*Article*

# Large-Scale Integration of Single-Cell RNA-Seq Data Reveals Astrocyte Diversity and Transcriptomic Modules across Six Central Nervous System Disorders

Zhenwei Qian [1], Jinglin Qin [1], Yiwen Lai [1], Chen Zhang [1,2,3,*] and Xiannian Zhang [1,*]

1 School of Basic Medical Sciences, Beijing Key Laboratory of Neural Regeneration and Repair, Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing 100069, China
2 Chinese Institute for Brain Research, Beijing 102206, China
3 State Key Laboratory of Translational Medicine and Innovative Drug Development, Nanjing 210000, China
* Correspondence: czhang@ccmu.edu.cn (C.Z.); xiannian@ccmu.edu.cn (X.Z.)

**Abstract:** The dysfunction of astrocytes in response to environmental factors contributes to many neurological diseases by impacting neuroinflammation responses, glutamate and ion homeostasis, and cholesterol and sphingolipid metabolism, which calls for comprehensive and high-resolution analysis. However, single-cell transcriptome analyses of astrocytes have been hampered by the sparseness of human brain specimens. Here, we demonstrate how large-scale integration of multi-omics data, including single-cell and spatial transcriptomic and proteomic data, overcomes these limitations. We created a single-cell transcriptomic dataset of human brains by integration, consensus annotation, and analyzing 302 publicly available single-cell RNA-sequencing (scRNA-seq) datasets, highlighting the power to resolve previously unidentifiable astrocyte subpopulations. The resulting dataset includes nearly one million cells that span a wide variety of diseases, including Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), multiple sclerosis (MS), epilepsy (Epi), and chronic traumatic encephalopathy (CTE). We profiled the astrocytes at three levels, subtype compositions, regulatory modules, and cell–cell communications, and comprehensively depicted the heterogeneity of pathological astrocytes. We constructed seven transcriptomic modules that are involved in the onset and progress of disease development, such as the M2 ECM and M4 stress modules. We validated that the M2 ECM module could furnish potential markers for AD early diagnosis at both the transcriptome and protein levels. In order to accomplish a high-resolution, local identification of astrocyte subtypes, we also carried out a spatial transcriptome analysis of mouse brains using the integrated dataset as a reference. We found that astrocyte subtypes are regionally heterogeneous. We identified dynamic cell–cell interactions in different disorders and found that astrocytes participate in key signaling pathways, such as NRG3-ERBB4, in epilepsy. Our work supports the utility of large-scale integration of single-cell transcriptomic data, which offers new insights into underlying multiple CNS disease mechanisms where astrocytes are involved.

**Keywords:** astrocytes; CNS disorder; scRNA-seq; spatial transcriptomics; transcriptomic modules; cell–cell communications

## 1. Introduction

Astrocytes, one of the most abundant glia cell types in the central nervous system (CNS), have been implicated in the pathogenesis of multiple CNS disorders [1]. Accumulating studies suggest that "reactive" astrocytes have a pathogenic role in both acute and chronic disorders, including Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), multiple sclerosis (MS), and behavioral neuropsychiatric disorders [1–3]. Their functions include remodeling of regular physiological processes including metabolism, synapse maintenance, neuroinflammatory regulation, blood–brain barrier

(BBB) integrity, and structural supports. Many studies link reactive astrocytes to inflammatory reactions to dying cells, immune cell dysfunction, and pathogenic proteins, including amyloid-β (Aβ), tau aggregation, and alpha-synuclein (αSyn) [3–5]. It is worth emphasizing that the tripartite synapse concept proposes that the communication between pre- and post-synaptic neurons is modulated by astrocytes, which constitute the third component of the synapse [6]. Astrocytes are key players in this process, as they modulate synaptic transmission and plasticity through various mechanisms, including the uptake and release of neurotransmitters and the modulation of synaptic signaling pathways [7]. The loss of astrocytic synaptic regulatory functions has been associated with the development of numerous neurological disorders, such as AD, PD, and epilepsy (Epi), underscoring the crucial role of these cells in maintaining proper brain function [7,8]. Therefore, understanding the intricate interactions between neurons and astrocytes in the regulation of synaptic function is crucial for elucidating the underlying mechanisms of brain function and dysfunction.

Among these investigations, high-throughput single-cell RNA sequencing (scRNA-seq) has been the leading method to characterize the cellular heterogeneity in neurological disorders, and its adoption has led to the generation of a plethora of datasets [9–12]. However, due to the inaccessibility of human brain specimens, most studies focus on a certain CNS disease from a specific brain region [9,13,14]. Moreover, astrocytes were typically underestimated in previous studies, accounting for just 3–18% of nuclei [11]. Therefore, astrocyte heterogeneity and reactivity are not fully demonstrated, leaving many unanswered questions. For example, what is the global landscape of astrocytes, what is the regional specificity of astrocytes, what are their transcriptomic modules, how are these modules regulated, and what intercellular signaling pathways are disturbed among different CNS diseases. We hypothesized that a single study may not be able to provide large enough samples to definitively characterize physiologically relevant yet lowly abundant astrocyte subpopulations.

