accomplish the projects. Without you I wouldn't be able to finish anything during last years. Muchas gracias Natalia!!!

I thank Roderic Guigo, my excellent PI. Roderic, I learned from you how to be on the edge of science with a brilliant scientific record and, on the other hand, be humble and friend to the people you work with. The way you look at the issues always surprises me. I need to thank you to accepting me in your lab, trusting me, and letting me learn from you and your people. We always say you are a good boss and at the same time friendly. Gracies Roderic!

I am grateful to all the people in the Roderic's lab, at CRG, Barcelona, who have helped me from the beginning of my project when I joined the lab. The amazing people including Emilio Palumbo, Manuel Munoz, Valentin Wucher, and the rest who have patiently tried to help and teach me.

I also thank Valeria D'Argenio from the lab in CEINGE, Napoli, Italy. Vale, you Are amazing. You are a person with a kind heart, patient and flexible who has been always willing to help me. You introduced me to the lab and provided a nice and friendly environment so I could learn from the people in the lab. Apart from this, you tried to help me with setting up in Naples, Italy so that as a foreigner I couldn't feel a stranger. Grazie mille Vale!

I need to also thank all the people in the D'Argenio lab, CEINGE, Naples, Italy. You have always tried to provie me a friendly environment and accepted me as one of them. You are warm, friendly, helpful, and trustworthy people who taught me how to enjoy my life. I appreciate whenever I was unhappy you tried to cheer me up. Grazie ragazzi, Maria, Marcella, Maria Valeria, Flavio and Serena.

This work is dedicated to two brothers who despite suffering most of their life due to harsh conditions, they never gave up fighting back to achieve the bests they want. -- I & A --

Summary

Neurodegenerative and neuropsychiatric disorders (NDD-NPDs) are multifactorial, polygenic and complex behavioral phenotypes caused by brain abnormalities. Most genetic studies have focused on understanding the genetic component of specific brain diseases. Several brain diseases also show similar clinical and pathological symptoms. In recen years, multiple studies have used next generation sequencing (NGS) technologies such as RNA sequencing (RNA-Seq) to investigate molecular signature of brain diseases. However, many studies have only focused on a particular disease and limited brain regions. By using the data from a broad range of cortical regions from multiple brain diseases, we will be able to dig deeper into the molecular basis of neurological diseases.

The main aim of this thesis was to examine the transcriptome-wide characterization of cortical brain regions across neurological disorders. We focused our research efforts on highlighting cross-disease shared molecular signatures, and exploring co-expression networks and cell-type-specific patterns for NDD-NPDs. By processing and analyzing RNA-Seq data using a set of computational tools and statistical tests, we performed transcriptomic profiling of brain samples from eight groups of patients with Alzheimer's disease (AD), Parkinson's disease (PD), Progressive Supranuclear Palsy (PSP), Pathological Aging (PA), Autism Spectrum Disorder (ASD), Schizophrenia (SZ), Major Depressive Disorder (MDD), and Bipolar Disorder (BP)-in comparison with 2,078 brain samples from matched control subjects.

In this thesis, we provide a transcriptomic framework to understand the molecular architecture of NPDs and NDDs through their shared- and specific gene expression in the brain.

Preface

Neuroscience, as the field of studying central nervous system and brain functions has been developing in recent years. A major focus of neuroscience is understanding the mechanisms underlying the pathobiology of brain diseases. The mechanism of many brain diseases is very complex involving both genetic and environmental factors. With the advent of recent advance technologies such as Next Generation Sequencing (NGS), many studies have tried to exploit them to explore the molecular basis of brain diseases. We hypothesized that there are potential overlapping molecular signatures across major brain diseases including NDDs and NPDs.

In this work, we contributed to 1) understanding the specific genes and mechanisms involved in the pathology of neurodegenerative diseases including Alzheimer's disease, Parkinson's disease, Progressive? Supranuclear Palsy and pre-clinical Alzheimer's disease, as well as psychiatric disorders including Schizophrenia, Autism Spectrum disorder, Major Depression disorder and Bipolar disorder, 2) gaining knowledge about shared molecular signatures across neurodegenerative and psychiatric disorders, 3) understanding the role of cortical regions in the pathobiology of diseaess, 4) unraveling cell-type-specific transcriptional changes across diseases, 5) understanding the relationship of the brain diseases at the molecular level, 6) offering suggestions and recommendations for the future research on the molecular basis of these diseases. Our research showed that, at the molecular level, some neurodegenerative diseases and psychiatric disorders have similarities. Moreover, the results presented in this work revealed that some cortical regions are involved in some specific diseases while others are more broadly changed across diseases. These findings could pave the way for future studies to understand the basis of these complex diseases and find a more efficient treatment.

The PhD candidate, apart from this main work, has mainly contributed to 6 scientific papers and one scientific congress that are listed at in the Annex Section.

Table of Contents

Table of Figures

List of Tables

Chapter1. Introduction

The human brain is the central organ of the nervous system and makes up the central nervous system (CNS) along with the spinal cord. The brain monitors major functions throughout the body by processing, integrating, and coordinating the information sent from the organs, and making decisions as to the instructions delivered to the organs (Mesulam and -marsel Mesulam, 2002) (Figure 1-1). The brain has three major parts including cerebrum, cerebellum, and brainstem, of these the cerebrum is the largest part, which can be divided into two symmetrical hemispheres. The outer layer of grey matter, covering the core of the white matter is called the cerebral cortex. The cortex is divided into the neocortex and the allocortex (Mai, Majtanik and Paxinos, 2015). The neocortex consists of six neuronal layers and the allocortex has nearly four. Each one of the hemispheres splits into four major lobes including the frontal, temporal, parietal, and occipital lobes (Marshall and Morriss-Kay, 2004).

The cerebrum is connected by the brainstem to the spinal cord (Ghosh, 2007). The cerebral cortex is positioned on the top of multiple structures, such as the thalamus, the epithalamus, the

pineal gland, the hypothalamus, the pituitary gland, and the subthalamus; the limbic structures, such as the amygdala and the hippocampus; the claustrum, the various nuclei of the basal ganglia; the basal forebrain structures, and the three circumventricular organs (Ghosh, 2007). Here, we are going to describe briefly the anatomy and function of major brain regions. These bigger regions or lobes are usually divided to subregions that are attributed to different functions of the CNS.

1-1- Major brain lobes and their function

1-1-1- **Frontal lobe**

The frontal lobe is positioned over and in front of the temporal lobe, at the front-head of the cerebral hemispheres and located in front of the parietal lobe (Figure 1-1). The precentral gyrus, shaping the posterior margin of the frontal lobe, harbors the primary motor cortex, which monitors voluntary movements of specific body organs (Stuss and Knight, 2013). The frontal lobe comprises most of the dopaminergic neurons in the cerebral cortex. The dopaminergic neurons are related to reward, attention, focus, working memory tasks, planning, and motivation. The frontal lobe contains the prefrontal cortex (PFC) positioned at the most anterior edge of the lobe (Stuss, 2011). This part of the lobe is vital in working memory and executive monitoring by facilitating goal maintenance and complex tasks organization (de Souza *et al.*, 2014).

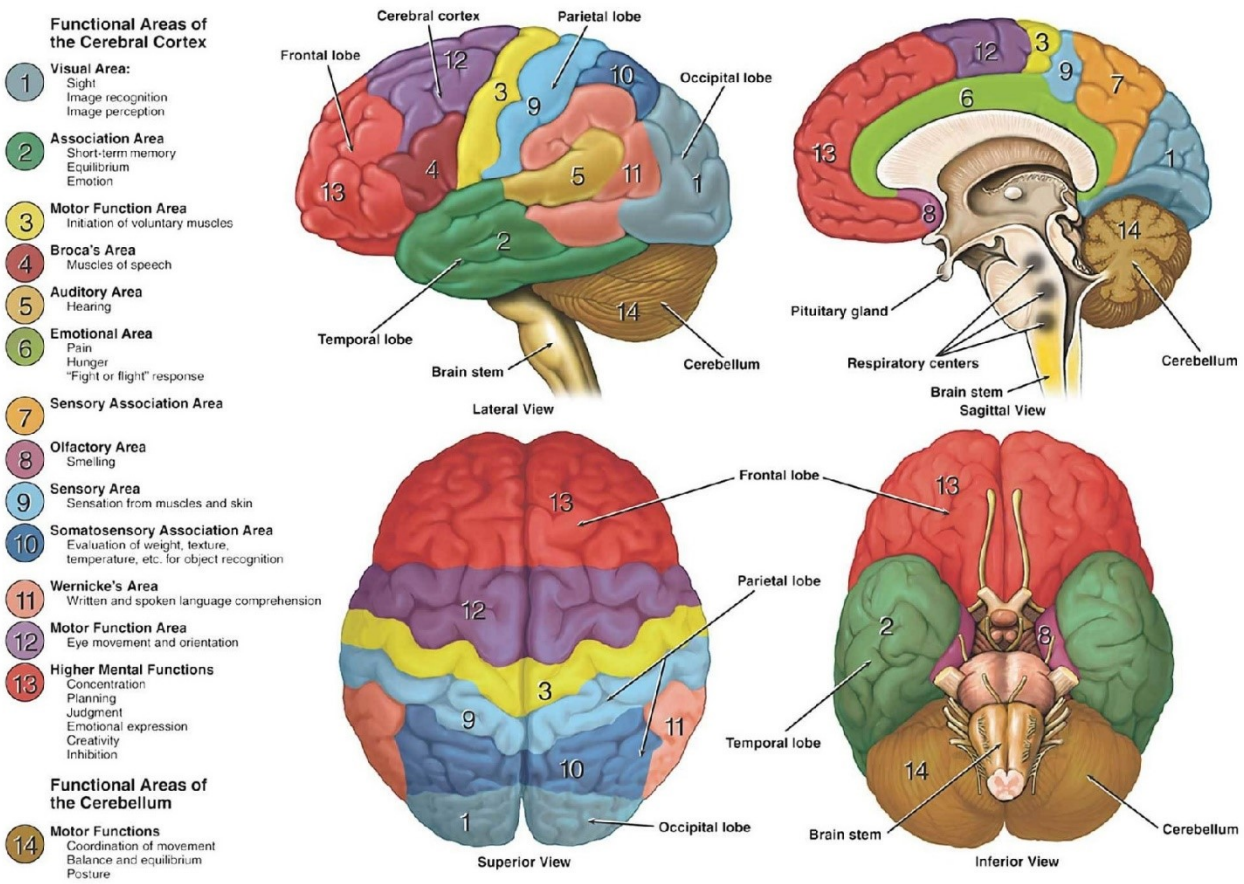

Figure 1-1 | A schematic presentation of brain regions and related functions. (Image adopted from Dana foundation URL)

The PFC splits into orbital, medial, and lateral PFC. The lateral PFC has two subregions including the dorsolateral and ventrolateral PFC. The dorsolateral PFC (dlPFC) is above the ventrolateral PFC (vlPFC) and is primarily involved in the executive control and guidance of memories that are regained from episodic memory (Fuster, 2015). The vlPFC is critical for the organization of meaningful stimuli that one experiences during lifetime, including names, images, and letters (Badre and Wagner, 2007). PFC impairment can lead to long-term and short-term memory damages, and changes in individuals behaviors and their planning and organization abilities (Mah, Arnold and Grafman, 2004).Many studies have shown that the impairment of PFC is associated with multiple diseases such as Alzheimer's (Sampath et al. 2017; Bakkour et al. 2013), Huntington's (Narayanan et al. 2014), Parkinson's (Narayanan et al. 2013; Schmidt 2005), depression (Hare and Duman 2020), and schizophrenia (Wible et al. 2001).

1-1-2- **Temporal lobe**

The temporal lobe is below the lateral fissure on both sides of the cerebral hemispheres of the human brain (Ramachandran, 2002). It is implicated in processing sensory inputs into derived meanings for the appropriate recall of visual memories, language comprehension, and linking emotions (Squire, Stark and Clark, 2004).

There is a region inside the temporal lobe called the hippocampus which is involved in memory formation and learning (Johnston and Amaral, 2004). Temporal lobe includes areas associated with the auditory, olfactory, vestibular, and visual senses, and in the perception of spoken and written language (Dawes *et al.*, 2009). In addition to cortex, the temporal lobe contains white matter, part of the lateral ventricle, the tail of the caudate nucleus, the stria

terminalis, the hippocampal formation, and the amygdala. The medial side of the temporal lobe contains areas involved in olfaction (the uncus and nearby cortex) and semantic memory (the hippocampal formation) (Squire, Stark and Clark, 2004; Dawes *et al.*, 2009). The nearby amygdala creates responses to perceived sensory stimuli that have been processed by other parts of the brain. These responses include largely involuntary ones, mediated by the autonomic and somatic motor systems, and mental functions, particularly those called feelings or emotions, that motivate decision and voluntary actions (Squire, Stark and Clark, 2004; Rajmohan and Mohandas, 2007).

Atrophy of the temporal lobe has been linked to frontotemporal dementia and Alzheimer's disease (Wible et al. 2001; Jack et al. 1998; Wolk et al. 2017; Nativio et al. 2018). Dysfunction of temporal lobe has also been associated with psychiatric disorders such as schizophrenia (Bobilev et al. 2020), depression (Galioto et al. 2017), and autism (Galioto et al. 2017; Valvo et al. 2016).

1-1-3- **Parietal lobe**

The parietal lobe is positioned on the top of the occipital lobe and behind the frontal lobe and central sulcus (Behrmann, Geng and Shomstein, 2004). This region combines sensory information among various modalities, including spatial sense and navigation, the main sensory receptive area for the sense of touch (mechanoreception) in the somatosensory cortex which is just posterior to the central sulcus in the postcentral gyrus, and the dorsal stream of the visual system (Behrmann, Geng and Shomstein, 2004; Culham and Valyear, 2006). The majority of sensory inputs from the skin (touch, temperature, and pain receptors) are amplified via the

thalamus to the parietal lobe. The parietal lobe is also involved in language processing. The superior and inferior parts of the parietal lobe are the regions for spatial awareness (Geranmayeh *et al.*, 2012).

Impairments in these regions are usually associated with hemineglect (Pouget and Sejnowski, 1997). Also, left parietal lobe damage causes limbic apraxia Limb apraxia that is a heterogeneous disorder of skilled action and tool use that has long perplexed clinicians and researchers. It occurs after damage to various loci in a densely interconnected network of regions in the left temporal, parietal, and frontal lobes (Buxbaum and Randerath 2018).

1-1-4- **Occipital lobe**

The occipital lobe is the center of the visual processing of the brain comprising the major part of the visual cortex (Larsson and Heeger, 2006). The primary visual cortex is Brodmann area 17 (BA17), which is also known as the visual one (V1). This region is positioned in the center, within the calcarine sulcus; the full extent of V1 often extends to the posterior pole of the occipital lobe. V1 is also known as the striate cortex due to the presence of a large stripe of myelin, the Stria of Gennari (Tong, 2003). The regions outside V1 are called the extrastriate cortex, functional for various visual tasks, such as visuospatial processing, color discrimination, and motion perception (Tootell and Nasr, 2017).

Damage to one side of the occipital lobe causes homonymous loss of vision with exactly the same field cut in both eyes. Disorders of the occipital lobe can cause visual hallucinations and illusions. Visual hallucinations (visual images with no external stimuli) can be caused by lesions to the occipital region or temporal lobe seizures (Tokida et al. 2018). Visual illusions (distorted

perceptions) can take the form of objects appearing larger or smaller than they actually are, objects lacking color or objects having abnormal coloring. Lesions in the parietal-temporal-occipital association area can cause word blindness with writing impairments (alexia and agraphia) (Liu et al. 2019).

1-1-5- **Limbic lobe**

The limbic lobe located at the inferomedial side of the cerebral hemispheres, contains two concentric gyri surrounding the corpus callosum (Rajmohan and Mohandas, 2007). The larger outer gyrus is named limbic gyrus and the smaller inner one the intralimbic gyrus. The limbic gyrus (limbic lobe) consists of the the paraterminal gyrus, the subcallosal area, the cingulate gyrus, the parahippocampal gyrus, the dentate gyrus, the hippocampus and the subiculum (and some authors include amygdala) (Hirai *et al.*, 2000).The cingulate gyrus dorsal to the corpus callosum is heavily interconnected with the association areas of the cerebral cortex. The parahippocampal gyrus in the medial temporal lobe contains several distinct regions, the most important being the entorhinal cortex (ERC) (Rajmohan and Mohandas, 2007). The ERC funnels highly processed cortical information to the hippocampal formation and serves as its major output pathway (Tsao *et al.*, 2018). Cingulate gyrus is associated with autonomic functions that regulate heart rate and blood pressure as well as cognitive, attentional and emotional processing. Parahippocampal gyrus is involved in spatial memory, hippocampus in long-term memory, amygdala in anxiety, aggression, fear conditioning; emotional memory and social cognition, hypothalamus in regulation of the autonomic nervous system via hormone production and release, regulation of blood pressure, heart rate, hunger, thirst, sexual arousal and the circadian

rhythm, and nucleus accumbens in reward and addiction (Karanian and Slotnick, 2017; Burke *et al.*, 2018).

Dysfunction of the limbic system is usually associated with many clinical manifestations, such as epilepsy, limbic encephalitis, dementia, anxiety disorder, schizophrenia, and autism. Degenerative changes in the limbic system likely have a role in the genesis of neurodegenerative diseases, particularly Pick's disease and Alzheimer's disease (Sakurai et al. 2019; Rajmohan and Mohandas 2007). Marked atrophy is found in the limbic system, most notably the dentate gyrus and hippocampus. In Alzheimer's disease, senile plaques and neurofibrillary tangles are dispersed throughout the cerebral cortex and basal ganglia, but the hippocampus and amygdala are often severely involved (Sakurai et al. 2019). It has been also shown that atrophy of limbic lobe is also associated with Parkinson’s disease (Gao et al. 2017).

Anxiety disorders may be the result of a failure of the anterior cingulate and hippocampus to modulate the activity of the amygdala (top-down regulation). A fear circuitry, involving the amygdala, prefrontal and anterior cingulate has been described (bottoms-up regulation) (Cannistraro and Rauch 2003).

Studies have shown reduced limbic volumes in schizophrenia. The evidence for this is the distortion of cortical neuronal organizations, decreased size of hippocampus and the reduced number of GABAergic cells in the cingulate and anterior thalamus with resultant glutamatergic excitotoxicity (Rajmohan and Mohandas 2007).

Studies have also shown variation in the volumes of the amygdala and hippocampus in affective disorders. Functional studies have revealed decreased prefrontal and anterior cingulate activity in affective disorders. The anterior cingulate is the center for integration of attentional

and emotional output and helps effortful control of emotional arousal (Stathis et al.; Kaschka 2002).

Autism and Asperger's syndrome involve the disproportionate impairment in specific aspects of social cognition. Limbic structures involved include the cingulate gyrus and amygdala, which mediate cognitive and affective processing. The basolateral circuit integral for social cognition is disrupted in autism spectrum disorders (Bachevalier and Loveland 2006; Patriquin et al. 2016).

1-1-6- **Insular cortex**

In primates, including humans, the insular cortex is located folded deep within the lateral sulcus of each hemisphere, hidden beneath parts of the frontal, parietal and temporal lobes (Gogolla, 2017). The human insular lobe splits into an anterior and a posterior part by the central insular sulcus. The extremities of these two parts differ substantially in their connectivity to other brain regions, while an intermediate middle insular zone exhibits mixed anterior and posterior connectivity features (Cunha-Cabral *et al.*, 2019). Interoception, the perception of bodily states, is a key function of the insular cortex. The insula receives topographically arranged afferents from discrete thalamic relay nuclei and combines information about blood pressure and oxygenation, the motility of the digestive system, the timing and strength of the heartbeat, as well as pain, hunger, nausea, tickle, itch and many more bodily sensations (Davidovic, Starck and Olausson, 2019). Furthermore, the insula also executes strong top-down monitoring of autonomic functions, for example regulation of the heartbeat rate, blood pressure, or gastric motility, probably via direct projections to the lateral hypothalamic area, the parabrachial nucleus, and the nucleus of the solitary tract (Hassanpour *et al.*, 2018). Hypometabolism and atrophy of the left

anterior insular cortex has been associated with progressive expressive aphasia, a condition of impairment of normal language function that leads individuals to lose the ability to communicate fluently while still being able to comprehend single words and intact other non-linguistic cognition (Namkung, Kim and Sawa, 2017; Balint *et al.*, 2019). This condition is present in some cases of neurodegenerative conditions such as Pick's disease, motor neuron disease, corticobasal degeneration, frontotemporal dementia, and Alzheimer's disease (Namkung, Kim and Sawa, 2017; Fathy *et al.*, 2019).

The insular cortex is suggested to be widely involved in mood disorder, panic disorder, post-traumatic stress disorder (PTSD), obsessive-compulsive disorder, eating disorders, and schizophrenia (Menon and Uddin 2010). Functional neuroimaging studies have discovered, in patients with schizophrenia, a reduction in the strength of the causal influences from the insula-centric salience network on the central executive network and the default mode network (Moran et al. 2013). The insula-mediated dynamic switching between the central executive network and the default mode network facilitates access to cognitive resources, such as attention and working memory, when a salient event is detected [20]. Thus, altered strength in the connectivity in these networks affects cognitive deficits in some cases of schizophrenia. This network deficit is also reported in some patients with autism spectrum disorder (Uddin et al. 2015), which may be consistent with the shared genetic and biological risks between autistic spectrum disorder and some cases of schizophrenia (Mihaylova et al. 2017). Blood marker examinations and brain-volume assessments have also shown a relationship between damage of limbic cortex in people with bipolar disorder (Bond et al. 2020; Delvecchio et al. 2019).

1-1-7- **Basal ganglia**

The basal ganglia (or basal nuclei) are a group of subcortical nuclei in the human brain, which are located at the root of the forebrain and top of the midbrain (Bostan, Dum and Strick, 2018). This region comprising the striatum, the globus pallidus, the substantia nigra, and the subthalamic nucleus, is engaged in various functions including eye movement, motivation, working memory, and decision making (Bogousslavsky and Tatu, 2017). The striatum, consisted of caudate, putamen, and ventral striatum (VS), is the main input of the basal ganglia, and the thalamus connects it to the cortex (Deffains *et al.*, 2016). The dopaminergic projections from the midbrain to the striatum are a crucial modulator of striatal processing of glutamatergic cortical and thalamic signals on the striatum principal neurons (Surmeier, Plotkin and Shen, 2009). A reduction of the dopaminergic projections to the striatum is strongly associated with Parkinson's disease (PD), characterized by motor and cognitive dysfunctions (Milosevic *et al.*, 2018). The basal ganglia mainly controlled by dopaminergic neurons, appears to play a role in the mediation of cognitive and motor modules to make an appropriate decision on a final action for the task being performed (Wessel *et al.*, 2016).

Dysfunction and damage of basal ganglia is associated with multiple diseases including both neurodegenerative and psychiatric disorders. Disruption of the basal ganglia network forms the basis for several movement disorders such as Parkinson's disease (Milosevic et al. 2018), Huntington disease (Singh-Bains et al. 2019). Although motor diseases are the most common associated with the basal ganglia, evidence show that basal ganglia dysfunctions can lead to other disorders such as obsessive–compulsive disorder (OCD) (Liu et al. 2020) and Tourette syndrome and Tic disorder (Caligiore et al. 2017).

1-1-8- **Brain cell types**

The brain is a mosaic made up of different cell types including neurons, glial cells, neural stem cells, and blood vessels. Neurons are divided into interneurons, pyramidal cells including Betz cells, motor neurons, and Purkinje cells (found in the cerebellum) (Tao and Zhang, 2016). The adult human brain approximately contains 90 billion neurons, with more or less the same number of non-neuronal cells. Almost 20% of the neurons are found in the cerebral cortex, and the rest of the neurons (80%) are located in the cerebellum (Chizhikov and Millen, 2020).

Glial cells are divided into astrocytes (including Bergmann glia), oligodendrocytes, ependymal cells (including tanycytes), radial glial cells, microglia, and a subtype of oligodendrocyte progenitor cells (Jäkel and Dimou, 2017). Astrocytes are star-shaped cells that surround neurons and support neuron functions. Astrocytes also help neurons signal to other neurons by passing chemicals from one neuron to another. Although microglia are the primary immune cells of the brain, astrocytes can also help microglia when the brain is in trouble (Jäkel and Dimou, 2017; Vasile, Dossi and Rouach, 2017). Oligodendrocyte which is a type of glial cells covers the axons of neurons, making up the myelin sheath. They separate the axon from surrounding and help neurons pass electrical signals at high speed over long distances (Bradl and Lassmann, 2010). Microglia, another type of glial cells, are the immune cells of the human brain. They circulate inside the brain and constantly communicate with other glial cells, to test the environment for any problem (Gremo *et al.*, 1997; Menassa and Gomez-Nicola, 2018). Mast cells are also white blood cells that communicate with the neuroimmune system in the brain (da Silva, Jamur and Oliver, 2014; Krystel-Whittemore, Dileepan and Wood, 2015) and mediate neuroimmune responses in inflammatory situations and help to keep the blood-brain barrier (Krystel-Whittemore, Dileepan and Wood, 2015).

1-2- Techniques for assessing transcriptome profiling of brain tissues

1-2-1- **Microarrays**

The Microarray technique has emerged as a high-throughput technology for large-scale gene expression quantifications. It principally includes a large number of hybridization reactions between labeled target cDNA and probe sequence fixed to a solid surface (Schulze and Downward, 2001). Currently, microarrays technology is able to analyze genome-wide transcriptome analysis, upto 30,000 different mRNA types (Li, 2016). Microarray has gained new features such as probing exon junctions to detect alternative splicing patterns (Clark, Sugnet and Ares, 2002). Before the advent of next generation sequencing (NGS) technologies, microarray has been widely harnessed for transcriptome profiling of brain tissues. Many research groups utilized it to analyze gene expressions in multiple brain diseases such as AD (Z. Wang *et al.*, 2018), PD (Feng and Wang, 2017), ASD (Tammimies *et al.*, 2015), Scz (Mirnics *et al.*, 2000; Tammimies *et al.*, 2015), BP (Logotheti *et al.*, 2013), and MDD (Urbach, Bruehl and Witte, 2006). However, microarray application is currently limited because of its use of indirect signal detection by hybridization. The limitations include (i) heavily dependent on known genome sequence data; (ii) using non-accurate hybridization reactions, and (iii) having a low range of signal detection due to background signals and the saturation of signals (Forster, Roy and Ghazal, 2003). These issues usually lead scientists to think of other more accurate high-throughput technologies such as RNA-Seq.

1-2-2- **RNA-Seq**

RNA-Seq is a high-throughput technology that emerged as an NGS platform for transcriptome profiling. It can be run on different platforms such as the Roche 454 Life Science Genome Sequencer FLX System, the Illumina Genome Analyzer and the Applied Biosystems SOLiD technology (Janitz, 2011). In RNA-Seq, RNAs are, in principle, reverse transcribed into cDNA, and adapters are added to both ends of the generated cDNA. Then, these sequences are individually sequenced either unidirectionally (single-end seq.) or bidirectionally (paired-end seq.). The raw sequencing data are then aligned to a reference genome or assembled de novo to generate a resulting map of the transcriptome profile (Schuster, 2008). RNA-Seq has become very popular because of its advantages over other methods including microarray. These advantages include being able to identify novel transcripts (Vera *et al.*, 2008), provide information about transcript structure, identifying genetic variations such as SNPs (Wang, Gerstein and Snyder, 2009), high sensibility of detecting molecules, and avoiding technical problems such as saturation and background noise signals (Mortazavi *et al.*, 2008).

The large amount of the data produced from RNA-Seq experiments requires vast computational resources and skills for analyses. A large number of bioinformatic tools such as ELAND, SOAP, MAQ, RMAP, Bowtie, TopHat, STAR and Cufflinks (Li *et al.*, 2008; Li, Ruan and Durbin, 2008; Smith, Xuan and Zhang, 2008) have emerged to overcome the obstacles such as alignment of short-reads to the reference genome. Some of these tools that are even able to identify the splice variants of different transcripts include STAR, Bowtie and TopHat (Trapnell, Pachter and Salzberg, 2009). In addition, many automated pipelines that combine data analysis tools are rising in order to facilitate the use and monitoring of RNA-Seq data processing, enabling researchers to analyze large-scale data simultaneously in a short time.

1-3- Brain diseases

We mentioned that the brain structure is divided into different regions and each region has several functions. Any abnormality or dysfunction of the brain regions can lead to one or several forms of the diseases which can debilitate one's living. Brain diseases are usually categorized differently based on the symptoms, clinical characteristics, genetic causes, and types of damages. Here, we will describe two common classes of brain illnesses including neurodegenerative diseases and neuropsychiatric disorders.

1-3-1- Neurodegenerative diseases

Neurodegenerative diseases (NDD), which are chronic and progressive illnesses, are identified by selective and symmetric loss of neurons in motor, sensory, and/or cognitive systems (Brittany N. Dugger, 2017). Description of the patterns of cell loss and the association of disease-specific cellular markers helped in their classification including senile plaques, neurofibrillary tangles, neuronal loss, and acetylcholine deficiency, Lewy bodies and depletion of dopamine, cellular inclusions, swollen motor axons, and g-aminobutyric acid-containing neurons of the neostriatum (Gan *et al.*, 2018). Here, we describe several common NDDs including Alzheimer's disease, Parkinsons' disease, progressive supranuclear palsy and pathological aging.

1-3-1-1- **Alzheimer's disease**

Alzheimer's disease (AD) is a well-known progressive, unremitting NDD that engages large regions of the cerebral cortex (particularly temporal lobe) and hippocampus (Bondi, Edmonds and Salmon, 2017). AD is characterized by the inclusion of insoluble forms of beta-amyloid (Aβ) plaques in extracellular spaces and in the walls of blood vessels, accumulation of the tau proteins in neurofibrillary tangles in neurons, synaptic loss and selective neuronal death (DeTure and Dickson, 2019). Aβ is generated via cleavage of amyloid precursor protein (APP) by a protein complex including γ-secretase, β-secretase enzymes, presenilin 1 (PS1; encoded by *PSEN1*), and PS2 (encoded by *PSEN2*) (Greenberg *et al.*, 2019). Sporadic AD is the most prevalent type with an average age of onset of 80 years. In this form, a failure to clear Aβ peptide occurs in the CNS. Less than 1% of patients have autosomal dominant inherited Alzheimer's disease (DIAD) with an average age of onset of 45 years (Hou *et al.*, 2019). These patients usually have mutations in the genes *APP*, *PS1*, and *PS2*, leading to production of toxic forms of Aβ (Hou *et al.*, 2019; Panza *et al.*, 2019). AD shares many characteristics with other molecularly defined NDDs, such as Parkinson's disease and the frontotemporal dementias, as well as with other brain disorders such as Down syndrome (Lott and Head, 2019). Currently, Aβ, *APOE*, and tau are the main risk factors frequently reported in association with AD.

Molecular basis of Alzheimer's disease

Aβ in Alzheimer's disease

Aβ is cleaved from APP into a group of N-terminally truncated peptides with different lengths (from 38 to 43 amino acids) (Golde, Eckman and Younkin, 2000). Evidence shows that

Aβ is highly expressed in neurons (and other cell types) as a normal protein from APP in a normal brain (Hou *et al.*, 2019). However, the exact mechanism of APP is unclear, but it is thought to be functional in synaptic plasticity. The first studies on patients with DIAD revealed that mutations in one of *APP*, *PSEN1*, and *PSEN2* genes are related to the involvement of Aβ in AD (Suárez-Calvet *et al.*, 2014). Mutations in these genes cause an increase of Aβ, particularly the 42 amino acid Aβ isoform (Aβ42), which seems to be more toxic than other forms (Gaiteri *et al.*, 2016). Also, mutations in *PSEN1* and *PSEN2* lead to elevated levels of Aβ42 (Kumar-Singh *et al.*, 2006). The type of mutation and level of Aβ42 predict the average age of onset of AD (Bateman *et al.*, 2012).

APOE and Alzheimer's disease

APOE on chromosome 19 is the strongest genetic risk factor for developing AD. APOE is involved in the normal catabolism of triglyceride rich lipoproteins. One of the first reports linking *APOE* to AD pathology was *APOE* immunoreactivity in Aβ deposits and neurofibrillary tangles, which are hallmarks of AD pathology (Namba *et al.*, 1991). In addition, polymorphisms in the transcriptional regulatory region of *APOE* have demonstrated association with AD (Artiga *et al.*, 1998). *APOE* has three isoforms named *APOE2*, *APOE3*, and *APOE4* with a prevalence of 7%, 78%, and 15%, respectively (Saunders *et al.*, 2011). *APOE* is highly expressed in the CNS, mainly in astrocytes and microglia. Based on reports, *APOE4* contributes to nearly 50% of sporadic AD (Ashford and Wesson Ashford, 2004), and raises the risk of early onset of AD. On the other hand, *APOE2* is usually associated with a lower risk of developing AD (Corder *et al.*, 1994).

Transcriptome profiling in Alzheimer's disease

RNA-Seq study on temporal and cerebellum samples from AD patients and controls showed that several SNPs influence the expression level of multiple genes detectable in the brain including *ABCA7*, *BIN1*, *CLU*, *MS4A4A*, *MS4A6A*, and *PICALM* (M. Allen *et al.*, 2012). Mostafavi et al. (Mostafavi *et al.*, 2018) used a cohort of 478 individuals with AD and controls to investigate aging-related molecular networks in the DLPFC. They observed the correlation of a group of genes, including *NPPL1* and *PLXNB1* that influence genes associated with β-amyloid biology in a co-expression module in astrocytes. A large-scale transcriptional analysis in the CA1 and CA3 brain regions of subjects with AD found candidate vulnerability genes including *ABCA1*, *MT1H*, *PDK4*, and *RHOBTB3* and protective genes *FAM13A1*, *LINGO2*, and *UNC13C* in CA3, compared with controls (Miller *et al.*, 2013). Also, using 1,647 postmortem brain tissues from late-onset AD (LOAD) patients and controls, Zhang et al. (Zhang *et al.*, 2013) observed a microglia-specific module that is dominated by genes involved in pathogen phagocytosis, including *TYROBP* as a key regulator. A recent study (Annese *et al.*, 2018) on hippocampus samples from a small group of AD patients and controls observed 2,064 genes, 47 lncRNAs and 4 miRNAs deregulated in the hippocampal region and involved in cognitive functions and long-term memory. we selected from a community-based neuropathological study. Combining brain tissue-specific protein interactomes with 414 expression profiles of symptomatic AD subjects and controls determined functionally distinct composite clusters of genes that reveal extensive changes in expression levels in AD, corresponding to synaptic transmission, metabolism, cell cycle, survival, and immune response (Canchi *et al.*, 2019). It was revealed that, particularly, the loss of *EGR3* regulation influences synaptic deficits by affecting the synaptic vesicle cycle.

1-3-1-2- **Parkinson disease**

Parkinson's disease (PD) is a common age-related NDD that has two forms of sporadic (99%) or familial (1%). PD is clinically identified by neuronal loss in certain brain regions including basal ganglia, substantia nigra and large intracellular aggregation of α-synuclein proteins (Poewe *et al.*, 2017). In early stages, the degeneration of dopaminergic neurons is limited to the ventrolateral substantia nigra, but over time grows wider across other parts of substantia nigra (Fearnley and Lees, 1991; Damier *et al.*, 1999). Multiple lines of evidence show that the onset of motor symptoms occurs after the loss of dopaminergic neurons (Dijkstra *et al.*, 2014). The accumulation of Lewy bodies happens first in cholinergic and monoaminergic brainstem neurons, as well as in neurons of the olfactory system. With the progression of the disease, these inclusion bodies also appear in limbic and neocortical areas.

Molecular basis of Parkinson's disease

Numerous point mutations and duplications of *SNCA*, the gene encoding α-synuclein, in heritable forms of PD firmly support the important role of α-synuclein as a risk factor in PD (Nalls *et al.*, 2014).

In addition, intracellular homeostasis of α-synuclein is maintained by the actions of the ubiquitin–proteasome system and the lysosomal autophagy system (LAS) (Kaushik and Cuervo, 2015). Multiple investigations revealed that perturbation of these degradation systems could lead to α-synuclein aggregation (Xilouri, Brekk and Stefanis, 2013). Increasing age — the greatest risk factor for PD — is associated with reduced LAS and ubiquitin-proteasome system functions, which is consistent with observations of increased levels of α-synuclein in nigral dopaminergic

neurons during normal ageing (Chu and Kordower, 2007). Moreover, a number of mutations in heritable forms of PD are associated with reduced LAS function. These mutations occur in the gene encoding *LRRK2* and in the gene encoding the lysosomal enzyme GBA, the most common genetic risk factor for PD (Sidransky *et al.*, 2009), that are linked to weakened LAS function (Fernandes *et al.*, 2016).

Transciptome profiling in Parkinson's disease

The first microarray studies on patients with PD showed a significant down-regulation of genes involved in mitochondrial function and energy synthesis and maintenance in PD patients (Vogt *et al.*, 2006; Patel *et al.*, 2008). Transcriptome profiling of cerebro-spinal fluid (CSF) from 27 PD patients and 30 healthy controls revealed several differentially expressed sequencing tags (DETs), including *Dnmt1*, *Ezh2*, *CCR3*, *SSTR5*, *PTPRC*, *UBC*, *NDUFV2*, *BMP7*, *SCN9*, *SCN9* antisense (*AC010127*.3), and long noncoding RNAs *AC079630* and *UC001lva*.4 (close to the *LRRK2* gene locus), as potential PD biomarkers (Hossein-nezhad *et al.*, 2016). A study conducted by Henderson et al.(Henderson-Smith *et al.*, 2016) on the posterior cingulate cortex of PD patients, showed that up-regulated genes in PD compared with controls were associated with the immune system, while down-regulated genes were involved in signal transduction and RNA processing. Also, RNA-Seq analysis on Braak Lewy body stages of brain regions from PD donors and controls revealed differential expression patterns of many genes, including *SNCA*, *ZNF184*, *BAP1*, *SH3GL2*, *ELOVL7*, and *SCARB2* that were associated with dopamine synthesis, the motor and immune system, and microglial activity (Keo *et al.*, no date). A recent meta-analysis by Kelly and colleagues using microarray datasets of PD and AD showed down-

regulation of 739 genes in PD. Top dysregulated genes were *YWHAZ*, *SCNA*, *DCLK1*, and *SLC18A2*, associated with mitochondrial dysfunction and oxidative stress (Kelly *et al.*, 2020).

1-3-1-3- **Progressive supranuclear palsy**

Progressive supranuclear palsy (PSP), another type of NDD which has similar characteristics to PD, is clinically characterized by behavioral, language, and movement aberrations (Williams and Lees, 2009). The classic type of PSP which is currently known as Richardson's syndrome (PSP-RS), was first defined in 1964 by Steele, Richardson, and Olszewski (Steele, Richardson and Olszewski, 2014). PSP, with a prevalence of 5 cases per 100000 people, is strongly linked to tau-protein abnormalities, both neuropathologically and genetically (G. Höglinger *et al.*, 2011; Coyle-Gilchrist *et al.*, 2016). PSP probably begins with a presymptomatic phase, in which neuropathological abnormalities begin to accumulate, continues with a suggestive-of-PSP (soPSP) phase, in which individuals show mild symptoms, and ends with a fully symptomatic stage that, eventually, meets the clinical criteria for PSP-RS or another variant of PSP (vPSP) (Respondek *et al.*, 2014). The neuropathological criteria for PSP involve the presence of neurofibrillary tangles, neuropil threads (made of tau protein), neuronal loss, gliosis, tufted astrocytes, and oligodendroglial coiled bodies in the basal ganglia and the brainstem (Respondek *et al.*, 2014). The regional pattern of tau pathology and neuronal loss gives result into pathologically and clinically heterogeneous PSP phenotypes (Williams *et al.*, 2007).

Molecular basis of progressive supranuclear palsy

MAPT polymorphisms and haplotypes, the locus most strongly associated with PSP is the gene encoding the microtubule-associated protein tau (MAPT) (Hinz and Geschwind, 2017). A rare coding *MAPT* variant alters microtubule assembly and is a strong risk factor for both PSP and frontotemporal dementia (Coppola *et al.*, 2012). Also, the first large GWA study (G. U. Höglinger *et al.*, 2011) on PSP individuals recognized new genes associated with PSP including *STX6*, *EIF2AK3*, and *MOBP*. An SNP near the *MOBP* gene correlates more strongly with enhanced expression of the *SLC25A38* gene, which encodes the protein appoptosin which in turn increases caspase-mediated tau cleavage, tau aggregation, synaptic dysfunction, and gait and balance deficits.45 Moreover, loss-of-function of *EIF2AK3* causes a severe childhood-onset disease with cerebral tau pathology named Wolcott-Rallison syndrome (Bruch *et al.*, 2015). Based on these studies, other genetic variations could also affect tau and, eventually, elevate the risk of PSP.

Transcriptome profiling in progressive supranuclear palsy

In a study on temporal cortex samples from 175 PSP patients, Allen et al. (M. Allen *et al.*, 2016) investigated associations between seven PSP risk variants and expression levels of 20 genes in-cis. They observed highly significant associations for PSP risk alleles of rs8070723 and rs1768208 with higher LRRC37A4 and MOBP brain levels, respectively. Also, methylation levels of one CpG in the 3' region of ARL17B were associated with rs242557 and rs8070723. In another study by the same group (Allen, Wang, Serie, *et al.*, 2018), the transcriptional associations with unique cell-specific tau pathologies was investigated in 268 PSP autopsies. They determined that NFT were positively correlated with a brain co-expression network

enriched for synaptic and PSP candidate risk genes including *MAPT*, *NSF*, *CRHR1*, *STX6*, *IER5*, *BMS1*, while tufted astrocytes (TA) were associated with a microglial gene-enriched immune network. Furthermore, they compared transcriptome profiles of PSP with AD in a recent meta-analysis study (Allen, Wang, Burgess, *et al.*, 2018). They found co-expression networks associated with myelination were lower in PSP temporal lobe in comparison with AD. However, these networks showed no difference between PSP and AD in cerebellum. These networks were significantly down-regulated in both AD and PSP compared with control subjects.

1-3-1-4- **Pathological aging**

Pathological aging (PA), also known as preclinical AD, is neuropathologically characterized by extracellular Aβ accumulation in senile plaques (Dickson, 1997) and intracellular tau in neurofibrillary tangles and neuropil threads. However, unlike in AD, dendritic and synaptic abnormalities, as well as neuronal loss, are absent in PA. Aβ deposits are found in the context of minimal cortical and limbic neurofibrillary pathology (Dickson *et al.*, 1992). Lately, it was discussed that subjects with PA may have protective or resistance factors against neuronal loss and cognitive impairment. First studies revealed that morphologic and immunohistochemical characteristics of Aβ deposits were different between PA and AD. Diffuse plaques, principally composed of Aβ42 (Iwatsubo *et al.*, 1994), were more abundant in pathological aging, and paired helical filament tau immunoreactivity or glial reaction was rarely observed in proximity to senile plaques (Dickson *et al.*, 1992). Aβ42 levels are surprisingly more abundant in PA compared with AD (Maarouf *et al.*, 2011), which might be due to higher levels of diffuse plaques in PA (Iwatsubo *et al.*, 1994; Moore *et al.*, 2012). A study on normal subjects, PA, and AD revealed

that Aβ oligomeric monomers and dimers in synaptoneurosomes may be lower in PA individuals in comparison with demented AD. By investigating synaptic densities around amyloid plaques in AD, a group of researchers observed deleterious effects of oligomeric Aβ (Koffie *et al.*, 2012), correlated with synapse loss in individuals with apolipoprotein E ε4. Given the similar characteristics of soluble oligomeric deposits in PA and AD (Xu *et al.*, 2016), and higher amount of synaptophysin and neprilysin levels in PA (Wang *et al.*, 2005), the existence of unknown protective factors may alleviate deleterious effects of toxic Aβ on neuronal loss in PA.

