Computational and Structural Biotechnology Journal 20 (2022) 4549–4561

COMPUTATIONAL ANDSTRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Brain transcriptomic profiling reveals common alterations across neurodegenerative and psychiatric disorders

Iman Sadeghi [a,b,c], Juan D. Gispert [a,d,e,f], Emilio Palumbo [b,d,e], Manuel Muñoz-Aguirre [b,g], Valentin Wucher [b], Valeria D'Argenio [c,h], Gabriel Santpere [d,e,i], Arcadi Navarro [a,b,d,j,k], Roderic Guigo [b,d,*,2], Natàlia Vilor-Tejedor [a,b,d,l,*,1]

[a] BarcelonaBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain
[b] Centre for Genomic Regulation (CRG), The Barcelona Institute for Science and Technology, Dr. Aiguader 88, Barcelona E-08003, Catalonia, Spain
[c] CEINGE-Biotecnologie Avanzate, via G. Salvatore 486, 80145 Naples, Italy
[d] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[e] IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
[f] Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain
[g] Universitat Politècnica de Catalunya. Departament d'Estadística i Investigació Operativa, Barcelona, Spain
[h] Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, Naples, Italy
[i] Neurogenomics Group, Research Programme on Biomedical Informatics (GRIB), Barcelona, Catalonia, Spain
[j] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain
[k] Institute of Evolutionary Biology (CSIC-UPF), Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain
[l] Erasmus MC University Medical Center Rotterdam, Department of Clinical Genetics, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 13 April 2022
Received in revised form 16 August 2022
Accepted 16 August 2022
Available online 19 August 2022

Keywords:
Transcriptome profiling
RNA-Seq
Neurodegeneration
Psychiatric disorder
Network analysis
Brain cell types

ABSTRACT

Neurodegenerative and neuropsychiatric disorders (ND-NPs) are multifactorial, polygenic and complex behavioral phenotypes caused by brain abnormalities. Large-scale collaborative efforts have tried to identify the genetic architecture of these conditions. However, the specific and shared underlying molecular pathobiology of brain illnesses is not clear. Here, we examine transcriptome-wide characterization of eight conditions, using a total of 2,633 post-mortem brain samples from patients with Alzheimer's disease (AD), Parkinson's disease (PD), Progressive Supranuclear Palsy (PSP), Pathological Aging (PA), Autism Spectrum Disorder (ASD), Schizophrenia (Scz), Major Depressive Disorder (MDD), and Bipolar Disorder (BP)–in comparison with 2,078 brain samples from matched control subjects.

Similar transcriptome alterations were observed between NDs and NPs with the top correlations obtained between Scz-BP, ASD-PD, AD-PD, and Scz-ASD. Region-specific comparisons also revealed shared transcriptome alterations in frontal and temporal lobes across NPs and NDs. Co-expression network analysis identified coordinated dysregulations of cell-type-specific modules across NDs and NPs. This study provides a transcriptomic framework to understand the molecular alterations of NPs and NDs through their shared- and specific gene expression in the brain.

1. Introduction

Neurodegenerative and neuropsychiatric disorders (ND-NPs) are multifactorial, polygenic, and complex behavioral phenotypes caused by changes in multiple underlying mechanisms [1,2]. In some NDs, nerve cells become unable to respond to changes in their internal and external environments, eventually resulting in an impairment of brain function [3–5]. At present, the most prevalent NDs [6,7] are Alzheimer's disease (AD), Parkinson's disease (PD), progressive supranuclear palsy (PSP), as well as early preclinical manifestations of those such as pathological aging (PA).

* Corresponding authors at: BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation. C.Wellington 30, 08005 Barcelona, Spain, (Natàlia Vilor-Tejedor); Center for Genomic Regulation (CRG). C. Doctor Aiguader 88, Edif. PRBB 08003 Barcelona, Spain (Roderic Guigo).

E-mail addresses: roderic.guigo@crg.eu (R. Guigo), nvilor@barcelonabeta.org (N. Vilor-Tejedor).

[1] ORCID: 0000-0003-4935.
[2] ORCID: 0000-0002-5738-4477.

https://doi.org/10.1016/j.csbj.2022.08.037

Because of the presence of amyloid plaques, but not tangles, and the absence of dementia, PA is considered to be either a prodrome of AD or a condition, in which there is resistance to the development of neurofibrillary tangles and/or dementia [8]. Alteration of neuronal communications has been implicated in NDs, as well as in NPs such as schizophrenia (Scz), bipolar disorder (BP), autism spectrum disorder (ASD), and major depressive disorder (MDD), which are also among the major contributors to disability worldwide [9]. The etiology and mechanisms of NDs and NPs are elusive and a broad spectrum of causative genetic and environmental factors have been proposed [10].

Several investigations have shed light on the genetic heterogeneity within brain conditions and the degree of molecular similarities between closely related disorders [11–13]. Patterns of converging clinical and biological characteristics across NDs such as AD, PD, and PA have been lately discussed [14–16]. NPs have also shown symptomatic overlaps [17,18]. This demands uncovering condition-specific and overlapping pathological mechanisms across ND-NPs, which has been partly revealed by recent large-scale genome-wide association studies (GWAS) [19–21]. For instance, the genetic correlation between eight neuropsychiatric disorders revealed 3 different sub-groups with high levels of genetic overlap as well as multiple pleiotropic loci related to genes involved in neurodevelopment [22]. Moreover, some effort has been made to relate shared genetic causes with shared transcriptomic alterations in the post-mortem brains in subsets of these disorders, revealing similar relationships among diseases even if not necessarily implicating the same genes [21,23].

Such integrative transcriptomic studies attempt to fill the functional gap and establish the degree of coupling between primary genetic causes and secondary events captured by the transcriptome in adult postmortem samples [24]. However, a comprehensive characterization of gene expression changes in brain regions from individuals with major brain NDs and NPs compared with healthy subjects is missing. To address this shortcoming, we have uniformly analyzed a large collection of bulk RNA-Seq samples from post-mortem brain regions of subjects with NDs and NPs produced by different studies. The results of our *meta*-analysis revealed similarities in transcriptomic alterations between NDs and NPs which have also been observed in brain regions such as frontal and temporal lobes across the conditions. We additionally found coordinated downregulation of neuron-specific co-expression modules in both NDs and NPs, while oligodendrocyte and astrocyte modules showed mainly upregulation across conditions.

2. Results

2.1. Samples characteristics and clustering

We analyzed 4,711 RNA-Seq samples produced by 19 different labs [8,21,25–42] from patients with AD (n = 906 samples), PD (n = 29), PA (n = 58), PSP (n = 168), Scz (n = 535), ASD (n = 187), MDD (n = 240), BP (n = 510), and non redundant controls (n = 2,078) pooled across all studies, obtained from seven major brain regions (Fig. 1 **& S1**, for further details see **Table S1 & S2**). To produce a uniform gene expression quantification that could be compared across different datasets, the samples were processed using the Grape RNA-Seq pipeline [43] and underwent normalization and quality control (see **Methods** & **Fig. S2-S9**). The cell-type proportion was also calculated for RNA-Seq data from each sample (**Fig. S10**) using BRETIGEA R package v. 1.0.3 [44] and the data were normalized for the proportion of the cell types using the same package. *t*-Distributed Stochastic Neighbor Embedding (tSNE) [45] analysis in combination with principal component analysis (PCA) was used to produce a latent representation of the samples from different conditions based on their gene expression profiles (see **Methods**). In general, some brain regions (e.g. cerebellum and basal ganglia) are clustered separately from other regions (Fig. 2**a**). Within the cerebellum cluster, ASD samples are clearly separated from other conditions (Fig. 2**a and Fig. S11a-b**). PA and PSP clustered together very distinctively, but this could be confounded because samples for these conditions come from the same regions. ASD clustered into two different groups, likely reflecting the brain region from which samples originated. This could indicate that ASD may correspond to different molecular conditions. We observed a better separation of condition samples compared to controls (**Fig. S11a**). This was confirmed by the higher correlation between the top five PCs of gene expression computed in the condition samples and the regions compared to that obtained for controls **(Fig. S11c).**To some extent, this also happens with MDD. Finally, BP shows partly clustering with Scz based on the origin of the regions (Fig. 2**a**).

2.2. Condition-specific differential gene expression (DGE)

Condition-specific differential gene expression (DGE) analyses were performed using a linear mixed-effect model (see **Methods**). These analyses provided insights regarding transcriptional changes for the pathobiology of each condition (**Supplementary Data 1 & Fig. S12**). As the density plot shows some conditions have DEGs with logFCs close to zero, we kept those genes with an absolute logFC > 0.58 (**Fig. S12**). In total, we found 2891 unique genes differentially expressed in at least one condition. The correlation test showed no significant relationship between sample size and the number of DEGs (p = 0.058). In addition, the results obtained from random downsampling of one of the datasets (PRJNA394722, Pantazatos et al [37]) showed that sample size does not affect DEG results **(Supplementary Data 2).** Also, significant overlap between condition-specific DEGs and those obtained from individual datasets represented the reproducibility of the results (**Fig. S13**). Most DEGs were exclusive of one single condition, and we did not find a single gene shared across all the conditions, but there were genes shared across several conditions **(**Fig. 2**b** & **Fig. S14**). For instance, *MPZL2, SERPINA3,* and heat shock protein-encoding genes *HSPA6* and *HSPB1* which have been associated with neuro-inflammation, and stress response in neuronal damage [46,47], were deregulated in at least five conditions. Gene enrichment analysis performed using DEGs for each condition showed that the top enriched pathways are condition-specific (Fig. 2**c**); however functions related to stress response, synapse, and immune response were shared among some conditions (FDR-corrected p-value < 0.05; Fig. 2**c** & **Fig. S15**).