Large-scale data integration can capture the comprehensive phenotypic features at single-cell resolution to empower the understanding of disease mechanisms [11,15,16]. One of the major issues in integrating a high-quality reference map from scRNA-seq datasets has been the reduction of technical bias and batch effects [17,18]. Recent developments in computational methods have addressed these issues and allowed the generation of large-scale reference atlases [19]. To portray the complete landscape of astrocytes in both physiological and pathological conditions, we gather and integrate large-scale scRNA-seq datasets from extensive human CNS disorders. We harmonize in total 302 human samples of 6 diseases from 15 studies spanning nearly 1 million cells. We identified eight astrocyte subtypes and explored their distribution specificities among brain regions. Meanwhile, we identified seven common expression modules from astrocytes, and module analysis exhibited pronounced heterogeneity across disorders at the transcriptomic level. We also validated these modules at the protein level by using a mass spectrometry dataset [20]. We further explored ligand-receptor signaling pathways and identify dynamic cell–cell interactions in different disorders, thus offering new insights into underlying CNS disease mechanisms.

## 2. Materials and Methods

### 2.1. Data Collection

We collected single-cell transcriptomic, spatial transcriptomic, and proteomic datasets originating from 17 well-generated studies published from June 2019 to March 2022. Some of these samples were donated by healthy individuals, and some were taken from autopsy tissues. In single-cell transcriptome data, except for Wheeler2020, which used scRNA-seq, all others used single-cell nuclear sequencing (snRNA-seq), which ensured the stability of the sequencing method. Similarly, except for Chancellor2021, which used the inDrops sequencing platform, the rest of the data were sequenced using the 10x Genomics platform. The age of the patients we collected ranged from 24 to 102 years old. As our study covers multiple brain regions, even the cortex alone includes more fine-grained subregions

such as occipital, occipitotemporal, somatosensory, entorhinal, and dorsolateral cortexes. However, since some of the papers did not specify the exact brain region, we grouped them into four categories, cortex (CTX), hippocampus (HIP), white matter (WM), and basal ganglia (BG), for the analysis. The basal ganglia mainly include the three brain regions of caudate, putamen, and substantia nigra. We named the original brain region sources as 'source.tissue', and the new classification is annotated as 'current.tissue' in the paper. See Table S1 for detailed information and metadata.

## 2.2. Preprocessing and Quality Control of Single-Cell RNA-Seq Data

We downloaded all raw scRNA-seq matrix data from different sources (Table S1). We followed Seurat's guide and processed the unique molecular identifier (UMI) count matrix using the R package Seurat (version 4.1.0) [21]. Several previous scRNA-seq studies have shown that dead cells and cell debris are excluded from datasets containing cells with <300 genes, >50,000 UMIs, or 5% mitochondrial genes [22,23]. We followed these criteria above. The final filtered matrix contained 990,317 cells and 36,590 genes.

## 2.3. Normalization and Initial Dimension Reduction of scRNA-Seq Data

After preprocessing and quality control, we merged all datasets using the Seurat::merge function. The UMI count matrix was log normalized, and the 3000 most variable features were identified. The linear dimensionality was reduced using the RunPCA function in Seurat. For graph-based clustering, we used the first 30 principal components determined by the ElbowPlot function, which displayed the *p*-value distribution for each principal component. Our initial results from K-nearest neighbor clustering, using FindNeighbors and FindClusters with the resolution = 0.2, and UMAP clustering, using RunUMAP with dim parameter set to 20, were 53 clusters. We found a significant batch effect in the merged datasets, so we performed a subsequent correction for the batch effect.

## 2.4. Batch Correction, Clustering, and Annotation of scRNA-Seq Data

To remove batch effects and non-biological technical biases, the Harmony algorithm in Seurat was used. We have used default values for the other parameters of Harmony [19]. This was followed by dimensional reduction using RunUMAP with dim parameter set to 20. Clustering was identified using the functions FindNeighbors and FindClusters with resolution 0.1. After integration, the batch effects were satisfactorily removed with no notable separation across different platforms or conditions. Raw clustering generated a total of 28 clusters. Major cell types were annotated based on the expression of known markers, i.e., *GFAP* and *AQP4* for astrocyte; *ABCB1* and *FLT1* for endothelial cells; *RBFOX3* and *GRIN1* for neurons; *CAMK2A* and *CBLN2* for excitatory neurons; *GAD1* and *GAD2* for inhibitory neurons; *DOCK8* and *CSF1R* for microglias; *MBP* and *PLP1* for oligodendrocytes; *PDGFRA* and *VCAN* for OPCs; and *PDGFRB* for pericytes.

## 2.5. Subcluster Analysis of Astrocytes

We re-clustered astrocytes (the analysis procedure is the same as above) and identified eight transcriptionally unique subpopulations. To identify the enriched genes in each astrocyte subtype, the FindAllMarkers function (log2FC > 0.25, using a Wilcoxon rank-sum test, adjusted *p*-value < 0.05 using the Bonferroni correction) was used to determine unique and/or highly enriched DEGs in one cluster compared to all other clusters. These cluster-specific features were used as reference genes of GO pathway analysis with the R package clusterProfiler (version 4.2.2, accessed on 11 April 2023) [24]. Data were visualized using Seurat package functions, including DimPlot, FeaturePlot, DotPlot, and DoHeatmap.