Transcriptome profiling in pathological aging

Several studies have investigated transcriptome profiles of aging brain tissues from human and animal models. In a study conducted in 1904 human brain samples from age-associated phenotypes, Rhinn and colleagues (Rhinn and Abeliovich, 2017) performed gene expression and GWAS analysis for cerebral cortex. Their findings revealed the association of TMEM106B and GRN loci with aging, which previously showed correlation with cortical dementia. The *TMEM106B* gene variants are correlated with neuronal loss and cognitive deficits. Meta-analysis of transcriptome profiles in the hippocampus has revealed significant expression changes of several genes in aged rats(Uddin and Singh, 2013). More than 30 genes showed up-regulation, including *APOE*, *Rela*, *S100b*, *Icam1*, *Igf1r*, *Abcc*, *Alf1*, *Cds*, *Actb* and *Cntn2*.4, which are associated with neuronal formation, survival, migration and axonogenesis. In addition, transcriptome profiling of the temporal cortex of the primate Microcebus murinus was measured for 6 young adults, 10 healthy old animals and 2 old, and AD-like animals that showed ß-amyloid plaques and cortical atrophy (Abdel Rassoul *et al.*, 2010). Differential expression analysis revealed 695 genes with altered expression in old and AD-like animals compared with young animals. Most of the genes that were up-regulated in old animals and down-regulated in AD-like

animals were associated with metabolic pathways, especially protein synthesis. The findings from this study suggest the presence of compensatory mechanisms during physiological brain aging that are absent in AD-like animals. Despite these studies on the brain tissues of aging-associated phenotypes, the transcriptome signature of PA (or preclinical AD) is poorly understood. Allen and colleagues have tried, in an effort, to profile the transcriptome of temporal cortex and cerebellum tissues obtained from patients with PA, AD, PSP and control subjects (Mariet Allen *et al.*, 2016). The data from this study was used in our analyses which we will describe later (see Methods and Results).

1-3-2- **Neuropsychiatric disorders**

Neuropsychiatric disorders (NPDs) are a group of sophisticated mental conditions with a wide range of clinical phenotypes (Hosseini *et al.*, 2019).These conditions may be caused by traumatic brain injury (TBI), genetic factors, environmental factors, drug side effects, and infection (Miyoshi and Morimura, 2010; Fleiss, Rivkees and Gressens, 2018). The important element of NPD is that the symptoms show the perturbation of brain function, emotion, and mood. These symptoms vary from issues with focus and attention, learning, sadness, irritability, memory problems, mood issues, depression, language problems, hallucinations, and social issues (Miyoshi and Morimura, 2010). People with NPD may have numerous risk factors in a heterogeneous genetic context that make candidate genes difficult to detect. Here we focus on some of the most common NPDs.

1-3-2-1- **Schizophrenia**

Schizophrenia (Scz) is a multifactorial and complex mental disorder, which is characterized by different categories of symptoms including positive, negative, and cognitive symptoms (Kahn *et al.*, 2015). Positive symptoms that are associated with decreased cortical thickness in the superior temporal gyrus (Early interventions to prevent psychosis: systematic review and meta-analysis, 2013) involve the loss of contact with reality such as delusions, hallucinations, and perturbation of speech and behavior (Stafford *et al.*, 2013). The negative category consists of social abandonment, the inability to feel pleasure, and energy decline. Multiple negative symptoms are associated with decreased thickness in the left medial orbitofrontal cortex (Patel *et al.*, 2014). Moreover, individuals with Scz usually demonstrate a wide range of cognitive dysfunctions (Kahn *et al.*, 2015).

Molecular mechanisms of schizophrenia

It is hypothesized that a combination of genetic factors and the environment may have causative effects in Scz (Weinberger and Levitt, 2011; Howes and Murray, 2014). The first neurophysiological studies demonstrated the dysfunction of the neurotransmitters dopamine, glutamate, and γ-aminobutyric acid (GABA) as pathogenetic factors (Fusar-Poli and Meyer-Lindenberg, 2013). Dopamimetic drugs cause paranoid moods and unusual behavior, while dopaminolytic medications are considered as the only available antipsychotic drugs (Kaalund *et al.*, 2014). A decrease in D1 receptors in the dorsolateral prefrontal cortex may also be responsible for deficits in working memory (Brisch *et al.*, 2014). The glutamate model is another assumption that correlates changes between glutamatergic neurotransmission and neural

oscillations that influence connections between the thalamus and the cortex (Hu *et al.*, 2015). N-methyl-d-aspartate (NMDA) is a subtype of glutamate receptors that has a crucial function in memory and learning processes. The antagonists for these receptors such as ketamine and phencyclidine cause psychosis-like dissociative conditions, cognitive impairment, and unusual behavior (Hu *et al.*, 2015). Unlike dopamine and glutamate, an unusual function of GABAergic neurons has been consistently reported (Lewis, 2014; Schmidt and Mirnics, 2015). For instance, glutamate decarboxylase 1 (*GAD1*), a gene involved in GABA synthesis, has shown consistent expressional changes. Although, GABA cell types and several brain areas are mainly involved in Scz pathology, the main focus has been placed on dysfunction of GABAergic neurons in PFC that express parvalbumin, as these neurons participate in major cortical activities, such as gamma frequency oscillations, that are perturbed in Scz brain (Reichenberg *et al.*, 2014). Some of these abnormalities are also present in several other NPDs, such as major depression, anxiety disorders, and autism (Schmidt and Mirnics, 2015).

Genetic studies have also identified many risk factors related to the pathology of Scz (Kahn *et al.*, 2015). GWAS studies, for instance, have found significant variants in several loci including *DRD2* (the gene encoding the dopamine D2 receptor), glutamate receptor components (*GRM3*, *GRIN2A* and *GRIA1*, the gene encoding metabotropic glutamate receptor 3 (mGluR3), *GluN2A* and *GluA1*, respectively), and *SRR* (encoding serine racemase, an enzyme for biosynthesis of an NMDA receptor allosteric site ligand) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). The genes *CRHR1* and *CRHBP* are also linked to suicidal behaviors in Scz patients (Salleh, 2004; Avramopoulos, 2018). These genes are functional in the control of the HPA axis. Stress-related situations may alter the function of the HPA axis, leading to long-term changes in regulation of emotion and behavior (Cherian, Schatzberg and Keller,

2019). Intriguingly, the majority of the Scz-linked genes are normally expressed during fetal development (Gulsuner *et al.*, 2013; Birnbaum *et al.*, 2014). This implies the overlap of genetics of Scz and that of brain development, as well as the onset of Scz in early life (Weinberger and Levitt, 2011).

Transcriptome profiling for schizophrenia

Gene expression has been largely investigated in Scz using conventional methods including Real-Time PCR to high-throughput technologies such as microarray, RNA-Seq and single-cell RNA-Seq. Anio et al (Arion *et al.*, 2007) have investigated gene expression differences in dorsolateral prefrontal cortex samples from 14 matched pairs of schizophrenia and control subjects using microarray. Out of >1800 genes, the overexpression of *SERPINA3*, *IFITM1*, *IFITM2*, *IFITM3*, *CHI3L1*, *MT2A*, *CD14*, *HSPB1*, *HSPA1B*, and *HSPA1A* in Scz patients was shown to be strongly correlated with synaptic changes, oligodendrocyte, and signal transduction. One of the initial studies using cDNA microarrays analyzed expression of >7800 genes in matched subjects (Mirnics *et al.*, 2000). It was identified that only a few groups of genes showed expressional changes in Scz patients, which were related to presynaptic secretory function (*PSYN* gene group). A broad investigation revealed down-regulation of the regulator of G-protein signaling 4 (RGS4) in Scz individuals (Mirnics *et al.*, 2000, 2001). RNA-Seq on samples from the anterior cingulate cortex, dorsolateral prefrontal cortex, and nucleus accumbens from Scz patients and control subjects identified disease-related differences in the anterior cingulate cortex, with down-regulation of the transcription factor EGR1(Ramaker *et al.*, 2017). Also, broad down-regulation of genes specific to neurons and concordant up-regulation of genes specific to

astrocytes was observed in Scz samples. This study also showed disruption of GABA levels in schizophrenia patients. Transcriptional changes in amygdala tissue have also been demonstrated in Scz patients using RNA-Seq (Chang *et al.*, 2017). Disruption of down-regulated genes related to synaptic transmission and behavior, and upregulated genes associated with immune response were observed in amygdala of Scz subjects. A combination of RNA-Seq and GWAS has shown wide-spread expression quantitative trait loci (eQTLs) in DLPFC of Scz patients compared with controls (Jaffe *et al.*, 2018). The evidence showed that 48% of risk variants for Scz are associated with adjacent expression loci. Also, dysregulation of 237 genes in Scz showed disturbance of synaptic processes in early development of the brain. A few studies have also used scRNA-Seq to investigate gene expression alterations in Scz. For instance, Sken and colleagues (Skene *et al.*, 2018) using a combination of GWAS data and scRNA-Seq observed gene sets that were consistently associated with schizophrenia—intolerant to loss-of-function variation, NMDA receptor complex, postsynaptic density, *PSD95* complex, *RBFOX* binding, *CELF4* binding and *FMRP*-associated genes—all had more specific expression in neocortical S1 and hippocampal CA1 pyramidal cells, medium spiny neurons from the dorsal striatum and cortical interneurons.

1-3-2-2- **Autism spectrum disorder**

Autism spectrum disorder (ASD), is one of the most common NPDs that usually shows heritable cases with heterogeneous cognitive conditions (Lord *et al.*, 2020). It is characterized by a broad range of symptoms including disability in social interaction, repetitive behaviors, abnormal body posturing and facial expressions, avoidance of eye contact, behavioral disturbances, deficits in language comprehension, and varying levels of intellectual disability

(Brugha *et al.*, 2016). Many of these symptoms are present in other psychiatric disorders such as attention-deficit/ hyperactivity disorder (ADHD), anxiety, major depression, and Scz (Risi *et al.*, 2006). Different medications are currently prescribed to treat some of the related symptoms of ASD, including irritability, and comorbidities, like depression and anxiety (Baxter *et al.*, 2015).

Pathobiology of ASD usually involves disruption of four social brain regions including the amygdala, orbitofrontal cortex (OFC), temporoparietal cortex (TPC), and insula (Bachevalier, 1994; Baron-Cohen *et al.*, 2000; Bachevalier and Loveland, 2006; McPartland, Coffman and Pelphrey, 2011; Pelphrey *et al.*, 2011). In addition, commonly structural and functional abnormalities are observed across other brain areas such as visual cortical regions, PFC, basal ganglia, hippocampus, sensorimotor cortex, cerebellum, and thalamus (Stanfield *et al.*, 2008; Nickl-Jockschat *et al.*, 2012; Patriquin *et al.*, 2016).

Molecular basis of autism

Genetic studies have revealed that ASD is a highly heritable disorder with an approximately 40%-90% genetic contribution (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013; Gaugler *et al.*, 2014). More than 100 genes and genomic regions have now been confidently associated with autism (Sanders *et al.*, 2015; Satterstrom *et al.*, 2020). At a population level, the contribution of de novo mutations to the risk of ASD is estimated to be nearly 3% (Gaugler *et al.*, 2014). Conservative estimates are that 10–20% of people with ASD harbour a de novo rare point mutation or CNV contributing to their presentation (Iossifov *et al.*, 2014; Sanders *et al.*, 2015). Many of the mutations are found in the genes related to other disorders. For example, mutations in the *ADNP* gene cause a syndrome called ADNP. Some of the other genes in which

rare mutations are associated with ASD are *ARID1B*, *ASH1L*, *CHD2*, *CHD8*, *DYRK1A*, *POGZ*, *SHANK3*, and *SYNGAP1* (Van Dijck *et al.*, 2016). The genes involved in syndromic ASD include *FMR1* (encoding fragile X mental retardation protein; fragile X syndrome)(Abbeduto, McDuffie and Thurman, 2014), *UBE3A* (encoding ubiquitin-protein ligase E3A; Angelman syndrome)(Trillingsgaard and ØStergaard, 2004), *TSC1* and *TSC2* (encoding hamartin and tuberin; tuberous sclerosis complex)(Vignoli *et al.*, 2015), *PTEN* (encoding phosphatase and tensin homologue) and *MECP2* (encoding methyl-CpG-binding protein 2; Rett syndrome) (Neul, 2012). Many of these ASD-linked genes are crucial in the brain development and function. They are functional in generation, growth, and organization of neurons, synaptic connections, cell-cell communication, dendritic projections (Trifonova, Khlebodarova and Gruntenko, 2017).

1-3-2-3-

Transcriptome profiling in autism

In an study conducted by Wright and colleagues (Wright *et al.*, 2017), a genome-wide expression analysis using RNA-Seq showed that histaminergic system (HS) genes including *HNMT*, *HRH1*, *HRH2*, and *HRH3* were significantly altered in DLPFC tissue of ASD patients compared with controls. HS has an important role in cognition, sleep and other behaviors. Another study on PFC samples from 63 ASD individuals showed alteration of four TFs (EGR1–4) that are involved in the regulation of neuronal function, including synaptic activity, neuronal plasticity, and neuronal cell death. Integrated system analysis using data from RNA-Seq exhibited a co-expression module enriched for transcriptional regulation, including ASD-associated transcription factors and chromatin remodelers (FOXP2, MECP2, and CHD8, etc.), and another module with many genes for synaptic transmission (*SHANK2*, *SHANK3*, *NLGN1*, *NLGN3*) (Li *et*

al., 2014). In addition, analysis of alternative splicing (AS) in the brain shows that neural microexons are frequently misregulated in the brains of individuals with ASD, which is associated with reduced levels of nSR100 a neuronal-specific splicing factor (Irimia *et al.*, 2014). Transcriptome profiling using scRNA-Seq has also shed light on genes involved in ASD pathobiology. Single-cell transcriptomics (P. Wang *et al.*, 2018) revealed that upregulated genes in multiple ASD cortex samples were enriched with genes highly expressed in inhibitory neurons, suggesting a potential increase of inhibitory neurons and an imbalance in the ratio between excitatory and inhibitory neurons in ASD brains. Also, the targets of several genes including *CHD8*, *EHMT1* and *SATB2* represented enriched expression in inhibitory neurons. Velmeshev and colleagues (Velmeshev *et al.*, 2019) conducted snRNA-seq on 41 post-mortem brain samples from PFC and anterior cingulate cortex, from 15 subjects with ASD and 16 controls. Unbiased clustering identified significant alterations in gene expression, especially in upper layer (layer 2/3) excitatory neurons, vasoactive intestinal polypeptide (VIP)-expressing interneurons, protoplasmic astrocytes and microglia.

1-3-2-4- **Bipolar disorder**

Bipolar disorders (BP) are a broad range of mood disorders affecting emotion, energy, thought, and behavior, and are identified by biphasic mood episodes of mania and depression. Individuals with BP usually demonstrate cognitive dysfunctions, particularly altered reaction time, verbal and visual memory and executive dysfunction (Cullen *et al.*, 2016; Vieta *et al.*, 2018). BP is prevalent in more than 1% of the worldwide population (Merikangas *et al.*, 2011). The subdivisions of BP show an approximate prevalence of 0.6% for BP I, 0.4% for BP II 1.4% for sub-threshold manifestations of BP and 2.4% for the BP spectrum (Merikangas *et al.*, 2011). The prevalence of BP I is the same in females and males, but BP II is more prevalent in women. BP co-occurs with

other NPDs such as anxiety disorders, substance use disorders, ADHD, depression, Scz and personality disorders (Merikangas *et al.*, 2007). Patients with BP demonstrated thinner cortical gray matter in frontal, temporal and parietal areas of each brain hemisphere, with the biggest impact on left pars opercularis, left fusiform gyrus and left rostral middle frontal cortex (Strakowski *et al.*, 2012).

Molecular basis of bipolar

It is hypothesized that an interaction of genetics and environmental factors contribute to the pathology of BP (Lichtenstein *et al.*, 2009; Mühleisen *et al.*, 2014). The heritability of bipolar disorders is approximately above 80% (Hunt, 2008). Using GWAS, researchers have found numerous genetic variants across the genome linked to the BP. The most common variants are usually observed in the genes *CACNA1C*, *ODZ4*, *TRANK1*, and *NCAN*, emerged as primary candidates for the disease in the GWAS (Hunt, 2008; Mühleisen *et al.*, 2014; Stahl *et al.*, 2019). The gene *CACNA1C* encodes an L-type voltage-gated ion channel with well-established roles in neuronal development and synaptic signaling (Gershon *et al.*, 2014). The gene ODZ4 encodes a large transmembrane protein that its structure is similar to the signal transduction molecules. During brain development, *ODZ4* is involved in the regulation of neuronal and synaptic connectivity (Heinrich *et al.*, 2013). The gene *NCAN* codes for a large secreted protein that is found predominantly in the extracellular space, the lumen of the Golgi apparatus, and the lysosomal cavities contributing to the modulation of cell adhesion, cell migration, and axon guidance (Schultz *et al.*, 2014). *TRANK1* encodes a large, mostly uncharacterized protein, highly

expressed in multiple tissues, especially the brain, and may play a role in maintenance of the blood–brain barrier (Jiang *et al.*, 2019).

1-3-2-5-

Transcriptome profiling in bipolar

Most transcriptome studies on BP have been performed on the frontal lobe and limbic cortex of the brain. Profiling of postmortem human dorsal striatum from 18 bipolar and 17 control subjects demonstrated fourteen differentially expressed genes, including immune response genes such as *NLRC5*, *S100A12*, *LILRA4* and *FCGBP* (Pacifico and Davis, 2017). This study showed the association of a co-expression module related to medium spiny neurons in dorsal striatum. Gene expression analysis of DLPFC in the brain of BP subjects has also exhibited DE of several transcripts including Prominin 1/CD133 and ATP-binding cassette-sub-family G-member2 (ABCG2) associated with neuroplasticity, and serine/arginine-rich splicing factor 5 (SRSF5) and regulatory factor X4 (RFX4), related to circadian rhythms (Akula *et al.*, 2014). In addition, large consortia including CommonMind consortium (CMC)(Hoffman *et al.*, 2019) and PsychEncode (PsychENCODE Consortium, 2018) are seeking to unravel the molecular structure of several psychiatric disorders including BP using transcriptome profiling of different parts of the brain from post-mortem subjects. However, transcriptome profiling at the single cell level is still missing to shed light on cell-type specific gene expression associated with pathology of BP.

1-3-2-6- **Major depressive disorder**

Major depressive disorder (MDD) is a common type of mood disorder defined by a wide range of symptoms such as lasting depression, significant alterations in mood, interests, pleasure, and cognitive function (Otte *et al.*, 2016). MDD symptoms usually co-occur with the symptoms of other NPDs such as in Scz and BP (Ramaker *et al.*, 2017). MDD affects nearly 6% of the adult population worldwide each year and is prevalent about twice as often in females than in males (Bromet *et al.*, 2011). The genetic contribution to MDD is estimated to be approximately 35%, or more in familial studies (Flint and Kendler, 2014). This implies the role of other genetic variations, like rare mutations, in pathology of MDD. Furthermore, environmental factors, such as sexual, physical or emotional abuse in childhood, have been linked to MDD development (Li, D'Arcy and Meng, 2016).

Smaller volumes of hippocampus, alterations in activity or connectivity of neural networks, such as the cognitive control network and the affective–salience network, are significant cortical changes related to the neurobiology of MDD (Etkin, Büchel and Gross, 2015). Cortical brain regions involved in MDD include the dorsal and medial PFC, the dorsal and ventral anterior cingulate cortex, the orbital frontal cortex and the insula (Pandya *et al.*, 2012; Zhang *et al.*, 2018). Also, hypometabolism of PFC, particularly in dorsolateral and dorsoventral brain regions, is repeatedly found in MDD (Zhang *et al.*, 2018; Ichikawa *et al.*, 2020). Moreover, the stress-response mechanisms, such as the HPA axis, the autonomic nervous system and the immune system, are significantly altered in MDD (Kupfer, Frank and Phillips, 2016).

Molecular basis of major depression

MDD is a heterogeneous disorder occurring more in familial cases. First-degree relatives of patients with MDD show a threefold increased risk of MDD and heritability for this disorder that has been quantified as approximately 35% (Geschwind and Flint, 2015). In addition, MDD has shown genetic similarity with other NPDs, including Scz and BP ('Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis', 2013; Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013). Large-scale GWA studies have linked 15 genetic loci to the risk of MDD development (Hyde *et al.*, 2016). Despite many limitations in the search of genetics of MDD, multiple candidate genes are frequently considered to be related to MDD. The Serotonin Transporter (5HTT/SLC6A4) and Serotonin Receptor 2A (HTR2A) (Lin *et al.*, 2015) genes involved in the serotonergic system are candidate genes associated with MDD given that the majority of antidepressant drugs interact with these systems (Schneider *et al.*, 2018). Multiple investigations have linked the genetic variations in the gene SLC6A4 to MDD (Mendonça *et al.*, 2019; Ran *et al.*, 2020). Using *in vitro* studies, it was shown that a 44-bp repeat polymorphism in the promoter region of the gene 5-HTTLPR affects expression levels of the serotonin transporter (Nonen *et al.*, 2016; Schneider *et al.*, 2018). Another candidate gene responsible for MDD is brain-derived neurotrophic factor (BDNF) that plays a role in neurogenesis (Hariri, 2011). Expression levels of BDNF have shown a decrease in the hippocampus of animals exposed to chronic stress (Phillips, 2017; Ferrer *et al.*, 2019).

The gene TPH2, encoding for Tryptophan hydroxylase, is another important gene involved in serotonin synthesis in the brain, that has shown linkage to MDD pathobiology (Tao *et al.*, 2018). It is situated on chromosome 12q, a region previously implicated in linkage studies of BP (Van Den Bogaert *et al.*, 2006). In addition, genetic variants have been associated with the risk

of developing MDD. Zhang et al. observed a functional mutation (Arg441His) that causes loss of function in serotonin synthesis. This mutation has also been observed in several patients with MDD, while it was absent in healthy controls (Zhang *et al.*, 2005). Furthermore, epigenetic studies have suggested that an interaction between genetics and environment could be the best model explaining the neurobiology of MDD. For instance, an interaction has been found between stress and DNA demethylation in glucocorticoid-response elements of *FKBP5* (Menke *et al.*, 2013; Humphreys *et al.*, 2019). This interplay causes elevated expression of *FKBP5*, resulting in glucocorticoid receptor resistance (Han *et al.*, 2017).

Transcriptome profiling in major depression

DLPFC gene expression analysis using RNA-Seq in MDD subjects demonstrated altered expression of immune related genes including humanin like-8 (*MTRNRL8*), interleukin-8 (*IL8*), and serpin peptidase inhibitor, clade H (*SERPINH1*) and chemokine ligand 4 (*CCL4*) compared with controls (Pantazatos *et al.*, 2017). Also, deficits in microglial, endothelial (blood-brain barrier), ATPase activity and astrocytic cell functions were shown to be contributing to MDD. Evidence shows that transcriptome signature is different in male and female. Labonte et al (Labonté *et al.*, 2017) profiled transcriptome of six brain regions, including Orbitofrontal cortex (BA11), DLPFC (BA8/9), cingulate gyrus 25 (BA25; vmPFC), anterior insula, nucleus accumbens, and ventral subiculum from MDD subjects and controls. They observed a major rearrangement of transcriptional patterns in MDD, with limited overlap between males and females, an effect seen in both depressed humans and stressed mice. In addition, they found decreased expression of the female-specific hub gene *Dusp6* in the mouse prefrontal cortex

mimicked stress susceptibility in females, but not males, by inducing ERK signaling and pyramidal neuron excitability. A recent study performed single nucleus RNA-Seq (snRNA-Seq) transcriptomics on ~80,000 nuclei from the DLPFC of male subjects with MDD and of healthy individuals (Nagy *et al.*, 2020). The authors observed 26 cellular clusters, of which 60% showed transcriptional alterations in MDD subjects. They reported that the greatest expression changes were found in deep layer excitatory neurons and immature oligodendrocyte precursor cells (OPCs).

1-3-3- **Shared molecular signature across brain diseases**

Previously, we have described the clinical symptoms and molecular mechanisms of several brain diseases. Many of these conditions share similar clinical characteristics and genetic alterations. For instance, Scz, BP, and MDD have frequently shown overlapping symptoms, suggesting that they could, at least, share similar pathobiology (American Psychiatric Association, 2013; Hafemeister, 2019).Several investigations have shed light on the heterogeneity within brain diseases and the degree of the similarities between closely related disorders (Matias, Morgado and Gomes, 2019; Pardiñas *et al.*, 2019; Strohäker *et al.*, 2019). Patterns of converging clinical and biological characteristics across NDDs such as AD, PD, and PA have been lately discussed (Parikshak, Gandal and Geschwind, 2015; Lynch *et al.*, 2016; Santiago, Bottero and Potashkin, 2017). This phenotypic overlap proposes potential similar molecular pathobiology, which is supported by recent large-scale genome-wide association studies (GWAS)(International Schizophrenia Consortium *et al.*, 2009; Cross-Disorder Group of the Psychiatric Genomics

Consortium, 2013; Gandal *et al.*, 2018). Similar examples can be found in genes that do not encode enzymes.

For example, among the four clinically discernible ciliopathies with overlapping features caused by *TMEM67* mutations, nephronophthisis with liver fibrosis (*NPHP11*) is distinguished from the other three by no or only mild neurological involvement, which can be explained by the hypomorphic nature of *NPHP11*-causing mutations55. Such observations indicate that different quantitative effects of mutations could result in either the presence or absence of specific symptoms or differing degrees of severity. Relevant examples for the latter might include mutations in enzymes required for O-mannosylation of proteins (for example, POMGNT1, POMT1 and POMT2) causing related types of muscular dystrophy-dystroglycanopathies distinguished by clinical severity56, and possibly *SCN1A* mutations causing familial febrile seizures (FEB, less severe), generalized epilepsy with febrile seizures plus (GEFS+, more severe) or Dravet syndrome (most severe). A group of researchers (Zhu *et al.*, 2014; Brainstorm Consortium *et al.*, 2018) observed a number of genes with particularly hot zone *de novo* mutations overlapped across multiple NPDs including Scz, ASD, intellectual disability (ID), and Epileptic encephalopathies (EE). Figure 1-2 illustrates some of these genes including *ALS2CL*, *CHD4*, *GNAO1*, *ITPR1*, *KIAA2018*, *PAQR8*, *SCN2A*, *TRRAP* and *ZBTB40*.

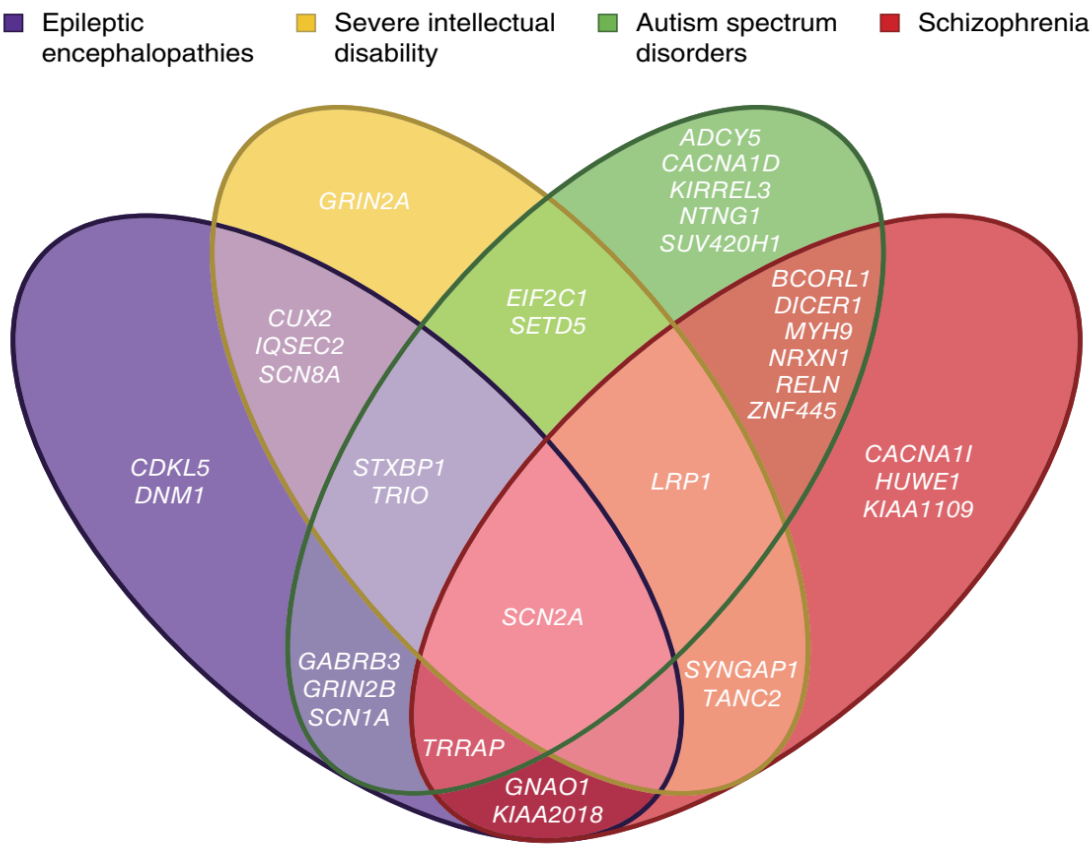

Figure 1-2 | The overlap of genes with de novo mutations across four mental illnesses (adopted from (Zhu et al. 2014))

In a large-scale GWAS study (Brainstorm Consortium *et al.*, 2018), genomic data for 265,218 patients and 784,643 control participants, as well as 17 phenotypes from a total of 1,191,588 subjects, was used to quantify the degree of overlap for genetic risk factors of common brain disorders (Figure 1-3). The authors found that common variant risks for psychiatric disorders correlate significantly between ADHD, BP, MDD, and Scz. On the other hand, neurological disorders were more distinct from one another and from the psychiatric disorders, except for migraine, which showed correlation with ADHD, MDD, and Tourette syndrome. The authors also observed significant genetic similarity between disorders and early life cognitive measures (e.g., years of education and college attainment) of the subjects, revealing positive correlation with multiples NPDs such as anorexia nervosa and BP, and negative correlation with some neurological disorders including AD and ischemic stroke.

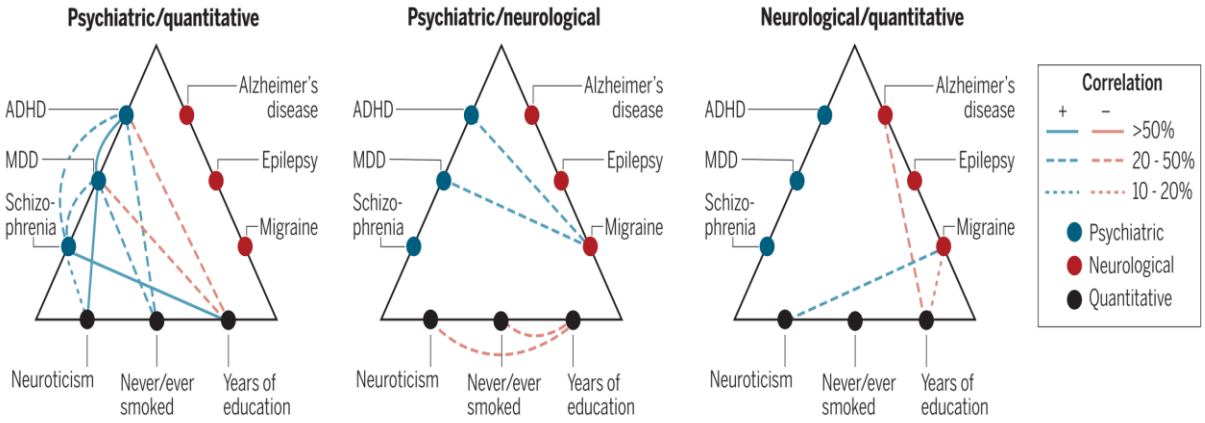

Figure 1-3 | Subsection of genetic risk correlations among brain disorders and quantitative phenotypes. Heritability analysis of brain disorders points to pervasive sharing of genetic risk among psychiatric disorders. These correlations are largely absent among neurological disorders but are present for both groups in relation to neurocognitive quantitative phenotypes. Only significant correlations shown. Line color and solidity indicate direction and magnitude of correlation, respectively (Brainstorm Consortium *et al.*, 2018).

Gandal and colleagues (Gandal *et al.*, 2018) used RNA-Seq and microarray to investigate transcriptome profiling of post-mortem samples from five major NPDs including ASD, Scz, MDD, BP, and alcoholism (AAD). They observed significant transcriptome overlap across ASD, Scz, BP, and MDD. Using co-expression network analysis, they further found that an astrocyte-specific co-expression module was up-regulated in ASD, BD, and SCZ. Also, multiple neuron-specific modules showed decreased expression in ASD, SCZ, and BD. Ramaker et al. (Ramaker *et al.*, 2017) also used RNA-Seq on brain samples from the anterior cingulate cortex (Acc), dlPFC, and nucleus accumbens from subjects with Scz, BP, MDD and controls. They found broadly down-regulated genes specific to neurons and up-regulation of astrocytes-specific genes in Scz and BP subjects compared with controls. One of the largest GWA studies was performed on 232,964 cases and 494,162 controls with anorexia nervosa, ADHD, ASD, BP, MDD, obsessive-compulsive disorder (OCD), Scz, and Tourette syndrome (Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu and Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). The authors found significant genetic correlations between three groups of interconnected disorders including mood disorders, neurodevelopmental disorders, and compulsive disorders (Figure 1-4). They detected 109 loci linked to at least two disorders, including 23 loci with pleiotropic impacts on four or more disorders and 11 loci with antagonistic effects on multiple disorders.

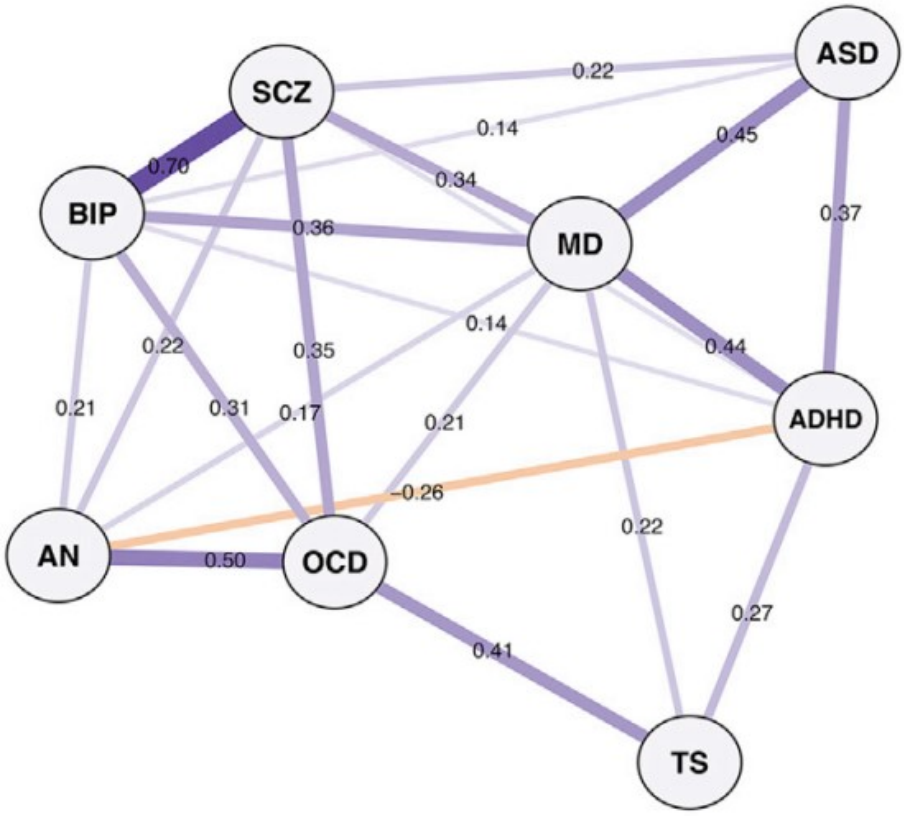

Figure 1-4 | SNP-based genetic correlations between eight disorders were depicted using an in-directed graph to reveal complex genetic relationships. Only significant genetic correlations after Bonferroni correction in (A) were displayed. Each node represents a disorder, with edges indicating the strength of the pairwise correlations. The width of the edges increases, while the length decreases, with the absolute values of r_g (Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu and Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019).

One of the main challenges in the genetic field of neurodevelopmental diseases is to uncover disease-specific and shared neurobiological mechanisms (Geschwind and Flint, 2015). Integrative transcriptomic studies focus on the functional gap, but a comprehensive characterization of gene expression changes in brain regions from individuals with major brain NDDs and NPDs compared with healthy subjects is missing. We hypothesized that there are potential overlapping molecular signatures across major brain diseases including NDDs and NPDs. To address this, we performed large-scale transcriptomic profiling using bulk RNA-Seq samples from post-mortem brain regions.

Objectives

Our main aim in this thesis was to explore the molecular signature of neurodegenerative and psychiatric disorders at the transcriptome level using analysis of RNA-Seq data by exploiting computational tools and statistical tests. The main objectives of the thesis were:

1) To study disease-specific transcriptional changes
    a) To investigate the genes responsible for the pathobiology of neurodegenerative and psychiatric diseases
    b) To explore similar transcriptome signature across NDDs and NPDs
    c) To study the molecular pathways involved in the pathology of brain diseases
2) To study cortical-specific transcriptional changes across NDD and NPDs
    a) To study the genes involved in the pathobiology of disease across regions
    b) To study shared molecular signature of cortical regions across diseases
3) To explore molecular signatures that are coordinatedly altered across diseases
    a) To find the genes that behave in the same manner across NDDs and NPDs
    b) To find correlation between brain cell types and gene modules that are involved in multiple NDDs and NPDs

Chapter2. Materials & Methods

2-1- Samples and raw data

RNA-Seq raw data were obtained from 4,711 post-mortem brain samples from subjects with AD (n = 906 samples), PD (n = 29), PA (n = 58), PSP (n = 168), SZ (n = 535), ASD (n = 187), MDD (n = 240), BP (n = 510), and matched controls pooled across alls studies (n = 2078), through 23 previously-published studies (Lipska et al., 2006; Wu et al., 2012; Akula et al., 2014; He et al., 2014; Li et al., 2014; Xiao et al., 2014; Corley et al., 2016; Dumitriu et al., 2016; Liu et al., 2016; Mariet Allen et al., 2016; Chang et al., 2017; Labonté et al., 2017; MacMullen, Fallahi and Davis, 2017; Pacifico and Davis, 2017; Pantazatos et al., 2017; Ramaker et al., 2017; Wright et al., 2017; Gandal et al., 2018; Jaffe et al., 2018; M. Wang et al., 2018) and consortia including CommonMind Consortium and PsychENCODE Consortium from Sage Synapse (https://www.synapse.org/) and the NCBI Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) (see Table 2-1 | RNA-Seq expression datasets obtained to use in this study .The samples from individual studies were processed separately and analyzed according to a harmonized pipeline as described below in data processing section.

Table 2-1 | RNA-Seq expression datasets obtained to use in this study

Diagnosis	Study/ Project	Accession ID	# Samples		Brain region	Ref
			Controls	Cases		
AD	Wang et al	syn3159438	280	742	Anterior prefrontal cortex, perirhinal cortex, superior temporal gyrus, pars opercularis	(M. Wang et al., 2018)
	Allen et al	syn5550404	155	164	Temporal cortex, cerebellum	(Mariet Allen et al., 2016)
PD	Dumitriu et al	PRJNA283498	44	29	BA9	(Dumitriu et al., 2016)
PA	Allen et al	syn5550404	Mentioned (155)	58	Temporal cortex, cerebellum	(Mariet Allen et al., 2016)
PSP	Allen et al	syn5550404	Mentioned (155)	168	Temporal cortex, cerebellum	(Mariet Allen et al., 2016)
Scz	Jaffe et al	syn12299750	320	175	Dorsolateral prefrontal cortex (DLPFC)	(Jaffe et al., 2018)
	Xiao et al	PRJNA235930	12	10	BA9, BA24	(Xiao et al., 2014)
	Chang et al	PRJNA379666	24	22	Amygdala	(Chang et al., 2017)
	Corley et al	PRJNA343829	19	19	DLPFC	(Corley et al., 2016)
	Wu et al	PRJEB2939	9	9	Superior temporal gyrus	(Wu et al., 2012)
	CommonMind consortium	syn18097439	294	47	DLPFC	(Fromer et al., 2016)
	CommonMind consortium-HBCC	syn18097439	167	87	DLPFC	(Fromer et al., 2016)
	BrainGVEX	syn4590909	257	95	Frontal cortex	(PsychENCODE Consortium et al., 2015)

	Ramaker et al	PRJNA319583	70	71	anterior cingulate cortex, DLPFC, nucleus accumbens	(Ramaker et al., 2017)
ASD	Wright et al	PRJNA398545	39	13	DLPFC	(Wright et al., 2017)
	Li J et al	PRJNA263196	6	6	Corpus callosum	(Li et al., 2014)
	Liu et al	PRJNA254971	38	34	Superior frontal gyrus	(Liu et al., 2016)
	Yale-ASD	syn4566141	30	15	DLPFC	(PsychENCODE Consortium et al., 2015)
	UCLA-ASD	syn4587609	126	119	BA4/6, BA38, BA7, BA17	(PsychENCODE Consortium et al., 2015)
MDD	Ramaker et al	PRJNA319583	Mentioned (70)	69	anterior cingulate cortex, DLPFC, nucleus accumbens	(Ramaker et al., 2017)
	Pantazatos et al	PRJNA394722	29	30	DLPFC (BA9)	(Pantazatos et al., 2017)
	Labonte et al	PRJNA398031	122	141	Orbitofrontal cortex (BA11), DLPFC (BA8/9), cingulate gyrus 25 (BA25; vmPFC), anterior insula, nucleus accumbens, ventral subiculum	(Labonté et al., 2017)
BP	Xiao et al	PRJNA235930	Mentioned (12)	14	BA9, BA24	(Xiao et al., 2014)
	CommonMind consortium	syn18097439	Mentioned (294)	257	DLPFC	(Fromer et al., 2016)

CommonMind consortium-HBCC	syn18097439	Mentioned (167)	58	DLPFC	(Fromer et al., 2016)
BrainGVEX	syn4590909	Mentioned (257)	73	Frontal cortex	(PsychENCODE Consortium et al., 2015)
Ramaker et al	PRJNA319583	Mentioned (70)	71	anterior cingulate cortex, DLPFC, nucleus accumbens	(Ramaker et al., 2017)
Pacifico et al	PRJNA318642	18	18	Dorsal striatum	(Pacifico and Davis, 2017)
McMullen et al	PRJNA321439	8	8	Putamen, Caudate nucleus	2-2- (MacMullen, Fallahi and Davis, 2017)
Akula et al	PRJNA231202	11	11	DLPFC	(Akula et al., 2014)

AD, Alzheimer's disease; PD, Parkinson's disease; PSP, progressive supranuclear palsy; PA, pathological aging; Scz, schizophrenia; ASD, autism spectrum disorder; MDD, majore depressive disorder; BP, bipolar disorder.

2-3- Transcriptomic Pipeline

2-3-1- **Data processing**

We used RNA-Seq *fastq* files as the initial source of analysis. The samples that were retrieved as *SRA* and *BAM* files were converted to *fastq* file formats using SRA-toolkit (Staff, SRAS, 2011) and SAM-tools(Ramirez-Gonzalez *et al.*, 2012), respectively. For further sample processing, the Grape pipeline (Knowles *et al.*, 2013) was used for RNA-Seq analysis, with Nextflow (Di Tommaso *et al.*, 2017) as the execution backend, the STAR aligner v.2.6.0a tool (Dobin *et al.*, 2013) for mapping reads to the human genome build hg19 with GENCODE v.28 annotations, and the RSEM tool (Li and Dewey, 2011) for isoform quantification (using default options). Nextflow is a workflow management system that uses Docker technology for the multi-scale handling of containerized computation. Nextflow was developed as a solution to the numerical instability problems that happen when data-analytic pipelines are used on different computational platforms. Next, post-alignment quality control (QC) was performed using STAR aligner statistics, Qualimap v.2.2.1 tools (Okonechnikov, Conesa and García-Alcalde, 2016), and Picard tools v1.8 (http://broadinstitute.github.io/picard/) by setting options to check for the total number of reads, total number of mapped reads, GC percentage, exonic rate, intronic rate, intergenic rate, duplication rate, and insertion/deletion rate. To control for differences in sequencing statistics, a matrix was generated from an aggregate of the above-mentioned QC tools according to the samples. Raw read counts were subsequently normalized against the read coverage, GC percentage, and gene length and $\log_2$-transformed using the *cqn* R package v.1.30.0 (Hansen, Irizarry and Wu, 2012). To filter out lowly-expressed genes, only genes with at least $\log_2$(FPKM) of 0.5 in half of the samples were kept for further analyses. The *sva* R package v.3.32.1(Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D., 2014) was used to correct for any batch effect of sequencing library

preparations. To remove sample outliers, standardized network connectivity Z-scores were measured and a cutoff of $Z < -2$ was set as the threshold (Shiffler, 1988; Cousineau and Chartier, 2010). The design of the study is illustrated in Figure 1-2.