2.3. Cross-condition transcriptome overlap observed across NDs and NPs

To investigate the similarity between transcriptome alterations underlying the NDs and NPs, we compared the genes' fold change (FC) in expression in cases vs controls between conditions, by explicitly computing the pairwise correlation of logFCs from 15,819 shared genes (See **Methods**, **Fig. S16a** and **Supplementary Data 3**). This set of genes showed significant overlap with the list of common genes across psychiatric disorders from Gandal et al [21]. study (odd ratio = 4.5, FDR-corrected p-value < 0.001). Scz and BP represented the strongest correlation with the overlap of both downregulated and upregulated genes (Fig. 3**a** & **Fig. S16b**). Also, performing correlation analysis including variably expressed genes across conditions showed the highest correlation between ASD-BP and AD-PD (**Fig. S16c**). According to the transcriptional alterations, ASD clustered together with NDs, rather than within other NPs

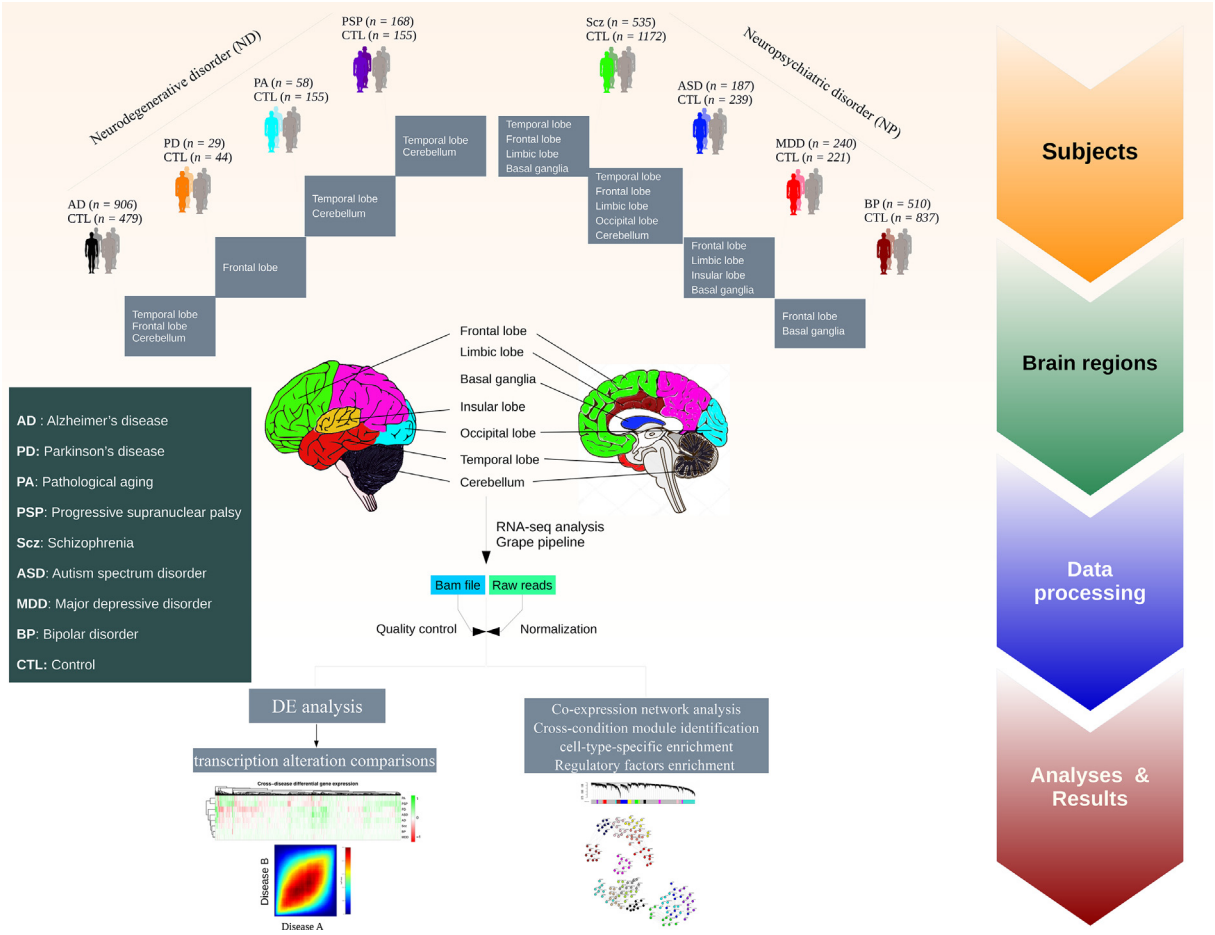

Fig. 1. Schematic of the study design and samples used for gene expression analysis via an RNA-Seq pipeline. Post-mortem brain RNA-Seq data were obtained from subjects with AD (n = 906 samples), PD (n = 29), PA (n = 58), PSP (n = 168), Scz (n = 535), ASD (n = 187), MDD (n = 240), BP (n = 510), and matched controls (n = 2078) (see Supplementary Table S1&S2 and Fig. S1).

(Fig. 3**b**). Within NPs, Scz and BP clustered closer, and within NDs, PA and PSP. MDD clustered separately from the rest of the conditions (Fig. 3**b**). These results were confirmed by computing correlations of the logFCs from the same set of genes across individual datasets to check for reproducibility (**Fig. S16d**).

2.4. Region-specific differential gene expression

Although some regions were present for only a limited number of conditions, we also considered analyzing differential expression for the samples from the same regions. Although after correcting for multiple tests, many regions did not show significant DEGs (**Fig. S17**), region-specific DGE revealed a set of overlapping genes that were frequently differentially expressed in multiple brain regions across conditions or vice versa (**Supplementary Data 4**). These genes included *GABRE* and *KCNE4* involved in neurotransmission [48,49], *SERPINA5* associated with neuropathies and the formation of amyloid-fibrils in condition [50–52], and HSPB1 associated with stress inflammation [53], and [54]. We built classifier models using the expression of DEGs from temporal and frontal regions (the most present regions across conditions) to explore the discrimination power of transcriptomic profiles between condition and control samples. High prediction accuracy was obtained for PD (78 %) and AD (79 %) in the frontal and temporal cortex, respectively (**Fig. S18**).

Analysis of the similarities in transcriptional alterations between different conditions in the individual brain regions separately generally recapitulated the findings in Fig. 3**a**. BP and Scz showed a high positive correlation in all the regions in which these conditions were assayed (Fig. 3**c** & **Fig. S19**). MDD showed little correlation with the other conditions across different regions.. Some conditions, however, showed similar or distinct transcriptome alterations depending on the region. Thus, transcriptomic changes underlying PSP and AD were highly correlated in the cerebellum but negatively associated in the temporal lobe (Fig. 3**c** & **Fig. S19**). In some cases, therefore, apparently similar phenotypic outcomes are the consequence of different molecular events in different brain regions.

2.5. Network analyses identified condition-specific and shared transcriptional signatures

To connect molecular alterations with the phenotypes of NDs and NPs, via their impact on the cellular composition of the brain, we constructed co-expression networks using combined normalized datasets for the same set of shared genes using weighted gene co-expression network analysis (rWGCNA) [55] and obtained seventeen co-expression modules (Fig. 4**a** & **Supplementary Data 5**). Module M0 comprises the genes that are not included in specific modules. We assigned cell types to each module based on the overlap between the genes in the module and the genes defining the cell type as in the PanglaoDB database [56]. Several modules were enriched for cells of a specific type: M2 for oligodendrocytes and Schwann cells, M5