## 2.6. The Signature Scores for General Characteristics of Astrocytes

We assessed signature gene scores of astrocytes using the AddModuleScore function in Seurat. These genes originate from previously published studies including PR, pan-reactive astrocytes (*ASPG, S1PR3, CXCL10, VIM, GFAP, CD44, CP, OSMR, TGM1,*

*FKBP5*, *TIMP1*, *HSBP1*, and *STEAP4*); A1, A1-reactive astrocytes (*C3*, *SRGN*, *GBP2*, *PSMB8*, *AMIGO2*, *FBLN5*, and *SERPING1*); A2, A2-reactive astrocytes (*EMP1*, *CD14*, *CD109*, *PTX3*, *STAT3*, *PTGS2*, *TM4SF1*, *CLCF1*, *SPHK1*, *B3GNT5*, *S100A10*, and *SLC10A6*); HM, homeostasis (*SLC1A2*, *SLC1A3*, *GLUL*, *SLC6A11*, *NRXN1*, *CADM2*, *PTN*, and *GPC5*); and PC, phagocytosis (*MEGF10* and *MERTK*) [25–27].

### 2.7. Analysis of Spatial Transcriptomic Data

We downloaded the recently published 10x Visium dataset from whole hemisphere saline or LPS-injected mice (N = 3 per condition) (data available at the Gene Expression Omnibus repository: series GSE165098) [10]. Data were analyzed using Seurat (version 4.1.0) [21]. In short, each section was normalized using SCTransform and merged based on variable features. Data were visualized using the SpatialFeaturePlot function in Seurat.

### 2.8. Regional Annotation of the Mouse Brain

We assessed module scores of classical region-specific genes using the AddModule-Score function in Seurat. These modules include cortex (CTX; *Mef2c*, *Mef2d*), hippocampus (HIP; *Cst3*, *Glul*), white matter (WM; *Plp1*, *Mbp*, and *Mog*), and ventricle modules (*Ccdc153*, *Foxj1*, *Enkur*, and *Ttr*) [28–32]. The midbrain (MB) module gene list was obtained from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb/, accessed on 11 April 2023).

### 2.9. Mouse Gene Signatures for Each Astrocyte Cluster

To highlight clusters identified in our single-cell RNA-seq data, we defined gene signatures for each astrocyte cluster using the top 50 marker genes (genes ranked by log2FC) from our earlier enriched genes in each astrocyte subtype (see Subcluster Analysis of Astrocytes and Table S2). To explore the expression of these gene signatures in the mouse dataset as well, we created mouse gene signatures comprising one-to-one orthologs of the human module genes, which were identified using the biomaRt R package (version 2.50.3) [33]. Signature scores were calculated using the AddModuleScore function in Seurat.

### 2.10. Identification of Astrocyte Transcriptomic Modules

The samples with less than 100 astrocytes were excluded. For each of the 294 samples with at least 100 astrocytes, we used non-negative matrix factorization (NMF) to extract 7 co-expression submodules with the NNLM R package (version 0.4.4) (https://github.com/linxihui/NNLM, accessed on 11 April 2023). The specific steps are as follows. We first filtered DEGs of each astrocyte cluster (log2FC > 0.6, adjusted *p*-value < 0.05). We used these genes to construct a normalized expression matrix (genes x cells) for each sample. If the number of astrocytes per sample was greater than 300, we set the number of decomposition rank to 8, and if it was between 100 and 300 then it was set to 6, which resulted in a total of 1556 submodules across the 294 samples. The 1556 submodules were compared by hierarchical clustering using the Pearson correlation coefficient over all gene scores as a distance metric. Based on the clustering dendrogram and cutree function in R, seven clusters of modules were identified manually. We used the top 20 genes of each module to define the module. For a single cell, if it expressed ≥70% of the genes within a given module, it was considered as a cell expressing the activated module, namely, a module cell.

### 2.11. Functional Analysis of Gene Modules

To functionally describe the astrocyte expression module and non-module cell subtypes, we performed the pathway analysis based on the hallmark and gene ontology gene sets of the MSigDB database and estimated the pathway activity of individual cells using the gene set variation analysis (GSVA) package (version 1.42.0) with standard settings [34]. To assess the differential activities of pathways between module cells

and non-module cells, activity scores were contrasted for each cell group using the Limma package (version 3.50.1) [35].

### *2.12. The Interaction Network of the M2 and M4 Modular Genes*

We used the top 50 M2 and M4 modular genes to construct the interaction networks by STRING analysis (https://string-db.org, accessed on 11 April 2023) and used Cytoscape (version 3.9.1) to visualize the networks.

### *2.13. Differential Expression Analysis and PCA of Different Disorders*

To avoid the spatial bias, the DEG analysis was firstly conducted between disease and controls from the same brain region. The number of unique and overlapping up- and down-regulated DEGs among different disorders are shown using UpsetR (version 1.4.0) [36]. We used genes shared in at least three out of six disorders to construct interaction networks by STRING analysis (https://string-db.org, accessed on 11 April 2023) and used Cytoscape (version 3.9.1) to visualize the networks. PCA was performed based on the top 200 DEGs for each disorder using the scatterplot3d package (version 0.3-41).

### *2.14. Proteomics Dataset Analysis*

We downloaded a recently published TMT-MS dataset, which has been normalized [20]. We performed strict quality control and removed proteins and samples containing missing values. After data processing and quality control, a total of 8619 proteins were retained for downstream analysis. We used the top 20 genes of each module to construct protein co-expression modules. At least more than half genes of each module, except M7, consisting of five proteins were paired, but the five proteins were highly correlated with cilium and could greatly represent the M7 module function. We then calculated the corresponding protein abundance scores for each module and assessed correlations within modules. Analysis of modular dependency at the protein level was estimated using Spearman correlations. Correlation coefficients and *p*-values were indicated by a heatmap.