2-3-2- **Clustering and tSNE analysis**

T-distributed Stochastic Neighbor Embedding (t-SNE) method (Maaten and Hinton, 2008) was used to visualize variation across datasets using all normalized datasets which were pooled together in an expression matrix. t-SNE is a machine learning algorithm for visualization developed by Laurens van der Maaten and Geoffrey Hinton. It is a nonlinear dimensionality reduction technique appropriate for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions. Specifically, it models each high-dimensional object by a two- or three-dimensional point in such a way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points with high probability.

Before using t-SNE, principal component analysis (PCA)(Monfreda, 2012) was used to reduce the number of dimensions and obtain a small number of principal components as input to the tSNE analysis. PCA is a data transformation technique that is used to reduce multidimensional data sets to a lower number of dimensions for further analysis. In PCA, a data set of interrelated variables is transformed to a new set of variables called principal components (PCs) in such a way that they are uncorrelated and the first few of these PCs retain most of the variation present in the entire data set. Thus, the first PC is a linear combination of all the actual variables in such a way

that it has the greatest amount of variation. Second PC is also a linear combination of the original variables in such a way that it has the most variation in the remaining PCs. Here, top two variable principal components were used for plotting. Sample clustering was performed by calculating the distance from the expression matrix ($\log_2$(FPKM)), clustering using hclust function, and building a dendrogram tree using the *dendextend* R package (Galili, 2015).

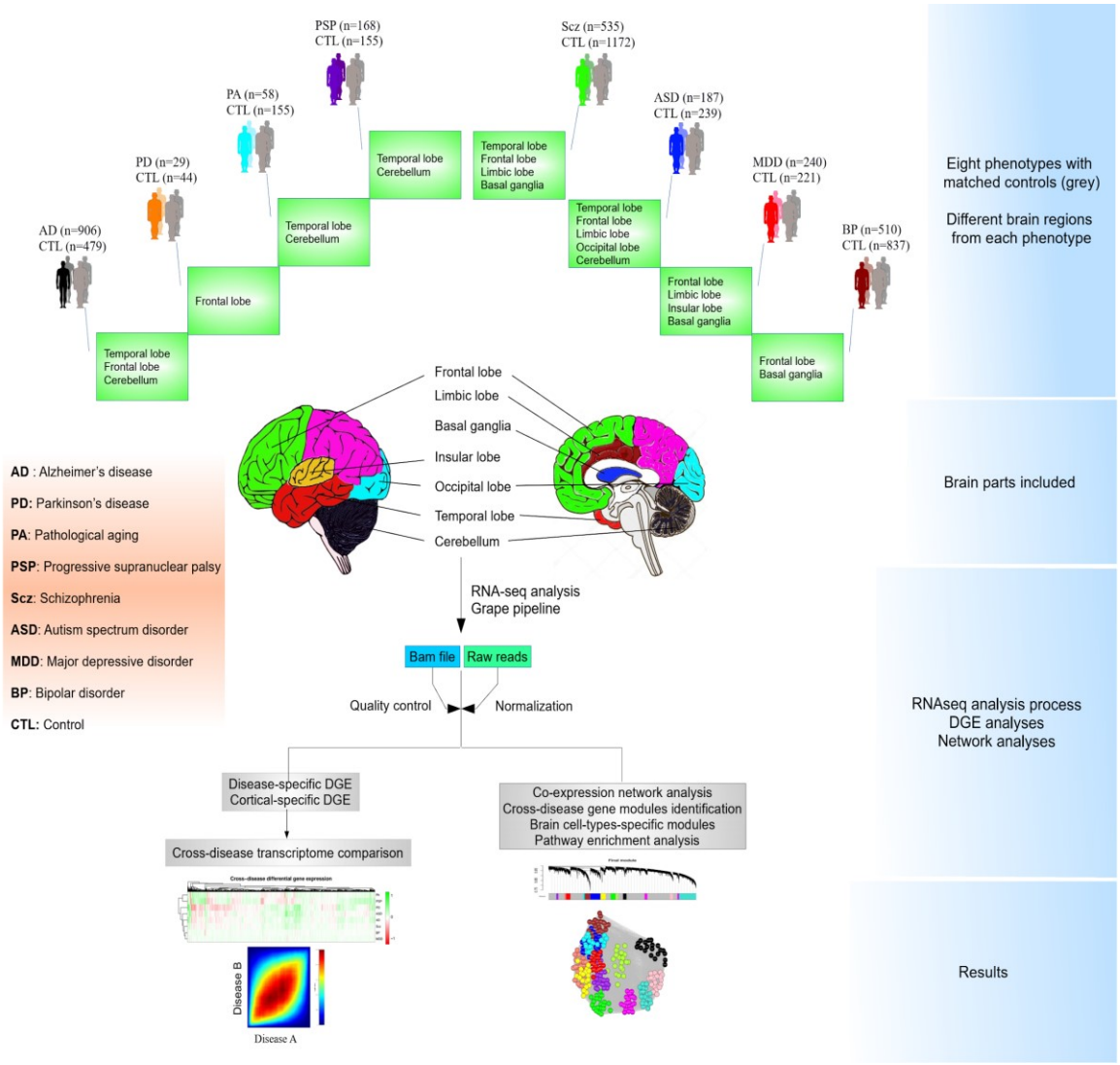

Figure 2-1 | Schematic of the study design and the samples used for gene expression analysis via a universal pipeline. Brain RNA-Seq data were obtained from subjects with AD (n =906 samples), PD (n = 29), PA (n = 58), PSP (n = 168), SZ (n = 535), ASD (n = 187), MDD (n = 240), BP (n = 510), and matched controls (n = 2078). Figure from *Sadeghi et al., 2020 (in preparation*)

2-4- Differential Gene Expression (DGE) analysis

For disease- and region-specific DGE analyses, normalized expression data from relevant studies were combined and DGE was calculated using linear mixed-effects models by the *nlme* R package v.3.1-140 (Pinheiro *et al.*, 2012), with fixed effects of diagnosis, age, sex, and study (and brain region when calculating disease-specific DE) and a random effect for subjects to fix for overlapping subjects between the studies (expression ~ diagnosis + age + sex+ study + 1 | subject). Linear mixed models are an extension of simple linear models to allow both fixed and random effects, and are particularly used when there is non independence in the data, such as arises from a hierarchical structure (Lindstrom and Bates, 1988). Mixed models are especially useful when working with a within-subjects design because it works around the ANOVA assumption that data points are independent of one another. In a within subjects' design, one participant provides multiple data points and those data will correlate with one another because they come from the same participant. Therefore, using a mixed model allows you to systematically account for item-level variability (within subjects) and subject-level variability (within groups) (Fitzmaurice, Laird and Ware, 2012).

The calculated $\log_2$ fold-change ($\log_2$FC) values were used for downstream analyses. Significantly differentially expressed genes (DEGs) for each disease compared to controls were filtered by using a p-value of < 0.05 and $|\log_2 FC| > 0.5$ as a significant threshold.

2-5- Prediction accuracy of regional transcriptome

To see which regions have the largest transcriptional difference between disease and control, normalized expressions of DEGs for each region were used to build classifier models using random forest algorithms (Breiman, 2001). Random forest classifier is an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Belgiu and Drăguţ, 2016).

Accuracy, sensitivity and specificity of the final models were used for comparing the results. To see discrimination of each disease from control whithin each region, top principal components were computed by a combination of PCA and tSNE using normalized expressions of DEGs.

2-6- Gene enrichment analysis for DEGs

A common approach to interpreting gene expression data is gene enrichment analysis based on the functional annotation of the differentially expressed genes (Werner 2008). This is useful for finding out if the differentially expressed genes are associated with a certain biological process or molecular function and may have an association with disease phenotypes (Curtis et al. 2005). The methods use statistical approaches to identify significantly enriched or depleted groups of genes. Transcriptomics technologies and proteomics results often identify thousands of genes which are used for the analysis.

The Gene Ontology, containing standardised annotation of gene products, is commonly used for this purpose. It works by comparing the frequency of individual annotations in the gene list (e.g differentially expressed genes) with a reference list (usually all genes on the microarray or in the genome) (Gene Ontology Consortium 2015). Enrichment of biological pathways supplied by KEGG, Ingenuity, Reactome or WikiPathways can be performed in a similar way (Harris et al. 2004).

Gene enrichment analysis for DEGs for each disease against controls was performed by retrieving data from KEGG and GO databases using the gprofiler2 R package. A p-value of > 0.05 followed by FDR correction was used to filter significant enrichments.

2-7- Analysis of transcriptome similarity

To analyze cross-disease transcriptome profile comparisons, we only kept the 10,313 genes that were common across all diseases. Pairwise gene expression comparisons were performed using Spearman's correlation over log_2FC values of the genes. Cohen's d effect sizes were also estimated based on the correlation statistics obtained from pairwise comparisons. Moreover, brain region-specific comparisons across diseases were performed only for the genes shared between the diseases.

In order to highlight the degree of overlap in gene signatures across diseases, as well as comparing disease pairs for shared brain regions, we performed an unbiased rank-rank hypergeometric overlap (RRHO) analysis using the *RRHO* R package v.1.24.0(Plaisier *et al.*, 2010). A one-sided version of the test only looking for over-enrichment was used. RRHO

difference maps were produced by calculating for each pixel the normal approximation of difference in the log odds ratio and standard error of overlap with expression data in the intersection list. P-values were calculated and FDR-corrected for multiple comparisons across pixels. In addition, a multidimensional scaling (MDS) analysis was performed using transcriptional correlations statistics obtained from log_2FC values of the 10313 common genes across diseases. To do this, correlation values obtained from transcriptional comparisons were first used to calculate topological proximity values. Next, the adjacency of the diseases was obtained using the Igraph R package (Han *et al.*, 2010).

2-8- Gene co-expression network analysis

In order to see at a more detailed level of the transcriptional overlap across diseases, we wanted to check the existence of the genes co-expressed or co-regulated together. To do so, we searched for co-expression modules acoss 10313 shared genes. We performed robust Weighted Gene Co-Expression Network Analysis (rWGCNA) using the *WGCNA* R package v.1.68 (Langfelder and Horvath, 2008) to identify co-expressed gene modules using expression data that were first normalized for different covariates. WGCNA is a systems biology method that was first developed for describing the correlation patterns among genes across microarray samples. It can be used for finding clusters (modules) of highly correlated genes, for obtaining such clusters using the module eigengene or an intramodular hub gene, for relating modules to one another and to external sample traits (using eigengene network methodology), and for obtaining module membership measures. Correlation networks facilitate network-based gene screening methods that

can be used to identify candidate biomarkers or therapeutic targets (Langfelder and Horvath, 2008).

The expression datasets from independent disease-specific DGE analyses were combined using the 10,313 genes common between all datasets. Batch effect correction for the studies was performed using the *ComBat* function from the *sva* R package.

Co-expression analysis was then performed using signed networks. Co-expression networks were built using the blockwise modules function. The network dendrogram was created using the "average" linkage hierarchical clustering of the topological overlap dissimilarity matrix to identify modules of highly co-regulated genes. To obtain an approximately scale-free weighted co-expression network, a power function with a soft-threshold of 10 was applied to the merged expression dataset. Modules were defined as branches of the dendrogram using the hybrid dynamic tree-cutting method, followed by a dynamic cut-tree algorithm to separate clustering dendrogram branches into gene modules. Modules were then summarized by their first principal component (ME, module eigengene) and those with high eigengene correlations were again merged.

Because the topological overlap between two genes reflects both their direct and indirect interactions through all other genes in the network, this approach helps to build more cohesive and more biologically meaningful modules. To ensure the robustness of the module, random resampling was performed from the initial set of samples 100 times followed by consensus network analysis. The final module was achieved using network parameters including biweight midcorrelation (bicor), a minimum module size of 100, deepsplit of 4, merge threshold of 0.1, and negative pamStage. Each module was assigned by a unique (and arbitrary) color identifier. Genes

with the highest intramodular connectivity (those with more connections at the core of the network) were considered as hub genes. Module (eigengene)-disease associations were evaluated using a linear mixed-effects model. Significance values were FDR-corrected to account for multiple comparisons. Genes within each module were prioritized based on their module membership (kME), defined as a correlation to the module eigengene.

2-9- Co-expression modules gene ontology analysis

Gene Ontology (GO) pathway enrichment for each gene module was performed using the *ClusterProfiler (Yu et al., 2012)* and the *gprofiler2* R packages. Only pathways that comprise 10 to 2000 genes were filtered for analysis. The top pathways with an FDR-corrected P-value < 0.05 were considered significantly related. Next, transcription factor binding sites (TFBSs) enrichment analysis was performed for the genes within modules using the *gprofiler2* R package, by using annotations from the TRANSFAC database (Wingender, 1996).

2-10-Cell-type-specific gene expression analysis

To analyze cell-type-specific gene expression within each module, we retrieved the single-cell data for human brain cell types from the PanglaoDB database (Franzén, Gan and Björkegren, 2019). PanglaoDB is a database that contains collections of single-cell experiments results and information of many tissues from human and mouse. The genes within each module were then compared to the marker genes for each brain cell type using the *GeneOverlap* R package v.1.20.0

(Shen, 2014). Fisher's exact test with an FDR-correction for p-values was used to analyze the gene overlap comparisons. To check the consistency of the results, another cell-type-specific expression dataset composed of five main brain cell types including neurons, astrocytes, oligodendrocytes, microglia, and endothelial cells was obtained from another single-cell RNAseq study(Zhang *et al.*, 2016). Gene symbols were mapped to Ensembl gene identifiers using the *biomaRt* R package. Specificity for the five brain cell types was calculated with the *specificity.index* function from *pSI* R package v.1.1 (Xu *et al.*, 2014). Fisher's exact test with FDR correction for p-values was applied to check for the significant cell-type specificity (FDR-corrected p-value <0.05 was considered statistically significant).

2-11-Brain enhancer RNAs co-expression analysis

In order to regulate the gene expression, there are many different types of regulatory factors involved at the level of epigenetics, genetics, transcriptional, and post-transcriptional level. Enhancer RNAs (eRNAs) are cis-regulatory elements in the genome that co-operate with promoters to regulate expression of target gene (Arnold, Wells and Li, 2019). Here, we aimed at searching for the relationship between gene co-expression modules and eRNA expression modules in the brain. An expression dataset for brain enhancer-RNAs (eRNA) was obtained from an independent study of human brain region-specific eRNAs co-expression analysis (Yao *et al.*, 2015). To estimate the co-expression of each gene module and each brain eRNA module Fisher's exact test was used. A P-value <0.05 followed by FDR correction was used to filter the significant coregulations.

2-12-Software and code availability

The R programming language version 3.5.0 (https://www.r-project.org/) was used for statistical analyses and data visualization. The functions and libraries used in this study are available as R packages: *WGCNA*, *nlme*, *RRHO*, *GeneOverlap*, *pSI*, *ggplot2, Rtsne, gprofiler2* at CRAN (http://cran.r-project.org/) and/or Bioconductor (https://bioconductor.org/). A list of packages and tools used in this study is provided in Table 2-2.

Table 2-2 | List of packages and tools used for data analysis and visualization

Package/tool	Usage	Ref.
WGCNA	Analysis of co-expression modules in a transcriptional dataset	(Langfelder and Horvath, 2008)
RRHO	A tool for measuring geometric overlap across gene expression datasets using pairwise comparisons	(Plaisier *et al.*, 2010)
GeneOverlap	Analysis of gene overlap between two or more expression datasets	(Shen, 2014)
pSI	Calculating specificity of gene expression for cell types	(Xu *et al.*, 2014)
ggplot2	Data visualization	(Wickham, 2011)
Rtsne	Analysis of t-SNE	(Krijthe, 2015)
caret	Machine learning and random forest	(Kuhn and Others, 2008)
igraph	Network visualization	(Csardi, Nepusz and Others, 2006)
gprofiler2	Gene ontology enrichment analysis	(Raudvere *et al.*, 2019)
nlme	Analysis of linear mixed effect model	(Pinheiro *et al.*, 2012)
limma	Linear model for differential gene expression analysis	(Ritchie *et al.*, 2015)
dendextend	Producing dendrogram and clustering tree	(Galili, 2015)
pheatmap	Heatmap visualization	(Kolde and Kolde, 2015)
ComplexHeatmap	Heatmap visualization	(Gu, Eils and Schlesner, 2016)

2-13-Data availability

2-13-1-**Raw data**

The raw data incorporated in this work were gathered from various resources. RNA-Seq raw data, metadata, and source files are available on the NCBI GEO database and Sage Synapse as described before.

2-13-2-**GitHub repository for the scripts**

The scripts used for the analyses in this thesis will be available at the following Github address: https://github.com/isadeghi87/Brain_RNAseq

Chapter3. Results

3-1- Sample processing and merging datasets

We analyzed 4711 RNA-Seq samples from patients with AD (n = 906 samples), PD (n = 29), PA (n = 58), PSP (n = 168), SZ (n = 535), ASD (n = 187), MDD (n = 240), BP (n = 510), and matched controls pooled across alls studies (n = 2078) (). These samples were obtained from 7 brain lobes. The majority of samples were collected from the frontal lobe (from DLPFC), temporal lobe and from the cerebellum.

The samples from each study were processed using the Grape RNA-Seq pipeline and underwent normalization and quality control. First, samples from each study were processed and the gene expression values were normalized and $\log_2$ transformed (as described in Materials and Methods). For each condition, the normalized expression datasets were then merged together. We performed statistical tests to see the difference of several covariates such as age (using ANOVA), sex (gender; using chi square), PMI (using ANOVA), and RIN (using ANOVA) between patients and controls groups.

3-1-1- **Alzheimer's data**

For AD, samples were obtained from three different studies. The number of AD samples were higher than that of controls. Statistical tests showed significant difference in age ($p = 2.5 \times 10^{-8}$), sex ($p = 9 \times 10^{-10}$), RIN ($p = 0.008$), PMI ($p = 0.003$), and study ($p = 2.3 \times 10^{-11}$). However, to correct for the effect of the variables, we included them in the model for differential gene expression. Multidimensional scaling (MDS) plots show that the samples are clustered into three groups which belong to three different studies. To remove the batch effect, we performed batch effect correction using the Combat function from the sva R package (described before). As is shown in Figure 3-2 after batch effect correction we do not see clustering of these three groups.

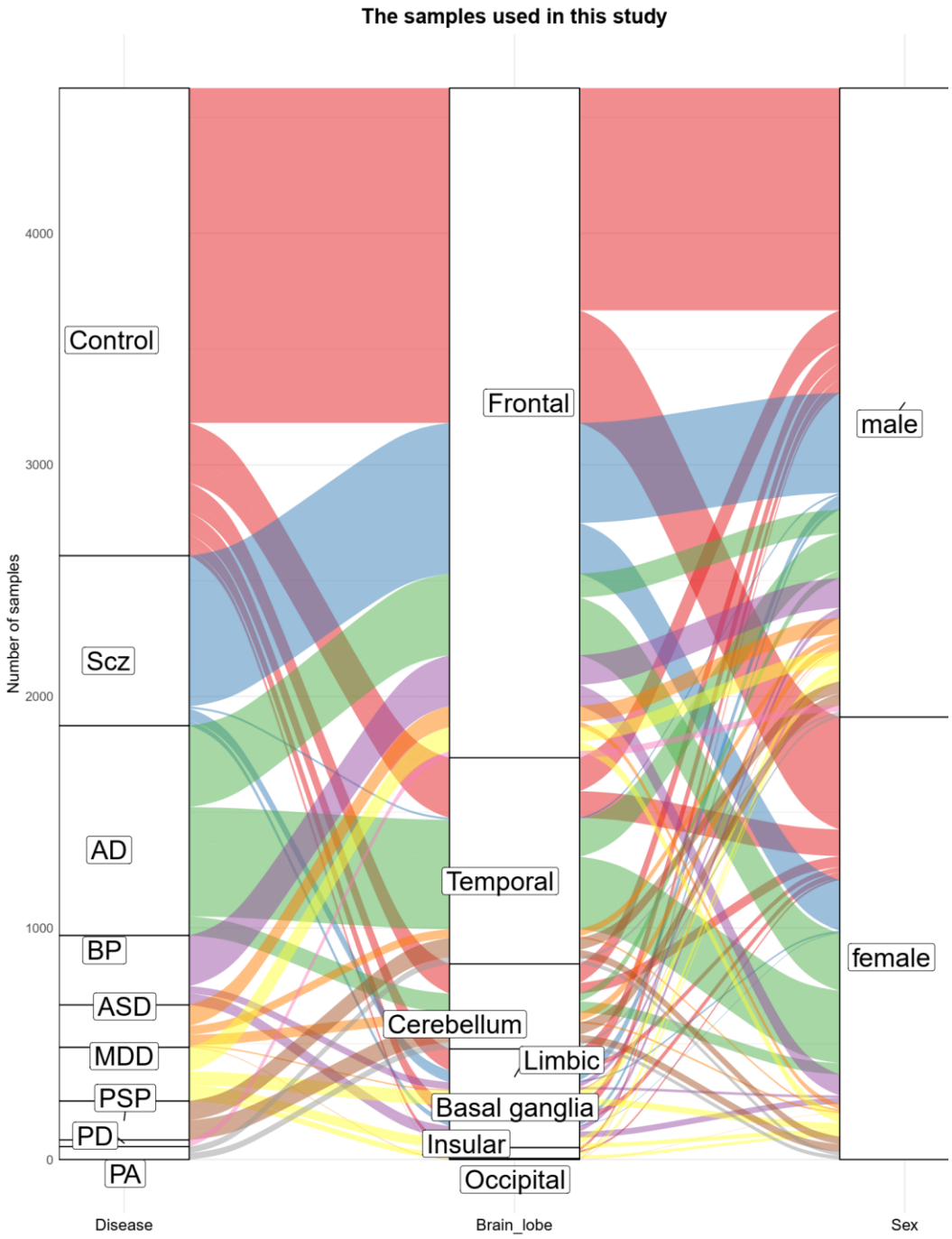

Figure 3-1 | A flowchart of the samples in this study. From left to right, the panels represent diagnosis, brain lobes, and brain subregions, with colors showing diagnosis.

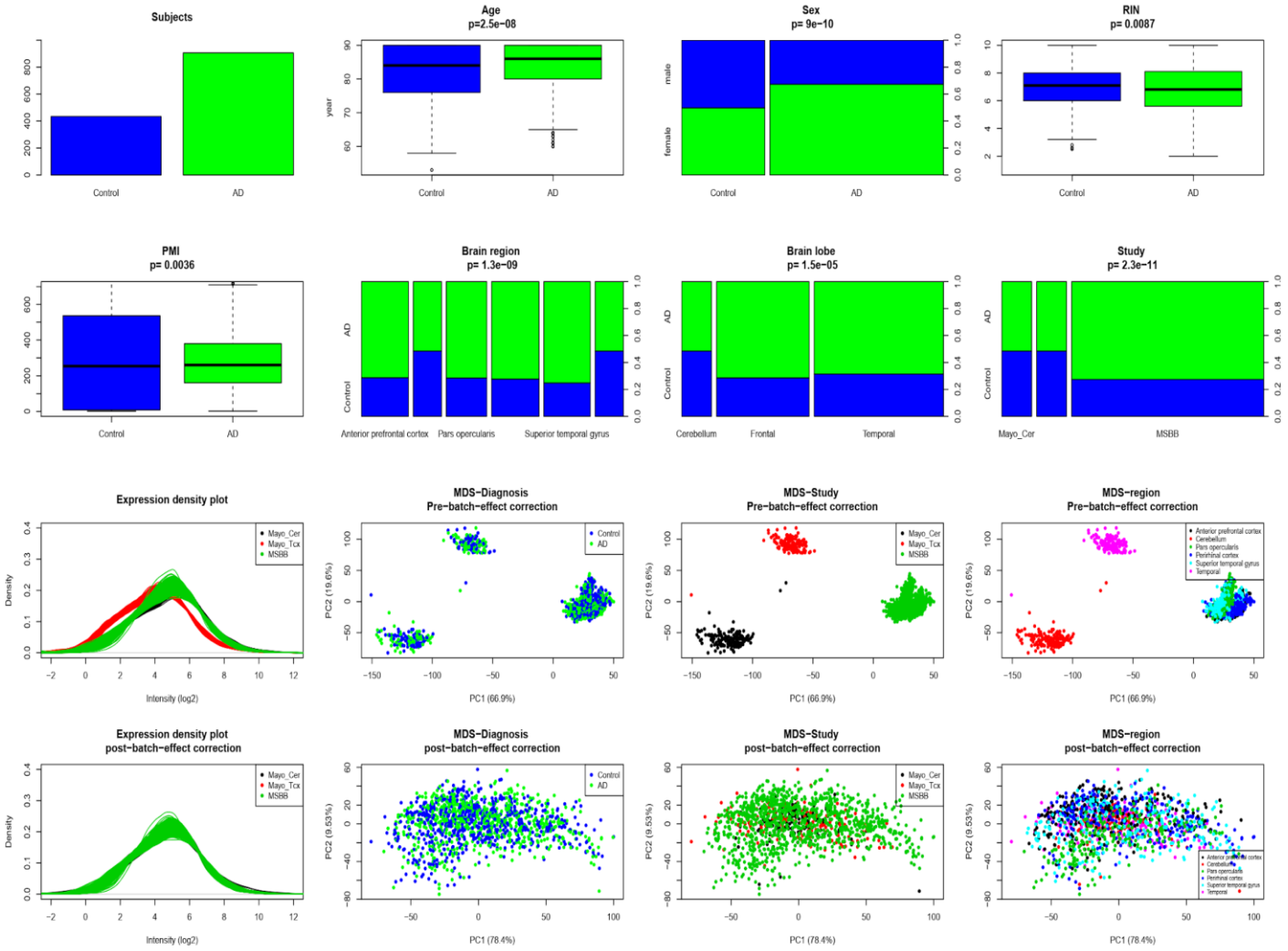

Figure 3-2. QC plots are shown for AD datasets. Normalized data from three cohorts analyzed in this study consisted of anterior prefrontal cortex, perirhinal cortex, superior temporal gyrus, pars opercularis, temporal cortex, and cerebellum brain samples from subjects with AD (n = 906) and controls (n = 479). Sample outliers were removed and batch-effects were corrected.

3-1-2- **Parkinson's data**

Samples for PD and matched controls were obtained from a single study. The number of samples for PD (n = 29) were lower than for controls (n = 44) (Figure 3-3). Statistical tests showed no significant differences between age ($p = 0.19$) and PMI ($p = 0.17$) of controls and PD patients, but a significant difference was observed for RIN ($p = 4.3\times10^{-5}$). MDS plot shows that the first principal component (PC1) catches most variability across the dataset.

3-1-3- **Pathological aging data**

Samples for PA and matched controls were obtained from two studies. The samples consisted of temporal cortex and cerebellum brain samples from subjects with PA (n =58) and controls (n =155) (Figure 3-4). The two groups showed significant differences in age ($p = 0.003$) and PMI ($p = 1\times10^{-4}$), but not in RIN ($p > 0.05$), sex ($p > 0.05$) and study ($p > 0.05$). MDS plot demonstrated separation of two clusters which contain samples from two different studies. Batch effect correction showed removal of the effect of study.

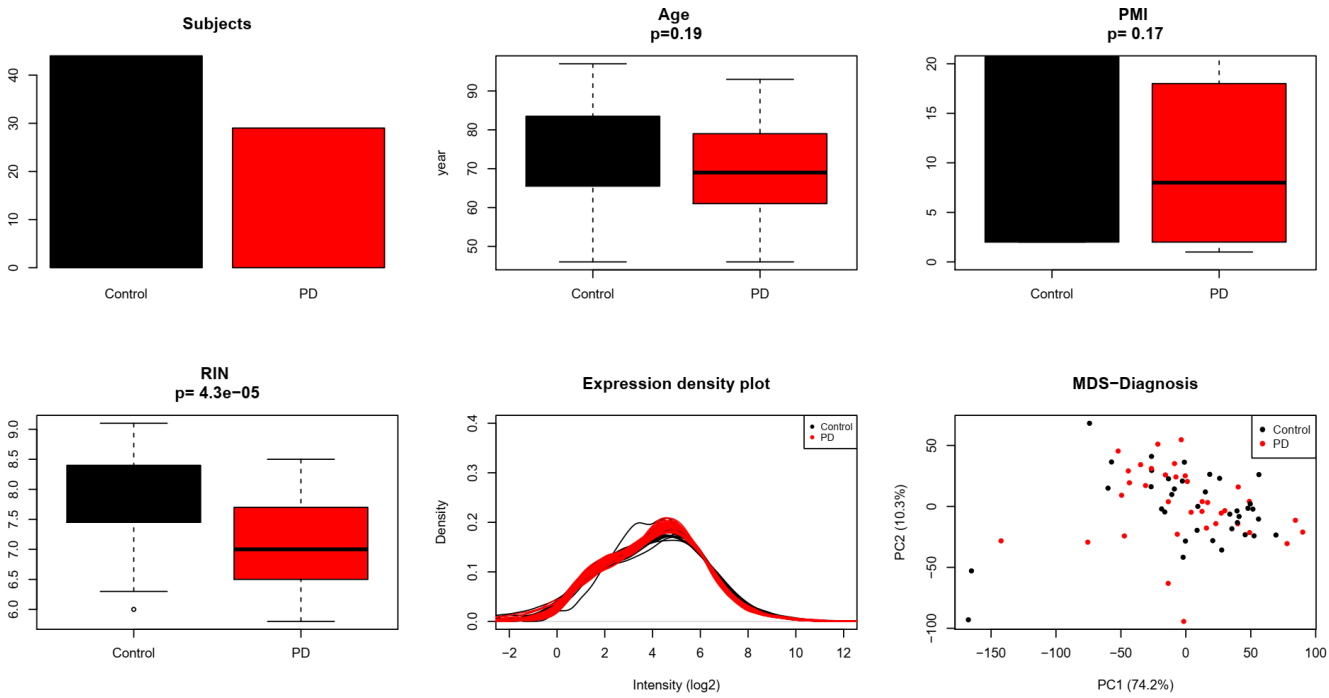

Figure 3-3 | QC plots for PD dataset. This dataset consisted of dorsolateral prefrontal cortex (BA9) brain samples from subjects with PD (n = 29) and controls (n = 44). Sample outliers were detected by standardized network connectivity z-scores and removed. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

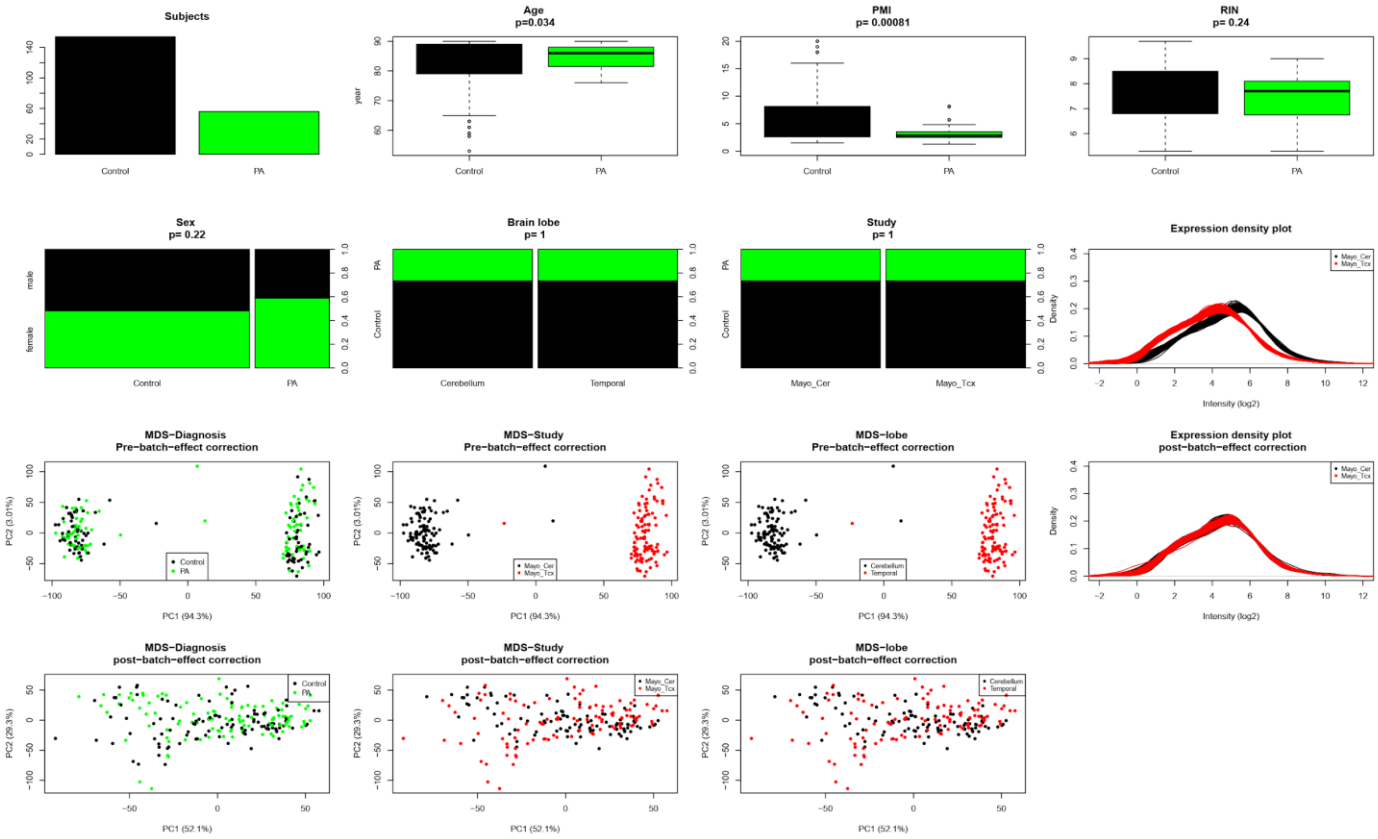

Figure 3-4 | QC plots are shown for PA datasets. Normalized data from two studies analyzed here consisted of temporal cortex and cerebellum brain samples from subjects with PA (n =58) and controls (n =155). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

3-1-4- **Pogressive supranuclear palsy data**

Like PA, PSP dataset was also obtained from two studies, consisting of temporal cortex and cerebellum brain samples from subjects with PSP (n = 168) and controls (n =155) (Figure 3-5). Between PA patients and controls we observed significant differences in age ($p = 4.9\times10^{-21}$), PMI ($p = 0.012$), and RIN ($p = 1\times10^{-15}$). Here, we also observed clustering of samples from two studies in MDS plot, which was removed after batch effect correction.

3-1-5- **Schizophrenia data**

For Scz, normalized data from nine studies analyzed here consisted of granular frontal area 9, DLPFC, anterior cingulate cortex, ventral anterior cingulate 25, amygdala, superior temporal gyrus, and nucleus accumbens brain samples from subjects with Scz (n = 535) and controls (n = 1172) (Figure 3-6). Age and sex did not show difference between Scz and control groups. But we observed a significant difference in study ($p = 1.5\times10^{-8}$) between the groups. MDS plots showed removal of batch effect from study.

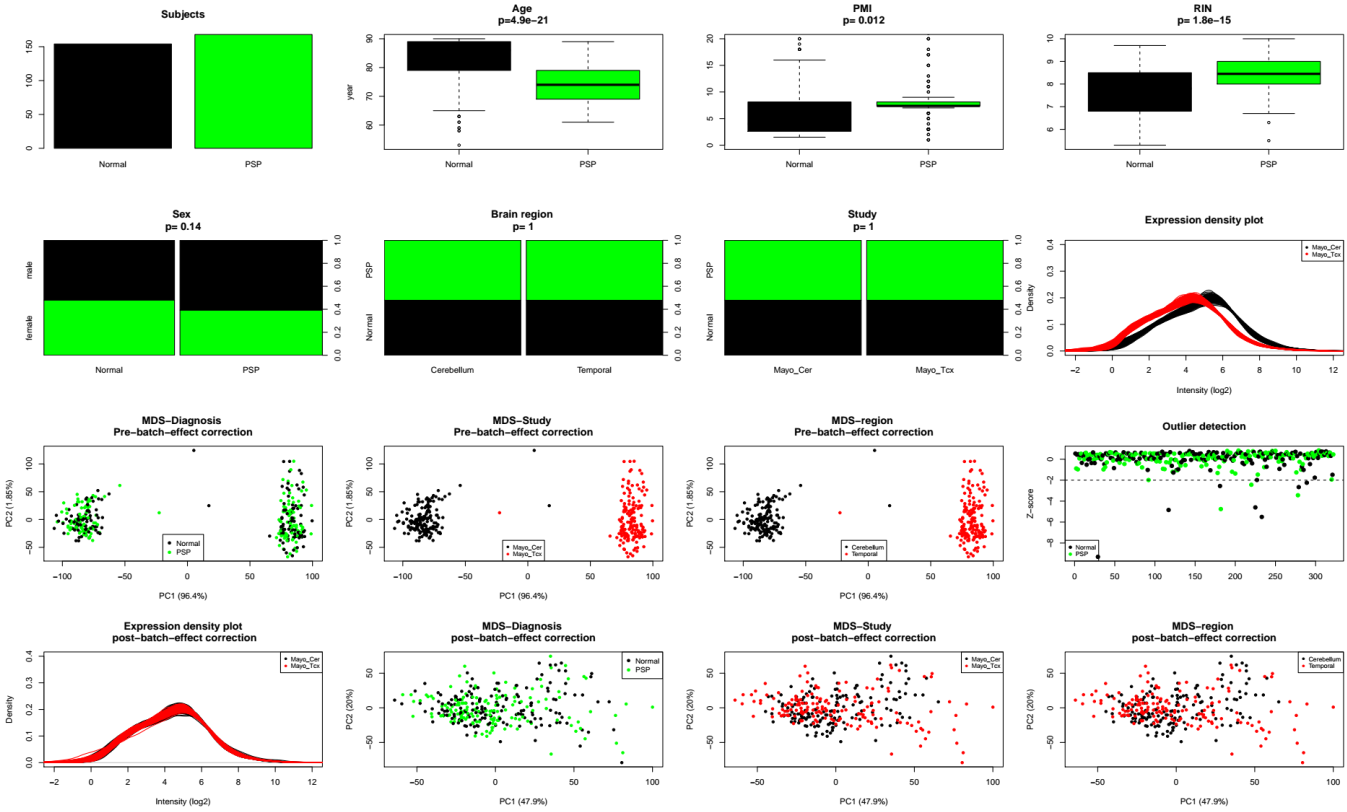

Figure 3-5 | QC plots are shown for PSP datasets. Normalized data from two studies analyzed here consisted of temporal cortex and cerebellum brain samples from subjects with PSP (n = 168) and controls (n =155). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

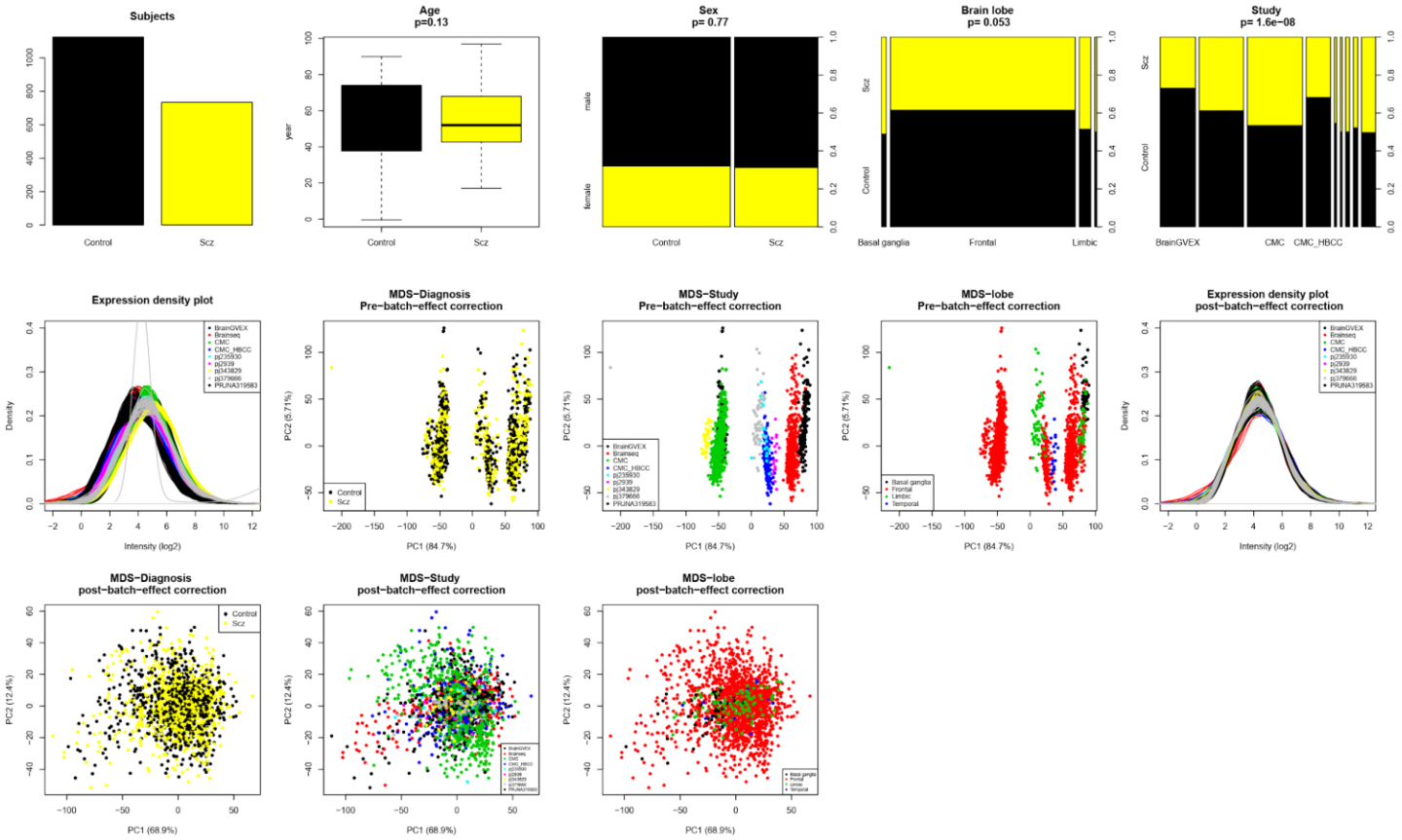

Figure 3-6 | QC plots shown for Scz datasets. Normalized data from nine studies analyzed here consisted of granular frontal area 9, DLPFC, anterior cingulate cortex, ventral anterior cingulate 25, amygdala, superior temporal gyrus, and nucleus accumbens brain samples from subjects with Scz (n = 535) and controls (n = 1172). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

3-1-6- **Austism data**

For ASD, samples obtained from five studies consisted of corpus callosum, inferior temporal cortex, auditory cortex, V1C (primary visual cortex), superior frontal gyrus, and DLPFC brain samples from subjects with ASD (n = 187) and controls (n = 239) (Figure 3-7). Statistical tests showed a difference in study ($p = 0.009$) between the groups.