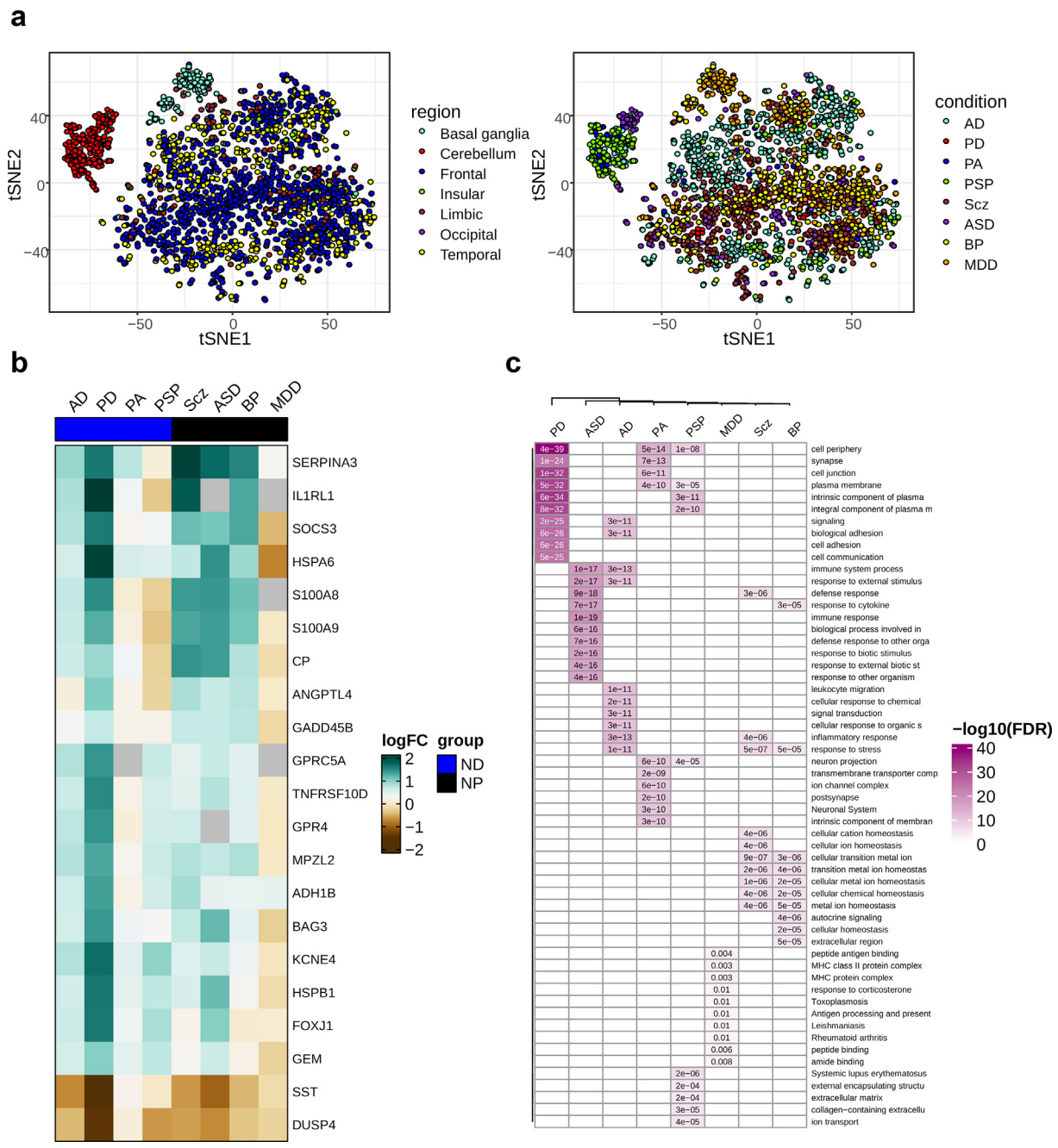

Fig. 2. Condition-specific transcriptome alterations. (a) tSNE visualization of the pooled samples as colored by region (left) and condition (right). (b) A heatmap of differentially expressed genes across neurodegenerative disorders (ND) and neuropsychiatric disorders (NP). The row labels represent the genes differentially expressed in at least 5 conditions. (FDR-corrected P < 0.05 & |log2FC| > 0.58). (c) Conditions-specific gene enrichment analysis. Top significantly enriched pathways are represented for significantly differentially expressed genes across conditions (FDR-corrected p-value < 0.05).

for astrocytes and Bergmann glia, M10 for microglia, M8 and M13 for neurons, and M14 for ependymal cells (Fig. 4**b**). These modules were the most distinct from the rest according to the correlation analysis and MDS (**Fig. S20**). The assignments were confirmed using an independent RNA-Seq dataset [57] composed of five main brain cell types including neurons, astrocytes, oligodendrocytes, microglia, and endothelial cells (**Fig. S21a**). We also identified the hub genes (the genes with the highest intramodular connectivity for each module (**Fig. S21b**). The functional categories enriched in cell-type-specific modules were broadly consistent with the cell-type assignments (Fig. 4**c** & **Supplementary Data 6**).

We next identified the differential expression of cell-type-specific modules in each condition from the expression of eigengenes (the first principal component of the expression matrix of the corresponding module, as a representative of an entire co-expression module) in each module using a linear mixed model (see **Methods,** Fig. 5**a** & **Supplementary Data 7**). In addition, the differential expression of top hub genes within these modules was measured for each condition (Fig. 5**b**). Then, we could relate NDs and NPs to cellular alterations in the brain.

Thus, neuron-specific modules (M8 and M13) were broadly downregulated across AD, PD, ASD, Scz, and BP, but upregulated in PA and PSP. The oligodendrocyte module M2 was upregulated

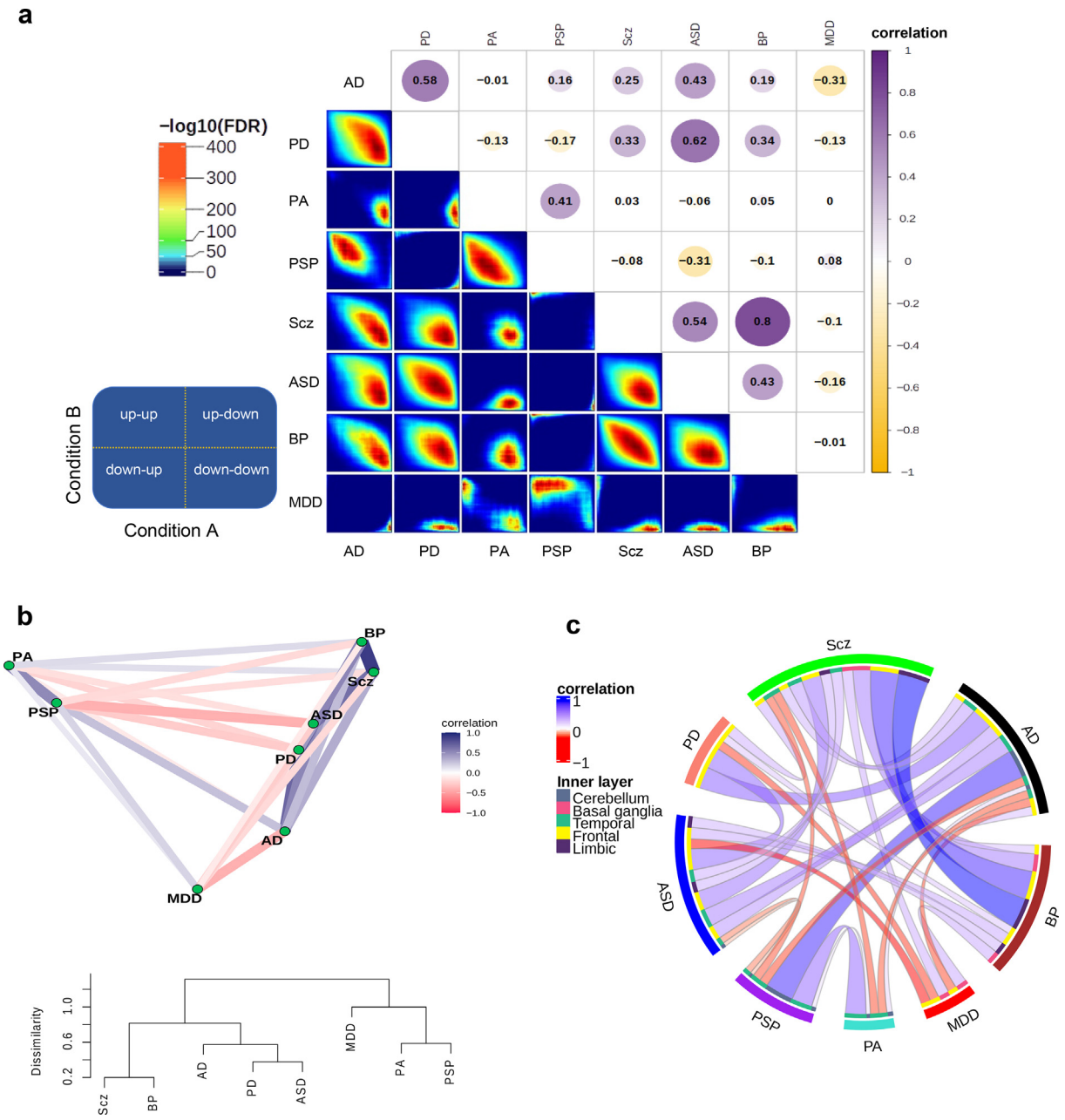

Fig. 3. Similarity of transcriptional alterations across conditions. (a) Correlation plot (top) shows transcriptome alterations overlap obtained by computing Spearman's correlations using logFC values of the shared genes between the conditions. Rank-rank hypergeometric overlap (RRHO; bottom) depicts the direction (upregulation and downregulation) of the logFC overlaps. The guide panel represents the cross-condition overlapping relationship. Signals in the upper left quadrant display an overlap for shared upregulated genes, while those in the bottom right quadrant depict shared downregulated genes. The color bar displays the degree of significance of the overlap (Fisher's exact test with FDR < 0.05). (b) Correlation network (top) and a tree dendrogram (bottom) obtained from pairwise correlations corresponding to a. show the relationship of the conditions based on transcriptome alterations. (c) A circos plot demonstrating correlations of transcriptional alterations across conditions. Only significant correlations after FDR correction (FDR < 0.05) with a cut-off of absolute correlation > 0.1 are displayed here (see Supplementary Fig. S19). The outer layer represents conditions, while the inner layer displays brain regions defined by colors.

in all conditions except PA and PSP (Fig. 5**a,b**), and consistently enriched with genes involved in oligodendrocyte and glial cell differentiation (Fig. 4**c**). The microglia-associated module (M10) which was upregulated in neurodegenerative diseases AD and PD, PA, and a psychiatric disorder ASD represented enrichment for genes involved in the immune response. These results are consistent with the reported microglial activation in AD [58], PD [59], and ASD [60] and also with the crucial role of microglia in CNS development and immunity [61,62]. The astrocyte-specific module (M5), broadly upregulated in AD, PD, PA, ASD, Scz, and BP (Fig. 5**a**), was enriched for metabolic genes. An increase in astrocytic reactivity has previously been reported in response to oxidative stress induced by amyloid-beta accumulation [63]. Additional analysis demonstrated a significant overlap of module M5 with an astrocyte-specific module from the Gandal et al. [21] study which represented coordinated upregulation in ASD, BP, and Scz (**Fig. S21c**). These results highlight the significant role of astrocytes in synaptic signalling, neuroprotection, and brain development [64–66].