### *2.15. Ligand-Receptor Analysis with CellChat*

Ligand-receptor analysis and visualization were performed using CellChat (version 1.1.3) [37]. CellChat offers reliable quantitative inference, analysis, and visualization of the intercellular communication network based on previously identified ligand/receptor pairings and other signaling cofactors. To dissect this global alteration in different disorders, we first calculated the relative information flows for each signaling pathway using the compareInteractions function, which is defined as the total communication probability among all the pairs of cell groups in the communication network. We also showed key signaling pathways using the netVisual_bubble function. Default values were used for the parameterization of each step.

## 3. Results

### *3.1. Large-Scale Unbiased Integration of Human scRNA-Seq Dataset*

The overall strategy includes collecting human scRNA-seq, spatial transcriptomic, and proteomic datasets to avoid the limitations by single methodology (Figure 1A–C). Briefly, we performed quality control and integration for scRNA-seq datasets from human CNS disease tissues (Figure 1A). To provide a high-resolution reference map for their regional preferences, a mouse spatial-transcriptomes dataset was included (Figure 1B). To validate these astrocyte subtypes and modules at the protein level, we used a published tandem mass tag mass spectrometry (TMT-MS) dataset in AD patient brain tissues (Figure 1C). In total, we curated 302 publicly available human brain scRNA-seq datasets that were generated by different platforms (10x v2/3 and inDrops v3) from June 2019 to March 2022 (Figure 1D). They originated from six CNS diseases, known as AD, PD, HD, MS, epilepsy (Epi), and chronic traumatic encephalopathy (CTE). After removing ambient RNA signatures and removing cells of low quality (Supplementary Figure S1A), we merged the

datasets to perform preliminary single-cell transcriptome analysis. The resulting dataset consisted of 990,317 cells and nuclei (Figure 2A and Table S1). Due to differences in the specimen sources, experimental conditions, and library preparation methods, substantial batch effects were evident in dimensionally reduced visualizations of these data (Figure 2A). After Harmony integration, the batch effects were satisfactorily removed with no notable separation across different platform or conditions (Figure 2B). For integrated datasets, network-based clustering analysis revealed 28 cell clusters, which were further manually categorized into 8 major cell types using canonical marker genes including astrocytes (Astro), endothelial cells (Endo), excitatory (ExN) and inhibitory (InN) neurons, microglia (Micro), oligodendrocytes (Oligo), oligodendrocyte progenitor cells (OPCs), and pericytes (Peri) (Figure 2C,D). Each dataset contained all the cell types (Supplementary Figure S1B,C). Thus, we have completed the integration of the datasets without bias and identified the main cell types for downstream analysis.
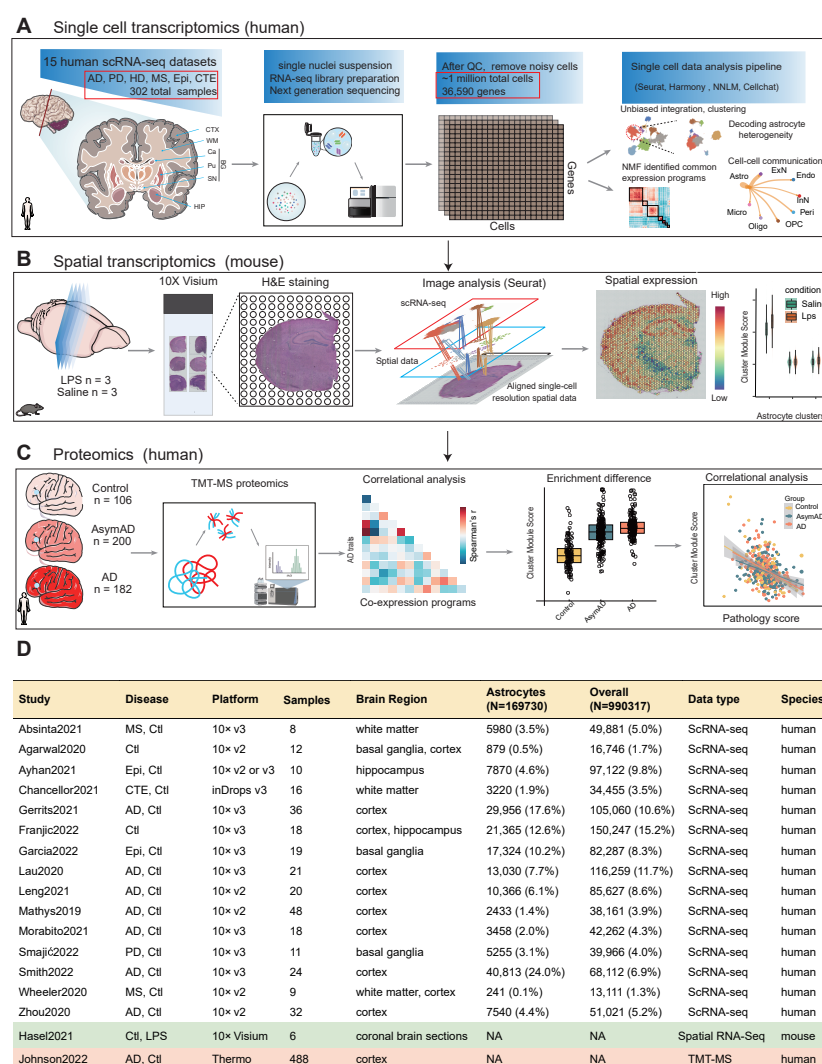