3-1-7- **Major depressive disorder data**

MDD dataset was obtained from normalized data from three studies consisting of anterior insula (aINS), nucleus accumbens (Nac), DLPFC, orbitofrontal cortex (OFC), ventral subiculum (vSub), cingulate gyrus 25 (Cg25), anterior cingulate cortex, brain samples from subjects with MDD (n = 240) and controls (n = 221) (Figure 3-8). The number of samples were balanced between the groups, as well as for age, PMI and brain lobe. There was a significant difference between sex of MDD compared with controls ($p = 0.002$). Pre-batch effect MDS plots show clustering of samples from three studies. This effect was removed after batch effect correction as is presented in post-batch effect MDS plots.

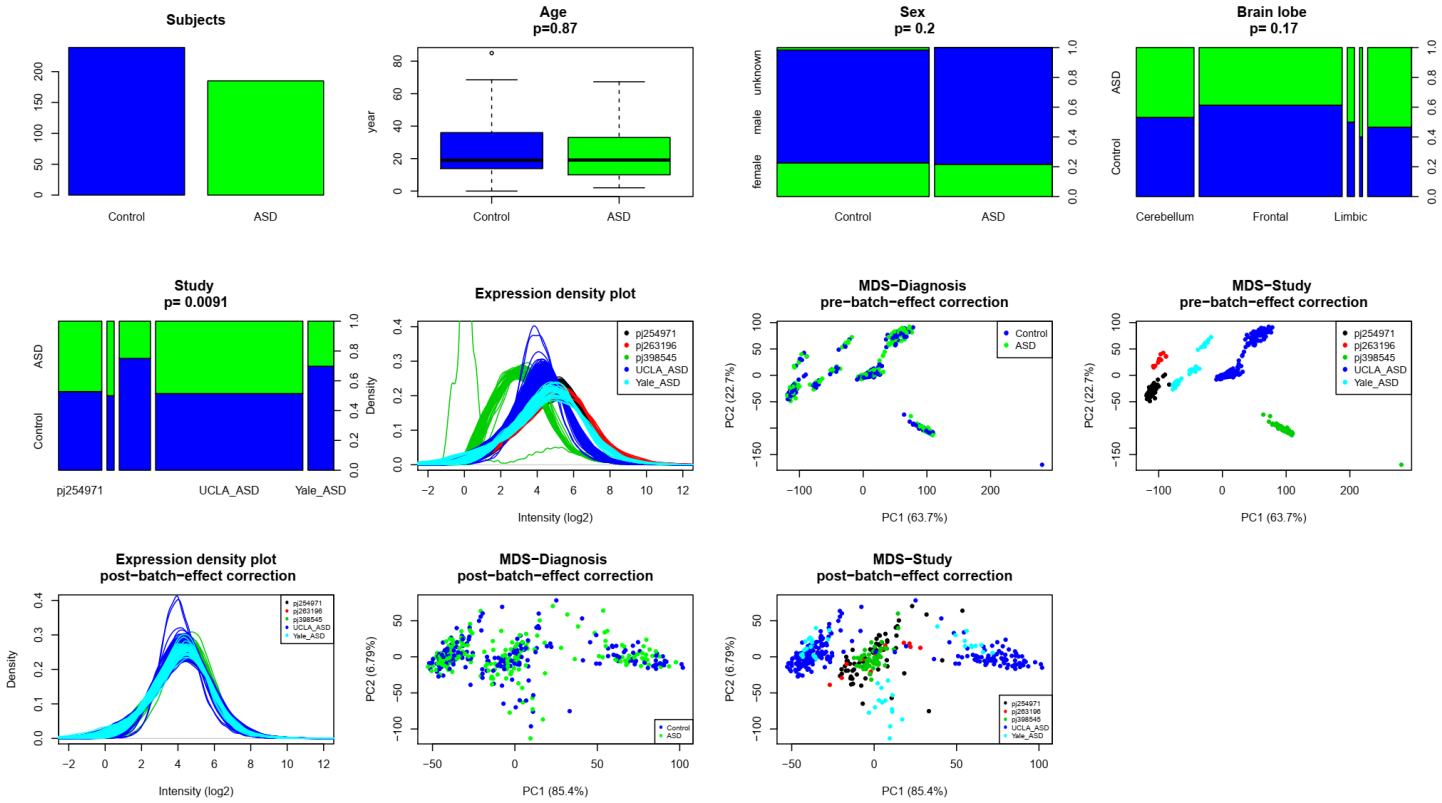

Figure 3-7 | QC plots are presented for ASD datasets. Normalized data from five studies analyzed here consisted of corpus callosum, inferior temporal cortex, auditory cortex, V1C (primary visual cortex), superior frontal gyrus, and DLPFC brain samples from subjects with ASD (n = 187) and controls (n = 239). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

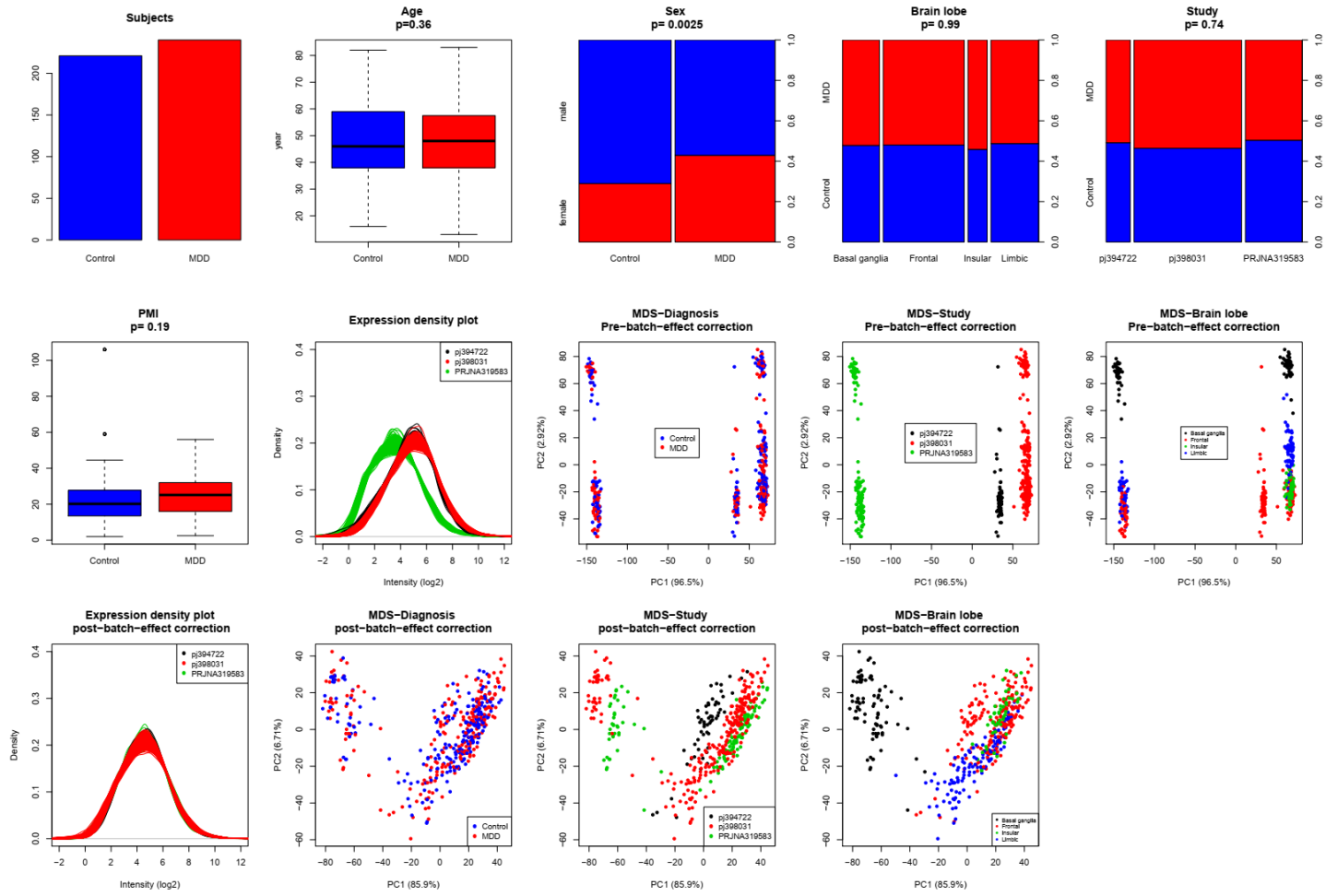

Figure 3-8 | QC plots are shown for MDD datasets. Normalized data from three studies analyzed here consisted of anterior insula (aINS), nucleus accumbens (Nac), DLPFC, orbitofrontal cortex (OFC), ventral subiculum (vSub), cingulate gyrus 25 (Cg25), anterior cingulate cortex, brain samples from subjects with MDD (n = 240) and controls (n = 221). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

3-1-8- **Bipolar disorder data**

BP dataset was obtained from combining normalized data from eight studies consisting of corpus callosum, inferior temporal cortex, auditory cortex, V1C (primary visual cortex), superior frontal gyrus, and DLPFC brain samples from subjects with BP (n = 511) and controls (n = 837) (Figure 3-9). Significant difference in age ($p = 1.8\times10^{-16}$), sex ($p = 0.02$), study ($p = 1.6\times10^{-19}$), and brain lobe ($p = 3.4\times10^{-13}$) between MDD patients and controls subjects. Batch effect of study was removed after correction as is shown in MDS plots.

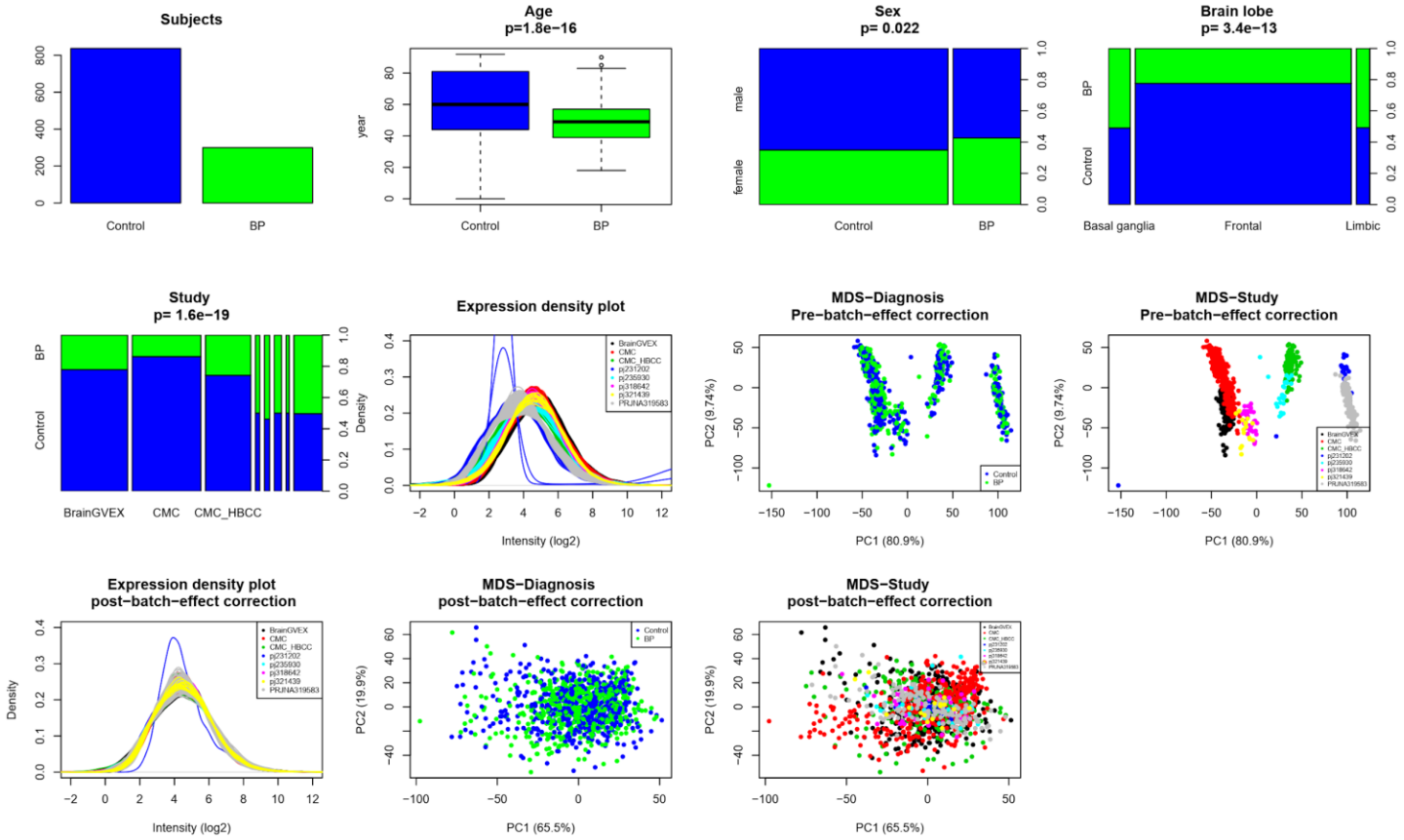

Figure 3-9 | A set of QC plots shown for BP datasets. Normalized data from eight studies analyzed here consisted of corpus callosum, inferior temporal cortex, auditory cortex, V1C (primary visual cortex), superior frontal gyrus, and DLPFC brain samples from subjects with BP (n = 511) and controls (n = 837). Sample outliers were detected by standardized network connectivity z-scores and removed. Batch effects were corrected for the studies. Multidimensional scaling (MDS) plots show sample clustering by the first two expression principal components. Groups were balanced by available covariates and potential confounding factors.

3-2- tSNE analysis and sample clustering

In order to see clustering of the samples which were merged from different datasets, tSNE analysis was used to visualize the variance across expression datasets. As described in **Chapter 2**, we first performed PCA analysis to reduce the dimensions. This analysis was performed in two ways: using shared genes across diseases (10313 genes) and using union of normalized genes (17997 genes). tSNE space was able to separate disease conditions into clusters. PA and PSP were clustered together which was likely due to their specific transcriptional architecture (Fig. 2). ASD was separated into three clusters that was explained by the variation across brain regions. Likewise, MDD, Scz and BP were separated into different clusters which was explained by variation across brain regions. t-SNE using union of normalized genes (Figure 3-10), showed more separation of phenotypes compared with the one obtained using shared genes (Figure 3-10). This is probably due to the genes present in one condition and absent in others.

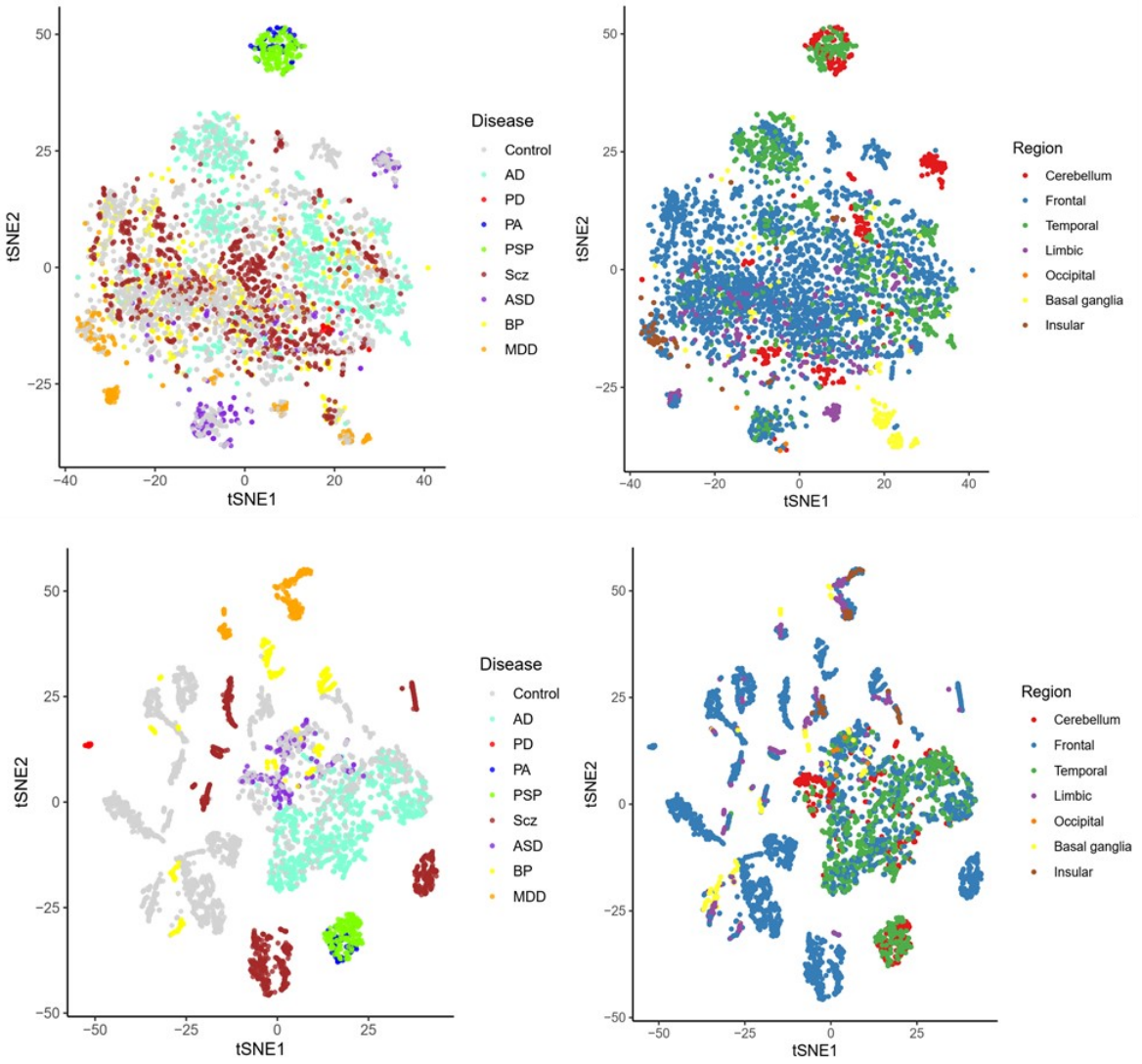

Figure 3-10 | Samples clustering across datasets. tSNE visualisation showing pooled samples, colored by disease condition (left) and brain regions (right). Plots on the top are obtained using 10313 shared genes across diseases, while plots at the bottom are obtained using union of 17997 normalized genes.

3-3- Disease-and brain region-specific DGE

Disease-specific DGE analyses were performed using a linear mixed-effects model. DGE results provided insights regarding transcriptional changes for the pathology of each disease. These results allowed us to determine the top differentially expressed genes (DEGs) across diseases (

Figure 3-11). PD, PSP, PA, and AD showed the highest number of DEGs (1188, 443, 123, and 81, respectively; $P < 0.05$), suggesting that NDDs underwent more transcriptional changes compared to NPDs. We observed that some of the genes were differentially expressed in at least seven diseases (Figure 3-12). Gene enrichment analysis performed using DEGs for each disease revealed perturbation of central nervous system development, mitochondrion, neuron part, axon mechanism, neuron development, and dendrite functions (empirical permutation test with false discovery rate (FDR)-corrected p-value < 0.05) (Figure 3-13).

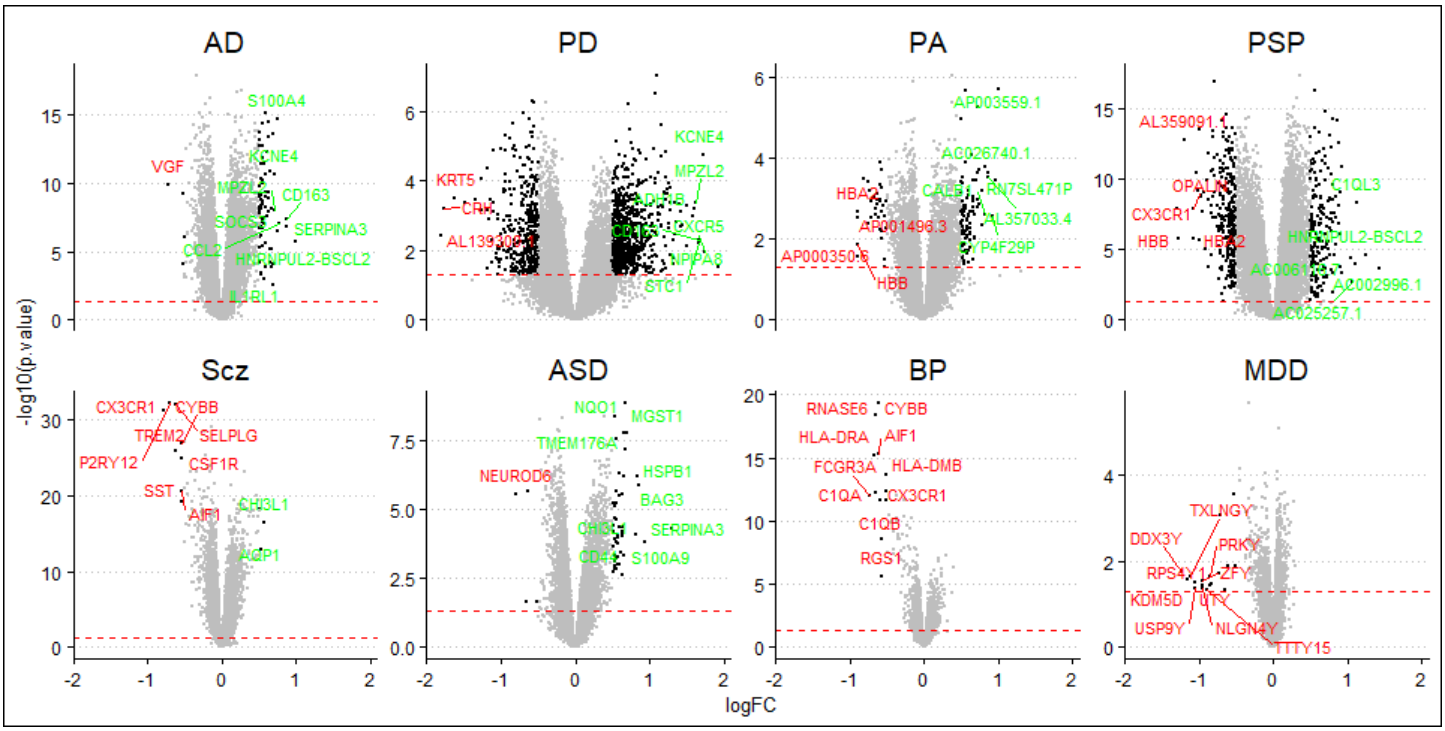

Figure 3-11 | Disease-specific DGE across eight diseases. Top differentially expressed genes (DEGs) are labeled for each disease compared to controls (green; upregulation, red; downregulation) ($P < 0.05$).

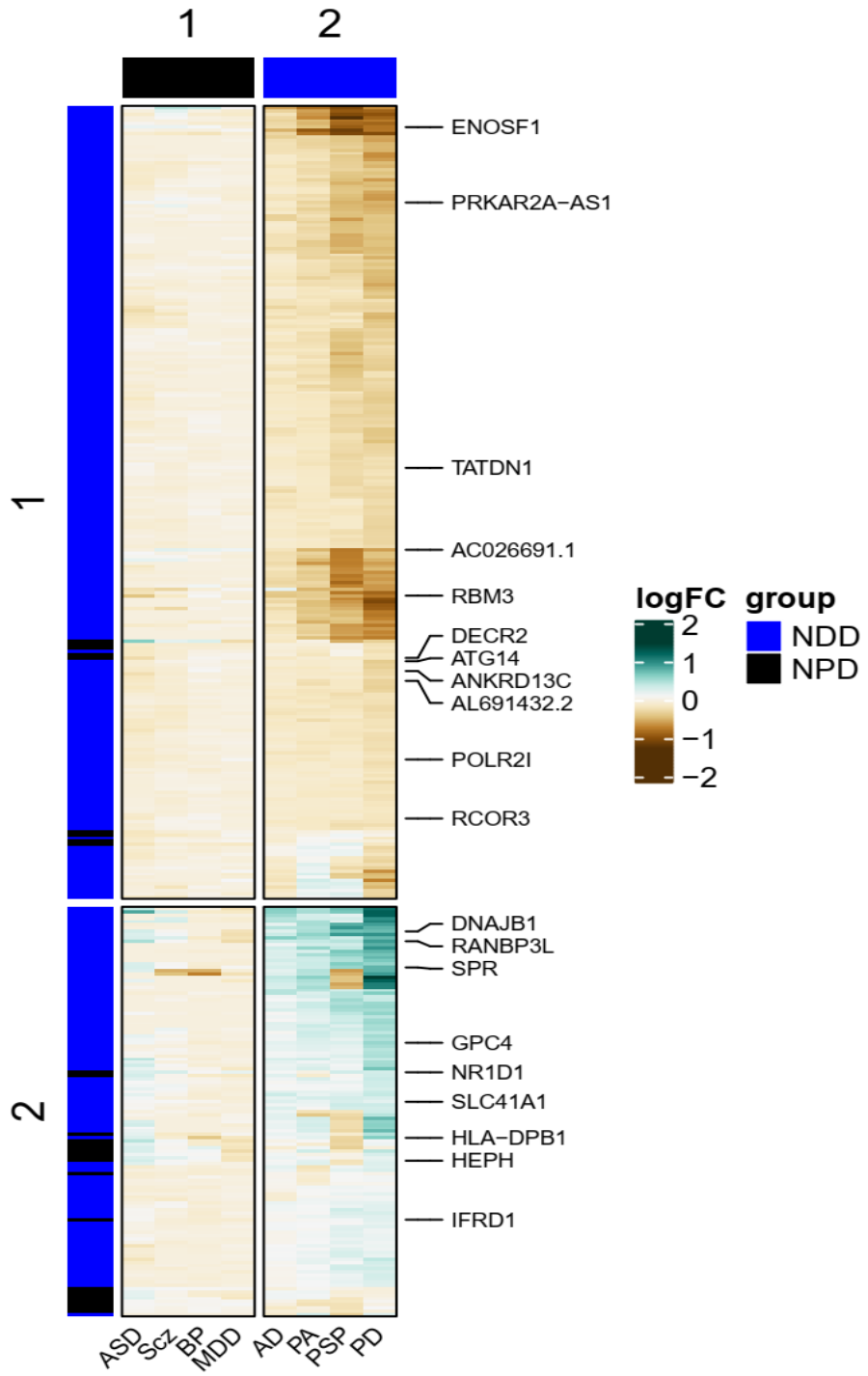

Figure 3-12 | A heatmap of differentially expressed genes across neurodegenerative disorders (NDD) and neuropsychiatric disorders (NPD). Each set of genes are differentially expressed either across NDD (blue rows) or across NPD (black rows). The row labels represent the genes differentially expressed in at least 7 diseases. ($P < 0.05$)

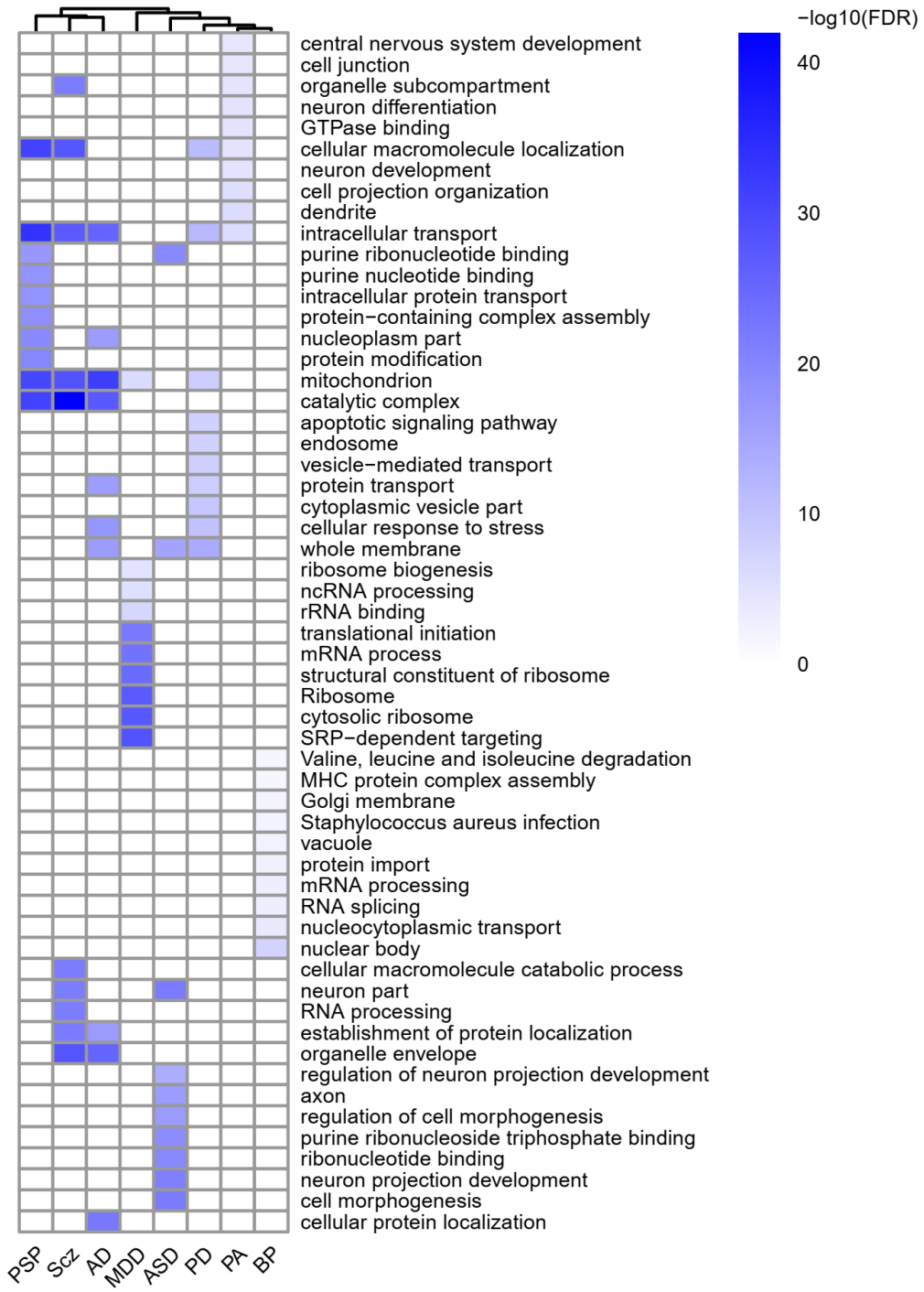

Figure 3-13 | Disease-specific gene enrichment analysis. Top significantly enriched pathways are represented for significantly differentially expressed genes across diseases (FDR-corrected p-value < 0.05).

3-4- Cortical-specific differential gene expression

Moreover, brain-region-specific DGE revealed the cerebellum had a high transcriptional alteration in AD, PSP, and PA (Figure 3-14 a, $P < 0.05$). Despite having small log_2FC values, temporal and frontal lobe showed broad changes across multiple diseases (Figure 3-14 and Figure 3-15). Limbic lobe showed greater changes for ASD and SZ.

3-4-1- Top overlapping differentially expressed genes

In order to see which genes are more differentially expressed across brain regions in different diseases, we compared DEGs obtained for each region in each disease. Top overlapping genes that were significantly differentially expressed across brain regions were *CX3CR1* with a developmental role in the migration of microglia (Gyoneva *et al.*, 2019), *CHI3L1* which activates astrocytes following amyloid-beta aggregation in the brain(Muszyński *et al.*, 2017; Gyoneva *et al.*, 2019), *NPAS4* with an inhibitory role in synapse development and synaptic plasticity (Spiegel *et al.*, 2014), *SERPINA3* and *BAG3* which are associated with PD, AD, and other neuropathies as they enhance the formation of amyloid-fibrils in disease(Kamboh *et al.*, 2006; Lei, Brizzee and Johnson, 2015; Cao *et al.*, 2017), and NR4A1 (Rothe *et al.*, 2017), a regulator of microglia activation (Figure 3-16).

a

	Cerebellum	Temporal	Frontal	Limbic	Occipital	Basal.ganglia	Insular
AD	844	248	10	–	–	–	–
ASD	12	387	146	75	33	–	–
BP	–	–	21	44	–	26	–
MDD	–	–	6	2	–	4	19
PA	92	1063	–	–	–	–	–
PD	–	–	1188	–	–	–	–
PSP	1228	561	–	–	–	–	–
Scz	–	28	15	112	–	80	–

b

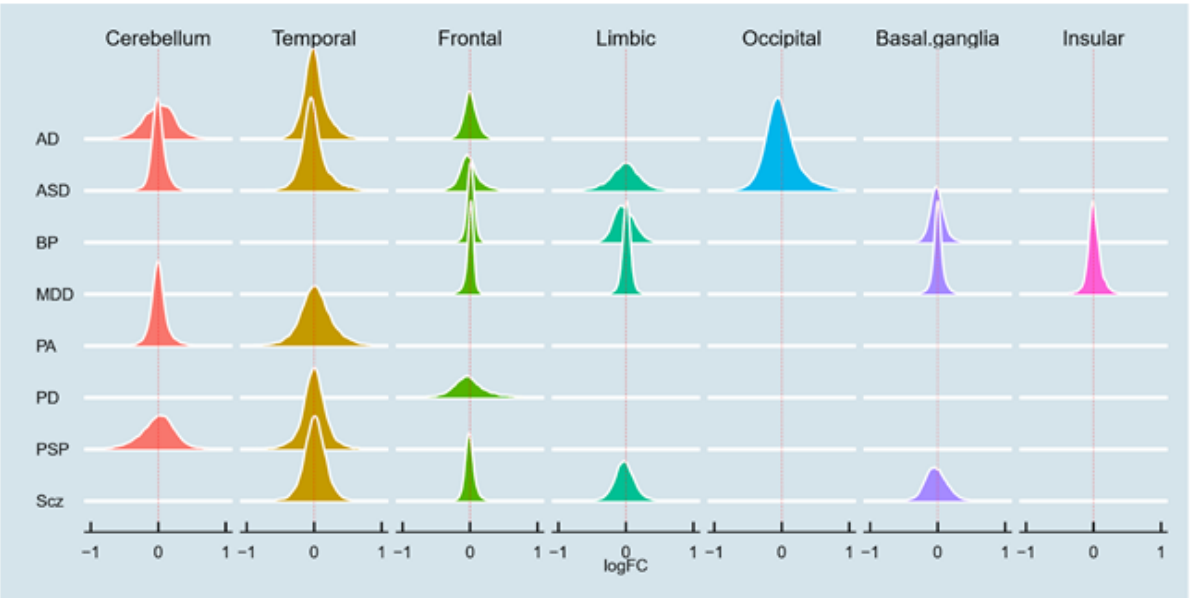

Figure 3-14 | The number of region-specific differentially expressed genes (DEGs) corresponding to each disease ($P < 0.05$). (**b**) Density plot illustrating regional gene expression (log_2FC) patterns for each disease.

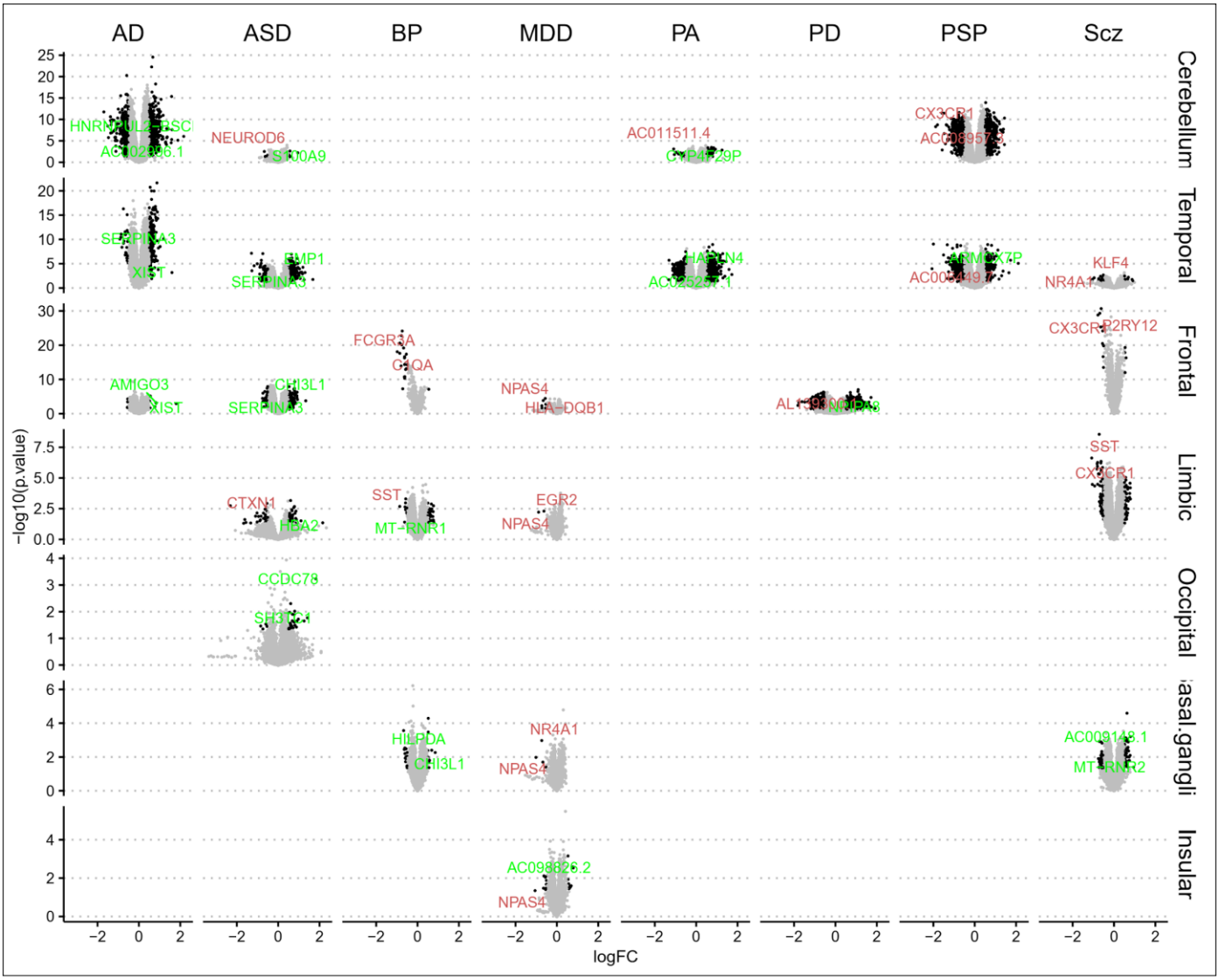

Figure 3-15 | Region-specific DGE. Top differentially expressed genes (DEGs) are labeled for each region across diseases (P < 0.05).

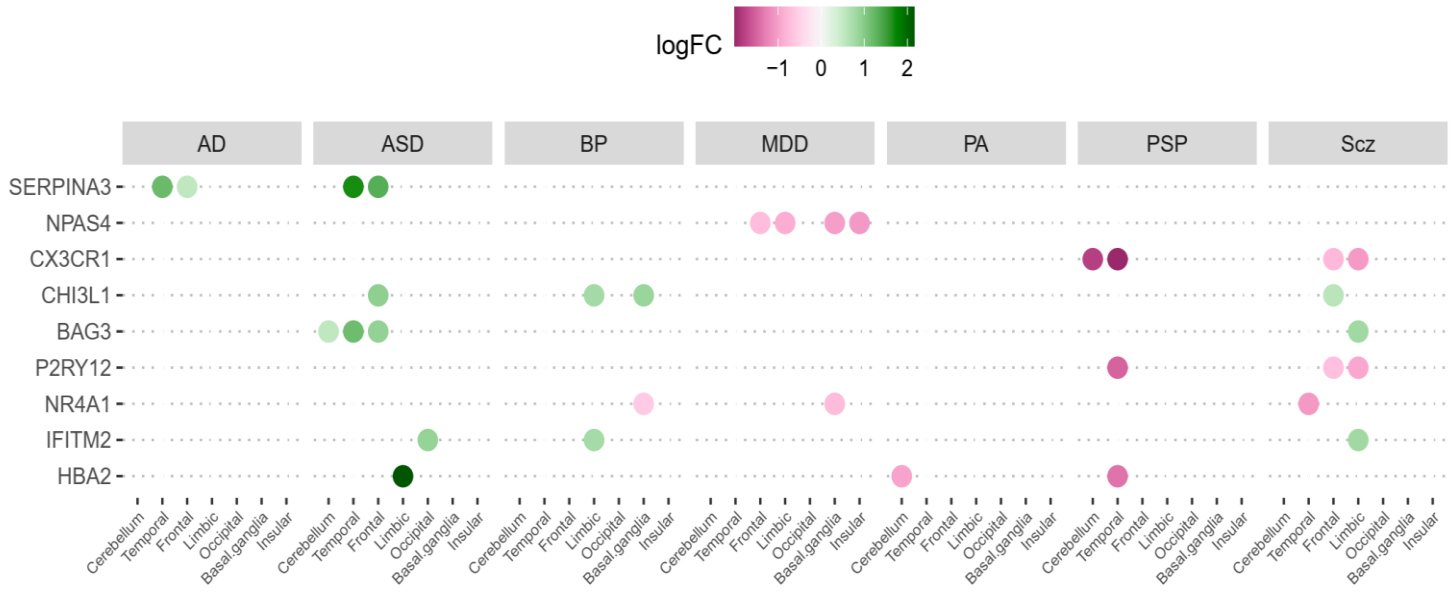

Figure 3-16 | Top overlapping genes with differential expression across diseases, with size representing frequency of overlap and color showing $\log_2$FC value.

3-5- Prediction accuracy of regional transcriptome

To assess which brain regions transcriptomic profiles were able to discriminate better between disease and control samples, we built classifier models using the expression of DEGs. High single-region accuracy for a disease was observed for temporal in Scz (88%), ASD (86%), PSP (80%) and AD (79%), for frontal in PD (86%) and for cerebellum in AD (80%) and PSP (79%) (Figure 3-17). Diseases with highest mean accuracy across regions were PD (86%), ASD (85%) and PSP (79%). MDD had the lowest accuracy (67%), suggesting its low transcriptional changes across diseases compared to controls (Figure 3-17 and Figure 3-18).

As is shown in Figure 3-18, the number of samples for some comparisons such as occipital and limbic lobe in ASD is very low. Accordingly, the values obtained for prediction accuracy and other factors are not reliable.

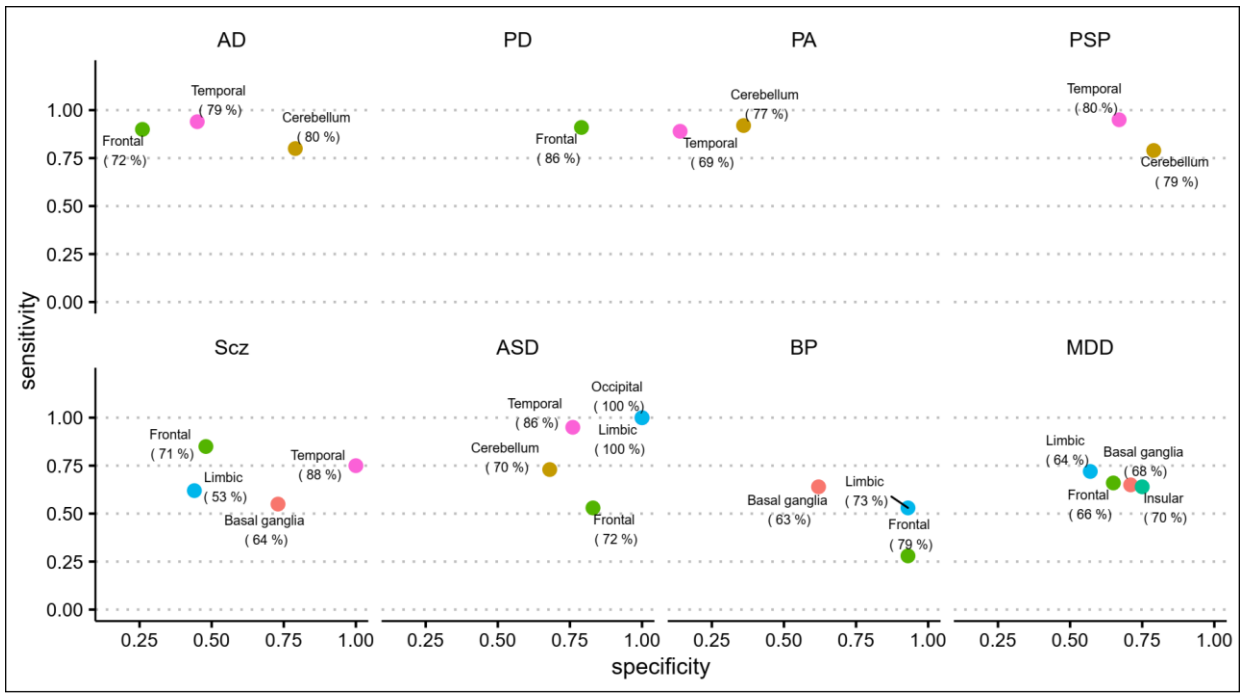

Figure 3-17 | Power of each region to discriminate between transcriptomes of disease and control subjects. Each model was built using normalized expression of differentially expressed genes for each region. x-axis shows specificity and y-axis shows sensitivity of the model. Prediction accuracy (accu) of regions across diseases is shown.