Furthermore, we investigated the role of enhancers in the regulation of the network modules. We used an independently derived dataset of brain enhancer RNAs or eRNA (a class of relatively long

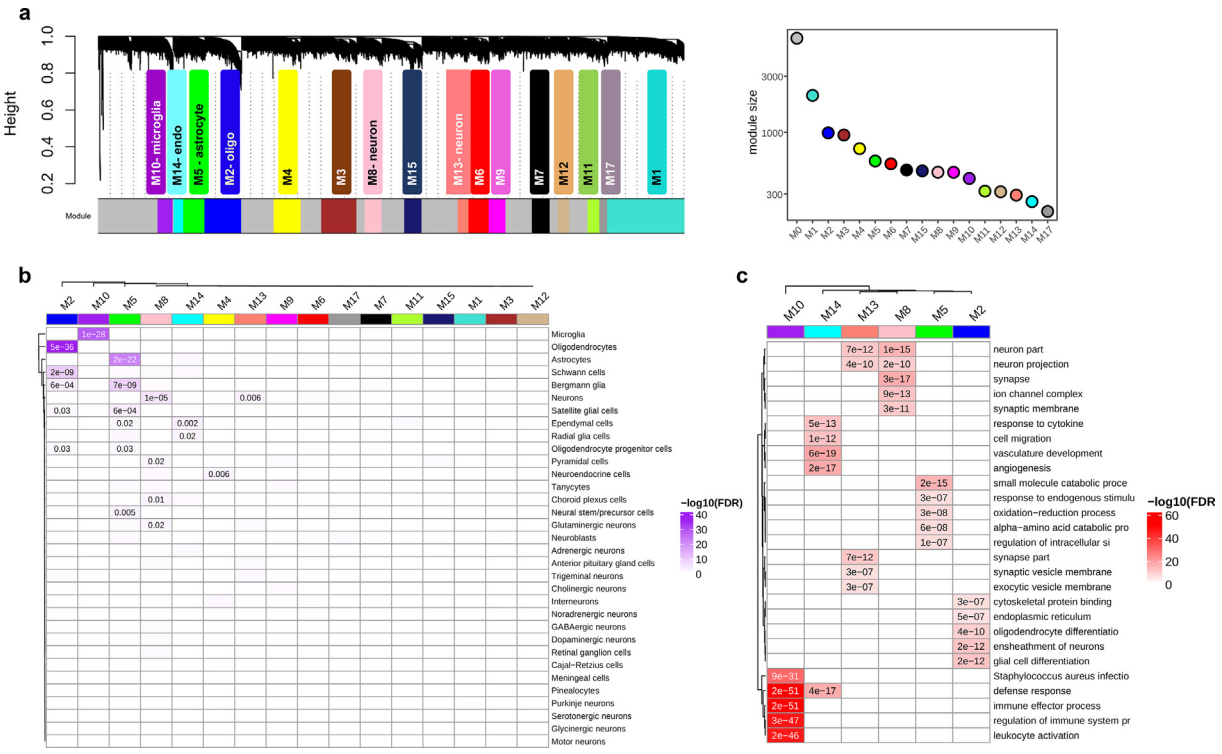

Fig. 4. Cross-condition co-expression modules identified by network analysis. (a) A dendrogram plot displaying co-expression modules obtained from the topological overlap of 15,819 shared genes between conditions. Each color represents an individual module and the grey color (M0) contains genes that are not included in a specific module. The corresponding plot on the right side shows the number of genes within each module. (b) Enrichment of co-expression modules for brain cell types, measured by comparing genes within each module to the brain single-cell dataset from PanglaoDB [56] (see also Supplementary Fig. S21a). (c) Heatmap plot of gene ontology enrichment for cell-type-specific modules using top five significant pathways for each module. The color key shows -$\log_{10}$(FDR).

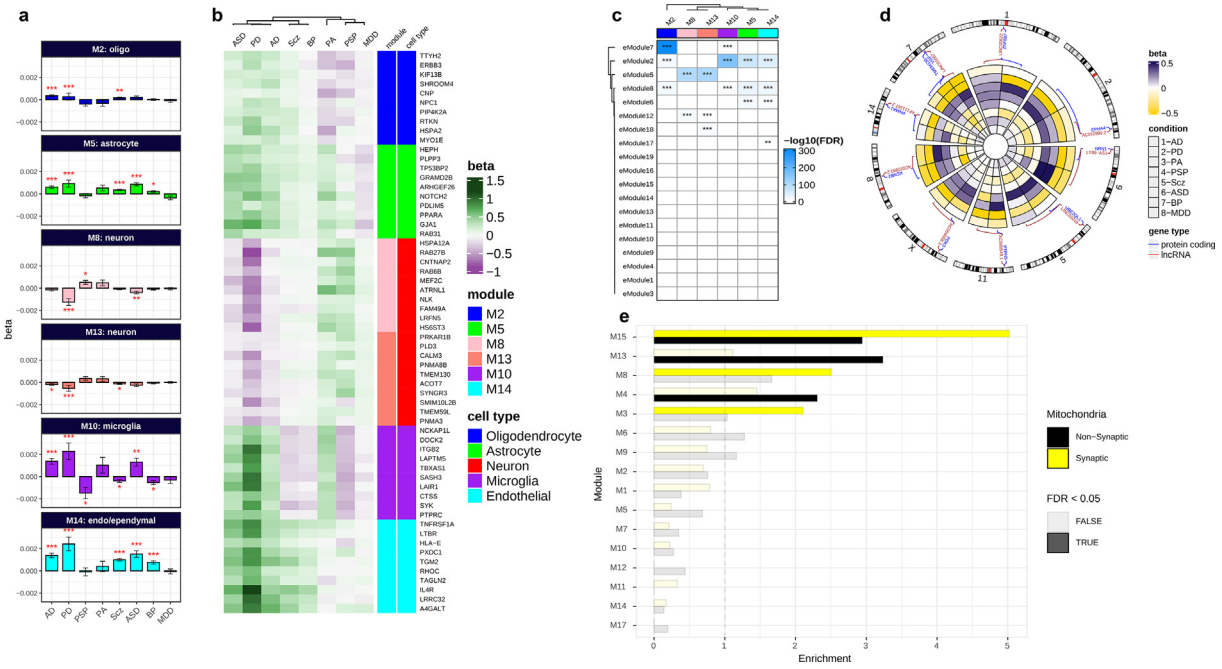

Fig. 5. Co-expression gene module characterizations. (a) Differential expression of cell-type-specific modules across conditions. β values on the y-axis computed by linear mixed effect model show the relationship of modules eigengenes with conditions. (b) Differential expression of top hub genes within cell-type-specific modules across conditions. Brain cell-type-specific modules are annotated with colors. (c) Enrichment of brain enhancer RNAs for cell-type-specific modules. The overlap between co-expression modules and eRNA modules from an independent dataset [67] was computed by Fisher's exact test (FDR < 0.05). Color key shows the -log10 (FDR-corrected p-values; *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001; see also Supplementary Fig. S22). (d) A circular heatmap showing expression of protein-coding and their flanking lncRNAs in neuron modules M8 and M13 across conditions (see also Supplementary Fig. S23). (e) The enrichment of co-expression modules for mitochondrial transcriptomes. An independent study that previously reported synaptic and nonsynaptic mitochondria co-expression modules was obtained and compared to co-expression modules in this study using Fisher's exact tests (FDR < 0.05). Yellow and black colors represent enrichment of synaptic and nonsynaptic mitochondrial transcriptomes for co-expression modules. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

non-coding RNAs) modules [67] (sets of eRNAs which are coordinately expressed in the brain). Results demonstrated the enrichment of multiple eRNA modules for the cell-type-specific modules (Fig. 5**c** & **Supplementary Fig. S22**), demonstrating the role of these enhancer modules in regulating network modules across brain conditions. Neuron-specific modules M8 and M13 showed enrichment for eRNA eModule5, which shows expression specificity for cerebral cortex [67]. In addition, the relationship between protein-coding genes and their adjacent lncRNAs in neuron modules shows their coordinated regulation across the conditions (Fig. 5**d** & **Fig. S23**). Oligodendrocyte and microglia modules (M2 and M10) were enriched for eModule7 which is specific for the thalamus- a region affected in a variety of systemic or metabolic diseases, degenerative diseases, and psychiatric conditions [68]. The astrocyte-specific module M5 is associated with eModule6, which is specific for medulla oblongata [67] (a part of the brainstem), a region associated with neurodegeneration and movement disorders [69–71]. These results demonstrate the regulation of co-expression modules across the conditions in the brain.

Eventually, as mitochondrial genes have been formerly linked to neuronal phenotypic diversity and brain conditions [72–75], we performed enrichment analysis of synaptic and nonsynaptic mitochondrial genes for each module using an independent dataset [75]. Multiple neuron-related modules (M15, M13, M8, M4, and M3) were enriched for mitochondrial genes [75], of which M15, M8 and M3 showed significant enrichment for synaptic mitochondria (Fig. 5**e** & **Supplementary Data 8**). These modules which have previously shown enrichment for neurons, contain genes mainly downregulated across the conditions (Fig. 5**a**). These results highlight the relationship between alteration of mitochondrial transcriptome and synaptic dysfunction in the brain across conditions [76].

3. Discussion

Leveraging the transcriptome profile of post-mortem tissues from several brain regions, for the first time to our knowledge, we highlighted the substantial overlapping molecular patterns across eight brain conditions including NDs and NPs. Dysregulation of overlapping genes such as *MPZL2*, *SERPINA, HSPA6* and *GABRE* across brain regions suggests shared perturbation of several mechanisms such as activation of microglia [77], inflammatory mechanisms, synapse development, and synaptic plasticity [54] across conditions. Microglia and astrocytes are vital in regulating neuronal activity and brain functioning during development and in the adult brain [78]. These results are consistent with previous well-established findings that molecular mechanisms in microglia and astrocytes are altered in ND-NPs [66,79].

In line with this, co-expression network results revealed mainly downregulation of neuron-specific modules across multiple conditions, reflecting neural dysfunction in both NDs [80] and NPs [81]. The microglial-related module showed mainly upregulation in NDs (i.e. AD and PD) reflecting activation of microglia during neurodegeneration and brain dysfunction [61,82]. Astrocyte- and oligodendrocyte-specific modules demonstrated broad upregulation across conditions including both NDs and NPs, representing activation of these cell types in neurogenesis, signaling, and cell development [21,64,83]. Also, coordinated downregulation of synaptic mitochondria-related modules across conditions suggests the importance of mitochondria for synaptic connections, neuronal survival, and function [75] in both NDs [84] such as PD [85] and AD [86], as well as in NPs [87].