| Study | Disease | Platform | Samples | Brain Region | Astrocytes (N=169730) | Overall (N=990317) | Data type | Species |
|---|---|---|---|---|---|---|---|---|
| Absinta2021 | MS, Ctl | 10× v3 | 8 | white matter | 5980 (3.5%) | 49,881 (5.0%) | ScRNA-seq | human |
| Agarwal2020 | Ctl | 10× v2 | 12 | basal ganglia, cortex | 879 (0.5%) | 16,746 (1.7%) | ScRNA-seq | human |
| Ayhan2021 | Epi, Ctl | 10× v2 or v3 | 10 | hippocampus | 7870 (4.6%) | 97,122 (9.8%) | ScRNA-seq | human |
| Chancellor2021 | CTE, Ctl | inDrops v3 | 16 | white matter | 3220 (1.9%) | 34,455 (3.5%) | ScRNA-seq | human |
| Gerrits2021 | AD, Ctl | 10× v3 | 36 | cortex | 29,956 (17.6%) | 105,060 (10.6%) | ScRNA-seq | human |
| Franjic2022 | Ctl | 10× v3 | 18 | cortex, hippocampus | 21,365 (12.6%) | 150,247 (15.2%) | ScRNA-seq | human |
| Garcia2022 | Epi, Ctl | 10× v3 | 19 | basal ganglia | 17,324 (10.2%) | 82,287 (8.3%) | ScRNA-seq | human |
| Lau2020 | AD, Ctl | 10× v3 | 21 | cortex | 13,030 (7.7%) | 116,259 (11.7%) | ScRNA-seq | human |
| Leng2021 | AD, Ctl | 10× v2 | 20 | cortex | 10,366 (6.1%) | 85,627 (8.6%) | ScRNA-seq | human |
| Mathys2019 | AD, Ctl | 10× v2 | 48 | cortex | 2433 (1.4%) | 38,161 (3.9%) | ScRNA-seq | human |
| Morabito2021 | AD, Ctl | 10× v3 | 18 | cortex | 3458 (2.0%) | 42,262 (4.3%) | ScRNA-seq | human |
| Smajić2022 | PD, Ctl | 10× v3 | 11 | basal ganglia | 5255 (3.1%) | 39,966 (4.0%) | ScRNA-seq | human |
| Smith2022 | AD, Ctl | 10× v3 | 24 | cortex | 40,813 (24.0%) | 68,112 (6.9%) | ScRNA-seq | human |
| Wheeler2020 | MS, Ctl | 10× v2 | 9 | white matter, cortex | 241 (0.1%) | 13,111 (1.3%) | ScRNA-seq | human |
| Zhou2020 | AD, Ctl | 10× v2 | 32 | cortex | 7540 (4.4%) | 51,021 (5.2%) | ScRNA-seq | human |
| Hasel2021 | Ctl, LPS | 10× Visium | 6 | coronal brain sections | NA | NA | Spatial RNA-Seq | mouse |
| Johnson2022 | AD, Ctl | Thermo | 488 | cortex | NA | NA | TMT-MS | human |