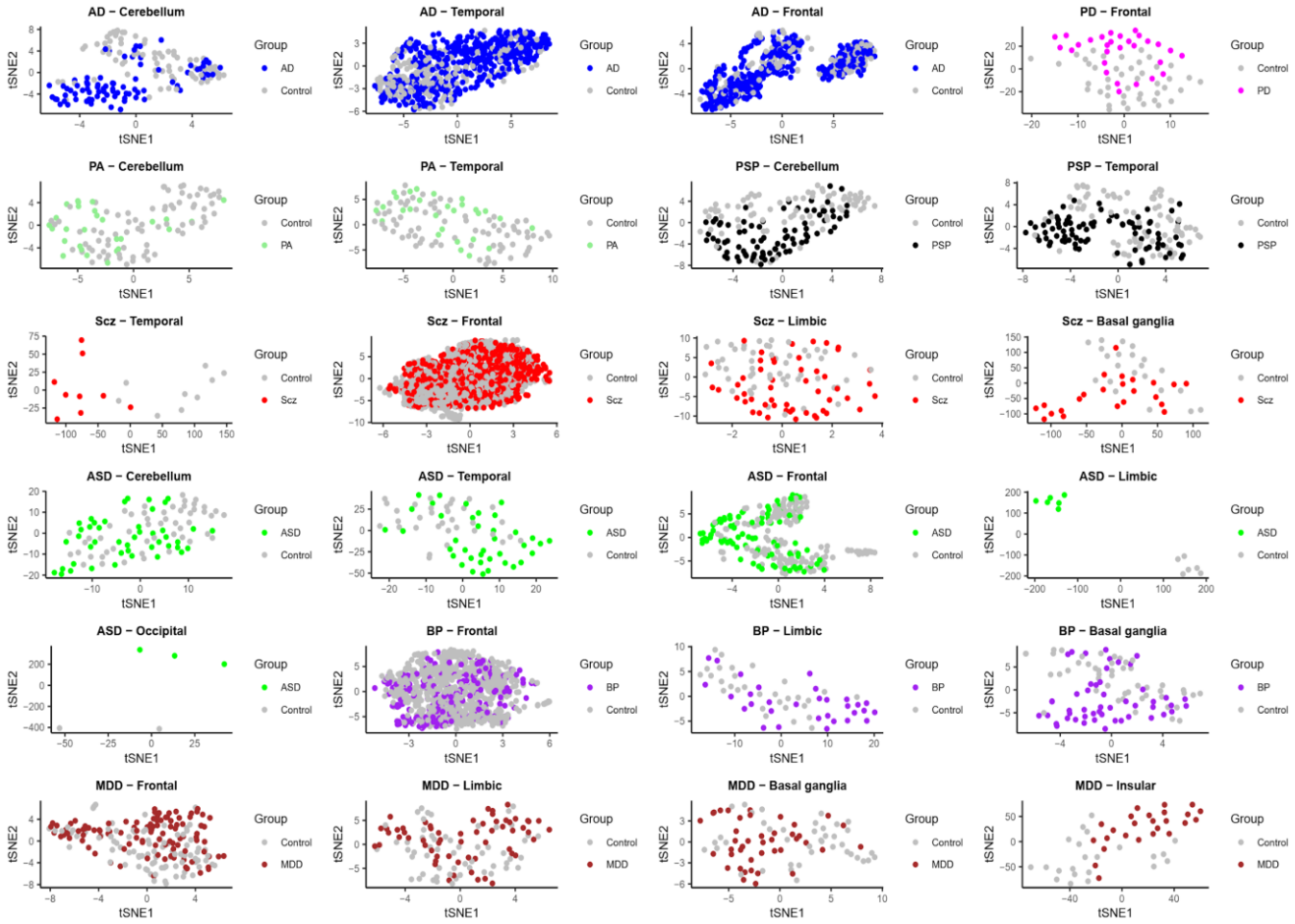

Figure 3-18 | tSNE visualization of gene expression of brain regions across diseases compared to controls.

Table 3-1 | Prediction accuracy of regional transcriptomes for each disease

Disease	region	accuracy	kappa	sensitivity	specificity
AD	Frontal	0.72	0.18	0.9	0.26
AD	Cerebellum	0.8	0.59	0.8	0.79
AD	Temporal	0.79	0.44	0.94	0.45
PD	Frontal	0.86	0.7	0.91	0.79
PA	Temporal	0.69	0.05	0.89	0.14
PA	Cerebellum	0.77	0.32	0.92	0.36
PSP	Cerebellum	0.79	0.57	0.79	0.79
PSP	Temporal	0.8	0.6	0.95	0.67
Scz	Frontal	0.71	0.35	0.85	0.48
Scz	Limbic	0.53	0.06	0.62	0.44
Scz	Temporal	0.88	0.75	0.75	1
Scz	Basal ganglia	0.64	0.27	0.55	0.73
ASD	Frontal	0.72	0.38	0.53	0.83
ASD	Temporal	0.86	0.72	0.95	0.76
ASD	Cerebellum	0.7	0.41	0.73	0.68
ASD	Occipital	1	1	1	1

ASD	Limbic	1	1	1	1
BP	Limbic	0.73	0.47	0.53	0.93
BP	Frontal	0.79	0.26	0.28	0.93
BP	Basal ganglia	0.63	0.27	0.64	0.62
MDD	Basal ganglia	0.68	0.36	0.65	0.71
MDD	Frontal	0.66	0.31	0.66	0.65
MDD	Insular	0.7	0.39	0.64	0.75
MDD	Limbic	0.64	0.29	0.72	0.57

AD, Alzheimer's disease; PD, Parkinson's disease; PSP, progressive supranuclear palsy; PA, pathological aging; Scz, schizophrenia; ASD, autism spectrum disorder; MDD, majore depressive disorder; BP, bipolar disorder.

3-6- Differential gene expression (DGE) analyses reveal cross-disease transcriptome overlap

Log_2FC values of 10,313 genes that were common across all diseases were used for cross-disease transcriptome overlap analyses. Cohen's *d* effect sizes were also estimated based on the correlation statistics obtained from pairwise comparisons. The results demonstrated significant transcriptome overlaps across diseases with the highest positive correlations of gene expressions between AD-PD, ASD-SZ, ASD-PD, PA-PSP, AD-SZ, AD-ASD, and SZ-PD (Spearman's ρ values ≥ 0.2, FDR-corrected $P < 0.05$, Cohen's $d > 0.2$) (Figure 3-19 and Table 3-2). BP did not reveal significant transcriptome correlation with other diseases except with MDD (Spearman's ρ = 0.13, Cohen's $d = 0.25$). This is likely due to smaller Log_2FC values in BP transcriptome (Fig. 3a).

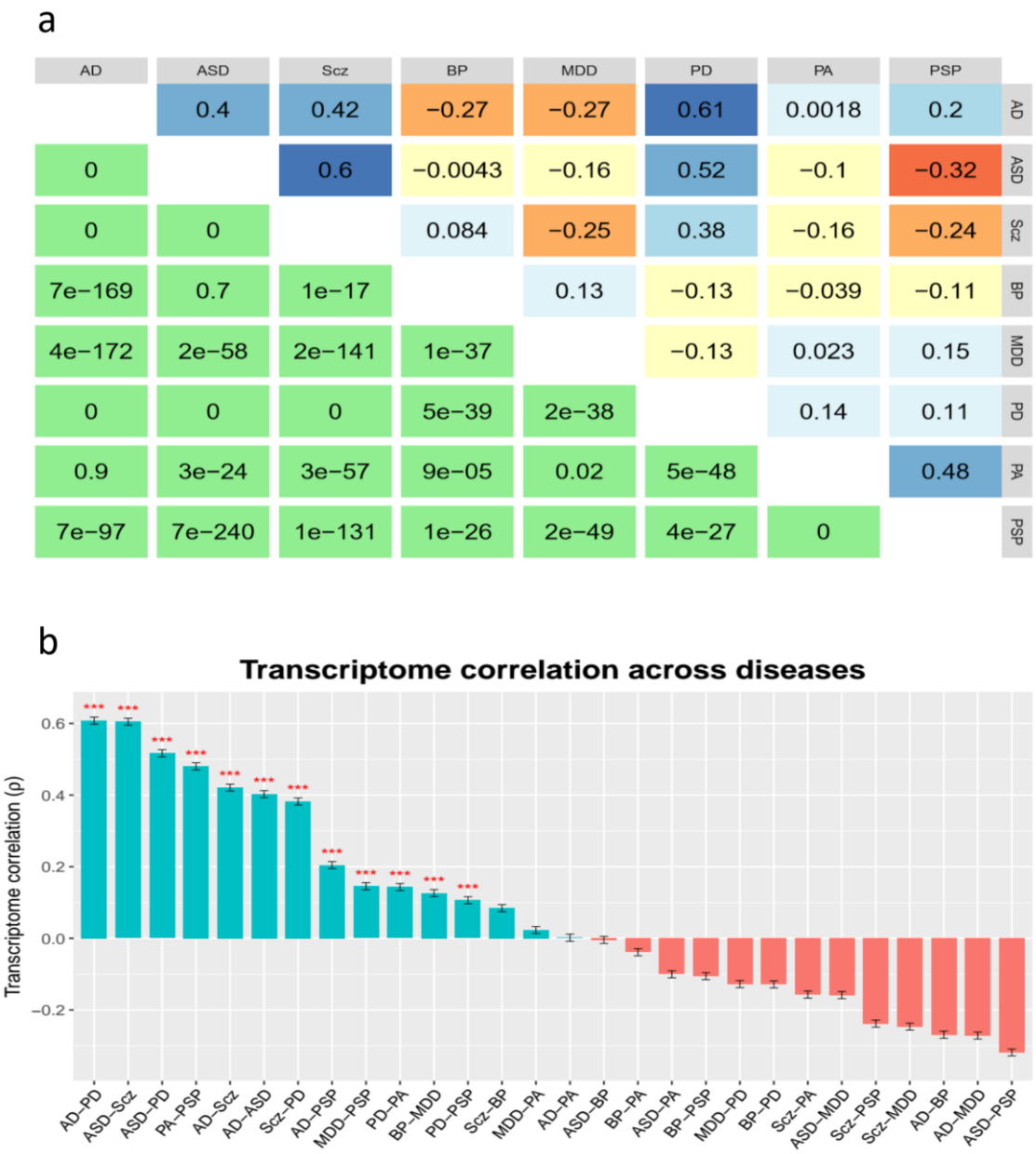

Figure 3-19 | Transcriptome similarities across diseases. (a) Transcriptome signature overlap obtained by Spearman's correlations using log2FC values of the shared genes across diseases. Upper panel shows correlation $\boldsymbol{\rho}$ values and the lower panel represents FDR-corrected p-values. (b) Barplot shows top pairwise transcriptome correlation across diseases measured by Spearman's correlation of log2FC values from common genes (* FDR-correct $P < 0.05$).

Table 3-2 | Pairwise transcriptome correlations across diseases.

	Correlation	SEM	FDR	Cohen_d	significance
AD-ASD	0.402	0.010	0	0.88	***
AD-Scz	0.421	0.010	0	0.93	***
AD-BP	-0.269	0.010	2E-168	-0.56	***
AD-MDD	-0.272	0.010	1E-171	-0.56	***
AD-PD	0.608	0.010	0	1.53	***
AD-PA	0.002	0.010	0.8533	0.00	
AD-PSP	0.205	0.010	1.4E-96	0.42	***
ASD-Scz	0.605	0.010	0	1.52	***
ASD-BP	-0.004	0.010	0.6899	-0.01	
ASD-MDD	-0.158	0.010	4.0E-58	-0.32	***
ASD-PD	0.517	0.010	0	1.21	***
ASD-PA	-0.100	0.010	3.3E-24	-0.20	***
ASD-PSP	-0.319	0.010	2.6E-239	-0.67	***
Scz-BP	0.084	0.010	1.5E-17	0.17	***
Scz-MDD	-0.247	0.010	5.2E-141	-0.51	***
Scz-PD	0.382	0.010	0	0.83	***

Scz-PA	-0.157	0.010	4.8E-57	-0.32	***
Scz-PSP	-0.238	0.010	3.1E-131	-0.49	***
BP-MDD	0.126	0.010	1.8E-37	0.25	***
BP-PD	-0.129	0.010	8.6E-39	-0.26	***
BP-PA	-0.039	0.010	0.0001	-0.08	***
BP-PSP	-0.106	0.010	1.3E-26	-0.21	***
MDD-PD	-0.128	0.010	2.4E-38	-0.26	***
MDD-PA	0.023	0.010	0.0208	0.05	*
MDD-PSP	0.146	0.010	3.1E-49	0.29	***
PD-PA	0.143	0.010	8.3E-48	0.29	***
PD-PSP	0.106	0.010	5.8E-27	0.21	***
PA-PSP	0.480	0.010	0	1.09	***

AD, Alzheimer's disease; PD, Parkinson's disease; PSP, progressive supranuclear palsy; PA, pathological aging; Scz, schizophrenia; ASD, autism spectrum disorder; MDD, majore depressive disorder; BP, bipolar disorder.

3-6-1- **Geometric overlap analysis**

In addition, an unbiased rank-rank hypergeometric overlap (RRHO (Plaisier *et al.*, 2010)) analysis was performed to highlight shared transcriptional changes across diseases. RRHO results revealed higher overlap of downregulated genes than upregulated ones across multiple diseases (Fisher's exact test with FDR < 0.05, Figure 3-20). PD demonstrated an overlap of both downregulated and upregulated transcriptome with AD, ASD, and SZ (Fisher's exact test, FDR < 0.05). The observed overlap between PSP and AD was mainly due to coordinated upregulation of the transcriptome (Figure 3-20).

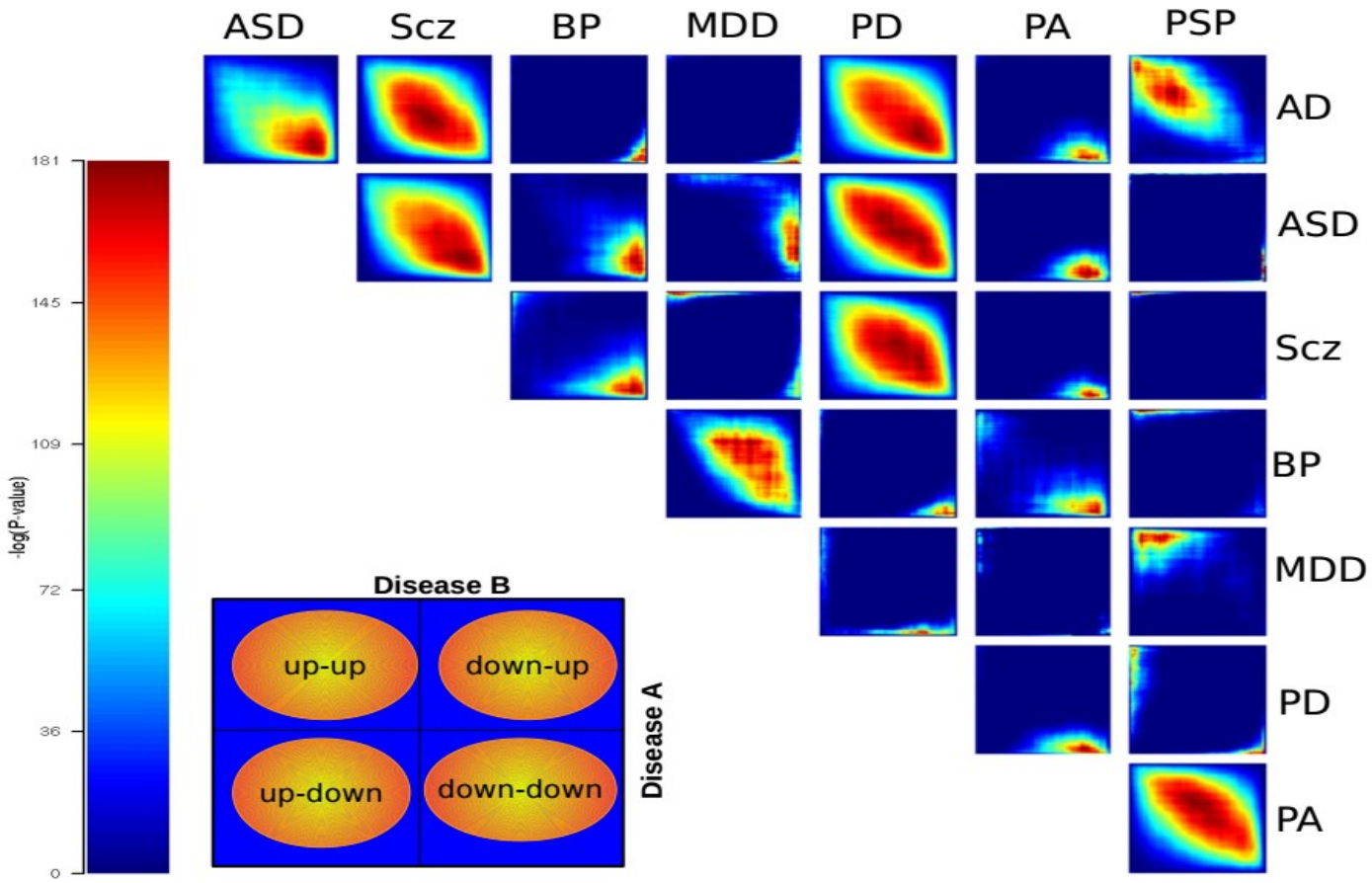

Figure 3-20 | RRHO maps of pairwise transcriptional comparisons across diseases. The bottom left guide panel represents the cross-disease overlapping relationship. Signals in the upper left quadrant display an overlap for shared upregulated genes, while signals in the bottom right quadrant depict shared downregulated genes. The color bar displays the degree of significance of the overlap between two diseases (Fisher's exact test with FDR < 0.05).

3-7- multidimensional scaling analysis

Moreover, a multidimensional scaling (MDS) analysis was performed by measuring distances from correlations of expressions (log_2FC) of shared genes across diseases. AD showed high relationships with other diseases including PD, ASD, SZ, and PSP (Figure 3-21). Notably, PA showed proximity with PSP and PD. Although the global structure of the network reveals a transcriptome relationship between NPDs (ASD, SZ, BP, and MDD) on one side and between NDDs (AD, PD, PSP, and PA) on the other side, ASD and SZ shared significant correlations with both NPDs and NDDs (Figure 3-21).

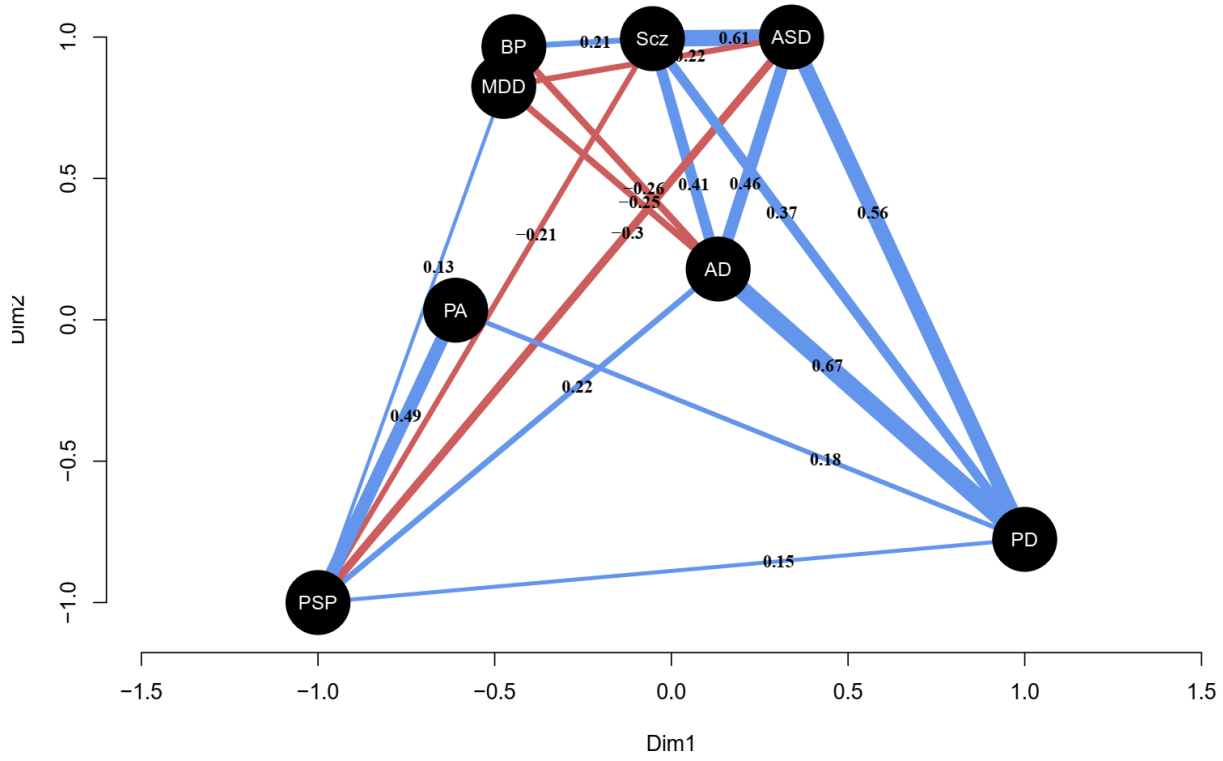

Figure 3-21 | Multidimensional scaling plot displays cross-disease transcriptome relationships using pairwise Spearman's correlations. Only significant transcriptome correlations after FDR correction are displayed. Nodes represent disease and edges depict the strength of the pairwise correlation. The width of the edges increases with the absolute $\boldsymbol{\rho}$ values.

3-8- Brain regions demonstrated transcriptome similarity across diseases

Brain region-specific DGE for each disease was performed using linear mixed effect models. Transcriptome comparisons for each region across diseases were performed using log_2FC values for the genes shared between the diseases. The limbic lobe demonstrated the highest correlation between SZ and BP (Spearman's $\rho = 0.9$, Figure 3-22, Figure 3-23 and Figure 3-24). The cerebellum revealed significant transcriptome correlation across AD, PSP, and PA (Fisher's exact test, FDR < 0.05) and not for ASD, suggesting its possible involvement in NDDs. On the other hand, the frontal and temporal lobe showed significant transcriptome overlap across AD, ASD, SZ, and PD (FDR < 0.05, Figure 3-22), indicating the impairment of the regions in a broader spectrum of diseases. A significant transcriptional overlap was also observed for basal ganglia between SZ and MDD and BP (FDR < 0.05, Fig. 4d). This suggests the involvement of basal ganglia in mood and psychotic disorders (Hwang *et al.*, 2006; Liu *et al.*, 2019; Macpherson and Hikida, 2019). These transcriptome similarities suggest engagement of each brain region in multiple diseases.

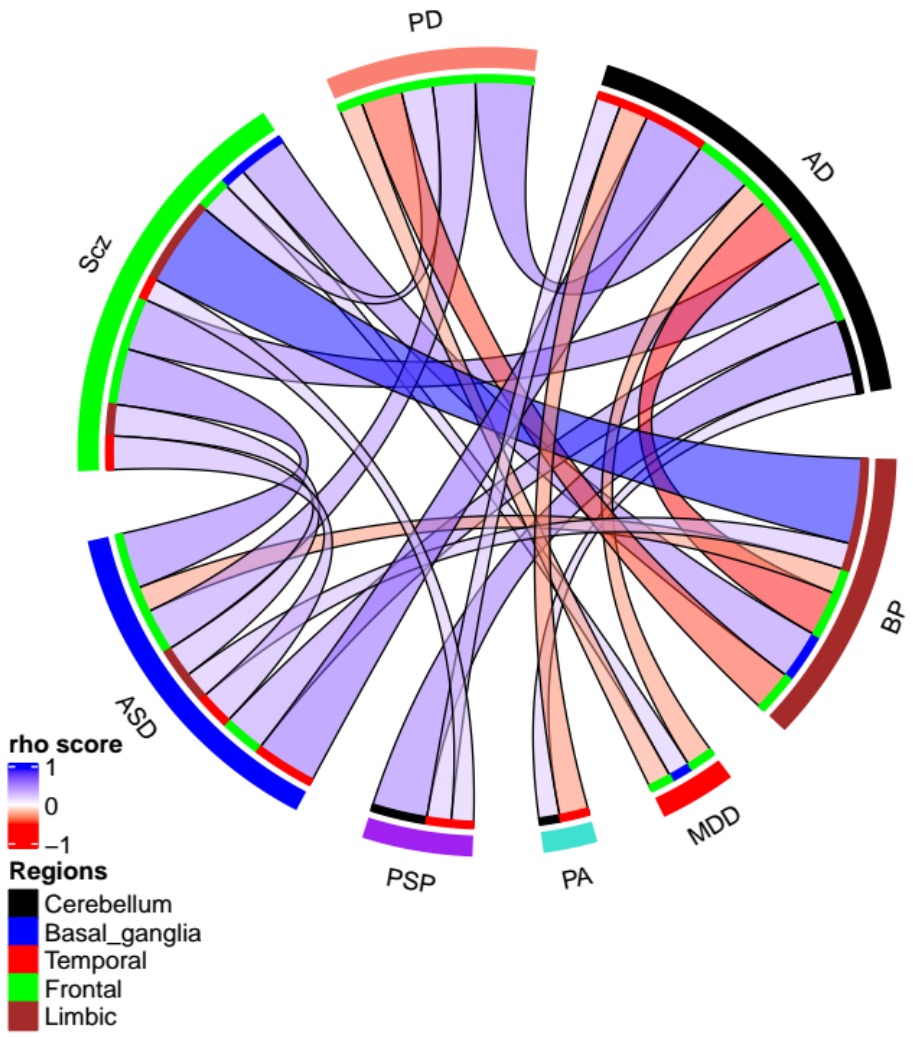

Figure 3-22 | | Brain region-specific transcriptome correlations across diseases. Only significant correlations after FDR correction (FDR < 0.05) are displayed. The outer layer represents diseases, while the inner layer displays brain regions defined by colors.

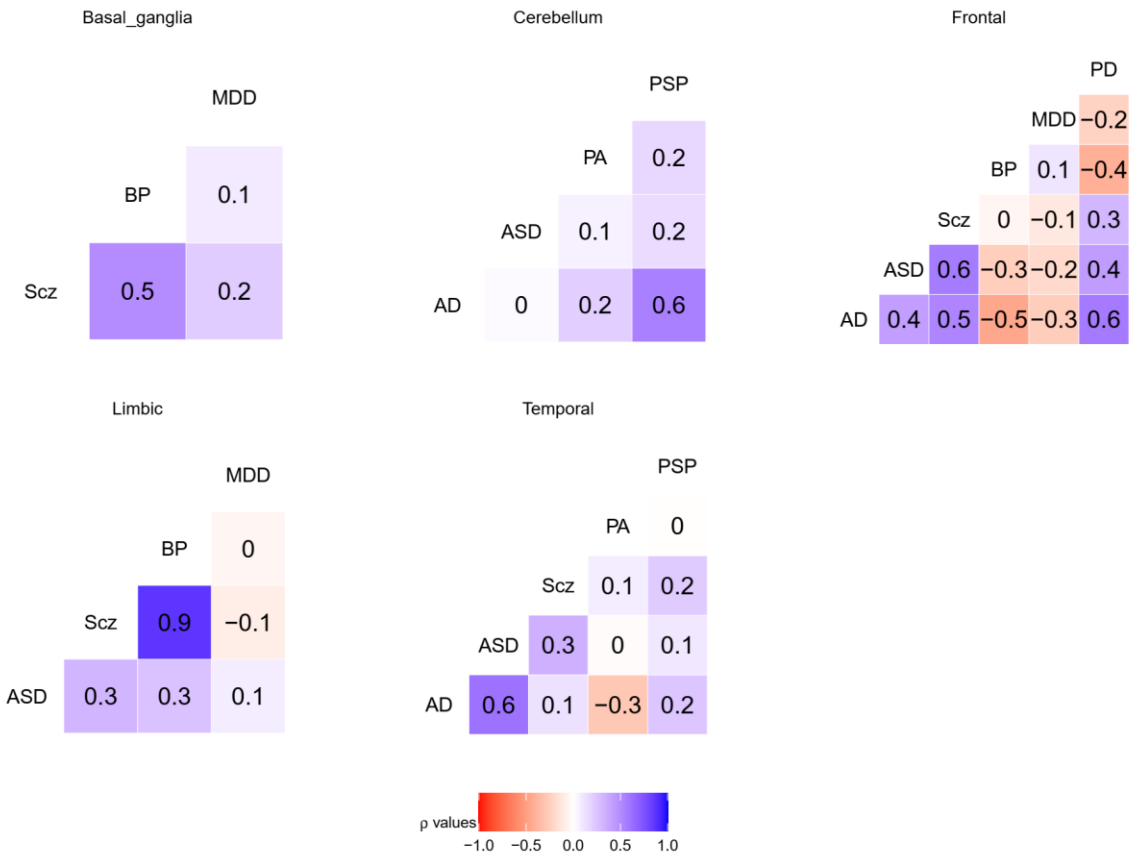

Figure 3-23 | Transcriptome signatures overlap within each region obtained by Spearman's correlations using log_2FC values of the shared genes across diseases

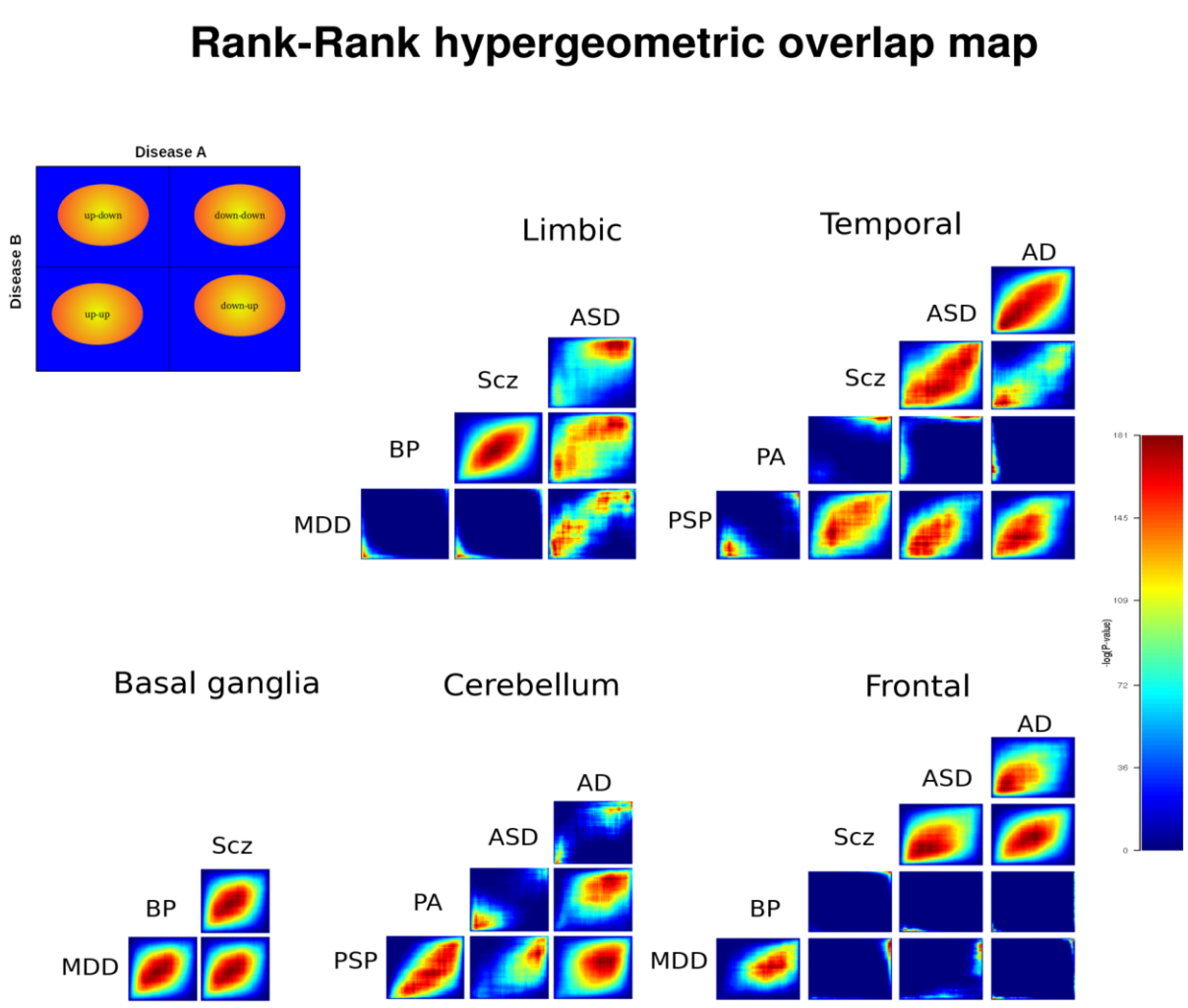

Figure 3-24 | RRHO maps of pairwise transcriptional comparisons for each region across diseases. The guide panel represents the cross-disease overlapping relationship. Signals in the upper left quadrant display an overlap for shared upregulated genes, while signals in the bottom right quadrant depict shared downregulated genes. The color bar displays the degree of significance of the overlap between two diseases (Fisher's exact test with FDR < 0.05).

3-9- Network analyses identified disease-specific and shared transcriptional signatures

To identify transcriptional signatures in the brain and gain insight into the underlying molecular mechanisms related to diseases, we constructed co-expression networks over all combined normalized datasets using weighted gene co-expression network analysis (rWGCNA) (Langfelder and Horvath, 2008). First, we tried a combination of different parameters to obrain specific modules (Figure 3-25). Our analysis identified eleven shared and disease-specific co-expression modules (Figure 3-26a). Each module comprised between 50 and 5073 genes (Figure 3-26b). Multidimensional scaling (Figure 3-27a) and correlation analysis (Figure 3-27b) revealed the relationship between the gene modules. Also, correlation analyses demonstrated a higher relationship between modules and disease (adjusted R^2 >0.2, FDR <0.05) than other covariates (Figure 3-28).

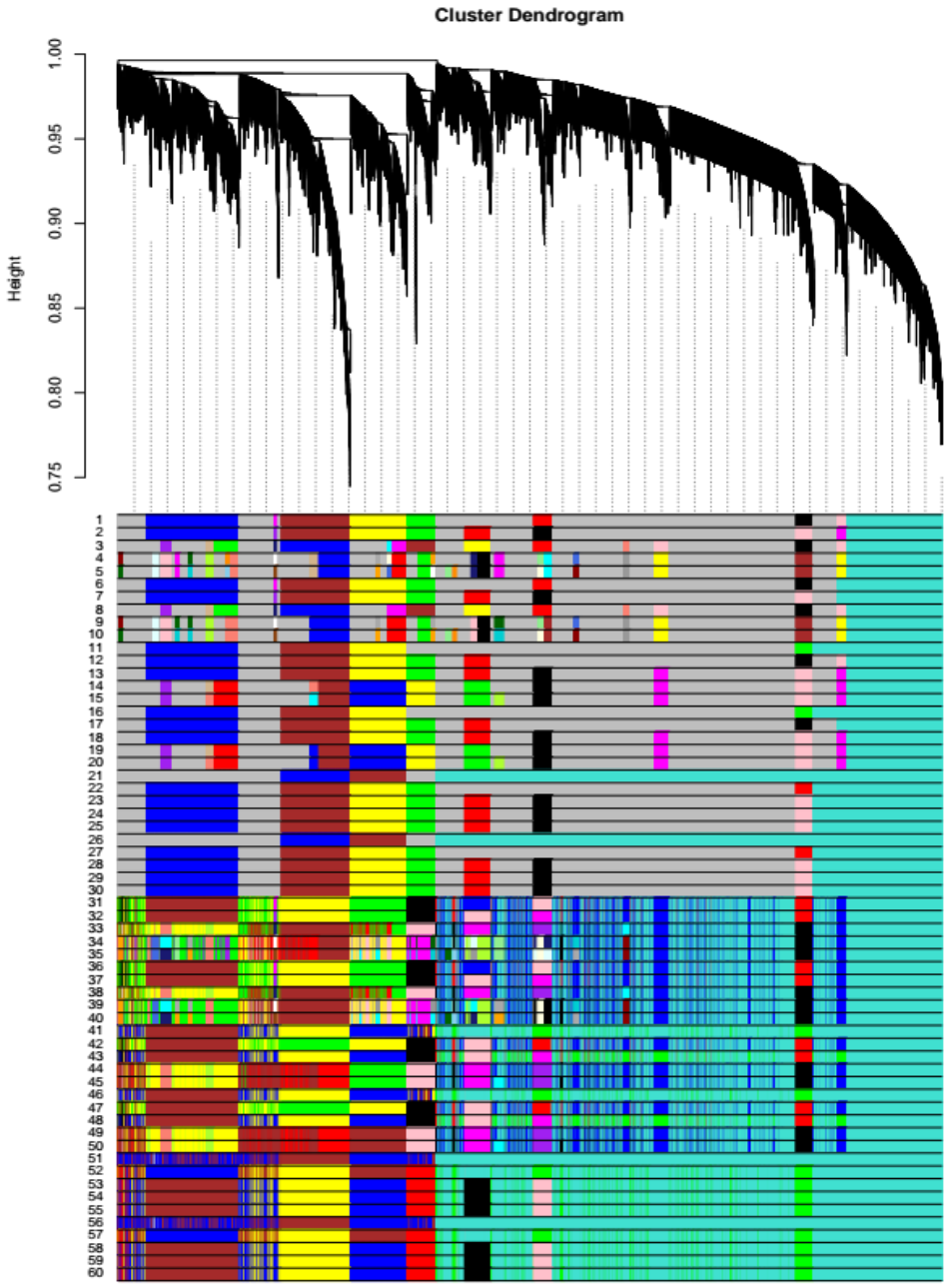

Figure 3-25 | Steps required for building robust gene co-expression network modules. (a) To obtain a final consensus network, co-expression network modules were built and tested using different network parameter sets.

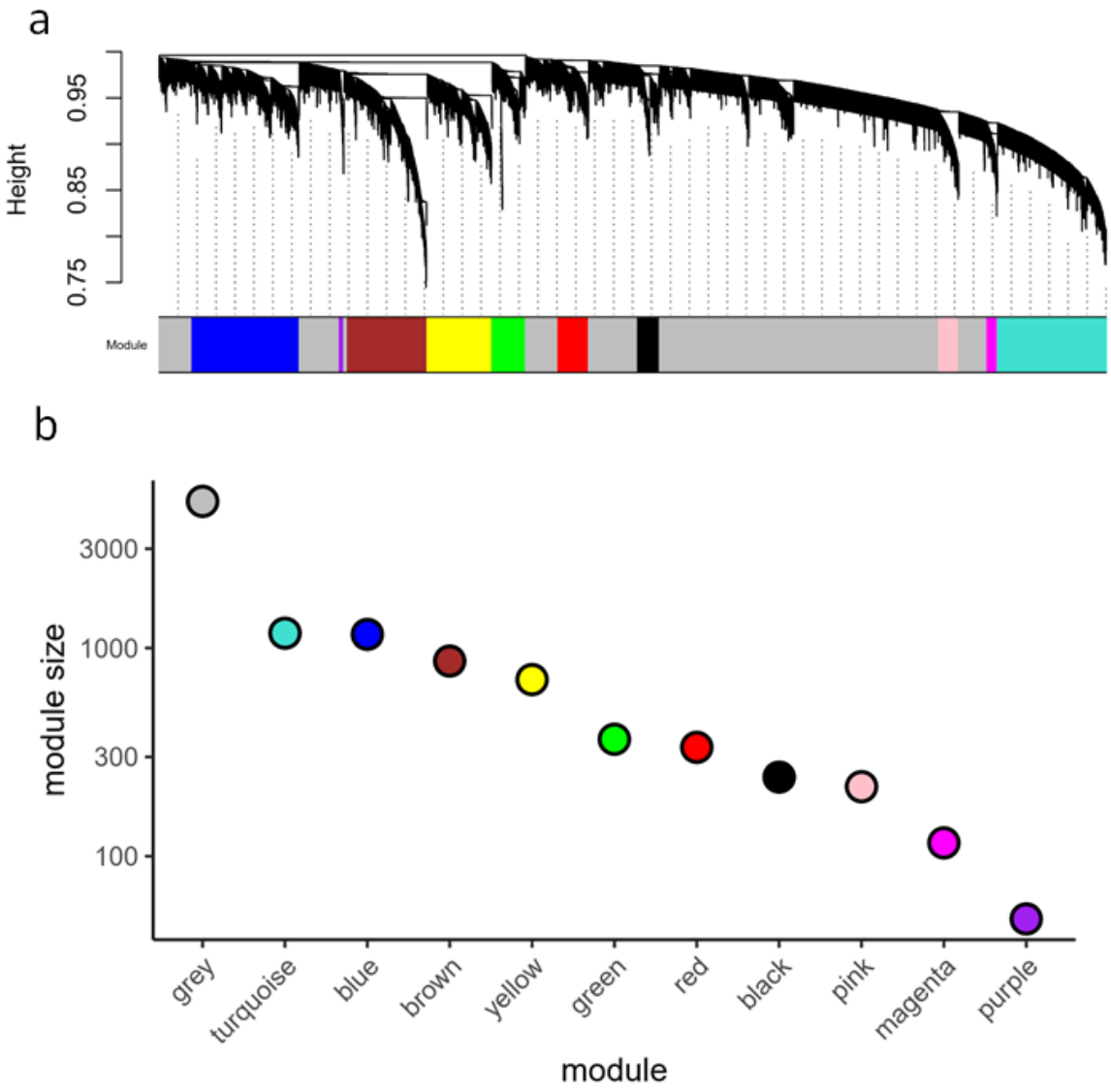

Figure 3-26 | Cross-disease gene co-expression modules identified by network analysis. (a) A dendrogram plot displaying co-expression modules obtained from topological overlap of genes across diseases. Each color represents an individual module. (b) The corresponding plot at the bottom shows the number of genes within each module.