Moreover, our results revealed shared transcriptional changes between neurodegenerative and psychiatric disorders. The findings from this study report new results about similar molecular changes between NDs and NPs, as well as supporting previous reports on shared transcriptional changes within each etiology (e.g., between NDs or NPs). Specifically, we observed similar transcriptional changes between PD-ASD, AD-ASD, PD-BP, PD-Scz, and AD-Scz. Moreover, within NDs, we found transcriptome overlap between AD-PD and PA-PSP, while within NPs, we found top transcriptome similarities between Scz-BP and Scz-ASD. The high correlations observed for the Scz-BP and Scz-ASD pairs support previous reports on the molecular similarity of these disorders [21,22]. Correlation of transcriptomic alterations across brain regions demonstrated the limbic lobe captures the majority of transcriptomic similarity between Scz and BP, supporting the role of this region in mood and psychotic disorders [88]. Although the only region studied for PD was the frontal lobe, similar transcriptional changes were observed between PD-ASD, particularly for the genes such as *CHI3L1, HSPA1,* and *HSPB1* involved in neuroinflammation pathway that has been recently linked to autism [89,90]. High frequency of parkinsonism has previously been reported in autistic cases [91], in which inflammatory mechanisms are seemingly involved in the pathobiology of the disease [92,93]. The correlation of transcriptional alterations between AD and ASD (p-value < 0.001) supports the evidence of neuroinflammation in autism [94,95–97]. This transcriptomic similarity between AD and ASD was mainly observed for temporal > frontal lobes. The overlapping expressional changes of genes such as *CHI3L1* between Scz and PD suggest perturbation of dopaminergic-glutamatergic balance in the brain, as described previously [98,99]. The Scz-AD similarity, which was mainly captured by the temporal lobe, is also consistent with the evidence of shared mechanisms between neurodegeneration and Scz [100,101]. Primary damaged regions in PSP and PA are reported to be brain stem (specifically substantia nigra) [102] and hippocampus [103], respectively. However, the significant transcriptome correlation between PSP and PA, mainly observed in the temporal lobe, suggests similar transcriptional changes in this brain area for the conditions. Although PSP is sometimes misdiagnosed as PD due to similar clinical symptoms [104], we did not observe transcriptome similarity here. This could be due to the lack of primary damaged regions for either condition and/or different pathobiology of the conditions [105].

In addition, our findings provide new insights into the shared pathobiology in ASD and Scz and their transcriptome similarity with NDs [106,107]. Also, the results did not show a significant correlation between AD and PA, suggesting their divergent molecular pathobiology [108]. Transcriptional relationships were not observed between MDD and other NPs such as Scz and ASD as expected, which could be due to its heterogeneous nature [21,22,109]. Of note, shared transcriptional changes observed here do not necessarily indicate similar pathobiology of the conditions, particularly in conditions such as PD, PSP, and PA which lack primary damaged regions and demand further investigation.

Comparisons of transcriptional changes across brain regions demonstrated similar molecular changes in the temporal and frontal lobe across NDs and NPs, implicating their possible impairment in the pathogenesis of a variety of brain diseases [110–116]. Similar molecular patterns of the cerebellum were observed across AD, PSP, and PA, supporting its emerging role in the pathobiology of NDs [117,118]. Despite the lack of samples for some of the conditions mentioned before, the 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 molecular changes of multiple brain regions (rather than one primary region) in one condition and, on the other hand, the involvement of one region in multiple conditions.

Of note, in the analysis of the overlapping transcriptomic alterations, we did not expect to capture modifications directly linked

to the underlying etiological mechanisms of the different conditions studied here. However, this was addressed in condition-specific gene expression analyses and will be more investigated in future analyses with the same database, along with the exploration of conditions sharing common mechanisms (i.e. cerebral proteinopathies) or between the different stages of the same disease (i.e. preclinical and clinical stages of AD) [119,120]. 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 [121]. Additional research could also benefit from further identification of sex-specific gene networks and transcription profiles to unravel the molecular mechanisms of brain diseases [122–124].

In addition, since tissue samples from all brain regions were not available for all 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 (i.e. basal ganglia in PD). Another limitation could be the potential impact of sample size on the DEG results. This could be avoided by using large-scale cohorts for all the conditions in future investigations. Also, even though batch-effect corrections have thoroughly been applied here, the inevitable effect of merging multiple datasets on the results could still be a limitation. Furthermore, the heterogeneity between brain regions could be a potential contributing factor for condition-specific differential expression analysis due to their cellular composition diversity. We tried to tackle this issue by performing analyses for each region. However, despite the limitations of the current analyses, molecular signatures here described across ND-NPs can provide target leads for the development of therapeutic interventions targeted to common pathological mechanisms which may overcome indications solely based on clinical manifestations, thus paving the way for the rational design of personalized and mechanistically-based therapies [125].

4. Methods

4.1. Samples and raw data

RNA-Seq raw data were obtained from 4,711 post-mortem brain samples from subjects with AD, PD, PSP, PA, Scz, MDD, ASD, BP, and controls through previously published studies [8,21,25–42] 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 S1**). The samples from individual datasets were processed separately (cases and controls from each study) and analyzed according to our RNA-Seq pipeline as described below.

4.2. Data processing

The data obtained from individual datasets were processed separately. We used RNA-Seq *fastq* files as the initial source of data processing. The samples that were retrieved as *SRA* and *BAM* files were converted to *fastq* file formats using SRA-toolkit [126] and SAM-tools [127], respectively. For further sample processing, the Grape pipeline [128] was used for RNA-Seq analysis, with Nextflow [129] as the execution backend, the STAR aligner v.2.6.0a tool [130] for mapping reads to the human genome build hg19 with GENCODE v.28 annotations, and the RSEM tool [131] for isoform quantification (using default options). Next, post-alignment quality control (QC) was performed using STAR aligner statistics, Qualimap v.2.2.1 tools [132], and Picard tools v1.8 (https://broadinstitute.github.io/picard/) to check for the total number of reads, the total number of mapped reads, GC percentage, exonic rate, intronic rate, intergenic rate, duplication rate, and insertion/deletion rate. Sequencing statistics were used to check for quality control and sample outliers for each study independently. The data were then normalized for library size using the *voom-limma* R package [133]. To filter out lowly-expressed genes, only genes with at least $\log_2$(-CPM) of 0.5 in 70 % of the samples were kept for further analyses. The *sva* R package v.3.32.1 [134] was used to correct for any batch effect of sequencing library preparations. For each study batch effect correction was performed for known batch effects such as library preparation, load date, and biobank as assessed by surrogate variables and PCA (data not shown). Unknown batch effects were also assessed by checking surrogate variables using the sva R package. For condition- and region-specific analyses, the data were corrected for study as a batch effect. To remove sample outliers, standardized network connectivity Z-scores were measured and a cutoff of Z < -2 was set as the threshold [135,136]. Normalized datasets were kept for downstream analyses.

4.3. Cell-type proportion analysis and cell-type normalization

To estimate the cell-type proportion of bulk tissue RNA-seq data used here, a cell-type deconvolution was performed on each sample using the brain cell–type marker signatures provided by the BRETIGEA R package v. 1.0.3 [137]. For each cell type, 500 genes were used from the human brain cell marker gene set (neurons, endothelial, oligodendrocytes, microglia, astrocytes, and OPCs) to generate all surrogate cell-type proportion (SPV) estimates. Normalization of the RNA-seq for brain cell type was also performed by BRETIGEA, using the default parameters and the calculated SPV values from the previous step.

4.4. Samples clustering and tSNE analysis

t-SNE [45] was used to produce a latent representation of the samples based on their gene expression profiles across datasets using all normalized datasets which were pooled together in an expression matrix using 15,819 genes common across the conditions. Before this, principal component analysis (PCA) [138] was first performed using the *prcomp* function in R to reduce the number of dimensions and obtain the top 50 principal components (PCs) as input to the tSNE analysis. The tSNE analysis was then performed by using *Rtsne* R package v. 0.15 using parameters dims = 2, initial_dims = 50, perplexity = 50, theta = 0, check_duplicates = TRUE, and PCA = FALSE. The first two *t*-SNE coordinates were used for visualization. To see clustering of the samples including controls, PCA and tSNE analyses were performed using the same parameters mentioned above (see **Fig. S10a**). In addition, tSNE analysis was performed for each region on the samples from different conditions (see **Fig. S10b**). We compared the correlation between the top five PCs of gene expression computed in control samples and the regions and those computed in the disease samples and the regions.

4.5. Differential gene expression (DGE) analysis and transcriptome comparisons

Differential expression analysis for each dataset was performed using *limma* with empirical Bayes moderated *t*-statistics, with the following model (expression ~ diagnosis + age + sex + RIN (RNA integrity) + PMI (postmortem interval)). Those variables with significant differences between cases and controls were included in the model. For the condition-specific DGE analyses, normalized expression data from relevant studies were combined and sample

outliers were removed as described previously. DGE was calculated using linear mixed-effects models by the *nlme* R package v.3.1–140 [139], with fixed effects of diagnosis, age, sex, brain region, and study. A random effect for subjects was used to fix for overlapping subjects between the studies (expression ~ diagnosis + age + sex + brain region + study + 1 | subject). The calculated log fold-change (logFC) values were used for downstream analyses. Significantly differentially expressed genes (DEGs) were filtered by using an FDR-corrected p-value of < 0.05. We then filtered those genes that had an absolute $\log_2$FC > 0.58.

In order to assess the similarity of transcriptional changes, we obtained an intersection of 15,819 genes that were present in all conditions after normalization. The logFCs of these genes were used to perform Spearman correlation tests. We also performed the same test with the inclusion of variably expressed genes across conditions (in total 26,366 genes).

4.6. Reproducibility of differential expression results

To check the reproducibility of the DGE results, the list of DEGs obtained from each individual dataset was compared to those obtained from condition-specific analysis by calculating the odd ratio using *GeneOverlap* R package v. 1.20.0 [140]. Fisher's exact test was used to calculate the p-values of each comparison. We also reperformed the DEG analysis for PRJNA394722, Pantazatos et al [37] by randomly downsampling. In addition, to avoid the effect of each study on the results from combined datasets, the DGE results from individual studies were also compared with each other for each condition (**Fig. S16d**).