**Figure 1.** Overview outline of astrocyte dataset categories and analysis pipeline. (**A**) Workflow of data generation, integration, and annotation for the human brains' single-cell transcriptome datasets. (**B**) Spatial transcriptomics profiling workflow of mouse brains. (**C**) Processing workflow of the TMT-MS proteomics dataset on control, asymptomatic AD (AsymAD), and AD brain tissues. (**D**) Metadata table for all datasets included in this study. Abbreviations: AD, Alzheimer's disease; PD, Parkinson's disease; HD, Huntington's disease; MS, multiple sclerosis; Epi, epilepsy; CTE, chronic traumatic encephalopathy; AsymAD, asymptomatic AD; scRNA-seq, single-cell RNA-sequencing; spatial RNA-seq, spatial RNA-sequencing; TMT-MS, tag mass tag spectrometry.
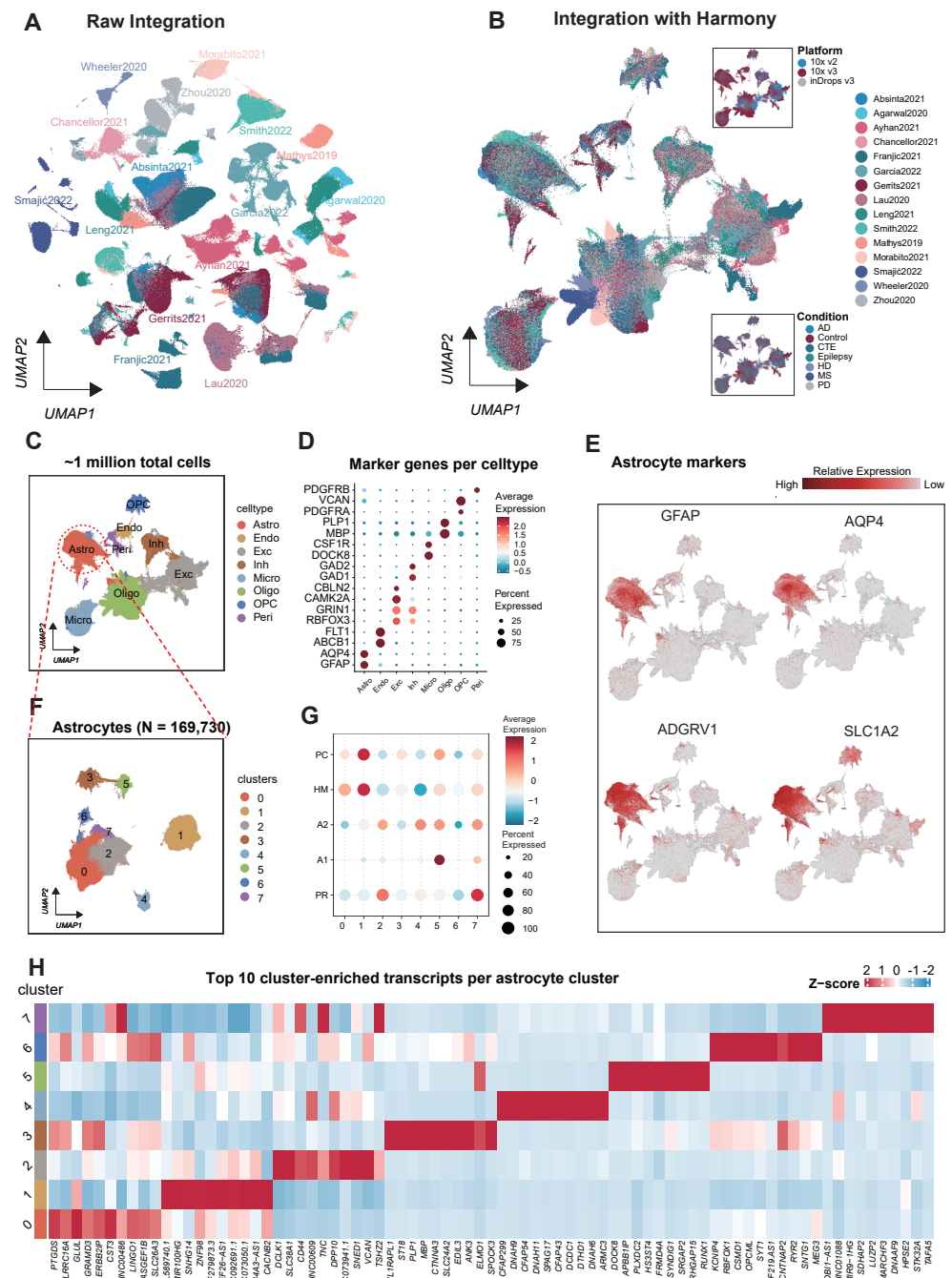
**Figure 2.** Large-scale integration of 302 single-cell RNA-seq samples reveals astrocyte subtypes in the human brain. (**A**) UMAP plot of the combined datasets before batch correction, colored by dataset source. (**B**) UMAP plot of integrated dataset after batch correction with Harmony, colored by dataset source. The corner insets are colored by scRNA-seq platforms (up) and disease types (down). (**C**) UMAP plot colored by major cell types in the human brain. (**D**) Expression of canonical marker genes of major cell types. Dot plot showing the expression level (color scale) and the percent of cells expressing (dot size) marker genes across cell types. (**E**) UMAP plots of all cells colored by expression levels of four classical astrocyte markers. (**F**) A total of 169,730 astrocytes were re-clustered into 8 astrocyte subtypes (cluster 0−7). (**G**) General characteristics of 8 subtypes using relative expression strength of 5 known astrocyte programs: phagocytosis (PC); homeostasis (HM); A1− and A2−reactive astrocytes (A1, A2); and pan-reactive astrocytes (PR). (**H**) Heatmap showing the normalized average expression of the top 10 enriched genes for each cluster.

### 3.2. Molecularly and Functionally Distinct Astrocyte Subtypes Are Identified

In total, 169,730 astrocytes were identified by showing high expression of markers such as *ADGRV1*, *AQP4*, *SLC1A2*, and *GFAP* (Figure 2E). We sought to determine whether the increased cell numbers enabled us to detect novel subpopulations that had previously been missed. Subcluster analysis of astrocytes generated eight transcriptionally unique subpopulations (Figure 2F–H). The biological relevance of each subtype was checked using several published functional gene signatures including phagocytosis (PC), homeostasis (HM), and A1-, A2-, and pan-reactive astrocytes (PR) [25–27] (Figure 2G). Their marker gene sets and functional annotations were listed (Supplementary Figure S3 and Table S2). We found that clusters 0 and 1 both highly expressed homeostasis signatures involved in amino acid transport, transportation across the blood–brain barrier, and positive regulation of synaptic transmission (e.g., *SLC1A2*, *SLC1A3*, and *GLUL*), suggesting their important putative protective roles in maintaining CNS homeostasis. In addition, cluster 0 was also involved in transition metal ion homeostasis (e.g., *MT3*, *FTL*/*FTH1*), and expressed *APOE* and *CLU* involved in amyloid metabolism and clearance [38–40]. Interestingly, some lncRNAs, *SNHG14* and *MIR100HG*, were enriched in cluster 1. *SNHG14* has previously been reported to provide neuroprotection by inhibiting inflammatory responses during AD progress [41]. In addition, phagocytosis-associated signatures (*MEGF10*, *MERTK*) were significantly enriched in cluster 1, which has previously been reported as a novel role for astrocytes in mediating synapse elimination to achieve precise neural connectivity [42,43]. Therefore, we hypothesize that there are subtype-specific astrocytes and not all cells perform this phagocytic function in the adult brain.