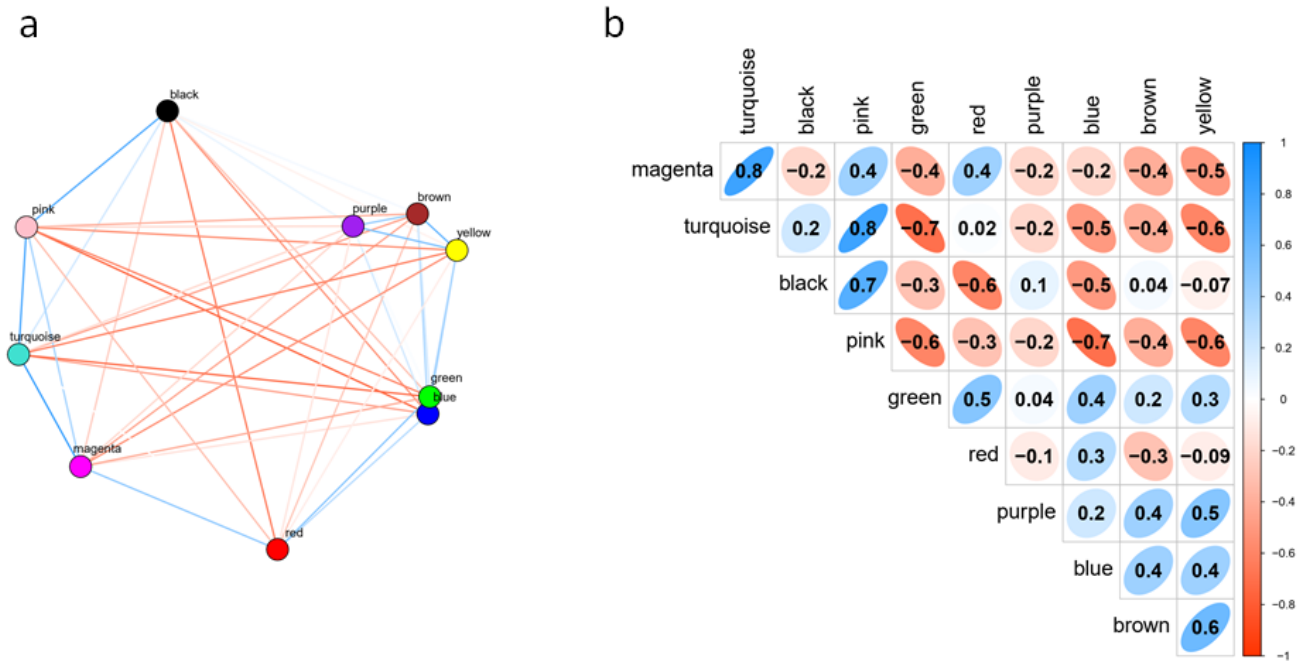

Figure 3-27 | Transciptome correlation of co-expression modules. (a) Multidimensional scaling plot depicts the modules relationship, with blue and red color representing positive and negative correlations, respectively. (b) Correlation plot corresponding to a, obtained from co-expression topological overlap of the modules.

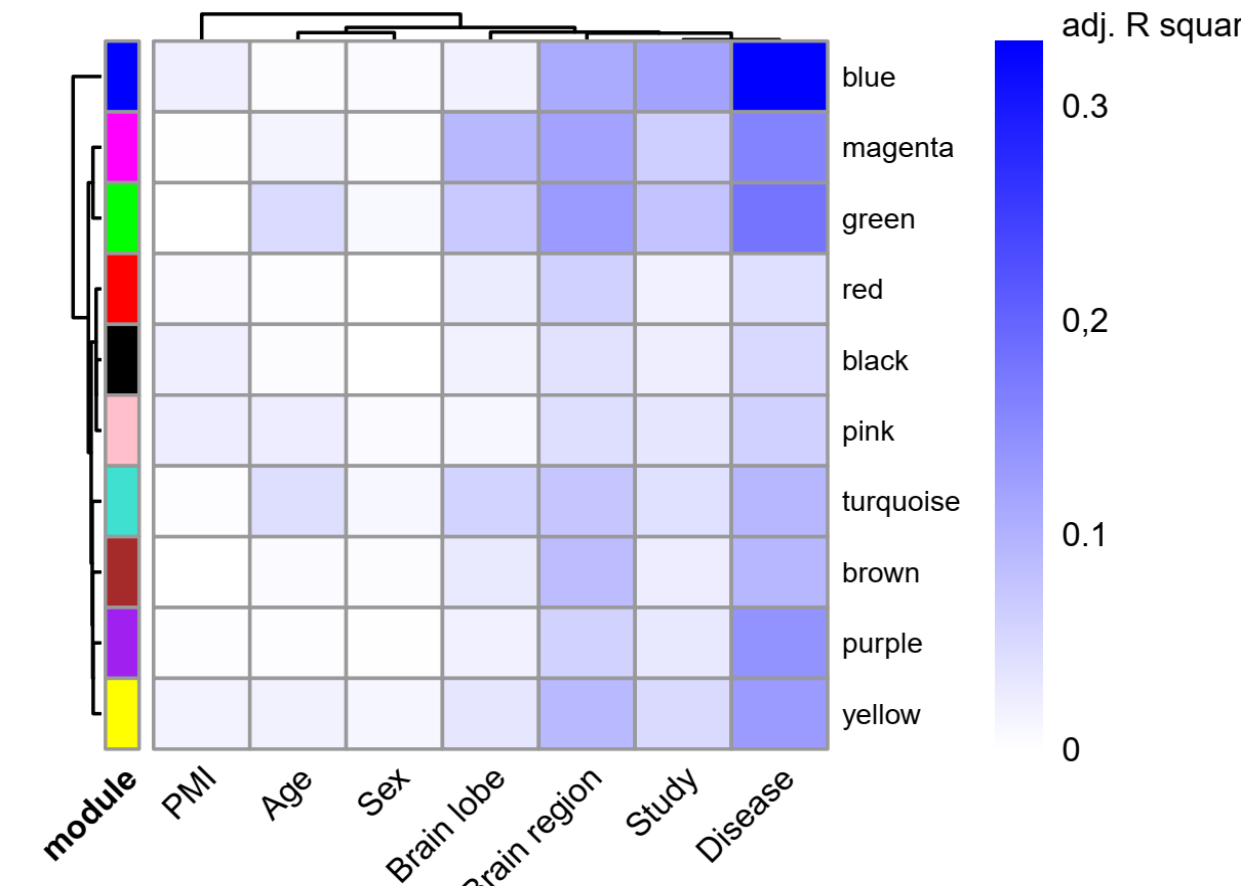

Figure 3-28 | A heatmap of relationship between covariates and modules eigengene using linear regression. Values show adjusted R^2 scores. (FDR < 0.05).

3-9-1- **Brain cell-type-specificity of co-expression modules**

To check for the brain cell-types specificity for each module, single-cell data for human brain cell types were obtained from the PanglaoDB database (Franzén, Gan and Björkegren, 2019). Gene overlap comparisons between modules and cell types were computed using Fisher's exact test. The results demonstrated the specificity of the brown module for oligodendrocytes (FDR = 2×10^{-20}) and Schwann cells (FDR = 5×10^{-20}), the yellow module for astrocytes (FDR = 6×10^{-9}) and Bergmann glia (FDR = 1×10^{-6}), the purple module for microglia (FDR = 1×10^{-7}), the magenta module for neurons (FDR = 7×10^{-4}) and pyramidal cells (FDR = 0.02), and the turquoise module for GABAergic neurons (FDR = 0.01, Figure 3-29).

Using another single-cell expression dataset (Zhang *et al.*, 2016) composed of five main brain cell types including neurons, astrocytes, oligodendrocytes, microglia, and endothelial cells, results confirmed the enrichment of the yellow module for astrocytes (FDR = 3×10^{-176}), the brown for oligodendrocytes (FDR = 5×10^{-152}), the turquoise and magenta for neurons (FDR = 3×10^{-77} and 9×10^{-31}, respectively), and the purple for microglia (FDR = 3×10^{-43}) (Figure 3-30).

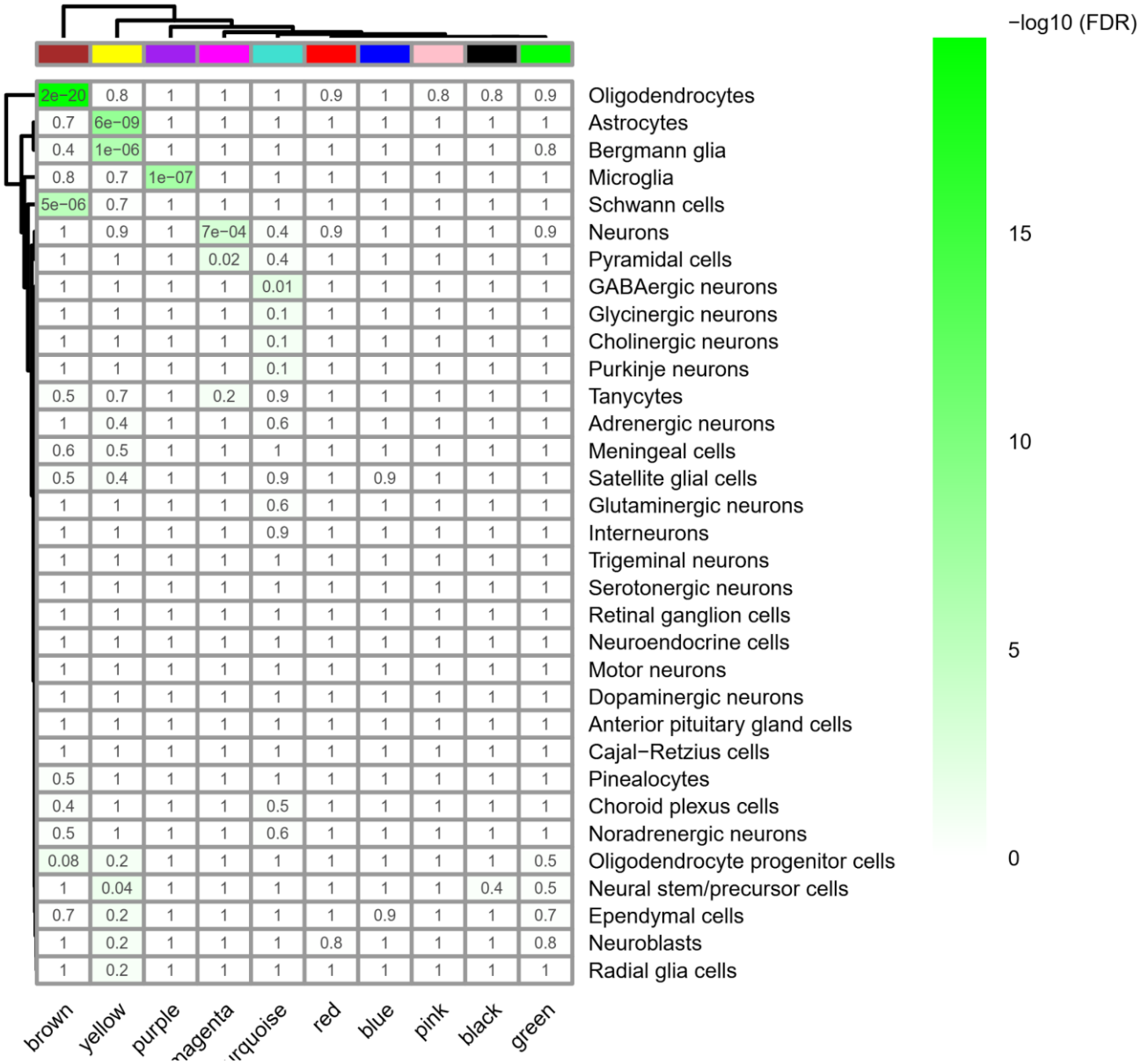

Figure 3-29 | Brain cell type-specific enrichment of modules measured by comparing genes within each module to the brain single-cell dataset from PanglaoDB (Franzén, Gan and Björkegren, 2019).

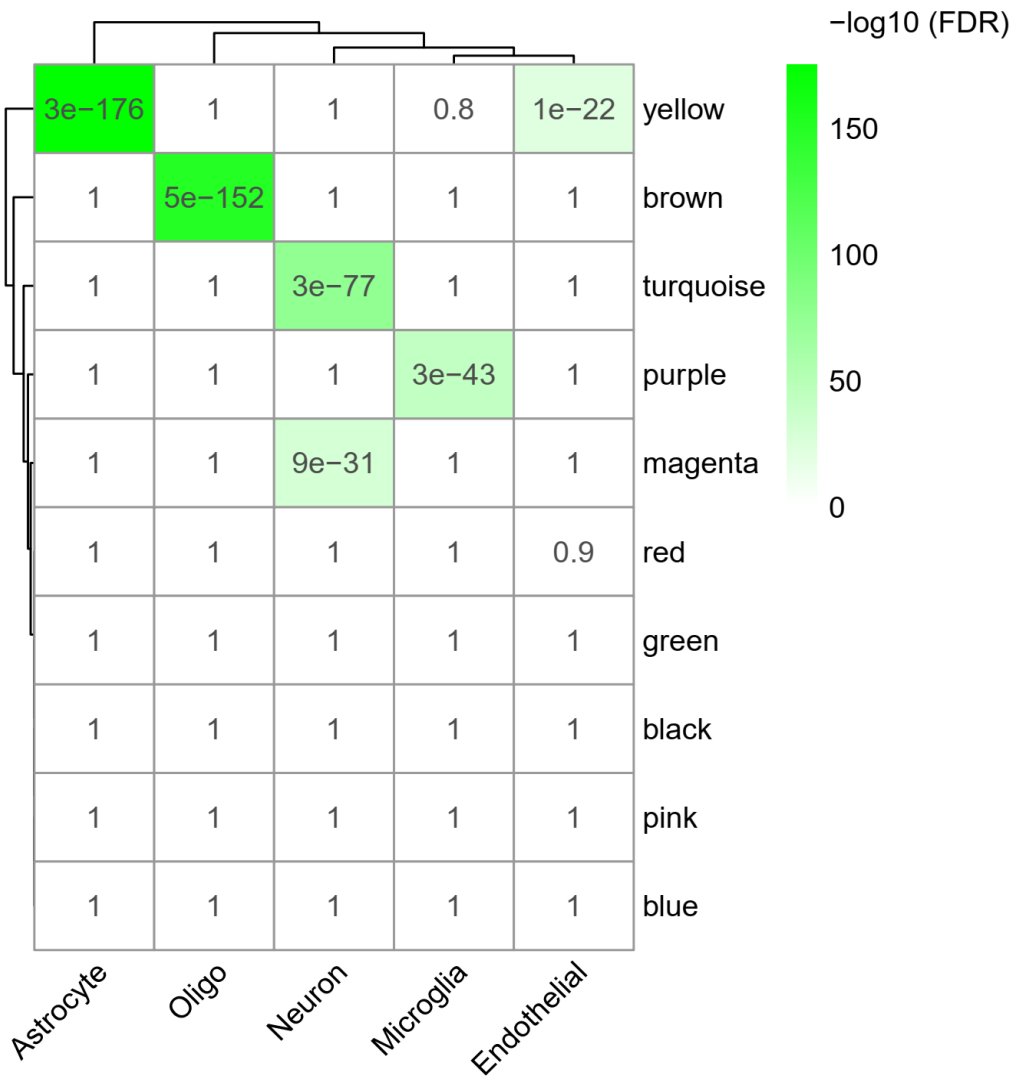

Figure 3-30 | Brain cell type-specific enrichment of co-expression modules measured using a single-cell expression dataset (Zhang *et al.*, 2016) composed of five main brain cell types including neurons, astrocytes, oligodendrocytes, microglia, and endothelial cells. Fisher's exact tests were used to perform the comparisons. Values show -$\log_{10}$(FDR-corrected p-values).

3-9-2- **Transcriptome correlations of cell-type-specific modules**

In addition, to see transcriptome correlations of cell-type-specific modules (yellow for astrocyte, brown for oligodendrocytes, purple for microglia, and magenta and turquoise for neurons), we compared gene expression of each module across diseases using Spearman's correlation test. Astrocyte and oligodendrocyte modules showed transcriptome similarity across PA, PSP and among Scz, PD, ASD, and AD (Spearman's correlation $\boldsymbol{\rho} > 0.2$ and FDR-corrected p-value < 0.05) (Figure 3-31). For microglia, we observed transcriptome relationships between PD, ASD and AD, among PA and PSP and between MDD, Scz and BP (Spearman's correlation $\boldsymbol{\rho} > 0.2$ and FDR-corrected p-value < 0.05). In neuron-specific modules, positive transcriptome correlations were observed across, PA, PSP and Scz and among PD, ASD and AD (Spearman's correlation $\boldsymbol{\rho} > 0.2$ and FDR-corrected p-value < 0.05).

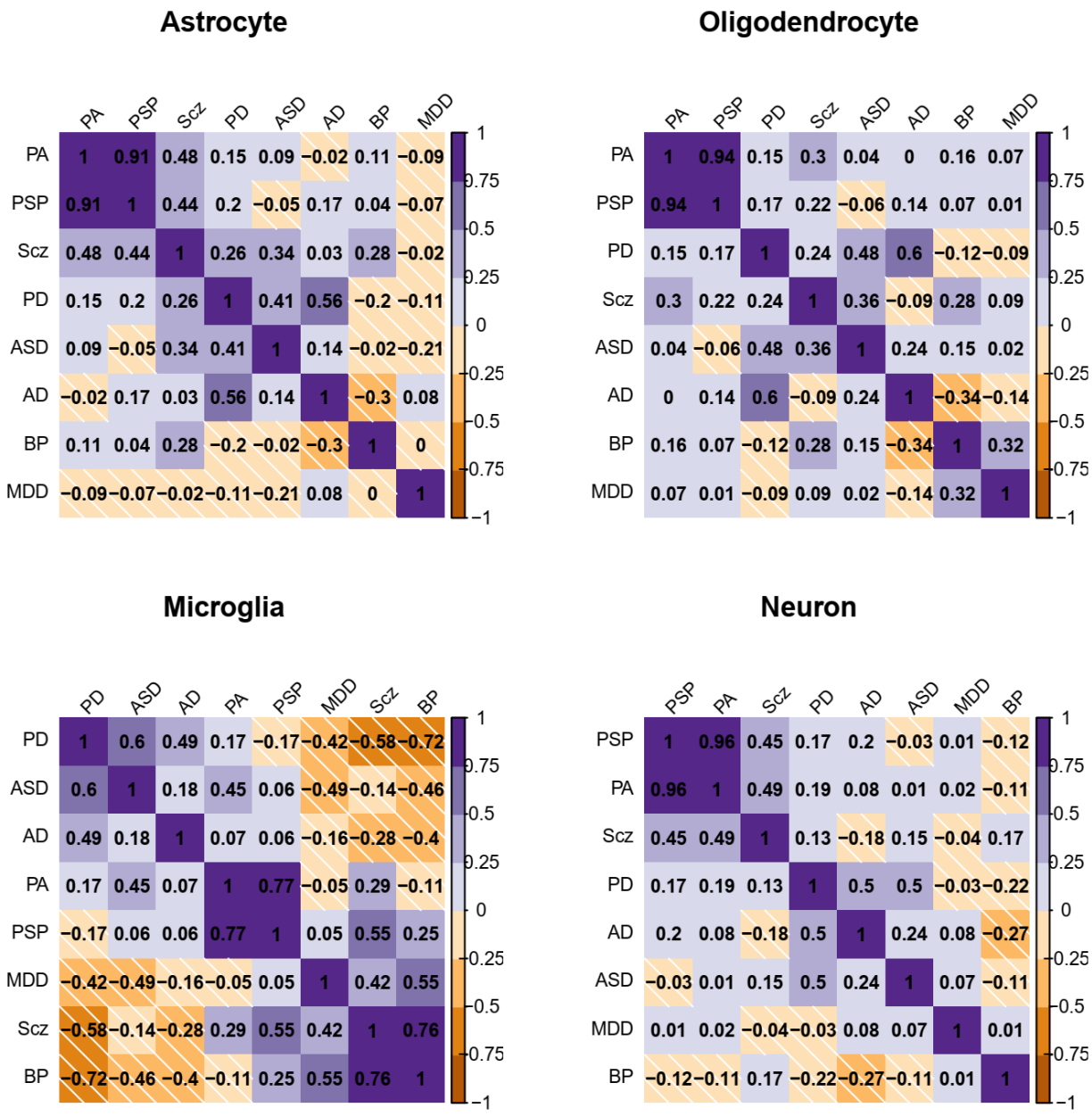

Figure 3-31 | Transcriptome correlations of cell-type specific modules (astrocyte: yellow, oligodendrocyte: brown; microglia: purple; neuron: magenta and turquoise) across diseases.

3-9-3- **Differential expression of co-expression modules**

Next, DGE for each module across diseases was computed using a linear mixed model. We observed that all modules except the red module are differentially expressed in AD. For ASD, five modules including the yellow, turquoise, red, purple, and magenta are dysregulated. All modules except the black module are up-regulated in MDD. PA and PSP show similarly down-regulation of the yellow module, turquoise, red, purple, magenta, brown and blue modules, and up-regulation of the black and green modules. PD shows up-regulation for the yellow, purple, green, brown and black modules and down-regulation of the turquoise and magenta modules. In Scz, we see mainly down-regulation of the modules (the turquoise, red, purple, magenta, and blue) and up-regulation of only black module. If we look at the behavior of each module across the diseases, we see that neuron-specific modules (magenta and turquoise) showed downregulation in all the diseases (FDR < 0.05) except in MDD (upregulation) and BP (no significant change observed) (Figure 3-32). These modules alongside the red module (with their hub genes; Figure 3-33 and Figure 3-34) were enriched for synapse organization, neuron part, and α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor activity which mediates rapid synaptic transmission in the brain (Jia *et al.*, 2020) (empirical permutation test with FDR < 0.05, Figure 3-35). An oligodendrocyte-related module (brown and hub genes *KIF13B*, *FMNL2*, *NDE1*, and *SLAIN1*), which was upregulated in AD, PD, and MDD and downregulated in PA and PSP (FDR < 0.05, Figure 3-32 and Figure 3-34) was enriched for neurogenesis, myelination, and cell projection (Figure 3-35). These results supported previous studies for the emerging role of oligodendrocytes in ND (Ahmed *et al.*, 2013; Liu and Zhou, 2013; Dean *et al.*, 2016; Krawczyk-Marć *et al.*, 2019). Microglia-associated module (purple and hub genes *INPP5D*, *ARHGDIB*, and *TYROBP*) revealed upregulation in AD, ASD, and PD, downregulation in BP, PA, PSP, and SZ (FDR < 0.05), and

enrichment for cell activation and regulation of the immune system. These findings were in line with the evidence for microglial changes in AD (Hemonnot *et al.*, 2019), ASD (Salter and Stevens, 2017), and PD (Ferreira and Romero-Ramos, 2018), and also the crucial role of microglia in the CNS development and immunity (Graeber and Stre'rt, 1990; Mosser *et al.*, 2017). An astrocyte-specific module (yellow and hub genes *YAP1* and *FoxO1*) enriched for circulatory system development, cell migration, and signal transduction (Figure 3-35) was upregulated in AD, ASD, MDD, and PD and downregulated in PA and PSP. An increase in astrocytic reactivity has previously been reported in response to amyloid-beta accumulation(Rodriguez-Vieitez *et al.*, 2016). Results confirmed previous evidence for the significant role of astrocytes in synapse formation, neuroprotection, and brain development (Kim *et al.*, 2015; Li *et al.*, 2019; Siracusa, Fusco and Cuzzocrea, 2019).

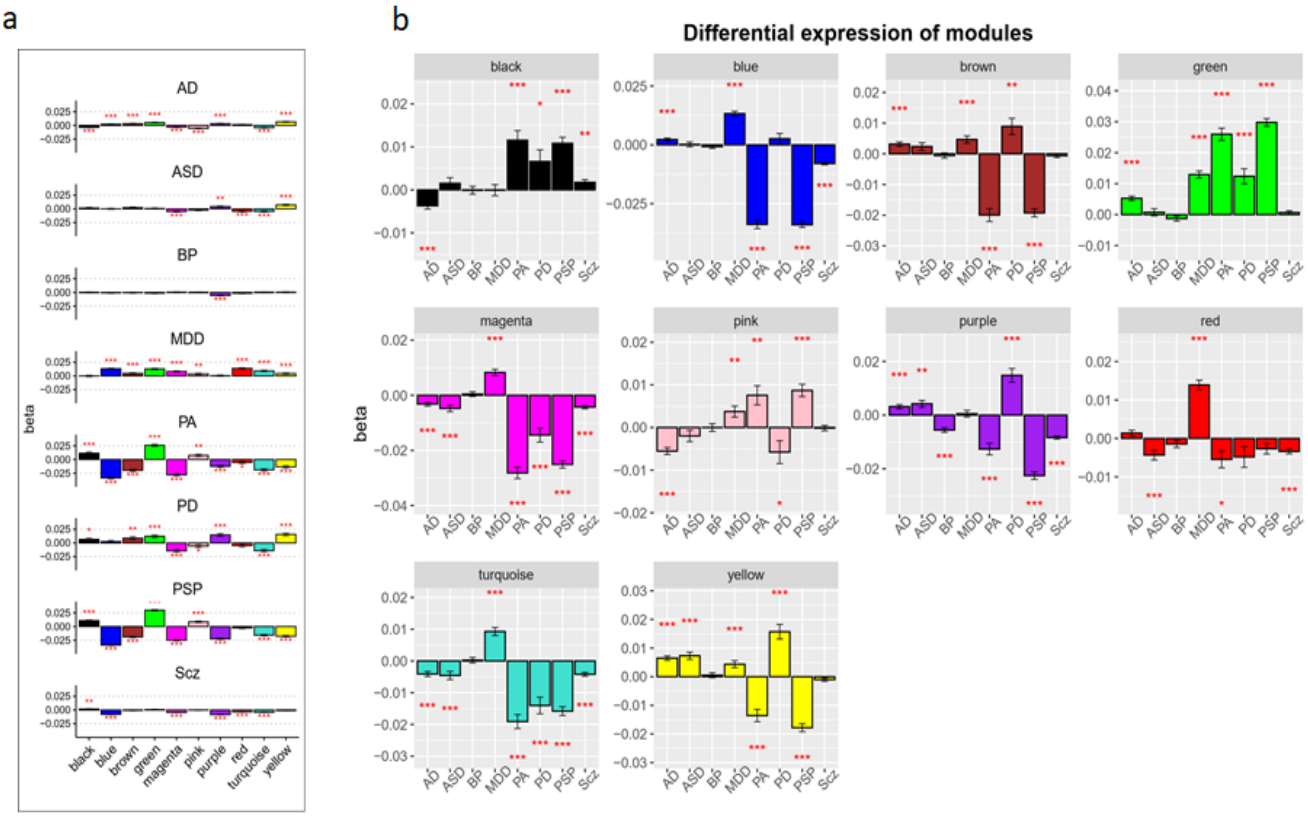

Figure 3-32 | Differential expression of modules across diseases shown by disease (a) and by module (b). β values on y-axis computed by linear mixed effect model show relationship of modules eigengenes with diseases (*FDR < 0.05, **FDR < 0.01, ***FDR < 0.001).

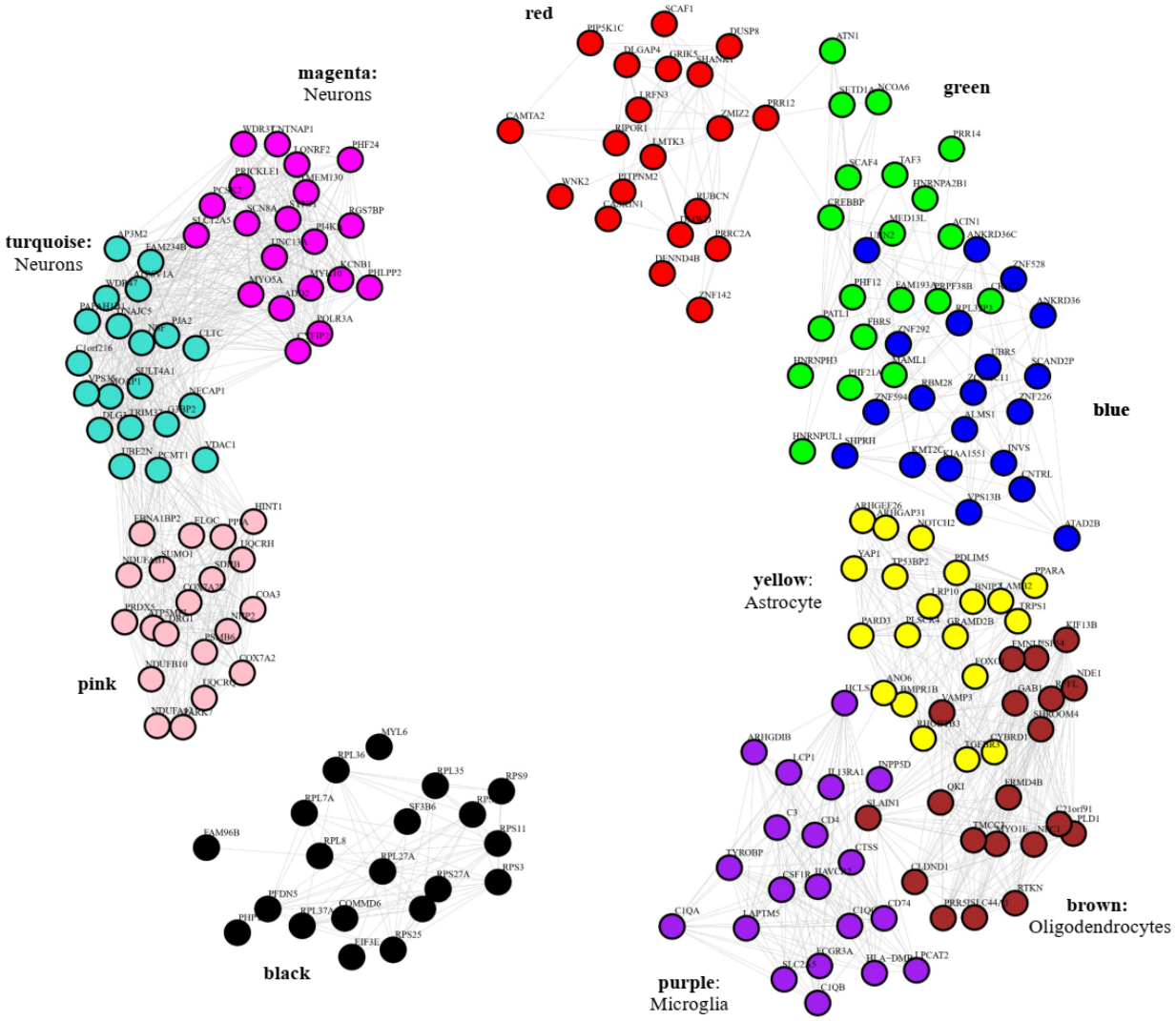

Figure 3-33 | A network of top hub genes (nodes) within each module with highest changes across diseases. Each color shows a module. Edges indicate gene-gene weighted-correlations.

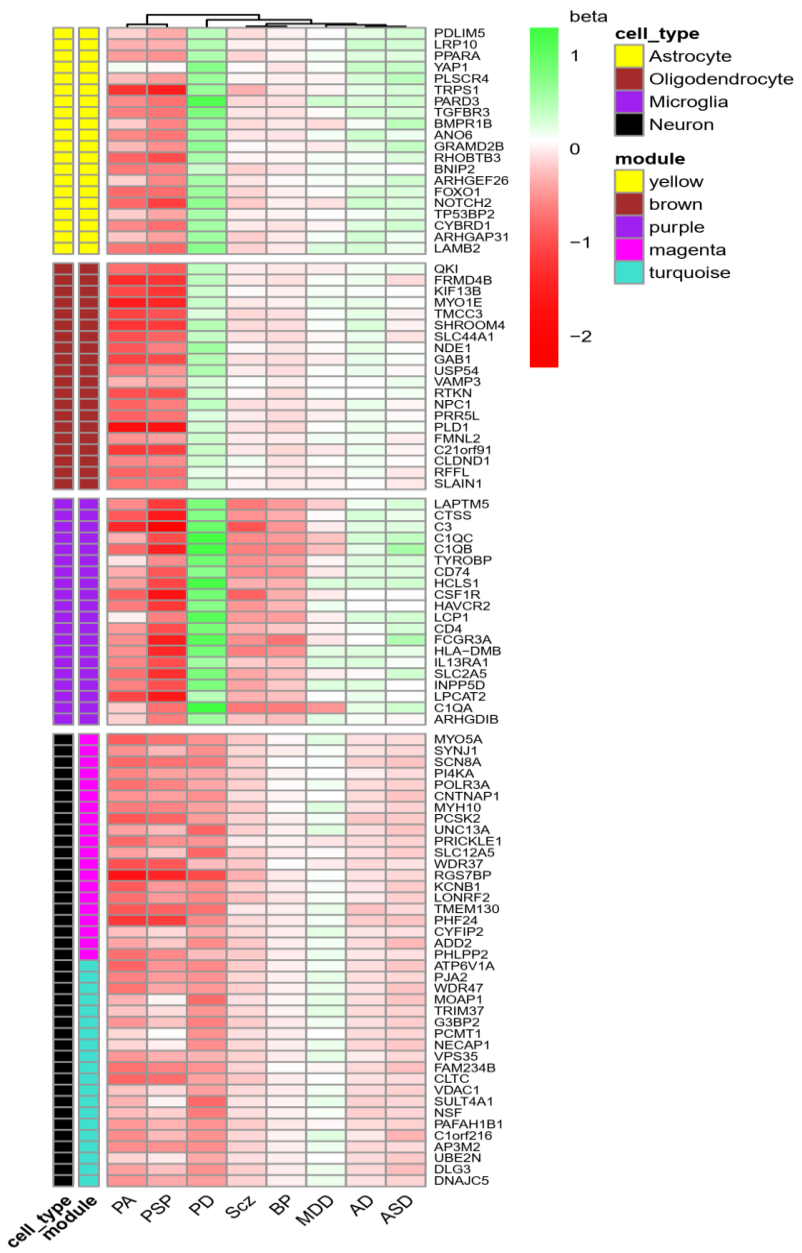

Figure 3-34 | The expression (log_2FC) of modules top hub genes across diseases. Brain cell-type-specific modules are annotated with colors.

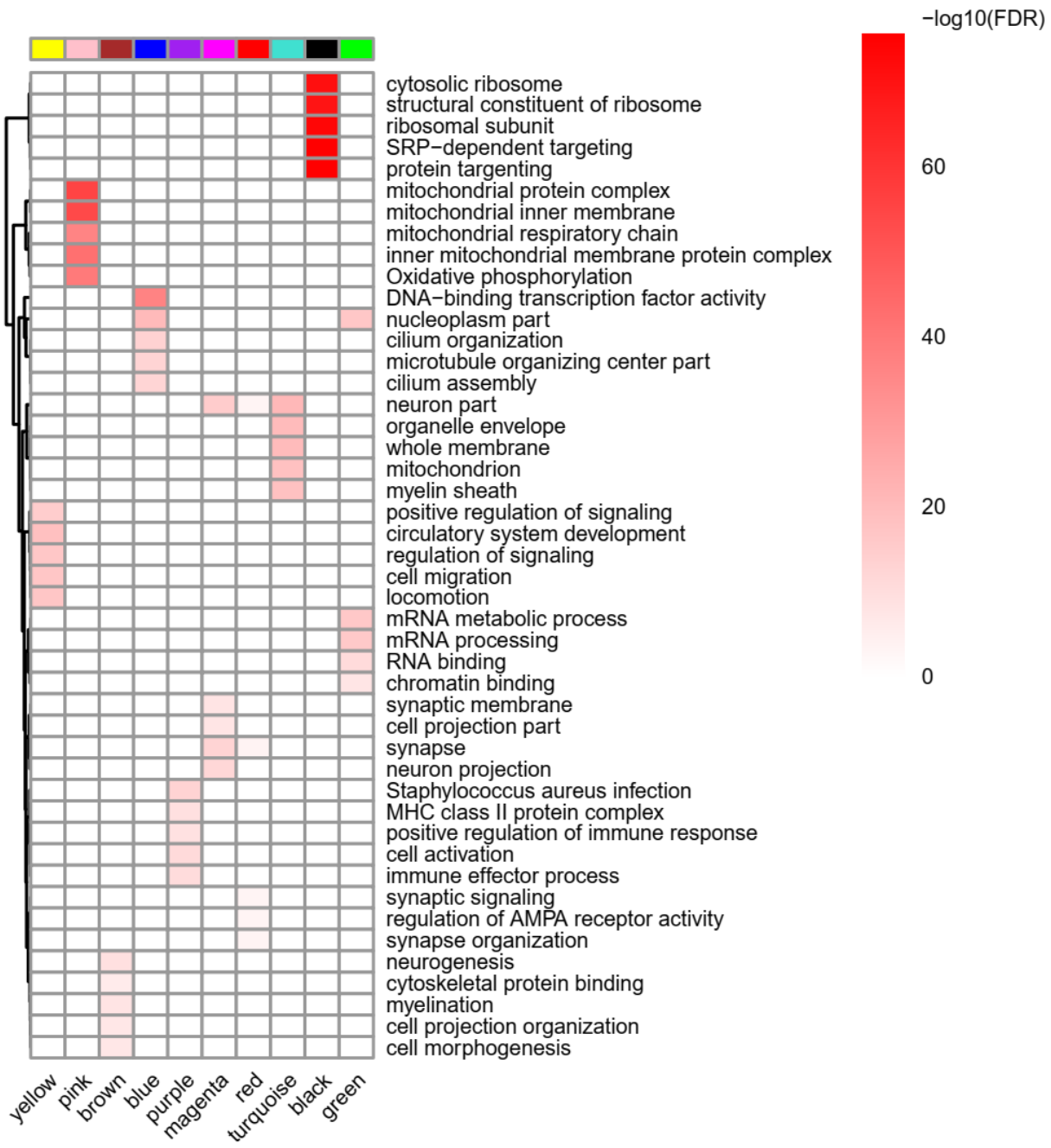

Figure 3-35 | Heatmap plot of gene ontology enrichment for modules using top five significant pathways for each module (color key shows -$\log_{10}$(FDR)).

3-9-4- **Enrichment analysis of transcription factors**

Finally, enrichment analysis of transcription factors and their binding motifs for each module revealed an overrepresentation of multiple motifs for several transcription factors including IRF-4 (empirical permutation test; FDR-corrected p-value = 1×10^{-6} involved in neural survival (Guo *et al.*, 2014), ELF1 (FDR-corrected p-value = 3.8×10^{-5}) important for neurite growth (Gao *et al.*, 1996), BRN1 (FDR-corrected p-value = 0.003) playing role in neurogenesis (Dominguez, Ayoub and Rakic, 2013), and PLAG1 (FDR-corrected p-value = 4.6×10^{-5}) involved in neocortical development (Sakai *et al.*, 2019)(Figure 3-36).

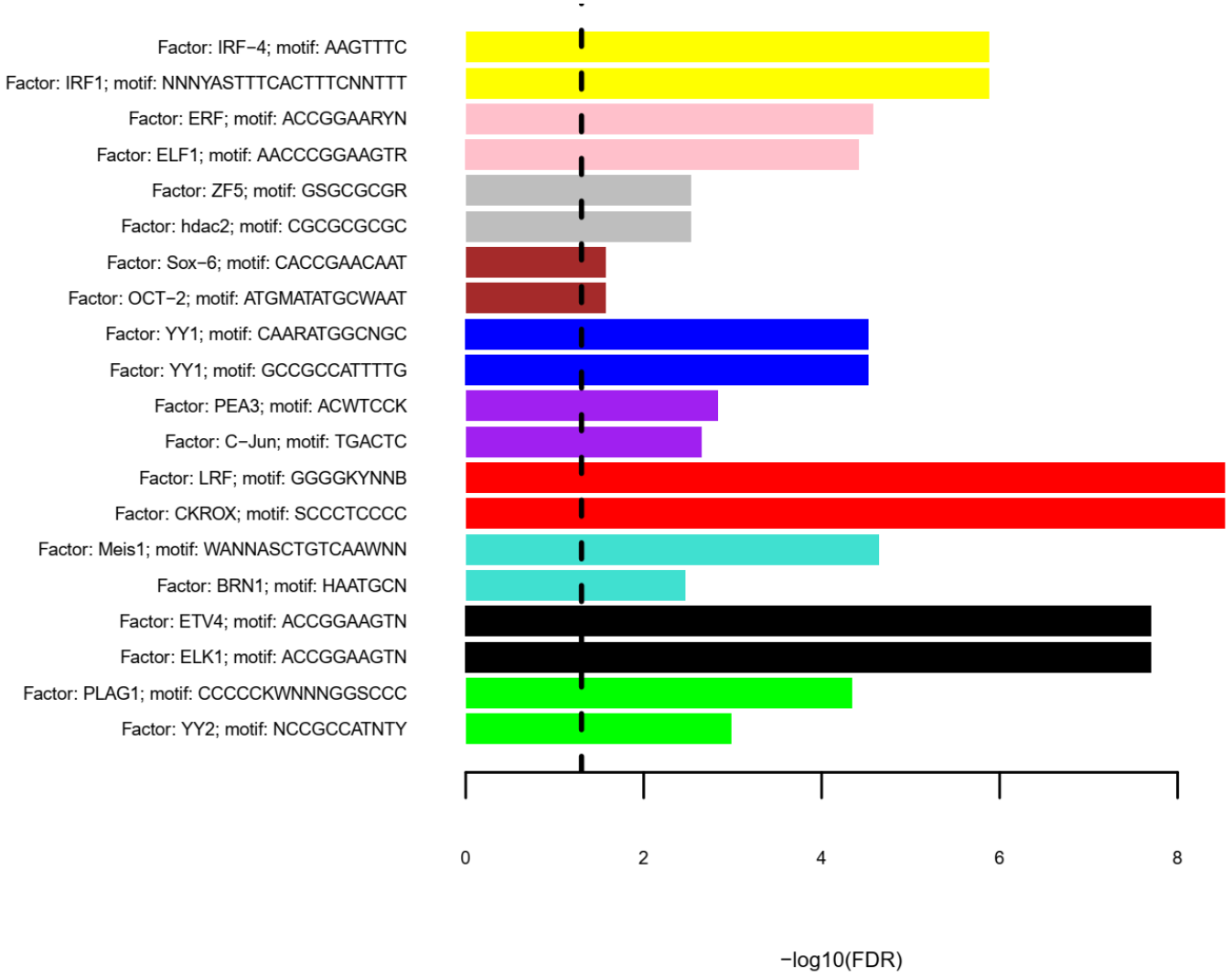

Figure 3-36 | Enrichment of transcription factor binding sites (top two motifs shown) for each module.

3-9-5- **Brain enhancer RNAs enrichment analysis**

Furthermore, to understand the relationship between the expression of brain enhancer RNAs (eRNA) and co-expression modules, enrichment analysis was performed using an independent dataset of eRNAs modules.(Yao *et al.*, 2015) Results demonstrated the enrichment of multiple eRNA modules for the brain cell-type-specific modules (brown, yellow, turquoise, purple, and magenta)(Fisher's exact test; FDR-corrected p-value <0.05, Figure 3-37), indicating their regulatory associations in brain circuits.

3-9-6- **Mitochondrial transcriptome enrichment analysis**

We have also performed mitochondrial transcriptome enrichment analysis for co-expression modules. In this analysis, the mitochondria were classified into synaptic and non-synaptic based on an independent study on mitochondrial modules. We observed that pink module was significantly enriched for non-synaptic, and turquoise module was enriched for both synaptic and non-synaptic mitochondria. These data show the involvement of pink and turquoise in mitochondrial function and their role in mitochondrial-related diseases such as neurodegeneratice diseases. AD and PD showed additional downregulation of a mitochondrial-related module (Figure 3-38Error! Reference source not found.), indicating the importance of mitochondria for synaptic connections, neuronal survival, and function (Schapira, 2008).