4.7. Building classifier models

To identify the prediction power of transcriptional alterations from the frontal and temporal lobe between cases and control samples, normalized expressions of DEGs for each region were used to build classifier models using random forests [141] The dataset containing cases and controls from each region was split into train and test partitions. The training was performed using the *trainControl* R function from caret R package v. 6.0–90 with a cross-validation method and 10-fold cross-validation to avoid overfitting, followed by using the *train* R function with a random forest method. Then, the *confusionMatrix* R function was used to obtain the accuracy, sensitivity, and specificity of the final models for comparing the results.

4.8. Comparisons of transcriptional alterations

To analyze cross-condition transcriptome profile comparisons, we only kept the 15,819 genes that were common across all diseases. Pairwise gene expression comparisons were performed using Spearman's correlation over logFC values of the genes. In addition, a correlation network and tree dendrogram was built using the correlations statistics to obtain a relative relationship of the conditions based on transcriptional alterations. To check the reproducibility of results, logFC values of the genes common across condition-specific analyses and individual datasets were compared using Spearman's correlation test. Moreover, brain region-specific comparisons across conditions were performed using logFC values of the genes shared between the present conditions for each region.

4.9. Rank-Rank hypergeometric overlap (RRHO) analysis

In order to highlight the degree of overlap in gene signatures across conditions, as well as compare condition pairs for shared brain regions, we performed an unbiased rank-rank hypergeometric overlap (RRHO) analysis using the *RRHO* R package v.1.24.0 [142]. 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.

4.10. Gene co-expression network analysis

We performed robust Weighted Gene Co-Expression Network Analysis (rWGCNA) using the *WGCNA* R package v.1.68 [55] to identify co-expressed gene modules using expression data that were first normalized for different covariates. The expression data from condition-specific DGE analyses were combined using the 15,819 genes common between all datasets. Batch effect correction for the studies was performed using the *ComBat* R function from the *sva* R package v. 3.42.0.

We first computed the soft threshold power using *pickSoftThreshold* R functions with a "signed" network type, "bicor" correlation function and a block size of "20,000". Co-expression networks were built using the *blockwisemodules* R function from the WGCNA R package. 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. A power function with a soft-threshold of 20 was applied to the merged expression dataset to obtain an approximately scale-free weighted co-expression network (scale-free R^2 > 0.9). 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 50, a deepsplit of 4, a merge threshold of 0.1, and negative pamStage. Each module was assigned a unique (and arbitrary) color identifier. Genes with the highest intramodular connectivity (those with more connections at the core of the network) were considered hub genes. Significance values were FDR-corrected to account for multiple comparisons. Top hub genes with the most connections were prioritized based on their module membership (kME), defined as a correlation to the module eigengene. Module-condition relationships were measured using a linear mixed-effects model, with a random effect for subjects to fix for overlapping subjects (expression ~ diagnosis + age + sex + 1 | subject).

4.11. Cell-type-specific enrichment 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 [56]. 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 [140]. 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

RNA-seq dataset of purified cell populations from healthy human brain samples [57]. 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 [143]. 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).

4.12. Gene ontology enrichment analysis

Gene Ontology (GO) pathway enrichment for each condition and gene module was performed using the *gprofiler2* R package v. 0.2.1. 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.

4.13. Brain enhancer RNAs co-expression analysis

To understand the relationship between regulatory factors in the brain and co-expression modules, an expression dataset for brain enhancer-RNAs (eRNA)- a non-coding RNA transcribed from active enhancers [67]- was obtained from an independent study of human brain region-specific eRNAs co-expression analysis [67]. To explore the co-expression of gene modules from our dataset and each brain eRNA module, the overlap of genes within each module and genes from eRNA modules was tested by Fisher's exact test. A p-value of < 0.05 followed by FDR correction was used to filter the significant enrichments.

4.14. Protein coding-lncRNA co-expression analysis

Within each co-expression module, genomic coordinates for the genes were obtained using the *BiomaRt* R package. Next, the genes were filtered to protein-coding genes and lncRNA pairs with a distance of < 10 Mb, as the *cis*-regulatory cutoff distance. The expression fold change (beta) values of the genes across the conditions were illustrated in a genomic circos plot using the *circlize* [144] R package.

4.15. Mitochondrial transcriptome enrichment analysis

To see whether co-expression modules are enriched for mitochondrial transcriptome, an independent dataset containing synaptic and nonsynaptic mitochondria modules was obtained [75,145]. Enrichment analysis for each module was performed by Fisher's exact test followed by FDR correction for p-values.

4.16. 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, caret, limma, pheatmap, ComplexHeatmap, circlize* at CRAN (https://cran.r-project.org/) and/or Bioconductor (https://bioconductor.org/).

5. Data availability

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 in **Supplementary Table S1**.

Funding

At the time of writing this manuscript, I.S. is supported by the European School of Molecular Medicine (SEMM) at CEINGE Biotecnologie Avanzate s.c.a.r.l, Naples, Italy. J.D.G. is supported by the Spanish Ministry of Science and Innovation (RYC-2013-13054). M.M.A. is supported by the FPU15/03635 grant from Ministerio de Educación, Cultura y Deporte. N.V.-T. is funded by a postdoctoral grant, Juan de la Cierva Programme (FJC2018-038085-I), Ministry of Science and Innovation–Spanish State Research Agency. N. V.-T. has received additional support from the Health Department of the Catalan Government (Health Research and Innovation Strategic Plan (PERIS) 2016-2020 grant# SLT002/16/00201) and "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300004. All CRG authors acknowledge the support of the Spanish Ministry of Science, Innovation, and Universities to the EMBL partnership, the Centro de Excelencia Severo Ochoa and the CERCA Programme / Generalitat de Catalunya.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

RNA-seq data (retrieved from Sage Synapse with accession number syn2759792, with access governed by NIMH Repository and Genomics Resource) were generated as part of the PsychENCODE Consortium supported by U01MH103339, U01MH103365, U01MH103392, U01MH103340, U01MH103346, R01MH105472, R01MH094714, R01MH105898, R21MH102791, R21MH105881, R21MH103877, and P50MH106934 awarded to Schahram Akbarian (Icahn School of Medicine at Mount Sinai), Gregory Crawford (Duke), Stella Dracheva (Icahn School of Medicine at Mount Sinai), Peggy Farnham (USC), Mark Gerstein (Yale), Daniel Geschwind (UCLA), Thomas M. Hyde (LIBD), Andrew Jaffe (LIBD), James A. Knowles (USC), Chunyu Liu (UIC), Dalila Pinto (Icahn School of Medicine at Mount Sinai), Nenad Sestan (Yale), Pamela Sklar (Icahn School of Medicine at Mount Sinai), Matthew State (UCSF), Patrick Sullivan (UNC), Flora Vaccarino (Yale), Sherman Weissman (Yale), Kevin White (UChicago) and Peter Zandi (JHU).

RNA-seq data (retrieved from Sage Synapse with accession number syn4921369 governed by NIMH Repository and Genomics Resource) were generated as part of the CommonMind Consortium supported by funding from Takeda Pharmaceuticals Company Limited, F. Hoffmann-La Roche Ltd and NIH grants R01MH085542, R01MH093725, P50MH066392, P50MH080405, R01MH097276, RO1-MH-075916, P50M096891, P50MH084053S1, R37MH057881, AG02219, AG05138, MH06692, R01MH110921, R01MH109677, R01MH109897, U01MH103392, and contract HHSN271201300031C through IRP NIMH. Brain tissue for the study was obtained from the following brain bank collections: the Mount Sinai NIH Brain and Tissue Repository, the University of Pennsylvania Alzheimer's Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories, and the NIMH Human Brain Collection Core. CMC Leadership: Panos Roussos, Joseph Buxbaum, Andrew Chess, Schahram Akbarian, Vahram Haroutunian (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Raquel Gur, Chang-Gyu Hahn (University of Pennsylvania), Enrico Domenici (University of Trento), Mette A. Peters, Solveig Sieberts (Sage Bionetworks), Thomas Lehner, Stefano Marenco, Barbara K. Lipska (NIMH).

Author contributions

I.S., V.D., R.G., and N.V.-T. conceived the study. I.S. performed the analyses. E.P., M.M.-A., G.S., A.N., R.G., and N.V.-T. provided statistical and analysis advice. J.D.G., V.W., M.M.-A., G.S., A.N., R.G., and N.V.-T. contributed to the interpretation of the results. I.S., J. D.G., R.G. and N.V.-T. wrote the manuscript. R.G. and N.V.-T. supervised the project. All authors read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.08.037.