Pan-reactive astrocyte signatures (e.g., *GFAP* and *CD44*) were enriched in cluster 2 and 7, representing astrocyte activation [44,45]. Cluster 2 highly expressed genes involved in oxidative stress (*FOS*, *JUN*/*JUNB*, *UBC*, and *ID3*) and the extracellular matrix (*TNC*, *TNR*, and *VCAN*) and associated with proteostasis (e.g., *HSP90AA1*, *HSPA1A*, and *HSPB1*, known as chaperones). These reactive genes presenting in the same astrocyte cluster suggest their putative close interaction. Cluster 5 was enriched for inflammation/A1 astrocyte signatures and highly expressed immune response genes including *DOCK8*, *APBB1IP*, *C3*, *SPP1*, *RUNX1*, and *PLXDC2*. Previously published studies have reported complement component 3 (C3) in a specific subtype of reactive astrocytes that respond to inflammation and mediate microglia–astrocyte crosstalk in a range of neurodegenerative diseases such as AD, HD, MS, and amyotrophic lateral sclerosis (ALS) [4,26,46]. Cluster 4 expressed many microtubule-related genes involved in the formation of the cytoskeleton. Interestingly, we found that cluster 3 expressed oligodendrocyte marker genes (*MBP* and *PLP1*) and higher *OLIG2* (oligodendrocyte transcription factor 2) compared with other clusters ($p = 0.022$, Table S2). We adopted the expression of *MBP* in microglia cells as a measurement of background *MBP* levels. We compared the total *MBP* expression level of all astrocyte cells or astrocyte cluster 3 to that of microglia. We found that *MBP* was expressed at much higher levels in astrocytes than in microglia, which indicates that we have ruled out the possibility of acellular free-RNA cross-contamination affecting the properties of astrocytes (Supplementary Figure S2C). Therefore, cluster 3 likely represented OLIG2$^+$ astrocyte subtypes involved in ensheathment of neurons [47,48]. Lastly, cluster 6 especially expressed synapse-associated genes (e.g., *SYT1*, *NLGN1*, and *GRIN2B*) and was involved in synaptic maintenance. Together, we identified eight transcriptionally distinct astrocyte subtypes, suggesting that integrated data enabled us to detect novel subpopulations that had previously been missed.

### 3.3. Astrocyte Subtypes Are Ogeneous in Human and Mouse Brains

Recent findings showed astrocytes varied across the CNS regions [10,11,49]. We next aimed to determine whether our diverse astrocyte subgroups are located in various parts of the brain. We found that some astrocyte cluster modules exhibited significantly region-specific expression in the mouse brain, in line with the origin of astrocytes in human brain tissue (Figure 3A). For example, cluster 1 was significantly enriched in the cerebral cortex

of humans where *RORB* and *TSHD7A* were enriched, which have previously been reported to show cortex-specific expression in astrocytes and neurons [49,50] (Figure 3B).
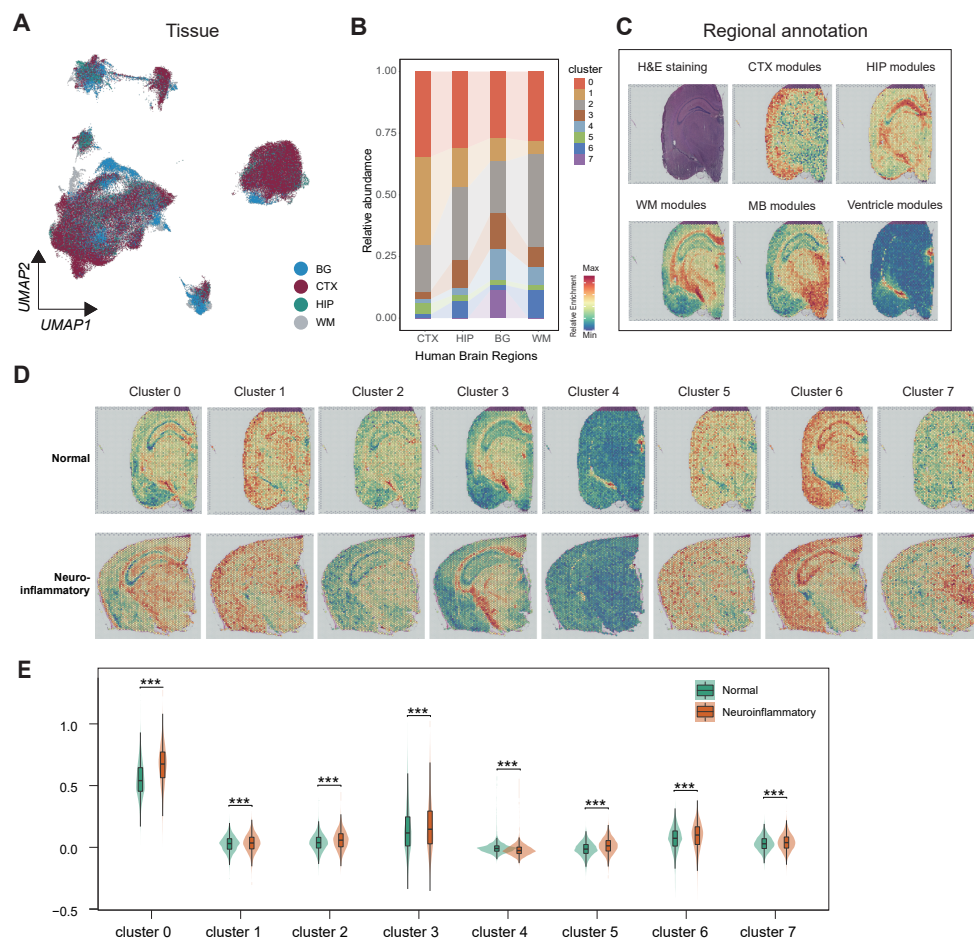