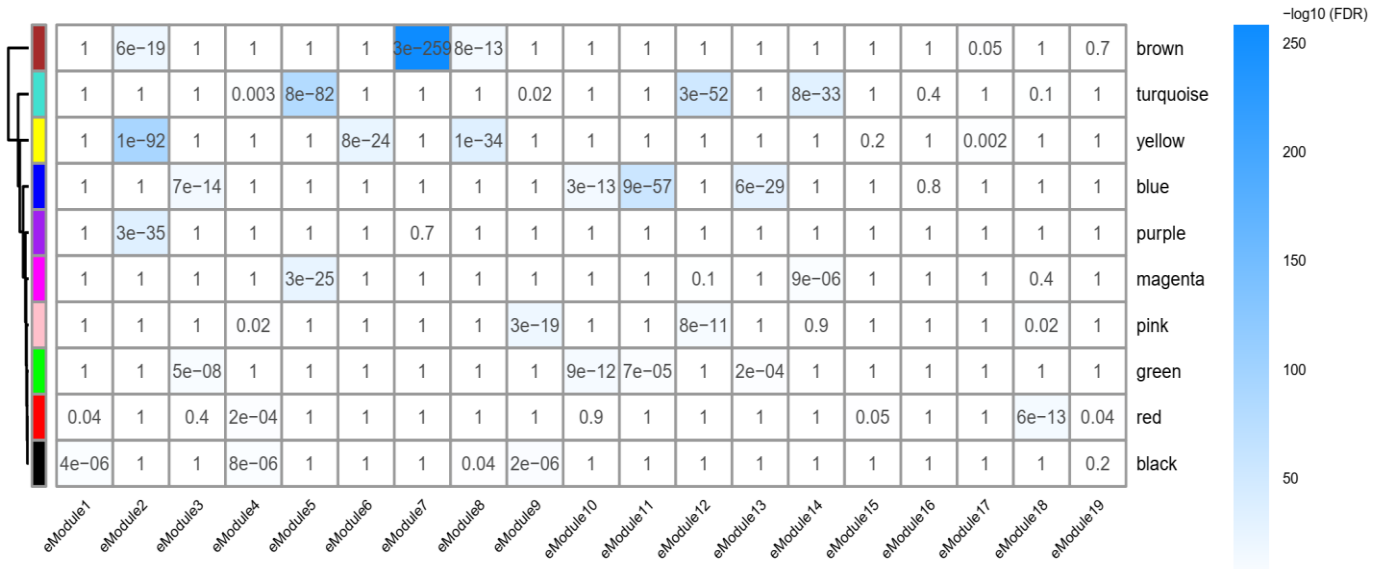

Figure 3-37 | Enrichment of brain enhancer RNAs for co-expression modules. The overlap between co-expression modules and eRNA modules from an independent dataset was computed by Fisher's exact test (FDR < 0.05).

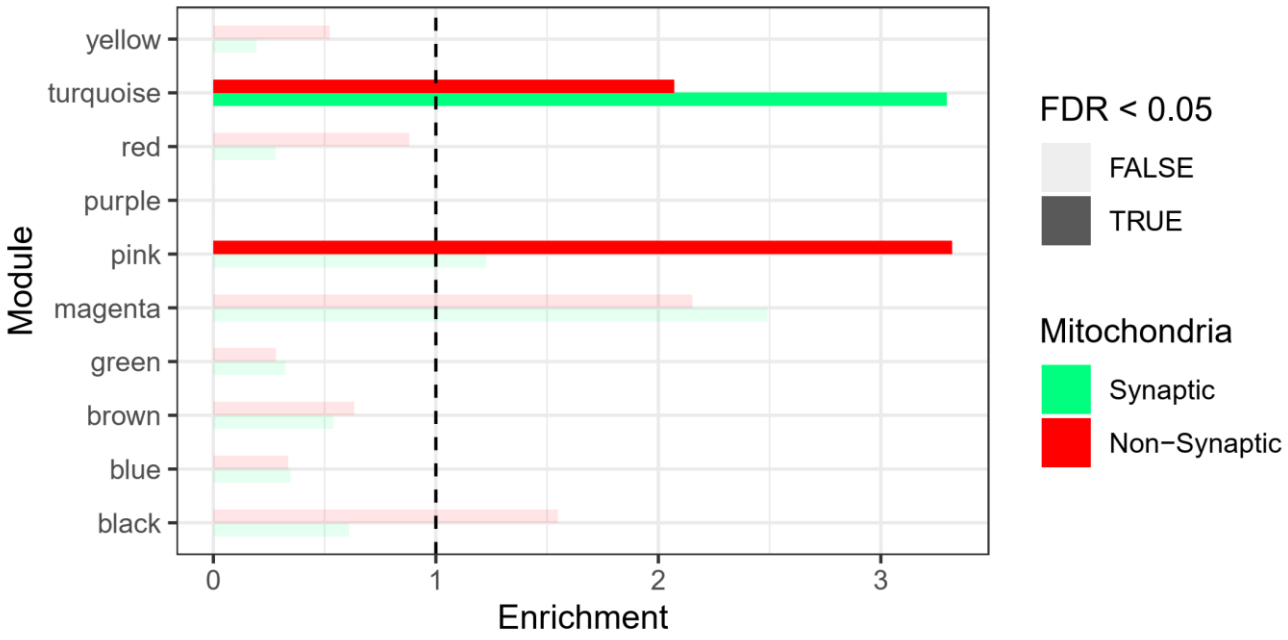

Figure 3-38 | The enrichment of co-expression modules for mitochondrial transcriptome. An independent study that previously reported mitochondrial co-expression modules was obtained and compared to co-expression modules in this study using Fisher's exact tests (FDR < 0.05). Green and red colors represent an enrichment of synaptic and non-synaptic mitochondrial co-expression modules.

Chapter4. Discussion

In this thesis, we aimed at exploring the molecular signature changes of cortical tissues from multiple brain diseases and understanding the similar and specific molecular pathology of neurodegenerative and psychiatric disorders. In order to achieve this, we exploited a large set of computational tools and statistical tests to analyze molecular signature of brain samples at the transcriptome level. The result found in this work are significantly strong that could provide a molecular framework to understanding pathogenesis of brain diseases.

4-1- Thesis contribution

Leveraging the transcriptome profile of post-mortem brain tissues, we highlighted the substantial overlapping molecular patterns across eight brain illnesses including NDDs and NPDs. Disease-specific DGE showed high transcriptional alterations in PSP, AD, SZ, and PD, suggesting NDDs underwent more transcriptional changes compared to NPDs. At the cortical level, the largest transcriptional changes were observed in temporal, cerebellum, and frontal lobe across diseases, even though our study was limited because of the lack of particular brain regions for a few diseases. However, the results were confirmed by classifier models built for each region across diseases using the expression of DEGs, suggesting their important role in the pathobiology of diseases. Dysregulation of overlapping genes such as *CX3CR1, CHI3L1, NPAS4, SERPINA3,* and *BAG3* across brain regions suggests shared perturbation of several mechanisms such as migration of microglia (Gyoneva *et al.*, 2019), astrocytes activation, synapse development, and synaptic plasticity (Spiegel *et al.*, 2014) across diseases. Microglia and astrocytes are vital in regulating neuronal activity and brain functioning during development and in the adult brain (Reemst *et al.*, 2016). These results support previous findings that dysfunction of molecular mechanisms in

microglia and astrocytes could contribute to neurodevelopmental diseases and potentially even late-onset neuropathology (Fakhoury, 2018; Li *et al.*, 2019).

Our new results also revealed shared transcription profiles between neurodegenerative and neurological diseases. Specifically, we observed similar transcriptional changes between AD-SZ, and AD-ASD, and between PD-ASD and PD-SZ. Moreover, within NDDs, we found transcriptome overlap between AD-PD and PA-PSP, while within NPDs, we found transcriptome similarity between SZ-ASD. These new findings provided insights about the shared cognitive impairment in ASD and SZ and their transcriptome similarity with NDDs (Sokol *et al.*, 2011; Gonatopoulos-Pournatzis *et al.*, 2020). Also, the results did not show a significant transcriptional correlation between AD and PA, suggesting their divergent molecular pathobiology (Murray and Dickson, 2014). Transcriptional relationships were not observed between MDD, BP, and other NPDs such as Scz and ASD as expected, which could be because of their heterogeneous nature (Gandal *et al.*, 2018; Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu and Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019).

Cortical transcriptome comparisons demonstrated similar transcriptome in temporal and frontal lobe across NDDs and NPDs, implicating their impairment in the pathogenesis of a variety of brain diseases (Cobia *et al.*, 2012; P. Allen *et al.*, 2012; Maidan *et al.*, 2016; Ng *et al.*, 2017; Wolk *et al.*, 2017; Matsuoka *et al.*, 2018; Cajanus *et al.*, 2019). Similar molecular patterns of the cerebellum were observed across AD, PSP, and PA, supporting its emerging role in the pathobiology of NDDs (Mormina *et al.*, 2017; Kaufmann *et al.*, 2019). Despite the lack of samples for all the diseases as mentioned before, basal ganglia and limbic lobe showed transcriptional similarities across Scz, ASD, BP and MDD, implying their involvement in several mood and

psychiatric disorders.These findings suggest, on one hand, the impairment of multiple brain regions (rather than one primary region) in disease and, on the other hand, the involvement of one region in multiple diseases.

Co-expression network results revealed mainly downregulation of neuron-specific modules across multiple diseases including AD, PA, PSP, PD, ASD, and Scz. The microglial-related module showed both upregulation and downregulation reflecting activation or deactivation of microglia in brain dysfunction(Mosser *et al.*, 2017; Salter and Stevens, 2017). Astrocyte- and oligodendrocyte-specific modules, similar to the microglial module, demonstrated broad dysregulation across diseases, representing the role of astrocytes and oligodendrocytes in neurogenesis, signaling, and cell development (Liu and Zhou, 2013; Gandal *et al.*, 2018; Siracusa, Fusco and Cuzzocrea, 2019). ASD showed coordinated upregulation of astrocyte- and microglia-specific modules on one hand, and downregulation of oligodendrocyte and neuron modules, on the other hand, in line with NDDs such as AD, PD, and PA, suggesting its similar molecular structure to that of NDDs (Kern *et al.*, 2013). AD and PD showed additional downregulation of a mitochondrial-related module (Fig. S16b), indicating the importance of mitochondria for synaptic connections, neuronal survival, and function (Schapira, 2008) .

4-2- Clinical prespective

In the last decades, using advanced technology, our understanding of the causative mechanisms and genetic factors associated with neurodegenerative and psychiatric disorders has been improving. Nevertheless, one of key challenges in the field of neuroscience is implementing molecular findings into clinical applications in order to develop better diagnosis and more efficient

treatments for brain diseases. Regarding this, using highly efficient high-throuput sequencing data such as RNA-Seq and genome-wide sequencing methods to profile molecular structure of tissues engaged in disease could be both informative and applicable in clinics. Although the findings in this thesis may not be directly applied in the clinical settings, in the long term they could pave the way, as a significant help, for finding novel diagnostic and treatment approaches. For instance, understanding the similarity or specificity of the mechanisms involved in the pathogenesis of neurogenerative diseases could provide insights about finding molecular targets that could be considered for treatment of one or multiple diseases.

4-3- Limitations and strengths

The work described in this thesis have its own limitations and strength. Although we tried to minimize the limitations as much as possible in order to provide more robust results, there were inevitable limitations. Fo instance, since tissue samples from all brain regions were not available for all the conditions, in some of them we might have missed the transcriptomic profile of the brain area in which the primary pathology is expected to be expressed (e.g. basal ganglia in PD). Additionaly, despite the large sample size included in the whole study, the number of samples for some diseases were lower compared to other diseases. Also, lack of some demographic data for the patients such as drug treatments could have provided us useful information as some medications are associated with transcriptomic changes after treatment.

However, the work described in this thesis includes, to our knowledge, the largest transcipromic profiling of post-mortem brain samples from a wide range of brain regions collected from eight NDD and NPDs. Despite the limitations of the current analyses, molecular signatures

described here across NDD-NPDs can provide target leads for the development of therapeutic interventions that overcome indications solely based on clinical manifestations, thus paving the way for the rational design of personalized and mechanistically-based therapies (Vilor-Tejedor *et al.*, 2018).

4-4- Future Research

Given the design of our analyses, focusing on overlapping transcriptomic alterations, we do not expect to capture modifications directly linked to the underlying etiological mechanisms of the different conditions here studied. This will be addressed in future analyses with the same database, along with the exploration of diseases sharing common mechanisms (i.e. cerebral proteinopathies) or between the different stages of the same disease (i.e. preclinical and clinical stages of AD)(Nazeri *et al.*, 2014; Yan *et al.*, 2017). In line with this, we anticipate that future research will also benefit from the integration of transcriptomics with other omics modalities, such as genomics, proteomics, metabolomics, and epigenomics. This promises to provide deeper insights into the causative pathways through which genes and environment interact during life and influence the human brain (Casamassimi *et al.*, 2017). Additional research could also benefit from further identification of sex-specific gene networks and transcription profiles to unravel the molecular mechanisms of brain diseases (Tiihonen *et al.*, 2019; Villa, Della Torre and Maggi, 2019; Kodama *et al.*, 2020).

Conclusions

The work presented in this work discusses transcriptional signatures of multiple neurodegeneratives diseases such as Alzheimer's, Parkinson's, Progressive supranuclear palsy, pre-clinical Alzheimer's and psychiatric disorders including schizophrenia, autism, major depressive disorder and bipolar disorder. The summary of our contributions could be listed as following:

- Neurodegenerative diseases show, in general, higher transcriptional changes compared to psychiatric disorders
- There are transcriptional similarities between neurodegenerative diseases and psychiatric disorders.
- Alzheimer', autism and schizophrenia show higher transcriptional similarties.
- Some cortical regions show higher transcriptional changes compared to other regions across diseases
- Multiple brain regions show transcriptional similarities between neurodegenerative and psychiatric disorders, while some regions such as Cerebellum show involvement in neurodegenerative diseases.
- There are some gene modules that coordinately change across brain diseases
- There are co-expression gene modules specific to brain cell types that alter across diseases.
- Transcriptome of some brain cell-types such as neurons change across both neurodegenerative and psychiatric disorders, while other cell-types (e.g. microglia) show transcriptional changes mostly in neurodegenerative diseases.

In this thesis, we provide a molecular framework to understand the molecular architecture of neurodegenerative and psychiatric disorders that could help the future studies and clinical approaches to find a better treatment.

REFERENCES

Abbeduto, L., McDuffie, A. and Thurman, A. J. (2014) 'The fragile X syndromeâ€"autism comorbidity: what do we really know?', Frontiers in Genetics. doi: 10.3389/fgene.2014.00355.

Abdel Rassoul, R. et al. (2010) 'Distinct transcriptome expression of the temporal cortex of the primate Microcebus murinus during brain aging versus Alzheimer's disease-like pathology', PloS one, 5(9). doi: 10.1371/journal.pone.0012770.

Ahmed, Z. et al. (2013) 'Identification and quantification of oligodendrocyte precursor cells in multiple system atrophy, progressive supranuclear palsy and Parkinson's disease', Brain pathology , 23(3), pp. 263–273.

Akula, N. et al. (2014) 'RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder', Molecular psychiatry, 19(11), pp. 1179–1185.

Allen, M. et al. (2012) 'Novel late-onset Alzheimer disease loci variants associate with brain gene expression', Neurology, 79(3), pp. 221–228.

Allen, M. et al. (2016) 'Gene Expression, Methylation and Neuropathology Correlations at Progressive Supranuclear Palsy Risk Loci', Acta neuropathologica. Acta Neuropathol, 132(2). doi: 10.1007/s00401-016-1576-7.

Allen, M. et al. (2016) 'Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases', Scientific Data. doi: 10.1038/sdata.2016.89.

Allen, M., Wang, X., Burgess, J. D., et al. (2018) 'Conserved brain myelination networks are altered in Alzheimer's and other neurodegenerative diseases', Alzheimer's & dementia: the journal of the Alzheimer's Association, 14(3), pp. 352–366.

Allen, M., Wang, X., Serie, D. J., et al. (2018) 'Divergent brain gene expression patterns associate with distinct cell-specific tau neuropathology traits in progressive supranuclear palsy', Acta neuropathologica, 136(5), pp. 709–727.

Allen, P. et al. (2012) 'Abnormal relationship between medial temporal lobe and subcortical dopamine function in people with an ultra high risk for psychosis', Schizophrenia bulletin, 38(5), pp. 1040–1049.

American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub.

Annese, A. et al. (2018) 'Whole transcriptome profiling of Late-Onset Alzheimer's Disease patients provides insights into the molecular changes involved in the disease', Scientific reports. Nature Publishing Group, 8(1), pp. 1–15.

Arion, D. et al. (2007) 'Molecular evidence for increased expression of genes related to immune and chaperone function in the prefrontal cortex in schizophrenia', Biological psychiatry, 62(7), pp. 711–721.

Arnold, P. R., Wells, A. D. and Li, X. C. (2019) 'Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate', Frontiers in cell and developmental biology, 7, p. 377.

Artiga, M. J. et al. (1998) 'Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene', FEBS letters, 421(2), pp. 105–108.

Ashford, J. W. and Wesson Ashford, J. (2004) 'APOE Genotype Effects on Alzheimer's Disease Onset and Epidemiology', Journal of Molecular Neuroscience, pp. 157–166. doi: 10.1385/jmn:23:3:157.

Avramopoulos, D. (2018) 'Recent Advances in the Genetics of Schizophrenia', Molecular neuropsychiatry, 4(1), pp. 35–51.

Bachevalier, J. (1994) 'Medial temporal lobe structures and autism: a review of clinical and experimental findings', Neuropsychologia, 32(6), pp. 627–648.

Bachevalier, J. and Loveland, K. A. (2006) 'The orbitofrontal–amygdala circuit and self-regulation of social–emotional behavior in autism', Neuroscience & Biobehavioral Reviews, pp. 97–117. doi: 10.1016/j.neubiorev.2005.07.002.

Badre, D. and Wagner, A. D. (2007) 'Left ventrolateral prefrontal cortex and the cognitive control of memory', Neuropsychologia. Elsevier, 45(13), pp. 2883–2901.

Bakkour, A. et al. (2013) 'The effects of aging and Alzheimer's disease on cerebral cortical anatomy: specificity and differential relationships with cognition', NeuroImage, 76, pp. 332–344.

Balint, B. et al. (2019) 'Left atrial microvascular endothelial dysfunction, myocardial inflammation and fibrosis after selective insular cortex ischemic stroke', International journal of cardiology. Elsevier, 292, pp. 148–155.

Baron-Cohen, S. et al. (2000) 'The amygdala theory of autism', Neuroscience & Biobehavioral Reviews, pp. 355–364. doi: 10.1016/s0149-7634(00)00011-7.

Bateman, R. J. et al. (2012) 'Clinical and biomarker changes in dominantly inherited Alzheimer's disease', The New England journal of medicine, 367(9), pp. 795–804.

Baxter, A. J. et al. (2015) 'The epidemiology and global burden of autism spectrum disorders', Psychological Medicine, pp. 601–613. doi: 10.1017/s003329171400172x.

Behrmann, M., Geng, J. J. and Shomstein, S. (2004) 'Parietal cortex and attention', Current opinion in neurobiology. Elsevier, 14(2), pp. 212–217.

Belgiu, M. and Drăguţ, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', ISPRS journal of photogrammetry and remote sensing: official publication of the International Society for Photogrammetry and Remote Sensing , 114, pp. 24–31.

Birnbaum, R. et al. (2014) 'Prenatal expression patterns of genes associated with neuropsychiatric disorders', The American journal of psychiatry, 171(7), pp. 758–767.

Bobilev, A. M., Perez, J. M. and Tamminga, C. A. (2020) 'Molecular alterations in the medial temporal lobe in schizophrenia', Schizophrenia research. Elsevier, 217, pp. 71–85.

Bogousslavsky, J. and Tatu, L. (2017) 'The history of basal ganglia anatomy', Revue neurologique. Elsevier, 173(10), p. 605.

Bond, D. J. et al. (2020) 'Serum epidermal growth factor, clinical illness course, and limbic brain volumes in early-stage bipolar disorder', Journal of affective disorders. Elsevier, 270, pp. 30–35.

Bondi, M. W., Edmonds, E. C. and Salmon, D. P. (2017) 'Alzheimer's Disease: Past, Present, and Future', Journal of the International Neuropsychological Society: JINS, 23(9-10), pp. 818–831.

Bostan, A. C., Dum, R. P. and Strick, P. L. (2018) 'Functional Anatomy of Basal Ganglia Circuits with the Cerebral Cortex and the Cerebellum', Progress in neurological surgery. karger.com, 33, pp. 50–61.

Bradl, M. and Lassmann, H. (2010) 'Oligodendrocytes: biology and pathology', Acta neuropathologica, 119(1), pp. 37–53.

Brainstorm Consortium et al. (2018) 'Analysis of shared heritability in common disorders of the brain', Science, 360(6395). doi: 10.1126/science.aap8757.

Breiman, L. (2001) 'Machine Learning', pp. 261–277. doi: 10.1023/a:1017934522171.

Brisch, R. et al. (2014) 'The Role of Dopamine in Schizophrenia from a Neurobiological and Evolutionary Perspective: Old Fashioned, but Still in Vogue', Frontiers in psychiatry / Frontiers Research Foundation. Frontiers, 5. doi: 10.3389/fpsyt.2014.00047.

Brittany N. Dugger, D. W. D. (2017) 'Pathology of Neurodegenerative Diseases', Cold Spring Harbor perspectives in biology. Cold Spring Harbor Laboratory Press, 9(7). doi: 10.1101/cshperspect.a028035.

Bromet, E. et al. (2011) 'Cross-national epidemiology of DSM-IV major depressive episode', BMC medicine, 9, p. 90.

Bruch, J. et al. (2015) 'Early Neurodegeneration in the Brain of a Child Without Functional PKR-like Endoplasmic Reticulum Kinase', Journal of neuropathology and experimental neurology, 74(8), pp. 850–857.

Brugha, T. S. et al. (2016) 'Epidemiology of autism in adults across age groups and ability levels', British Journal of Psychiatry, pp. 498–503. doi: 10.1192/bjp.bp.115.174649.

Burke, S. N. et al. (2018) 'Shared Functions of Perirhinal and Parahippocampal Cortices: Implications for Cognitive Aging', Trends in neurosciences. Elsevier, 41(6), pp. 349–359.

Buxbaum, L. J. and Randerath, J. (2018) 'Limb apraxia and the left parietal lobe', Handbook of clinical neurology. Elsevier, 151, pp. 349–363.

Cajanus, A. et al. (2019) 'The Association Between Distinct Frontal Brain Volumes and Behavioral Symptoms in Mild Cognitive Impairment, Alzheimer's Disease, and Frontotemporal Dementia', Frontiers in neurology, 10, p. 1059.

Caligiore, D. et al. (2017) 'Dysfunctions of the basal ganglia-cerebellar-thalamo-cortical system produce motor tics in Tourette syndrome', PLoS computational biology, 13(3), p. e1005395.

Canchi, S. et al. (2019) 'Integrating Gene and Protein Expression Reveals Perturbed Functional Networks in Alzheimer's Disease', Cell Reports, pp. 1103–1116.e4. doi: 10.1016/j.celrep.2019.06.073.

Cannistraro, P. A. and Rauch, S. L. (2003) 'Neural circuitry of anxiety: evidence from structural and functional neuroimaging studies', Psychopharmacology bulletin, 37(4), pp. 8–25.

Cao, Y.-L. et al. (2017) 'A role of BAG3 in regulating SNCA/α-synuclein clearance via selective macroautophagy', Neurobiology of aging, 60, pp. 104–115.

Casamassimi, A. et al. (2017) 'Transcriptome Profiling in Human Diseases: New Advances and Perspectives', International journal of molecular sciences, 18(8). doi: 10.3390/ijms18081652.

Chang, X. et al. (2017) 'RNA-seq analysis of amygdala tissue reveals characteristic expression profiles in schizophrenia', Translational psychiatry, 7(8), p. e1203.

Cherian, K., Schatzberg, A. F. and Keller, J. (2019) 'HPA axis in psychotic major depression and schizophrenia spectrum disorders: Cortisol, clinical symptomatology, and cognition', Schizophrenia research, 213, pp. 72–79.

Chizhikov, V. V. and Millen, K. J. (2020) 'Neurogenesis in the cerebellum', in Patterning and Cell Type Specification in the Developing CNS and PNS. Elsevier, pp. 349–367.

Chu, Y. and Kordower, J. H. (2007) 'Age-associated increases of α-synuclein in monkeys and humans are associated with nigrostriatal dopamine depletion: Is this the target for Parkinson's disease?', Neurobiology of Disease, pp. 134–149. doi: 10.1016/j.nbd.2006.08.021.

Clark, T. A., Sugnet, C. W. and Ares, M., Jr (2002) 'Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays', Science, 296(5569), pp. 907–910.

Cobia, D. J. et al. (2012) ‘Longitudinal progression of frontal and temporal lobe changes in schizophrenia’, Schizophrenia research, 139(1-3), pp. 1–6.

Coppola, G. et al. (2012) ‘Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer’s diseases’, Human molecular genetics, 21(15), pp. 3500–3512.

Corder, E. H. et al. (1994) ‘Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease’, Nature Genetics, pp. 180–184. doi: 10.1038/ng0694-180.

Corley, S. M. et al. (2016) ‘Transcriptomic Analysis Shows Decreased Cortical Expression of NR4A1, NR4A2 and RXRB in Schizophrenia and Provides Evidence for Nuclear Receptor Dysregulation’, PloS one, 11(12), p. e0166944.

Cousineau, D. and Chartier, S. (2010) ‘Outliers detection and treatment: a review’, International Journal of Psychological Research, p. 58. doi: 10.21500/20112084.844.

Coyle-Gilchrist, I. T. S. et al. (2016) ‘Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes’, Neurology, 86(18), pp. 1736–1743.

Cross-Disorder Group of the Psychiatric Genomics Consortium et al. (2013) ‘Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs’, Nature genetics, 45(9), pp. 984–994.

Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) ‘Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis’, The Lancet, 381(9875), pp. 1371–1379.

Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu and Cross-Disorder Group of the Psychiatric Genomics Consortium (2019) ‘Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders’, Cell, 179(7), pp. 1469–1482.e11.

Csardi, G., Nepusz, T. and Others (2006) ‘The igraph software package for complex network research’, InterJournal, complex systems, 1695(5), pp. 1–9.

Culham, J. C. and Valyear, K. F. (2006) ‘Human parietal cortex in action’, Current opinion in neurobiology. Elsevier, 16(2), pp. 205–212.

Cullen, B. et al. (2016) ‘Prevalence and correlates of cognitive impairment in euthymic adults with bipolar disorder: A systematic review’, Journal of affective disorders, 205, pp. 165–181.

Cunha-Cabral, D. et al. (2019) ‘Neurosurgical anatomy of the insular cortex’, Clinical neurology and neurosurgery. Elsevier, 186, p. 105530.

Curtis, R. K., Oresic, M. and Vidal-Puig, A. (2005) ‘Pathways to the analysis of microarray data’, Trends in biotechnology. Elsevier, 23(8), pp. 429–435.

Damier, P. et al. (1999) ‘The Substantia Nigra of the Human Brain. II. Patterns of Loss of Dopamine-Containing Neurons in Parkinson’s Disease’, Brain: a journal of neurology. Brain, 122 (Pt 8). doi: 10.1093/brain/122.8.1437.

Davidovic, M., Starck, G. and Olausson, H. (2019) ‘Processing of affective and emotionally neutral tactile stimuli in the insular cortex’, Developmental cognitive neuroscience, 35, pp. 94–103.

Dawes, P. et al. (2009) ‘Temporal auditory and visual motion processing of children diagnosed with auditory processing disorder and dyslexia’, Ear and hearing. journals.lww.com, 30(6), pp. 675–686.

Dean, D. C. et al. (2016) ‘Alterations of Myelin Content in Parkinson’s Disease: A Cross-Sectional Neuroimaging Study’, PLOS ONE, p. e0163774. doi: 10.1371/journal.pone.0163774.

Deffains, M. et al. (2016) ‘Subthalamic, not striatal, activity correlates with basal ganglia downstream activity in normal and parkinsonian monkeys’, eLife. elifesciences.org, 5. doi: 10.7554/eLife.16443.

Delvecchio, G. et al. (2019) ‘Structural and metabolic cerebral alterations between elderly bipolar disorder and behavioural variant frontotemporal dementia: A combined MRI-PET study’, The Australian and New Zealand journal of psychiatry. journals.sagepub.com, 53(5), pp. 413–423.

DeTure, M. A. and Dickson, D. W. (2019) ‘The neuropathological diagnosis of Alzheimer’s disease’, Molecular neurodegeneration, 14(1), p. 32.

Dickson, D. W. et al. (1992) ‘Identification of normal and pathological aging in prospectively studied nondemented elderly humans’, Neurobiology of aging, 13(1), pp. 179–189.

Dickson, D. W. (1997) ‘The Pathogenesis of Senile Plaques’, Journal of Neuropathology and Experimental Neurology, pp. 321–339. doi: 10.1097/00005072-199704000-00001.

Dijkstra, A. A. et al. (2014) ‘Stage-dependent nigral neuronal loss in incidental Lewy body and Parkinson’s disease’, Movement disorders: official journal of the Movement Disorder Society, 29(10), pp. 1244–1251.

Di Tommaso, P. et al. (2017) ‘Nextflow enables reproducible computational workflows’, Nature biotechnology, 35(4), pp. 316–319.

Dobin, A. et al. (2013) ‘STAR: ultrafast universal RNA-seq aligner’, Bioinformatics , 29(1), pp. 15–21.

Dominguez, M. H., Ayoub, A. E. and Rakic, P. (2013) ‘POU-III transcription factors (Brn1, Brn2, and Oct6) influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the cerebral cortex’, Cerebral cortex , 23(11), pp. 2632–2643.

Dumitriu, A. et al. (2016) 'Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease', BMC medical genomics, 9, p. 5.

'Early interventions to prevent psychosis: systematic review and meta-analysis' (2013) BMJ, pp. f762–f762. doi: 10.1136/bmj.f762.

Etkin, A., Büchel, C. and Gross, J. J. (2015) 'The neural bases of emotion regulation', Nature Reviews Neuroscience, pp. 693–700. doi: 10.1038/nrn4044.

Fakhoury, M. (2018) 'Microglia and Astrocytes in Alzheimer's Disease: Implications for Therapy', Current neuropharmacology, 16(5), pp. 508–518.

Fathy, Y. Y. et al. (2019) 'Differential insular cortex sub-regional atrophy in neurodegenerative diseases: a systematic review and meta-analysis', Brain Imaging and Behavior. doi: 10.1007/s11682-019-00099-3.

Fearnley, J. M. and Lees, A. J. (1991) 'Ageing and Parkinson's disease: substantia nigra regional selectivity', Brain: a journal of neurology, 114 (Pt 5), pp. 2283–2301.

Feng, Y. and Wang, X. (2017) 'Systematic analysis of microarray datasets to identify Parkinson's disease-associated pathways and genes', Molecular Medicine Reports, pp. 1252–1262. doi: 10.3892/mmr.2017.6124.

Fernandes, H. J. R. et al. (2016) 'ER Stress and Autophagic Perturbations Lead to Elevated Extracellular α-Synuclein in GBA-N370S Parkinson's iPSC-Derived Dopamine Neurons', Stem Cell Reports, pp. 342–356. doi: 10.1016/j.stemcr.2016.01.013.

Ferreira, S. A. and Romero-Ramos, M. (2018) 'Microglia Response During Parkinson's Disease: Alpha-Synuclein Intervention', Frontiers in Cellular Neuroscience. doi: 10.3389/fncel.2018.00247.

Ferrer, A. et al. (2019) 'BDNF genetic variants and methylation: effects on cognition in major depressive disorder', Translational psychiatry, 9(1), p. 265.

Fitzmaurice, G. M., Laird, N. M. and Ware, J. H. (2012) Applied Longitudinal Analysis. John Wiley & Sons.

Fleiss, B., Rivkees, S. A. and Gressens, P. (2018) 'Early origins of neuropsychiatric disorders', Pediatric research. Nature Publishing Group, 85(2), pp. 113–114.

Flint, J. and Kendler, K. S. (2014) 'The Genetics of Major Depression', Neuron, p. 1214. doi: 10.1016/j.neuron.2014.02.033.

Forster, T., Roy, D. and Ghazal, P. (2003) 'Experiments using microarray technology: limitations and standard operating procedures', The Journal of endocrinology. ncbi.nlm.nih.gov, 178(2), pp. 195–204.

Franzén, O., Gan, L.-M. and Björkegren, J. L. M. (2019) 'PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data', Database: the journal of biological databases and curation, 2019. doi: 10.1093/database/baz046.

Fromer, M. et al. (2016) 'Gene expression elucidates functional impact of polygenic risk for schizophrenia', Nature neuroscience, 19(11), pp. 1442–1453.

Fusar-Poli, P. and Meyer-Lindenberg, A. (2013) 'Striatal Presynaptic Dopamine in Schizophrenia, Part II: Meta-Analysis of [18F/11C]-DOPA PET Studies', Schizophrenia Bulletin, pp. 33–42. doi: 10.1093/schbul/sbr180.

Fuster, J. (2015) The Prefrontal Cortex. Academic Press.

Gaiteri, C. et al. (2016) 'Genetic variants in Alzheimer disease — molecular and brain network approaches', Nature reviews. Neurology. Nature Publishing Group, 12(7), pp. 413–427.

Galili, T. (2015) 'dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering', Bioinformatics , 31(22), pp. 3718–3720.

Galioto, R. et al. (2017) 'Depressive Symptoms Contribute to Executive Deficits in Temporal Lobe Epilepsy', The Journal of neuropsychiatry and clinical neurosciences. Am Neuropsych Assoc, 29(2), pp. 135–141.

Gandal, M. J. et al. (2018) 'Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap', Science, 359(6376), pp. 693–697.

Gan, L. et al. (2018) 'Converging pathways in neurodegeneration, from genetics to mechanisms', Nature neuroscience. Nature Publishing Group, 21(10), pp. 1300–1309.

Gao, P. P. et al. (1996) 'Regulation of topographic projection in the brain: Elf-1 in the hippocamposeptal system', Proceedings of the National Academy of Sciences of the United States of America, 93(20), pp. 11161–11166.

Gao, Y. et al. (2017) 'Changes of brain structure in Parkinson's disease patients with mild cognitive impairment analyzed via VBM technology', Neuroscience letters. Elsevier, 658, pp. 121–132.

Gaugler, T. et al. (2014) 'Most genetic risk for autism resides with common variation', Nature genetics, 46(8), pp. 881–885.

Gene Ontology Consortium (2015) 'Gene Ontology Consortium: going forward', Nucleic acids research. academic.oup.com, 43(Database issue), pp. D1049–56.

Geranmayeh, F. et al. (2012) 'The contribution of the inferior parietal cortex to spoken language production', Brain and language. Elsevier, 121(1), pp. 47–57.

Gershon, E. S. et al. (2014) 'A rare mutation of CACNA1C in a patient with bipolar disorder, and decreased gene expression associated with a bipolar-associated common SNP of CACNA1C in brain', Molecular psychiatry, 19(8), pp. 890–894.

Geschwind, D. H. and Flint, J. (2015) ‘Genetics and genomics of psychiatric disease’, Science, pp. 1489–1494. doi: 10.1126/science.aaa8954.

Ghosh, B. (2007) ‘Functional Areas of Cerebrum’, Human Anatomy for Students, pp. 304–304. doi: 10.5005/jp/books/10364_89.

Gogolla, N. (2017) ‘The insular cortex’, Current biology: CB. Elsevier, 27(12), pp. R580–R586.

Golde, T. E., Eckman, C. B. and Younkin, S. G. (2000) ‘Biochemical detection of Aβ isoforms: implications for pathogenesis, diagnosis, and treatment of Alzheimer’s disease’, Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, pp. 172–187. doi: 10.1016/s0925-4439(00)00043-0.

Gonatopoulos-Pournatzis, T. et al. (2020) ‘Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions’, Molecular Cell. doi: 10.1016/j.molcel.2020.01.006.

Graeber, M. B. and Stre’rt, W. J. (1990) ‘Microglia: Immune Network in the CNS’, Brain Pathology, pp. 2–5. doi: 10.1111/j.1750-3639.1990.tb00630.x.

Greenberg, S. M. et al. (2019) ‘Cerebral amyloid angiopathy and Alzheimer disease — one peptide, two pathways’, Nature reviews. Neurology. Nature Publishing Group, 16(1), pp. 30–42.

Gremo, F. et al. (1997) ‘Features and Functions of Human Microglia Cells’, in Filogamo, G. et al. (eds) Brain Plasticity: Development and Aging. Boston, MA: Springer US, pp. 79–97.

Gulsuner, S. et al. (2013) ‘Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network’, Cell, 154(3), pp. 518–529.

Guo, S. et al. (2014) ‘IRF4 is a novel mediator for neuronal survival in ischaemic stroke’, Cell death and differentiation, 21(6), pp. 888–903.

Gu, Z., Eils, R. and Schlesner, M. (2016) ‘Complex heatmaps reveal patterns and correlations in multidimensional genomic data’, Bioinformatics , 32(18), pp. 2847–2849.

Gyoneva, S. et al. (2019) ‘Cx3cr1-deficient microglia exhibit a premature aging transcriptome’, Life Science Alliance, p. e201900453. doi: 10.26508/lsa.201900453.

Hafemeister, T. L. (2019) ‘Mental Disorders and Criminal Behavior’, Criminal Trials and Mental Disorders, pp. 7–42. doi: 10.18574/nyu/9781479804856.003.0002.

Han, K.-M. et al. (2017) ‘Influence of FKBP5 polymorphism and DNA methylation on structural changes of the brain in major depressive disorder’, Scientific reports, 7, p. 42621.

Hansen, K. D., Irizarry, R. A. and Wu, Z. (2012) ‘Removing technical variability in RNA-seq data using conditional quantile normalization’, Biostatistics , 13(2), pp. 204–216.

Han, W.-S. et al. (2010) ‘iGraph: a framework for comparisons of disk-based graph indexing techniques’, Proceedings of the VLDB Endowment International Conference on Very Large Data Bases. VLDB Endowment, 3(1-2), pp. 449–459.

Hare, B. D. and Duman, R. S. (2020) ‘Prefrontal cortex circuits in depression and anxiety: contribution of discrete neuronal populations and target regions’, Molecular psychiatry. doi: 10.1038/s41380-020-0685-9.

Hariri, A. (2011) ‘Faculty Opinions recommendation of How the serotonin transporter 5-HTTLPR polymorphism influences amygdala function: the roles of in vivo serotonin transporter expression and amygdala structure’, Faculty Opinions – Post-Publication Peer Review of the Biomedical Literature. doi: 10.3410/f.13369993.14740107.

Harris, M. A. et al. (2004) ‘The Gene Ontology (GO) database and informatics resource’, Nucleic acids research. academic.oup.com, 32(Database issue), pp. D258–61.

Hassanpour, M. S. et al. (2018) ‘The Insular Cortex Dynamically Maps Changes in Cardiorespiratory Interoception’, Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology. nature.com, 43(2), pp. 426–434.

Heinrich, A. et al. (2013) ‘The risk variant in ODZ4 for bipolar disorder impacts on amygdala activation during reward processing’, Bipolar disorders, 15(4), pp. 440–445.

Hemonnot, A.-L. et al. (2019) ‘Microglia in Alzheimer Disease: Well-Known Targets and New Opportunities’, Frontiers in aging neuroscience, 11, p. 233.

Henderson-Smith, A. et al. (2016) ‘Next-generation profiling to identify the molecular etiology of Parkinson dementia’, Neurology. Genetics, 2(3), p. e75.

He, Z. et al. (2014) ‘Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation’, RNA , 20(7), pp. 1103–1111.

Hinz, F. I. and Geschwind, D. H. (2017) ‘Molecular Genetics of Neurodegenerative Dementias’, Cold Spring Harbor Perspectives in Biology, p. a023705. doi: 10.1101/cshperspect.a023705.

Hirai, T. et al. (2000) ‘Limbic lobe of the human brain: evaluation with turbo fluid-attenuated inversion-recovery MR imaging’, Radiology, 215(2), pp. 470–475.

Hoffman, G. E. et al. (2019) ‘CommonMind Consortium provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder’, Scientific data, 6(1), p. 180.

Höglinger, G. et al. (2011) ‘V37 Common variants affect risk for the tauopathy progressive supranuclear palsy’, Basal Ganglia, p. 14. doi: 10.1016/j.baga.2011.01.006.

Höglinger, G. U. et al. (2011) ‘Identification of Common Variants Influencing Risk of the Tauopathy Progressive Supranuclear Palsy’, Nature genetics. Nat Genet, 43(7). doi: 10.1038/ng.859.

Hosseini, E. et al. (2019) 'The importance of long non-coding RNAs in neuropsychiatric disorders', Molecular aspects of medicine, 70, pp. 127–140.

Hossein-nezhad, A. et al. (2016) 'Transcriptomic Profiling of Extracellular RNAs Present in Cerebrospinal Fluid Identifies Differentially Expressed Transcripts in Parkinson's Disease', Journal of Parkinson's Disease, pp. 109–117. doi: 10.3233/jpd-150737.

Hou, Y. et al. (2019) 'Ageing as a risk factor for neurodegenerative disease', Nature reviews. Neurology. Nature Publishing Group, 15(10), pp. 565–581.

Howes, O. D. and Murray, R. M. (2014) 'Schizophrenia: an integrated sociodevelopmental-cognitive model', The Lancet, 383(9929), pp. 1677–1687.

Humphreys, K. L. et al. (2019) 'DNA methylation of HPA-axis genes and the onset of major depressive disorder in adolescent girls: a prospective analysis', Translational psychiatry, 9(1), p. 245.

Hunt, J. (2008) 'Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression, Second Edition', Journal of the American Academy of Child & Adolescent Psychiatry, pp. 1208–1209. doi: 10.1097/01.chi.0000313989.41405.94.

Hu, W. et al. (2015) 'The glutamate hypothesis of schizophrenia: evidence from human brain tissue studies', Annals of the New York Academy of Sciences. NIH Public Access, 1338(1), p. 38.

Hwang, J. et al. (2006) 'Basal ganglia shape alterations in bipolar disorder', The American journal of psychiatry, 163(2), pp. 276–285.

Hyde, C. L. et al. (2016) 'Identification of 15 genetic loci associated with risk of major depression in individuals of European descent', Nature Genetics, pp. 1031–1036. doi: 10.1038/ng.3623.

Ichikawa, N. et al. (2020) 'Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants', Scientific reports, 10(1), p. 3542.

'Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis' (2013) The Lancet, pp. 1371–1379. doi: 10.1016/s0140-6736(12)62129-1.

International Schizophrenia Consortium et al. (2009) 'Common polygenic variation contributes to risk of schizophrenia and bipolar disorder', Nature, 460(7256), pp. 748–752.

Iossifov, I. et al. (2014) 'The contribution of de novo coding mutations to autism spectrum disorder', Nature, pp. 216–221. doi: 10.1038/nature13908.

Irimia, M. et al. (2014) 'A highly conserved program of neuronal microexons is misregulated in autistic brains', Cell, 159(7), pp. 1511–1523.

Iwatsubo, T. et al. (1994) ‘Visualization of Aβ42(43) and Aβ40 in senile plaques with end-specific Aβ monoclonals: Evidence that an initially deposited species is Aβ42(43)’, Neuron, pp. 45–53. doi: 10.1016/0896-6273(94)90458-8.

Jack, C. R., Jr et al. (1998) ‘Rate of medial temporal lobe atrophy in typical aging and Alzheimer’s disease’, Neurology. AAN Enterprises, 51(4), pp. 993–999.

Jaffe, A. E. et al. (2018) ‘Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis’, Nature neuroscience, 21(8), pp. 1117–1125.

Jäkel, S. and Dimou, L. (2017) ‘Glial Cells and Their Function in the Adult Brain: A Journey through the History of Their Ablation’, Frontiers in cellular neuroscience, 11, p. 24.

Janitz, M. (2011) Next-Generation Genome Sequencing: Towards Personalized Medicine. John Wiley & Sons.

Jiang, X. et al. (2019) ‘Sodium valproate rescues expression of TRANK1 in iPSC-derived neural cells that carry a genetic variant associated with serious mental illness’, Molecular psychiatry, 24(4), pp. 613–624.

Jia, Y. et al. (2020) ‘Limbic encephalitis associated with AMPA receptor and CRMP5 antibodies: A case report and literature review’, Brain and behavior, p. e01528.

Johnston, D. and Amaral, D. G. (2004) ‘Hippocampus’, The synaptic organization of the brain., 5th ed. New York, NY, US: Oxford University Press, xiv, 5, pp. 455–498.

Kaalund, S. S. et al. (2014) ‘Contrasting changes in DRD1 and DRD2 splice variant expression in schizophrenia and affective disorders, and associations with SNPs in postmortem brain’, Molecular Psychiatry, pp. 1258–1266. doi: 10.1038/mp.2013.165.