References

[1] Gratten J, Wray NR, Keller MC, Visscher PM. Large-scale genomics unveils the genetic architecture of psychiatric disorders. Nat Neurosci 2014;17:782–90.
[2] Hosseini E, Bagheri-Hosseinabadi Z, De Toma I, Jafarisani M, Sadeghi I. The importance of long non-coding RNAs in neuropsychiatric disorders. Mol Aspects Med 2019;70:127–40.
[3] Gerfen CR. Indirect-pathway neurons lose their spines in Parkinson disease. Nat Neurosci 2006;9:157–8.
[4] Surmeier DJ, Obeso JA, Halliday GM. Selective neuronal vulnerability in Parkinson disease. Nat Rev Neurosci 2017;18:101–13.
[5] Menzies FM, Fleming A, Rubinsztein DC. Compromised autophagy and neurodegenerative diseases. Nat Rev Neurosci 2015;16:345–57.
[6] Hodges JR. Alzheimer's disease and other dementias. Oxford Medicine Online 2010. https://doi.org/10.1093/med/9780199204854.003.2442.
[7] Möller T. Huntington Disease, Parkinson Disease, and Other Neurodegenerative Diseases. Oxford Medicine Online 2013. https://doi.org/10.1093/med/9780199794591.003.0065.
[8] Allen M et al. Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. Sci Data 2016;3.
[9] World Health Organization. World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals. World Health Organization; 2016.
[10] Insel TR, Cuthbert BN. Brain disorders? Precisely *Science* 2015;348:499–500.
[11] Pardiñas AF et al. Publisher Correction: Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 2019;51:1193.
[12] Matias I, Morgado J, Gomes FCA. Astrocyte Heterogeneity: Impact to Brain Aging and Disease. Front Aging Neurosci 2019;11.
[13] Strohäker T et al. Structural heterogeneity of α-synuclein fibrils amplified from patient brain extracts. Nat Commun 2019;10:5535.
[14] Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. Nat Rev Genet 2015;16:441–58.
[15] Lynch MA, Hardiman O, Elamin M, Kirby J, Rowland LP. Common Themes in the Pathogenesis of Neurodegeneration. Neurodegenerative Disorders 2016;1–12. https://doi.org/10.1007/978-3-319-23309-3_1.
[16] Santiago JA, Bottero V, Potashkin JA. Dissecting the Molecular Mechanisms of Neurodegenerative Diseases through Network Biology. Front Aging Neurosci 2017;9:166.
[17] . (American Psychiatric Pub 2013.
[18] Hafemeister TL. Mental Disorders and Criminal Behavior. Criminal Trials and Mental Disorders 2019;7–42. https://doi.org/10.18574/nyu/9781479804856.003.0002.
[19] Consortium IS et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009;460:748–52.
[20] Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 2013;381:1371–9.
[21] Gandal MJ et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science 2018;359:693–7.
[22] Cross-Disorder Group of the Psychiatric Genomics Consortium. Electronic address: plee0@mgh.harvard.edu & Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. **179**
[23] Li M et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science 2018;362.
[24] Ahmadi A, Gispert JD, Navarro A, Vilor-Tejedor N, Sadeghi I. Single-Cell Transcriptional Changes In Neurodegenerative Diseases. Neuroscience 2021. https://doi.org/10.1016/j.neuroscience.2021.10.025.
[25] Labonté B et al. Sex-specific transcriptional signatures in human depression. Nat Med 2017;23:1102–11.
[26] Jaffe AE et al. Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat Neurosci 2018;21:1117–25.
[27] Wu JQ et al. Transcriptome sequencing revealed significant alteration of cortical promoter usage and splicing in schizophrenia. PLoS ONE 2012;7:e36351.
[28] Pacifico R, Davis RL. Transcriptome sequencing implicates dorsal striatum-specific gene network, immune response and energy metabolism pathways in bipolar disorder. Mol Psychiatry 2017;22:441–9.
[29] MacMullen CM, Fallahi M, Davis RL. Novel PDE10A transcript diversity in the human striatum: Insights into gene complexity, conservation and regulation. Gene 2017;606:17–24.
[30] Akula N et al. RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. Mol Psychiatry 2014;19:1179–85.
[31] Lipska BK et al. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. Biol Psychiatry 2006;60:650–8.
[32] Xiao Y et al. The DNA methylome and transcriptome of different brain regions in schizophrenia and bipolar disorder. PLoS ONE 2014;9:e95875.
[33] Chang X et al. RNA-seq analysis of amygdala tissue reveals characteristic expression profiles in schizophrenia. Transl Psychiatry 2017;7:e1203.
[34] Corley SM, Tsai S-Y, Wilkins MR, Shannon Weickert C. Transcriptomic Analysis Shows Decreased Cortical Expression of NR4A1, NR4A2 and RXRB in Schizophrenia and Provides Evidence for Nuclear Receptor Dysregulation. PLoS ONE 2016;11:e0166944.
[35] Wright C et al. Altered expression of histamine signaling genes in autism spectrum disorder. Transl Psychiatry 2017;7:e1126.
[36] Li J et al. Integrated systems analysis reveals a molecular network underlying autism spectrum disorders. Mol Syst Biol 2014;10:774.
[37] Pantazatos SP et al. Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity. Mol Psychiatry 2017;22:760–73.
[38] Dumitriu A et al. Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease. BMC Med Genomics 2016;9:5.
[39] Wang M et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. Sci Data 2018;5.
[40] He Z, Bammann H, Han D, Xie G, Khaitovich P. Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation. RNA 2014;20:1103–11.
[41] Liu X et al. Disruption of an Evolutionarily Novel Synaptic Expression Pattern in Autism. PLoS Biol 2016;14:e1002558.
[42] Ramaker RC et al. Post-mortem molecular profiling of three psychiatric disorders. Genome Med 2017;9:72.
[43] guigolab. GitHub - guigolab/grape-nf: An automated RNA-seq pipeline using Nextflow. https://github.com/guigolab/grape-nf.
[44] Andrew McKenzie, M. W. A. B. Z. BRETIGEA: Brain Cell Type Specific Gene Expression Analysis. *R package version 1.0.2.* (2019).
[45] van der Maaten L, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res 2008;9:2579–605.
[46] Matute-Blanch C et al. Chitinase 3-like 1 is neurotoxic in primary cultured neurons. Sci Rep 2020;10:7118.
[47] Sanfilippo C, Malaguarnera L, Di Rosa M. Chitinase expression in Alzheimer's disease and non-demented brains regions. J Neurol Sci 2016;369:242–9.
[48] Paschou P, Fernandez TV, Sharp F, Heiman GA, Hoekstra PJ. Genetic susceptibility and neurotransmitters in Tourette syndrome. Int Rev Neurobiol 2013;112:155–77.
[49] Lucas TA, Zhu L, Buckwalter MS. Spleen glia are a transcriptionally unique glial subtype interposed between immune cells and sympathetic axons. Glia 2021;69:1799–815.
[50] Kamboh MI et al. Alpha-1-antichymotrypsin (ACT or SERPINA3) polymorphism may affect age-at-onset and disease duration of Alzheimer's disease. Neurobiol Aging 2006;27:1435–9.
[51] Lei Z, Brizzee C, Johnson GVW. BAG3 facilitates the clearance of endogenous tau in primary neurons. Neurobiol Aging 2015;36:241–8.
[52] Cao Y-L et al. A role of BAG3 in regulating SNCA/α-synuclein clearance via selective macroautophagy. Neurobiol Aging 2017;60:104–15.
[53] Muranova LK, Sudnitsyna MV, Strelkov SV, Gusev NB. Mutations in HspB1 and hereditary neuropathies. Cell Stress Chaperones 2020;25:655–65.
[54] Spiegel I et al. Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. Cell 2014;157:1216–29.
[55] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinf 2008;9.
[56] Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database 2019;2019.
[57] Zhang Y et al. Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron 2016;89:37–53.
[58] Hemonnot A-L, Hua J, Ulmann L, Hirbec H. Microglia in Alzheimer Disease: Well-Known Targets and New Opportunities. Front Aging Neurosci 2019;11:233.
[59] Ferreira SA, Romero-Ramos M. Microglia Response During Parkinson's Disease: Alpha-Synuclein Intervention. Front Cell Neurosci 2018;12.
[60] Rodriguez JI, Kern JK. Evidence of microglial activation in autism and its possible role in brain underconnectivity. Neuron Glia Biol 2011;7:205–13.
[61] Mosser C-A, Baptista S, Arnoux I, Audinat E. Microglia in CNS development: Shaping the brain for the future. Prog Neurobiol 2017;149–150:1–20.
[62] *Brain Pathology* vol. 1 2–5 (1990).