Kahn, R. S. et al. (2015) ‘Schizophrenia’, Nature Reviews Disease Primers. Nature Publishing Group, 1(1), pp. 1–23.

Kamboh, M. I. et al. (2006) ‘Alpha-1-antichymotrypsin (ACT or SERPINA3) polymorphism may affect age-at-onset and disease duration of Alzheimer’s disease’, Neurobiology of Aging, pp. 1435–1439. doi: 10.1016/j.neurobiolaging.2005.07.015.

Karanian, J. M. and Slotnick, S. D. (2017) ‘False memory for context and true memory for context similarly activate the parahippocampal cortex’, Cortex; a journal devoted to the study of the nervous system and behavior. Elsevier, 91, pp. 79–88.

Kaschka, W. P. (2002) Perspectives in Affective Disorders. Karger Medical and Scientific Publishers.

Kaufmann, T. et al. (2019) ‘Common brain disorders are associated with heritable patterns of apparent aging of the brain’, Nature neuroscience, 22(10), pp. 1617–1623.

Kaushik, S. and Cuervo, A. M. (2015) ‘Proteostasis and aging’, Nature medicine. Nature Publishing Group, 21(12), pp. 1406–1415.

Kelly, J. et al. (2020) ‘Gene Expression Meta-Analysis of Parkinson’s disease and its Relationship with Alzheimer’s disease’, Prime Archives in Molecular Biology. doi: 10.37247/pamb.1.2020.32.

Keo, A. et al. (no date) ‘Transcriptomic signatures of brain regional vulnerability to Parkinson’s disease’. doi: 10.1101/664771.

Kern, J. K. et al. (2013) ‘Evidence of neurodegeneration in autism spectrum disorder’, Translational neurodegeneration, 2(1), p. 17.

Kim, D.-Y. et al. (2015) ‘Functional regulation of FoxO1 in neural stem cell differentiation’, Cell Death & Differentiation, pp. 2034–2045. doi: 10.1038/cdd.2015.123.

Knowles, D. G. et al. (2013) ‘Grape RNA-Seq analysis pipeline environment’, Bioinformatics, pp. 614–621. doi: 10.1093/bioinformatics/btt016.

Kodama, L. et al. (2020) ‘Microglial microRNAs mediate sex-specific responses to tau pathology’, Nature neuroscience, 23(2), pp. 167–171.

Koffie, R. M. et al. (2012) ‘Apolipoprotein E4 effects in Alzheimer’s disease are mediated by synaptotoxic oligomeric amyloid-β’, Brain: a journal of neurology, 135(Pt 7), pp. 2155–2168.

Kolde, R. and Kolde, M. R. (2015) ‘Package “pheatmap”’, R Package. mran.microsoft.com, 1(7), p. 790.

Krawczyk-Marć, I. et al. (2019) ‘Oligodendrocytes: Morphology, functions and involvement in neurodegenerative diseases’, Medycyna Weterynaryjna, pp. 6258–2019. doi: 10.21521/mw.6258.

Krijthe, J. H. (2015) ‘Rtsne: T-distributed stochastic neighbor embedding using Barnes-Hut implementation’, R package version 0. 13, URL https://github. com/jkrijthe/Rtsne.

Krystel-Whittemore, M., Dileepan, K. N. and Wood, J. G. (2015) ‘Mast Cell: A Multi-Functional Master Cell’, Frontiers in immunology, 6, p. 620.

Kuhn, M. and Others (2008) ‘Building predictive models in R using the caret package’, Journal of statistical software, 28(5), pp. 1–26.

Kumar-Singh, S. et al. (2006) ‘Mean age-of-onset of familial alzheimer disease caused by presenilin mutations correlates with both increased Aβ42 and decreased Aβ40’, Human Mutation, pp. 686–695. doi: 10.1002/humu.20336.

Kupfer, D. J., Frank, E. and Phillips, M. L. (2016) ‘Major Depressive Disorder: New Clinical, Neurobiological, and Treatment Perspectives’, Focus , 14(2), pp. 266–276.

Labonté, B. et al. (2017) ‘Sex-specific transcriptional signatures in human depression’, Nature medicine, 23(9), pp. 1102–1111.

Langfelder, P. and Horvath, S. (2008) ‘WGCNA: an R package for weighted correlation network analysis’, BMC Bioinformatics. doi: 10.1186/1471-2105-9-559.

Larsson, J. and Heeger, D. J. (2006) ‘Two Retinotopic Visual Areas in Human Lateral Occipital Cortex’, Journal of Neuroscience, pp. 13128–13142. doi: 10.1523/jneurosci.1657-06.2006.

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2014) ‘sva: Surrogate Variable Analysis’, R package version 3.10. 0., 10, p. B9.

Lei, Z., Brizzee, C. and Johnson, G. V. W. (2015) ‘BAG3 facilitates the clearance of endogenous tau in primary neurons’, Neurobiology of aging, 36(1), pp. 241–248.

Lewis, D. A. (2014) ‘Inhibitory neurons in human cortical circuits: substrate for cognitive dysfunction in schizophrenia’, Current opinion in neurobiology, 26, pp. 22–26.

Li, B. and Dewey, C. N. (2011) ‘RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome’, BMC Bioinformatics. doi: 10.1186/1471-2105-12-323.

Lichtenstein, P. et al. (2009) ‘Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study’, The Lancet, pp. 234–239. doi: 10.1016/s0140-6736(09)60072-6.

Li, H., Ruan, J. and Durbin, R. (2008) ‘Mapping short DNA sequencing reads and calling variants using mapping quality scores’, Genome research, 18(11), pp. 1851–1858.

Li, J. et al. (2014) ‘Integrated systems analysis reveals a molecular network underlying autism spectrum disorders’, Molecular systems biology, 10, p. 774.

Li, K. et al. (2019) ‘Reactive Astrocytes in Neurodegenerative Diseases’, Aging and disease, p. 664. doi: 10.14336/ad.2018.0720.

Li, M., D’Arcy, C. and Meng, X. (2016) ‘Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions’, Psychological medicine, 46(4), pp. 717–730.

Lin, C.-X. et al. (2015) ‘Association between HTR2A T102C polymorphism and major depressive disorder: a meta-analysis in the Chinese population’, International journal of clinical and experimental medicine, 8(11), pp. 20897–20903.

Lindstrom, M. J. and Bates, D. M. (1988) ‘Newton—Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data’, Journal of the American Statistical Association. Taylor & Francis, 83(404), pp. 1014–1022.

Li, P. C. H. (2016) ‘Overview of Microarray Technology’, Methods in Molecular Biology, pp. 3–4. doi: 10.1007/978-1-4939-3136-1_1.

Lipska, B. K. et al. (2006) ‘Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia’, Biological psychiatry, 60(6), pp. 650–658.

Li, R. et al. (2008) ‘SOAP: short oligonucleotide alignment program’, Bioinformatics , 24(5), pp. 713–714.

Liu, C. et al. (2019) ‘Basal ganglia volumetric changes in psychotic spectrum disorders’, Journal of affective disorders, 255, pp. 150–157.

Liu, G. T., Volpe, N. J. and Galetta, S. L. (2019) ‘12 - Visual Hallucinations and Illusions’, in Liu, G. T., Volpe, N. J., and Galetta, S. L. (eds) Liu, Volpe, and Galetta’s Neuro-Ophthalmology (Third Edition). Elsevier, pp. 395–413.

Liu, W. et al. (2020) ‘Disrupted pathways from frontal-parietal cortex to basal ganglia and cerebellum in patients with unmedicated obsessive compulsive disorder as observed by whole-brain resting-state effective connectivity analysis – a small sample pilot study’, Brain Imaging and Behavior. doi: 10.1007/s11682-020-00333-3.

Liu, X. et al. (2016) ‘Disruption of an Evolutionarily Novel Synaptic Expression Pattern in Autism’, PLOS Biology, p. e1002558. doi: 10.1371/journal.pbio.1002558.

Liu, Y. and Zhou, J. (2013) ‘Oligodendrocytes in neurodegenerative diseases’, Frontiers in Biology, pp. 127–133. doi: 10.1007/s11515-013-1260-4.

Logotheti, M. et al. (2013) ‘A comparative genomic study in schizophrenic and in bipolar disorder patients, based on microarray expression profiling meta-analysis’, TheScientificWorldJournal. hindawi.com, 2013, p. 685917.

Lord, C. et al. (2020) ‘Autism spectrum disorder’, Nature Reviews Disease Primers. Nature Publishing Group, 6(1), pp. 1–23.

Lott, I. T. and Head, E. (2019) ‘Dementia in Down syndrome: unique insights for Alzheimer disease research’, Nature reviews. Neurology. Nature Publishing Group, 15(3), pp. 135–147.

Lynch, M. A. et al. (2016) ‘Common Themes in the Pathogenesis of Neurodegeneration’, Neurodegenerative Disorders, pp. 1–12. doi: 10.1007/978-3-319-23309-3_1.

Maarouf, C. L. et al. (2011) ‘Alzheimer’s Disease and Non-Demented High Pathology Control Nonagenarians: Comparing and Contrasting the Biochemistry of Cognitively Successful Aging’, PLoS ONE, p. e27291. doi: 10.1371/journal.pone.0027291.

Maaten, L. van der and Hinton, G. (2008) ‘Visualizing Data using t-SNE’, Journal of machine learning research: JMLR. jmlr.org, 9(Nov), pp. 2579–2605.

MacMullen, C. M., Fallahi, M. and Davis, R. L. (2017) ‘Novel PDE10A transcript diversity in the human striatum: Insights into gene complexity, conservation and regulation’, Gene, 606, pp. 17–24.

Macpherson, T. and Hikida, T. (2019) 'Role of basal ganglia neurocircuitry in the pathology of psychiatric disorders', Psychiatry and clinical neurosciences, 73(6), pp. 289–301.

Mah, L., Arnold, M. C. and Grafman, J. (2004) 'Impairment of social perception associated with lesions of the prefrontal cortex', The American journal of psychiatry. Am Psychiatric Assoc, 161(7), pp. 1247–1255.

Maidan, I. et al. (2016) 'The Role of the Frontal Lobe in Complex Walking Among Patients With Parkinson's Disease and Healthy Older Adults: An fNIRS Study', Neurorehabilitation and neural repair, 30(10), pp. 963–971.

Mai, J. K., Majtanik, M. and Paxinos, G. (2015) Atlas of the Human Brain. Academic Press.

Marshall, J. C. and Morriss-Kay, G. M. (2004) 'Functional anatomy of the human brain', Journal of Anatomy, pp. 415–415. doi: 10.1111/j.0021-8782.2004.00360.x.

Matias, I., Morgado, J. and Gomes, F. C. A. (2019) 'Astrocyte Heterogeneity: Impact to Brain Aging and Disease', Frontiers in Aging Neuroscience. doi: 10.3389/fnagi.2019.00059.

Matsuoka, K. et al. (2018) 'Left dorsolateral prefrontal cortex atrophy is associated with frontal lobe function in Alzheimer's disease and contributes to caregiver burden', International journal of geriatric psychiatry, 33(5), pp. 703–709.

McPartland, J. C., Coffman, M. and Pelphrey, K. A. (2011) 'Recent advances in understanding the neural bases of autism spectrum disorder', Current opinion in pediatrics, 23(6), pp. 628–632.

Menassa, D. A. and Gomez-Nicola, D. (2018) 'Microglial Dynamics During Human Brain Development', Frontiers in immunology, 9, p. 1014.

Mendonça, M. S. et al. (2019) 'Epigenetic variation at the SLC6A4 gene promoter in mother-child pairs with major depressive disorder', Journal of affective disorders, 245, pp. 716–723.

Menke, A. et al. (2013) 'Genetic variation in FKBP5 associated with the extent of stress hormone dysregulation in major depression', Genes, Brain and Behavior, pp. 289–296. doi: 10.1111/gbb.12026.

Menon, V. and Uddin, L. Q. (2010) 'Saliency, switching, attention and control: a network model of insula function', Brain structure & function, 214(5-6), pp. 655–667.

Merikangas, K. R. et al. (2007) 'Lifetime and 12-Month Prevalence of Bipolar Spectrum Disorder in the National Comorbidity Survey Replication', Archives of General Psychiatry, p. 543. doi: 10.1001/archpsyc.64.5.543.

Merikangas, K. R. et al. (2011) 'Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative', Archives of general psychiatry, 68(3), pp. 241–251.

Mesulam, M.-M. and -marsel Mesulam, M. (2002) 'Brain Anatomy and Networks*', Encyclopedia of the Human Brain, pp. 469–480. doi: 10.1016/b0-12-227210-2/00069-8.

Mihaylova, M. et al. (2017) ‘Benign, pathogenic and copy number variations of unknown clinical significance in patients with congenital malformations and developmental delay’, Balkan Journal of Medical Genetics, pp. 5–12. doi: 10.1515/bjmg-2017-0010.

Miller, J. A. et al. (2013) ‘Genes and pathways underlying regional and cell type changes in Alzheimer’s disease’, Genome medicine. BioMed Central, 5(5), pp. 1–17.

Milosevic, L. et al. (2018) ‘Neuronal inhibition and synaptic plasticity of basal ganglia neurons in Parkinson’s disease’, Brain: a journal of neurology. academic.oup.com, 141(1), pp. 177–190.

Mirnics, K. et al. (2000) ‘Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex’, Neuron, 28(1), pp. 53–67.

Mirnics, K. et al. (2001) ‘Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia’, Molecular psychiatry, 6(3), pp. 293–301.

Miyoshi, K. and Morimura, Y. (2010) ‘Clinical Manifestations of Neuropsychiatric Disorders’, in Neuropsychiatric Disorders. Springer, Tokyo, pp. 1–14.

Monfreda, M. (2012) ‘Principal Component Analysis: A Powerful Interpretative Tool at the Service of Analytical Methodology’, Principal Component Analysis. doi: 10.5772/36929.

Moore, B. D. et al. (2012) ‘Overlapping profiles of Abeta peptides in the Alzheimer’s disease and pathological aging brains’, Alzheimer’s Research & Therapy, p. 18. doi: 10.1186/alzrt121.

Moran, L. V. et al. (2013) ‘Disruption of anterior insula modulation of large-scale brain networks in schizophrenia’, Biological psychiatry, 74(6), pp. 467–474.

Mormina, E. et al. (2017) ‘Cerebellum and neurodegenerative diseases: Beyond conventional magnetic resonance imaging’, World journal of radiology, 9(10), pp. 371–388.

Mortazavi, A. et al. (2008) ‘Mapping and quantifying mammalian transcriptomes by RNA-Seq’, Nature methods, 5(7), pp. 621–628.

Mosser, C.-A. et al. (2017) ‘Microglia in CNS development: Shaping the brain for the future’, Progress in neurobiology, 149-150, pp. 1–20.

Mostafavi, S. et al. (2018) ‘A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer’s disease’, Nature neuroscience. Nature Publishing Group, 21(6), pp. 811–819.

Mühleisen, T. W. et al. (2014) ‘Genome-wide association study reveals two new risk loci for bipolar disorder’, Nature communications, 5, p. 3339.

Murray, M. E. and Dickson, D. W. (2014) ‘Is pathological aging a successful resistance against amyloid-beta or preclinical Alzheimer’s disease?’, Alzheimer’s Research & Therapy, p. 24. doi: 10.1186/alzrt254.

Muszyński, P. et al. (2017) 'YKL-40 as a Potential Biomarker and a Possible Target in Therapeutic Strategies of Alzheimer's Disease', Current Neuropharmacology. doi: 10.2174/1570159x15666170208124324.

Nagy, C. et al. (2020) 'Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons', Nature neuroscience, 23(6), pp. 771–781.

Nalls, M. A. et al. (2014) 'Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease', Nature genetics. Nature Publishing Group, 46(9), pp. 989–993.

Namba, Y. et al. (1991) 'Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease', Brain research, 541(1), pp. 163–166.

Namkung, H., Kim, S.-H. and Sawa, A. (2017) 'The insula: an underestimated brain area in clinical neuroscience, psychiatry, and neurology', Trends in neurosciences. NIH Public Access, 40(4), p. 200.

Narayanan, M. et al. (2014) 'Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases', Molecular systems biology. embopress.org, 10, p. 743.

Narayanan, N. S., Rodnitzky, R. L. and Uc, E. Y. (2013) 'Prefrontal dopamine signaling and cognitive symptoms of Parkinson's disease', Reviews in the Neurosciences. doi: 10.1515/revneuro-2013-0004.

Nativio, R. et al. (2018) 'Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease', Nature neuroscience. nature.com, 21(4), pp. 497–505.

Nazeri, A. et al. (2014) 'Imaging proteomics for diagnosis, monitoring and prediction of Alzheimer's disease', NeuroImage, 102 Pt 2, pp. 657–665.

Neul, J. L. (2012) 'The relationship of Rett syndrome and MECP2 disorders to autism', Dialogues in clinical neuroscience, 14(3), pp. 253–262.

Ng, B. et al. (2017) 'Distinct alterations in Parkinson's medication-state and disease-state connectivity', NeuroImage: Clinical, pp. 575–585. doi: 10.1016/j.nicl.2017.09.004.

Nickl-Jockschat, T. et al. (2012) 'Brain structure anomalies in autism spectrum disorder--a meta-analysis of VBM studies using anatomic likelihood estimation', Human brain mapping, 33(6), pp. 1470–1489.

Nonen, S. et al. (2016) 'Polymorphism of rs3813034 in Serotonin Transporter Gene SLC6A4 Is Associated With the Selective Serotonin and Serotonin-Norepinephrine Reuptake Inhibitor Response in Depressive Disorder', Journal of Clinical Psychopharmacology, pp. 27–31. doi: 10.1097/jcp.0000000000000454.

Okonechnikov, K., Conesa, A. and García-Alcalde, F. (2016) 'Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data', Bioinformatics , 32(2), pp. 292–294.

Otte, C. et al. (2016) 'Major depressive disorder', Nature reviews. Disease primers, 2, p. 16065.

Pacifico, R. and Davis, R. L. (2017) 'Transcriptome sequencing implicates dorsal striatum-specific gene network, immune response and energy metabolism pathways in bipolar disorder', Molecular psychiatry, 22(3), pp. 441–449.

Pandya, M. et al. (2012) 'Where in the brain is depression?', Current psychiatry reports, 14(6), pp. 634–642.

Pantazatos, S. P. et al. (2017) 'Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity', Molecular psychiatry, 22(5), pp. 760–773.

Panza, F. et al. (2019) 'A critical appraisal of amyloid-β-targeting therapies for Alzheimer disease', Nature reviews. Neurology. Nature Publishing Group, 15(2), pp. 73–88.

Pardiñas, A. F. et al. (2019) 'Publisher Correction: Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection', Nature genetics, 51(7), p. 1193.

Parikshak, N. N., Gandal, M. J. and Geschwind, D. H. (2015) 'Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders', Nature reviews. Genetics, 16(8), pp. 441–458.

Patel, K. R. et al. (2014) 'Schizophrenia: Overview and Treatment Options', Pharmacy and Therapeutics. MediMedia, USA, 39(9), p. 638.

Patel, S. et al. (2008) 'Gene expression profiles of mouse striatum in control and maneb + paraquat-induced Parkinson's disease phenotype: validation of differentially expressed energy metabolizing transcripts', Molecular biotechnology, 40(1), pp. 59–68.

Patriquin, M. A. et al. (2016) 'Neuroanatomical and neurofunctional markers of social cognition in autism spectrum disorder', Human brain mapping, 37(11), pp. 3957–3978.

Pelphrey, K. A. et al. (2011) 'Research review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder', Journal of child psychology and psychiatry, and allied disciplines, 52(6), pp. 631–644.

Phillips, C. (2017) 'Brain-Derived Neurotrophic Factor, Depression, and Physical Activity: Making the Neuroplastic Connection', Neural plasticity, 2017, p. 7260130.

Pinheiro, J. et al. (2012) 'nlme: Linear and nonlinear mixed effects models', R package version, 3(0).

Plaisier, S. B. et al. (2010) ‘Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures’, Nucleic Acids Research, pp. e169–e169. doi: 10.1093/nar/gkq636.

Poewe, W. et al. (2017) ‘Parkinson disease’, Nature Reviews Disease Primers. Nature Publishing Group, 3(1), pp. 1–21.

Pouget, A. and Sejnowski, T. J. (1997) ‘A new view of hemineglect based on the response properties of parietal neurones’, Philosophical transactions of the Royal Society of London. Series B, Biological sciences. royalsocietypublishing.org, 352(1360), pp. 1449–1459.

PsychENCODE Consortium et al. (2015) ‘The PsychENCODE project’, Nature neuroscience, 18(12), pp. 1707–1712.

PsychENCODE Consortium (2018) ‘Revealing the brain’s molecular architecture’, Science, 362(6420), pp. 1262–1263.

Rajmohan, V. and Mohandas, E. (2007) ‘The limbic system’, Indian journal of psychiatry. ncbi.nlm.nih.gov, 49(2), pp. 132–139.

Ramachandran, V. S. (2002) Encyclopedia of the Human Brain, Four-Volume Set. Academic Press.

Ramaker, R. C. et al. (2017) ‘Post-mortem molecular profiling of three psychiatric disorders’, Genome medicine, 9(1), p. 72.

Ramirez-Gonzalez, R. H. et al. (2012) ‘bio-samtools: Ruby bindings for SAMtools, a library for accessing BAM files containing high-throughput sequence alignments’, Open Research Computation, p. 1. doi: 10.1186/2042-5767-1-1.

Ran, L. et al. (2020) ‘Rare variants in SLC6A4 cause susceptibility to major depressive disorder with suicidal ideation in Han Chinese adolescents and young adults’, Gene, 726, p. 144147.

Raudvere, U. et al. (2019) ‘g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update)’, Nucleic Acids Research, pp. W191–W198. doi: 10.1093/nar/gkz369.

Reemst, K. et al. (2016) ‘The Indispensable Roles of Microglia and Astrocytes during Brain Development’, Frontiers in human neuroscience, 10, p. 566.

Reichenberg, A. et al. (2014) ‘The Course and Correlates of Everyday Functioning in Schizophrenia’, Schizophrenia research. Cognition, 1(1), pp. e47–e52.

Respondek, G. et al. (2014) ‘The phenotypic spectrum of progressive supranuclear palsy: a retrospective multicenter study of 100 definite cases’, Movement disorders: official journal of the Movement Disorder Society, 29(14), pp. 1758–1766.

Rhinn, H. and Abeliovich, A. (2017) 'Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes', Cell systems, 4(4), pp. 404–415.e5.

Risi, S. et al. (2006) 'Combining Information From Multiple Sources in the Diagnosis of Autism Spectrum Disorders', Journal of the American Academy of Child & Adolescent Psychiatry, pp. 1094–1103. doi: 10.1097/01.chi.0000227880.42780.0e.

Ritchie, M. E. et al. (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies', Nucleic acids research, 43(7), p. e47.

Rodriguez-Vieitez, E. et al. (2016) 'Diverging longitudinal changes in astrocytosis and amyloid PET in autosomal dominant Alzheimer's disease', Brain, pp. 922–936. doi: 10.1093/brain/awv404.

Rothe, T. et al. (2017) 'The Nuclear Receptor Nr4a1 Acts as a Microglia Rheostat and Serves as a Therapeutic Target in Autoimmune-Driven Central Nervous System Inflammation', Journal of immunology , 198(10), pp. 3878–3885.

Sakai, H. et al. (2019) 'Plag1 regulates neuronal gene expression and neuronal differentiation of neocortical neural progenitor cells', Genes to Cells, pp. 650–666. doi: 10.1111/gtc.12718.

Sakurai, K. et al. (2019) 'Characteristic asymmetric limbic and anterior temporal atrophy in demented patients with pathologically confirmed argyrophilic grain disease', Neuroradiology. Springer, 61(11), pp. 1239–1249.

Salleh, M. R. (2004) 'The genetics of schizophrenia', The Malaysian journal of medical sciences: MJMS, 11(2), pp. 3–11.

Salter, M. W. and Stevens, B. (2017) 'Microglia emerge as central players in brain disease', Nature Medicine, pp. 1018–1027. doi: 10.1038/nm.4397.

Sampath, D., Sathyanesan, M. and Newton, S. (2017) 'Cognitive dysfunction in major depression and Alzheimer's disease is associated with hippocampus–prefrontal cortex dysconnectivity', Neuropsychiatric Disease and Treatment, pp. 1509–1519. doi: 10.2147/ndt.s136122.

Sanders, S. J. et al. (2015) 'Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci', Neuron, 87(6), pp. 1215–1233.

Santiago, J. A., Bottero, V. and Potashkin, J. A. (2017) 'Dissecting the Molecular Mechanisms of Neurodegenerative Diseases through Network Biology', Frontiers in aging neuroscience, 9, p. 166.

Satterstrom, F. K. et al. (2020) 'Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism', Cell, 180(3), pp. 568–584.e23.

Saunders, A. M. et al. (2011) ‘Association of apolipoprotein E allele 4 with late-onset familial and sporadic Alzheimer’s disease’, Neurology, pp. 950–950. doi: 10.1212/01.wnl.0000405291.74999.10.

Schapira, A. H. V. (2008) ‘Mitochondrial Dysfunction in Neurodegenerative Diseases’, Neurochemical Research, pp. 2502–2509. doi: 10.1007/s11064-008-9855-x.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) ‘Biological insights from 108 schizophrenia-associated genetic loci’, Nature, 511(7510), pp. 421–427.

Schmidt, M. J. and Mirnics, K. (2015) ‘Neurodevelopment, GABA system dysfunction, and schizophrenia’, Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology, 40(1), pp. 190–206.

Schmidt, W. J. (2005) ‘Dopamine and Glutamate in Motor and Cognitive Symptoms of Parkinson’s Disease’, Dopamine and Glutamate in Psychiatric Disorders, pp. 523–535. doi: 10.1007/978-1-59259-852-6_22.

Schneider, I. et al. (2018) ‘Association of Serotonin Transporter Gene AluJb Methylation with Major Depression, Amygdala Responsiveness, 5-HTTLPR/rs25531 Polymorphism, and Stress’, Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology, 43(6), pp. 1308–1316.

Schultz, C. C. et al. (2014) ‘Common variation in NCAN, a risk factor for bipolar disorder and schizophrenia, influences local cortical folding in schizophrenia’, Psychological medicine, 44(4), pp. 811–820.

Schulze, A. and Downward, J. (2001) ‘Navigating gene expression using microarrays--a technology review’, Nature cell biology, 3(8), pp. E190–5.

Schuster, S. C. (2008) ‘Next-generation sequencing transforms today’s biology’, Nature methods. nature.com, 5(1), pp. 16–18.

Shen, L. (2014) ‘GeneOverlap: An R package to test and visualize gene overlaps’, R Package.

Shiffler, R. E. (1988) ‘Maximum Z Scores and Outliers’, The American Statistician, p. 79. doi: 10.2307/2685269.

Sidransky, E. et al. (2009) ‘Multicenter analysis of glucocerebrosidase mutations in Parkinson’s disease’, The New England journal of medicine, 361(17), pp. 1651–1661.

da Silva, E. Z. M., Jamur, M. C. and Oliver, C. (2014) ‘Mast cell function: a new vision of an old cell’, The journal of histochemistry and cytochemistry: official journal of the Histochemistry Society, 62(10), pp. 698–738.

Singh-Bains, M. K. et al. (2019) ‘Cerebellar degeneration correlates with motor symptoms in Huntington disease’, Annals of neurology, 85(3), pp. 396–405.

Siracusa, R., Fusco, R. and Cuzzocrea, S. (2019) ‘Astrocytes: Role and Functions in Brain Pathologies’, Frontiers in pharmacology, 10, p. 1114.

Skene, N. G. et al. (2018) ‘Genetic identification of brain cell types underlying schizophrenia’, Nature genetics, 50(6), pp. 825–833.

Smith, A. D., Xuan, Z. and Zhang, M. Q. (2008) ‘Using quality scores and longer reads improves accuracy of Solexa read mapping’, BMC bioinformatics, 9, p. 128.

Sokol, D. K. et al. (2011) ‘Autism, Alzheimer disease, and fragile X: APP, FMRP, and mGluR5 are molecular links’, Neurology, 76(15), pp. 1344–1352.

de Souza, L. C. et al. (2014) ‘Frontal lobe neurology and the creative mind’, Frontiers in psychology. frontiersin.org, 5, p. 761.

Spiegel, I. et al. (2014) ‘Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs’, Cell, 157(5), pp. 1216–1229.

Squire, L. R., Stark, C. E. L. and Clark, R. E. (2004) ‘The medial temporal lobe’, Annual review of neuroscience. annualreviews.org, 27, pp. 279–306.

Stafford, M. R. et al. (2013) ‘Early interventions to prevent psychosis: systematic review and meta-analysis’, BMJ, pp. f185–f185. doi: 10.1136/bmj.f185.

Staff, SRAS (2011) ‘Using the sra toolkit to convert. sra files into other formats’, National Center for Biotechnology Information (US).

Stahl, E. A. et al. (2019) ‘Genome-wide association study identifies 30 loci associated with bipolar disorder’, Nature genetics, 51(5), pp. 793–803.

Stanfield, A. C. et al. (2008) ‘Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies’, European psychiatry: the journal of the Association of European Psychiatrists, 23(4), pp. 289–299.

Stathis, P. et al. (no date) ‘Connections of the basal ganglia with the limbic system: implications for neuromodulation therapies of anxiety and affective disorders’, Operative Neuromodulation, pp. 575–586. doi: 10.1007/978-3-211-33081-4_67.

Steele, J., Richardson, J. and Olszewski, J. (2014) ‘Progressive Supranuclear Palsy: A Heterogeneous Degeneration Involving the Brain Stem, Basal Ganglia and Cerebellum With Vertical Gaze and Pseudobulbar Palsy, Nuchal Dystonia and Dementia’, Seminars in Neurology, pp. 129–150. doi: 10.1055/s-0034-1377058.

Strakowski, S. M. et al. (2012) ‘The functional neuroanatomy of bipolar disorder: a consensus model’, Bipolar disorders, 14(4), pp. 313–325.

Strohäker, T. et al. (2019) ‘Structural heterogeneity of α-synuclein fibrils amplified from patient brain extracts’, Nature communications, 10(1), p. 5535.

Stuss, D. T. (2011) ‘Functions of the frontal lobes: relation to executive functions’, Journal of the International Neuropsychological Society: JINS. unige.ch, 17(5), pp. 759–765.

Stuss, D. T. and Knight, R. T. (2013) Principles of Frontal Lobe Function. OUP USA.

Suárez-Calvet, M. et al. (2014) ‘Autosomal-dominant Alzheimer’s Disease Mutations at the Same Codon of Amyloid Precursor Protein Differentially Alter Aβ Production’, Journal of neurochemistry. J Neurochem, 128(2). doi: 10.1111/jnc.12466.

Surmeier, D. J., Plotkin, J. and Shen, W. (2009) ‘Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection’, Current opinion in neurobiology, 19(6), pp. 621–628.

Tammimies, K. et al. (2015) ‘Molecular Diagnostic Yield of Chromosomal Microarray Analysis and Whole-Exome Sequencing in Children With Autism Spectrum Disorder’, JAMA: the journal of the American Medical Association. jamanetwork.com, 314(9), pp. 895–903.

Tao, S. et al. (2018) ‘TPH-2 Gene Polymorphism in Major Depressive Disorder Patients With Early-Wakening Symptom’, Frontiers in neuroscience, 12, p. 827.

Tao, Y. and Zhang, S.-C. (2016) ‘Neural Subtype Specification from Human Pluripotent Stem Cells’, Cell Stem Cell, pp. 573–586. doi: 10.1016/j.stem.2016.10.015.

Tiihonen, J. et al. (2019) ‘Sex-specific transcriptional and proteomic signatures in schizophrenia’, Nature communications, 10(1), p. 3933.

Tokida H. et al. (2018) ‘[A case of various illusion, and hallucination caused by occipital lobe infarction]’, Rinsho shinkeigaku = Clinical neurology. ncbi.nlm.nih.gov, 58(9), pp. 556–559.

Tong, F. (2003) ‘Primary visual cortex and visual awareness’, Nature reviews. Neuroscience. nature.com, 4(3), pp. 219–229.

Tootell, R. B. H. and Nasr, S. (2017) ‘Columnar Segregation of Magnocellular and Parvocellular Streams in Human Extrastriate Cortex’, The Journal of neuroscience: the official journal of the Society for Neuroscience. Soc Neuroscience, 37(33), pp. 8014–8032.

Trapnell, C., Pachter, L. and Salzberg, S. L. (2009) ‘TopHat: Discovering Splice Junctions With RNA-Seq’, Bioinformatics . Bioinformatics, 25(9). doi: 10.1093/bioinformatics/btp120.

Trifonova, E. A., Khlebodarova, T. M. and Gruntenko, N. E. (2017) ‘Molecular mechanisms of autism as a form of synaptic dysfunction’, Russian Journal of Genetics: Applied Research, 7(8), pp. 869–877.

Trillingsgaard, A. and ØStergaard, J. R. (2004) ‘Autism in Angelman syndrome: an exploration of comorbidity’, Autism: the international journal of research and practice, 8(2), pp. 163–174.

Tsao, A. et al. (2018) ‘Integrating time from experience in the lateral entorhinal cortex’, Nature. nature.com, 561(7721), pp. 57–62.

Uddin, L. Q. et al. (2015) ‘Brain State Differentiation and Behavioral Inflexibility in Autism’, Cerebral cortex , 25(12), pp. 4740–4747.

Uddin, R. K. and Singh, S. M. (2013) ‘Hippocampal gene expression meta-analysis identifies aging and age-associated spatial learning impairment (ASLI) genes and pathways’, PloS one, 8(7), p. e69768.

Urbach, A., Bruehl, C. and Witte, O. W. (2006) ‘Microarray-based long-term detection of genes differentially expressed after cortical spreading depression’, The European journal of neuroscience, 24(3), pp. 841–856.

Valvo, G. et al. (2016) ‘Temporal lobe connects regression and macrocephaly to autism spectrum disorders’, European child & adolescent psychiatry. Springer, 25(4), pp. 421–429.

Van Den Bogaert, A. et al. (2006) ‘Association of brain-specific tryptophan hydroxylase, TPH2, with unipolar and bipolar disorder in a Northern Swedish, isolated population’, Archives of general psychiatry, 63(10), pp. 1103–1110.

Van Dijck, A. et al. (2016) ‘ADNP-Related Intellectual Disability and Autism Spectrum Disorder’, in Adam, M. P. et al. (eds) GeneReviews®. Seattle (WA): University of Washington, Seattle.

Vasile, F., Dossi, E. and Rouach, N. (2017) ‘Human astrocytes: structure and functions in the healthy brain’, Brain structure & function, 222(5), pp. 2017–2029.

Velmeshev, D. et al. (2019) ‘Single-cell genomics identifies cell type-specific molecular changes in autism’, Science, 364(6441), pp. 685–689.

Vera, J. C. et al. (2008) ‘Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing’, Molecular ecology, 17(7), pp. 1636–1647.

Vieta, E. et al. (2018) ‘Bipolar disorders’, Nature reviews. Disease primers, 4, p. 18008.

Vignoli, A. et al. (2015) ‘Autism spectrum disorder in tuberous sclerosis complex: searching for risk markers’, Orphanet journal of rare diseases, 10, p. 154.

Villa, A., Della Torre, S. and Maggi, A. (2019) ‘Sexual differentiation of microglia’, Frontiers in neuroendocrinology, 52, pp. 156–164.

Vilor-Tejedor, N. et al. (2018) ‘Strategies for integrated analysis in imaging genetics studies’, Neuroscience and biobehavioral reviews, 93, pp. 57–70.

Vogt, I. R. et al. (2006) ‘Transcriptional changes in multiple system atrophy and Parkinson’s disease putamen’, Experimental neurology, 199(2), pp. 465–478.

Wang, D.-S. et al. (2005) ‘Decreased Neprilysin Immunoreactivity in Alzheimer Disease, but Not in Pathological Aging’, Journal of Neuropathology & Experimental Neurology, pp. 378–385. doi: 10.1093/jnen/64.5.378.

Wang, M. et al. (2018) ‘The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer’s disease’, Scientific Data. doi: 10.1038/sdata.2018.185.

Wang, P. et al. (2018) ‘Enriched expression of genes associated with autism spectrum disorders in human inhibitory neurons’, Translational psychiatry, 8(1), p. 13.

Wang, Z. et al. (2018) ‘Identifying circRNA-associated-ceRNA networks in the hippocampus of Aβ1-42-induced Alzheimer’s disease-like rats using microarray analysis’, Aging, 10(4), pp. 775–788.

Wang, Z., Gerstein, M. and Snyder, M. (2009) ‘RNA-Seq: a revolutionary tool for transcriptomics’, Nature reviews. Genetics, 10(1), pp. 57–63.

Weinberger, D. R. and Levitt, P. (2011) ‘Neurodevelopmental Origins of Schizophrenia’, Schizophrenia, pp. 393–412. doi: 10.1002/9781444327298.ch19.

Werner, T. (2008) ‘Bioinformatics applications for pathway analysis of microarray data’, Current Opinion in Biotechnology, pp. 50–54. doi: 10.1016/j.copbio.2007.11.005.

Wessel, J. R. et al. (2016) ‘Surprise disrupts cognition via a fronto-basal ganglia suppressive mechanism’, Nature communications. nature.com, 7, p. 11195.

Wible, C. G. et al. (2001) ‘Prefrontal cortex, negative symptoms, and schizophrenia: an MRI study’, Psychiatry Research: Neuroimaging, pp. 65–78. doi: 10.1016/s0925-4927(01)00109-3.

Wickham, H. (2011) ‘ggplot2’, Wiley Interdisciplinary Reviews: Computational Statistics, 3(2), pp. 180–185.

Williams, D. R. et al. (2007) ‘Pathological Tau Burden and Distribution Distinguishes Progressive Supranuclear Palsy-Parkinsonism From Richardson’s Syndrome’, Brain: a journal of neurology. Brain, 130(Pt 6). doi: 10.1093/brain/awm104.

Williams, D. R. and Lees, A. J. (2009) ‘Progressive supranuclear palsy: clinicopathological concepts and diagnostic challenges’, Lancet neurology, 8(3), pp. 270–279.

Wingender, E. (1996) ‘TRANSFAC: a database on transcription factors and their DNA binding sites’, Nucleic Acids Research, pp. 238–241. doi: 10.1093/nar/24.1.238.

Wolk, D. A. et al. (2017) ‘Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer’s disease’, Neurobiology of aging, 49, pp. 204–213.

Wright, C. et al. (2017) ‘Altered expression of histamine signaling genes in autism spectrum disorder’, Translational psychiatry, 7(5), p. e1126.

Wu, J. Q. et al. (2012) 'Transcriptome sequencing revealed significant alteration of cortical promoter usage and splicing in schizophrenia', PloS one, 7(4), p. e36351.

Xiao, Y. et al. (2014) 'The DNA methylome and transcriptome of different brain regions in schizophrenia and bipolar disorder', PloS one, 9(4), p. e95875.

Xilouri, M., Brekk, O. R. and Stefanis, L. (2013) 'Alpha-synuclein and Protein Degradation Systems: a Reciprocal Relationship', Molecular Neurobiology, pp. 537–551. doi: 10.1007/s12035-012-8341-2.

Xu, J. et al. (2016) 'Neuregulin 1 improves cognitive deficits and neuropathology in an Alzheimer's disease model', Scientific reports, 6, p. 31692.

Xu, X. et al. (2014) 'Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders', The Journal of neuroscience: the official journal of the Society for Neuroscience, 34(4), pp. 1420–1431.

Yan, J. et al. (2017) 'IDENTIFICATION OF DISCRIMINATIVE IMAGING PROTEOMICS ASSOCIATIONS IN ALZHEIMER'S DISEASE VIA A NOVEL SPARSE CORRELATION MODEL', Biocomputing 2017. doi: 10.1142/9789813207813_0010.

Yao, P. et al. (2015) 'Coexpression networks identify brain region-specific enhancer RNAs in the human brain', Nature neuroscience, 18(8), pp. 1168–1174.

Yu, G. et al. (2012) 'clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters', OMICS: A Journal of Integrative Biology, pp. 284–287. doi: 10.1089/omi.2011.0118.

Zhang, B. et al. (2013) 'Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease', Cell, pp. 707–720. doi: 10.1016/j.cell.2013.03.030.

Zhang, F.-F. et al. (2018) 'Brain structure alterations in depression: Psychoradiological evidence', CNS neuroscience & therapeutics, 24(11), pp. 994–1003.

Zhang, X. et al. (2005) 'Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression', Neuron, 45(1), pp. 11–16.

Zhang, Y. et al. (2016) 'Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse', Neuron, 89(1), pp. 37–53.

Zhu, X. et al. (2014) 'One gene, many neuropsychiatric disorders: lessons from Mendelian diseases', Nature neuroscience, 17(6), pp. 773–781.

Annex

List of publications

1. **Iman Sadeghi**, Juan D. Gispert, Emilio Palumbo, Manuel Muñoz-Aguirre, Valentin Wucher, Valeria D'Argenio, Lucio Pastore, Roderic Guigo, Natàlia Vilor-Tejedor**.** "Brain transcriptomic profiling reveals common patterns across neurodegenerative and psychiatric disorders". *(Ready to submit). [First Author]*
2. Shirvani-Farsani, Zeinab, Zahra Bagheri-Hossein-abadi and **Iman Sadeghi.** "DNA-methylation as a biomarker of Alzheimer's and other neurodegenerative diseases". *Alzheimer's & Dementia (Under Review). [Corresponding Author]*
3. Ahmadi, Amirhossein, Ilario De Toma, Natàlia Vilor-Tejedor, Mohammad Reza Eftekhariyan Ghamsari, and **Iman Sadeghi**. "Transposable elements in brain health and disease." Ageing Research Reviews (2020): 101153. *[Corresponding Author]*
4. Hosseini, Ebrahim, Zahra Bagheri-Hosseinabadi, Ilario De Toma, Moslem Jafarisani, and **Iman Sadeghi**. "The importance of long non-coding RNAs in neuropsychiatric disorders"**.** *Molecular aspects of medicine* 70 (2019): 127-140. *[Corresponding Author]*
5. Abbasifard, Mitra, Zahra Kamiab, Zahra Bagheri-Hosseinabadi, and **Iman Sadeghi**. "The role and function of long non-coding RNAs in osteoarthritis." *Experimental and molecular pathology* 114 (2020): 104407. *[Corresponding Author]*
6. Zadehbagheri, Fatemeh, Ebrahim Hosseini, Zahra Bagheri-Hosseinabadi, Hossein Moradi Rekabdarkolaee, and **Iman Sadeghi**. "Profiling of miRNAs in serum of children with attention-deficit hyperactivity disorder shows significant alterations." *Journal of psychiatric research* 109 (2019): 185-192. *[Corresponding Author]*

Presentations at Meetings

Iman Sadeghi, Emilio Palumbo, Manuel Muñoz-Aguirre, Valentin Wucher, Juan Domingo Gispert, Roderic Guigó, Natalia Vilor-Tejedor. “Brain transcriptomic profiling reveals common patterns across neurological and neuropsychiatric disorders”. *Bioinformatics and Genomics Symposium, Barcelona, Spain. 17 December 2019. [Presenting Author]*

Iman Sadeghi, Emilio Palumbo, Manuel Muñoz-Aguirre, Valentin Wucher, Juan Domingo Gispert, Roderic Guigó, Natalia Vilor-Tejedor. “Brain transcriptomic profiling reveals common patterns across neurological and neuropsychiatric disorders*”. Internal lab meetings at Computational Biology of RNA Processing group - Center for Genomic Regulation, Barcelona, Spain. [Presenting Author]*

Iman Sadeghi, Emilio Palumbo, Manuel Muñoz-Aguirre, Valentin Wucher, Juan Domingo Gispert, Roderic Guigó, Natalia Vilor-Tejedor. “Brain transcriptomic profiling reveals common patterns across neurological and neuropsychiatric disorders”. *Internal lab meetings at the Precision Epidemiology group at the Departmetn of Clinical Genetics (ERASMUS University Medical Center). [Presenting Author]*

Github profile: isadeghi87 (Iman Sadeghi)