[63] Rodriguez-Vieitez E et al. Diverging longitudinal changes in astrocytosis and amyloid PET in autosomal dominant Alzheimer's disease. Brain 2016;139:922–36.
[64] Siracusa R, Fusco R, Cuzzocrea S. Astrocytes: Role and Functions in Brain Pathologies. Front Pharmacol 2019;10:1114.
[65] Kim D-Y, Hwang I, Muller FL, Paik J-H. Functional regulation of FoxO1 in neural stem cell differentiation. Cell Death Differ 2015;22:2034–45.
[66] Li K, Li J, Zheng J, Qin S. Reactive Astrocytes in Neurodegenerative Diseases. Aging and disease 2019;10:664.
[67] Yao P et al. Coexpression networks identify brain region-specific enhancer RNAs in the human brain. Nat Neurosci 2015;18:1168–74.
[68] Elvsåshagen T et al. The genetic architecture of the human thalamus and its overlap with ten common brain disorders. Nat Commun 2021;12:1–9.
[69] Grinberg LT, Rueb U, Heinsen H. Brainstem: neglected locus in neurodegenerative diseases. Front Neurol 2011;2:42.
[70] Grabher P, Blaiotta C, Ashburner J, Freund P. Relationship between brainstem neurodegeneration and clinical impairment in traumatic spinal cord injury. Neuroimage Clin 2017;15:494–501.
[71] Hemali Phatnani TM. Astrocytes in Neurodegenerative Disease. Cold Spring Harb Perspect Biol 2015;7.
[72] Lunnon K et al. Mitochondrial genes are altered in blood early in Alzheimer's disease. Neurobiol Aging 2017;53:36–47.
[73] Iwamoto K, Bundo M, Kato T. Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis. Hum Mol Genet 2005;14:241–53.
[74] Vawter MP et al. Mitochondrial-related gene expression changes are sensitive to agonal-pH state: implications for brain disorders. Mol Psychiatry 2006;11:663–79.
[75] Winden KD et al. The organization of the transcriptional network in specific neuronal classes. Mol Syst Biol 2009;5:291.
[76] Olesen MA, Torres AK, Jara C, Murphy MP, Tapia-Rojas C. Premature synaptic mitochondrial dysfunction in the hippocampus during aging contributes to memory loss. Redox Biol 2020;34:101558.
[77] Gyoneva S et al. Cx3cr1-deficient microglia exhibit a premature aging transcriptome. Life Science Alliance 2019;2:e201900453.
[78] Reemst K, Noctor SC, Lucassen PJ, Hol EM. The Indispensable Roles of Microglia and Astrocytes during Brain Development. Front Hum Neurosci 2016;10:566.
[79] Fakhoury M. Microglia and Astrocytes in Alzheimer's Disease: Implications for Therapy. Curr Neuropharmacol 2018;16:508–18.
[80] Saxena S, Caroni P. Selective neuronal vulnerability in neurodegenerative diseases: from stressor thresholds to degeneration. Neuron 2011;71:35–48.
[81] Wang X, Christian KM, Song H, Ming G-L. Synaptic dysfunction in complex psychiatric disorders: from genetics to mechanisms. Genome Med 2018;10:9.
[82] Salter MW, Stevens B. Microglia emerge as central players in brain disease. Nat Med 2017;23:1018–27.
[83] Liu Y, Zhou J. Oligodendrocytes in neurodegenerative diseases. Frontiers in Biology 2013;8:127–33.
[84] Sheng Z-H, Cai Q. Mitochondrial transport in neurons: impact on synaptic homeostasis and neurodegeneration. Nat Rev Neurosci 2012;13:77–93.
[85] González-Rodríguez P et al. Disruption of mitochondrial complex I induces progressive parkinsonism. Nature 2021. https://doi.org/10.1038/s41586-021-04059-0.
[86] Stojakovic A et al. Partial inhibition of mitochondrial complex I ameliorates Alzheimer's disease pathology and cognition in APP/PS1 female mice. Communications Biology 2021;4:1–20.
[87] Manji H et al. Impaired mitochondrial function in psychiatric disorders. Nat Rev Neurosci 2012;13:293–307.
[88] Benes FM. The development of 'mis-wired' limbic lobe circuitry in schizophrenia and bipolar disorder. Neurodevelopment and Schizophrenia 2004;295–309. https://doi.org/10.1017/cbo9780511735103.018.
[89] Patel S et al. Maternal immune conditions are increased in males with autism spectrum disorders and are associated with behavioural and emotional but not cognitive co-morbidity. Transl Psychiatry 2020;10.
[90] Eissa N, Sadeq A, Sasse A, Sadek B. Role of Neuroinflammation in Autism Spectrum Disorder and the Emergence of Brain Histaminergic System. Lessons Also for BPSD? Front Pharmacol 2020;11:886.
[91] Starkstein S, Gellar S, Parlier M, Payne L, Piven J. High rates of parkinsonism in adults with autism. J Neurodev Disord 2015;7:29.
[92] Suzuki K et al. Microglial Activation in Young Adults With Autism Spectrum Disorder. JAMA Psychiatry 2013;70:49–58.
[93] Pajares M, Rojo IA, Manda G, Boscá L, Cuadrado A. Inflammation in Parkinson's disease: Mechanisms and therapeutic implications. Cells 2020;9:1687.
[94] Kern JK, Geier DA, Sykes LK, Geier MR. Evidence of neurodegeneration in autism spectrum disorder. Transl Neurodegener 2013;2:17.
[95] Pearson BL et al. Identification of chemicals that mimic transcriptional changes associated with autism, brain aging and neurodegeneration. Nat Commun 2016;7:11173.
[96] Liao X, Yang J, Wang H, Li Y. Microglia mediated neuroinflammation in autism spectrum disorder. J Psychiatr Res 2020;130:167–76.
[97] Matta SM, Hill-Yardin EL, Crack PJ. The influence of neuroinflammation in Autism Spectrum Disorder. Brain Behav Immun 2019;79:75–90.
[98] Riederer P, Lange KW, Kornhuber J, Danielczyk W. Glutamatergic-dopaminergic balance in the brain. Its importance in motor disorders and schizophrenia. Arzneimittelforschung 1992;42:265–8.
[99] McCutcheon RA, Krystal JH, Howes OD. Dopamine and glutamate in schizophrenia: biology, symptoms and treatment. World Psychiatry 2020;19:15–33.
[100] Archer T. Neurodegeneration in schizophrenia. Expert Rev Neurother 2010;10:1131–41.
[101] Ashe PC, Berry MD, Boulton AA. Schizophrenia, a neurodegenerative disorder with neurodevelopmental antecedents. Prog Neuropsychopharmacol Biol Psychiatry 2001;25:691–707.
[102] Ahmed Z, Asi YT, Lees AJ, Revesz T, Holton JL. Identification and quantification of oligodendrocyte precursor cells in multiple system atrophy, progressive supranuclear palsy and Parkinson's disease. Brain Pathol 2013;23:263–73.
[103] Dubois B et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimers Dement 2016;12:292–323.
[104] Coughlin DG, Litvan I. Progressive supranuclear palsy: Advances in diagnosis and management. Parkinsonism Relat Disord 2020;73:105–16.
[105] Dickson DW, Rademakers R, Hutton ML. Progressive Supranuclear Palsy: Pathology and Genetics. Brain Pathol 2007;17:74–82.
[106] Gonatopoulos-Pournatzis T et al. Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions. Mol Cell 2020. https://doi.org/10.1016/j.molcel.2020.01.006.
[107] Sokol DK, Maloney B, Long JM, Ray B, Lahiri DK. Autism, Alzheimer disease, and fragile X: APP, FMRP, and mGluR5 are molecular links. Neurology 2011;76:1344–52.
[108] Murray ME, Dickson DW. Is pathological aging a successful resistance against amyloid-beta or preclinical Alzheimer's disease? Alzheimer's Research & Therapy 2014;6:24.
[109] Penninx BWJH, Lamers F, Milaneschi Y. Clinical heterogeneity in major depressive disorder. Eur Neuropsychopharmacol 2018;28:S59–60.
[110] Cajanus A. *et al.* The Association Between Distinct Frontal Brain Volumes and Behavioral Symptoms in Mild Cognitive Impairment, Alzheimer's Disease, and Frontotemporal Dementia. *Front. Neurol.* **10**, 1059 (2019)
[111] Matsuoka K et al. Left dorsolateral prefrontal cortex atrophy is associated with frontal lobe function in Alzheimer's disease and contributes to caregiver burden. Int J Geriatr Psychiatry 2018;33:703–9.
[112] Maidan I et al. The Role of the Frontal Lobe in Complex Walking Among Patients With Parkinson's Disease and Healthy Older Adults: An fNIRS Study. Neurorehabil Neural Repair 2016;30:963–71.
[113] Ng B et al. Distinct alterations in Parkinson's medication-state and disease-state connectivity. NeuroImage: Clinical 2017;16:575–85.
[114] Wolk DA et al. Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. Neurobiol Aging 2017;49:204–13.
[115] Allen P et al. Abnormal relationship between medial temporal lobe and subcortical dopamine function in people with an ultra high risk for psychosis. Schizophr Bull 2012;38:1040–9.
[116] Cobia DJ, Smith MJ, Wang L, Csernansky JG. Longitudinal progression of frontal and temporal lobe changes in schizophrenia. Schizophr Res 2012;139:1–6.
[117] Mormina E et al. Cerebellum and neurodegenerative diseases: Beyond conventional magnetic resonance imaging. World J Radiol 2017;9:371–88.
[118] Kaufmann T et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nat Neurosci 2019;22:1617–23.
[119] Yan J et al. Identification of discriminative imaging proteomics associations in alzheimer's disease via a novel sparse correlation model. Biocomputing 2017;2017. https://doi.org/10.1142/9789813207813_0010.
[120] Nazeri A et al. Imaging proteomics for diagnosis, monitoring and prediction of Alzheimer's disease. Neuroimage 2014;102(Pt 2):657–65.
[121] Casamassimi A, Federico A, Rienzo M, Esposito S, Ciccodicola A. Transcriptome profiling in human diseases: new advances and perspectives. Int J Mol Sci 2017;18.
[122] Kodama L et al. Microglial microRNAs mediate sex-specific responses to tau pathology. Nat Neurosci 2020;23:167–71.
[123] Villa A, Della Torre S, Maggi A. Sexual differentiation of microglia. Front Neuroendocrinol 2019;52:156–64.
[124] Tiihonen J et al. Sex-specific transcriptional and proteomic signatures in schizophrenia. Nat Commun 2019;10:3933.
[125] Vilor-Tejedor N et al. Strategies for integrated analysis in imaging genetics studies. Neurosci Biobehav Rev 2018;93:57–70.
[126] Staff SRAS. Using the sra toolkit to convert. sra files into other formats. National Center for Biotechnology Information (US); 2011.
[127] Ramirez-Gonzalez RH, Bonnal R, Caccamo M, MacLean D. bio-samtools: Ruby bindings for SAMtools, a library for accessing BAM files containing high-throughput sequence alignments. Open Research Computation 2012;1:1.
[128] Knowles DG, Roder M, Merkel A, Guigo R. Grape RNA-Seq analysis pipeline environment. Bioinformatics 2013;29:614–21.
[129] Di Tommaso P et al. Nextflow enables reproducible computational workflows. Nat Biotechnol 2017;35:316–9.
[130] Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[131] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf 2011;12.

[132] Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics 2016;32:292–4.
[133] Ritchie ME et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[134] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey J. D. sva: Surrogate Variable Analysis. R package version 3.10. 0. 2014;10:B9.
[135] Cousineau D, Chartier S. Outliers detection and treatment: a review. International Journal of Psychological Research 2010;3:58.
[136] Shiffler RE. Maximum Z Scores and Outliers. The American Statistician 1988;42:79.
[137] McKenzie AT et al. Brain Cell Type Specific Gene Expression and Co-expression Network Architectures. Sci Rep 2018;8.
[138] Monfreda M. Principal Component Analysis: A Powerful Interpretative Tool at the Service of Analytical Methodology. Principal Component Analysis 2012. https://doi.org/10.5772/36929.
[139] Pinheiro J, Bates D, DebRoy S, Sarkar D. & Team, R Core. nlme: Linear and nonlinear mixed effects models. R package version **3**, 2012.
[140] Shen LG. An R package to test and visualize gene overlaps. R Package 2014.
[141] Breiman L. Machine Learning 2001;45:261–77.
[142] Plaisier SB, Taschereau R, Wong JA, Graeber TG. Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. Nucleic Acids Res 2010;38:e169–e.
[143] Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD. Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders. J Neurosci 2014;34:1420–31.
[144] Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics 2014;30:2811–2.
[145] Schapira AHV. Mitochondrial Dysfunction in Neurodegenerative Diseases. Neurochem Res 2008;33:2502–9.