**Figure 3.** Astrocyte subtypes are regionally heterogeneous. (**A**) UMAP plot of astrocytes, colored by the human brain regions including basal ganglia (BG), cortex (CTX), hippocampus (HIP), and white matter (WM). (**B**) Relative proportions of astrocyte subtypes for samples from different brain regions. BG, N = 40; CTX, N = 209; HIP, N =16; WM, N = 29. (**C**) Identification of brain region locations using module scores of classical region-specific marker genes: CTX, HIP, WM, midbrain (MB), and ventricle modules. (**D**) The highlighted spatial distribution of 8 astrocyte subtype signatures in sections from normal (N = 3) and neuroinflammatory (N = 3) brains using Visium spatial transcriptomics. (**E**) Quantification of signature scores of the top 50 enriched genes for each cluster, across all spots from normal and neuroinflammatory samples. (\*\*\* $p < 0.001$, Kruskal–Wallis test).

In order to examine regional variations in our astrocyte subtypes in the human and mouse brain, we used available spatial transcriptomics datasets [10] (Figure 3C,D). To precisely annotate the spatial transcriptomics datasets, we labeled the brain slices with the expression level of known regional specific markers (Figure 3C) including cortex (CTX; *Mef2c*, *Mef2d*), hippocampus (HIP; *Cst3*, *Glul*), white matter (WM; *Plp1*, *Mbp*, and *Mog*), midbrain (MB, see methods), and ventricle modules (*Ccdc153*, *Foxj1*, *Enkur*, and *Ttr*) [28–30,30–32] (Figure 3C). We then generated the spatial expression of assembled marker genes for each cluster to estimate the likely location of each astrocyte population (Figure 3D). These cluster signatures have diversified brain region distributions. Cluster 1 signatures are enriched in the cortex (Figure 3D). Notably, we discovered that cluster 6 was considerably enriched in the mouse cortex and hippocampal regions, which is in line with the involved synapse formation function (Figure 3D). The cluster 4 signature was exclusively observed in regions surrounding the 3rd and lateral ventricles of the mice, suggesting that cluster 4 might interact with some factors carried by the cerebrospinal

fluid [10,51]. Meanwhile, we also validated the expression of tanycyte markers such as COL23A1 and SLC16A2 in cluster 4, and we did not detect the expression of these two genes in cluster 4 (Supplementary Figure S2B) [52]. Therefore, we excluded the possibility that cluster 4 represents tanycytes. Interestingly, we found that cluster 7 from human basal ganglion (BG) tissues was significantly enriched in MB of the mouse brain. In addition to expressing genes involved in cell–cell and cell–ECM (extracellular matrix) interaction (*TNC*, *CD44*, and *APOE*), cluster 7 also highly expressed *LUZP2* (log2FC = 2.48) compared to other clusters (Table S2). This is consistent with the report that an LUZP2$^+$ astrocyte subset exists in the mouse medial amygdala, part of the BG [53]. These results suggested that cluster 7 may represent a novel astrocyte subtype unique to the basal ganglion in human and mouse brains. In addition to the conserved subtypes shared between mouse and human brains, we also found species-specific differences, such as the large proportion of cluster 0 distributed in the cortex of the human brain, which we found difficult to detect in the mouse cortex. This may be due to the lack of interlaminar and varicose projection astrocytes in rodent animals, which has been reported by many studies [54,55].

In addition, we also explored whether neuroinflammation induced by the endotoxin lipopolysaccharide (LPS) affects the expression of module signatures. We found that all the module signatures, except for cluster 4, were significantly upregulated in neuroinflammatory mice compared to normal controls (Figure 3E). Together, we discovered notable and robust brain area preferences for astrocyte clusters that are shared between the human and mouse datasets, such as clusters 1 and 7, indicating that the astrocyte subtypes we found may be preserved between species at the transcriptome level. Meanwhile, we also found differences in the transcriptional profiles of some astrocyte subtypes between human and mouse brains, such as cluster 0, which indicates species-specific variation.

### 3.4. Astrocyte Subtypes Are Partially Associated with CNS Disease

To investigate how astrocytes contribute to disease onset, we preliminarily analyzed the relationship between astrocyte subtypes and disease (Supplementary Figure S4). We found clusters 2 and 7 showed an increase in their relative abundance in PD compared with control groups in the same regional brain. In contrast, clusters 3 and 5 showed a sharp decrease in PD. Furthermore, we found that multiple astrocyte subtypes were involved in the development of epilepsy and multiple sclerosis, such as clusters 1, 2, 3, and 5 in epilepsy and clusters 3 and 4 in MS. Interestingly, we also identified two astrocyte subtypes, clusters 0 and 6, associated only with HD and CTE, respectively, likely indicating that only the sole astrocyte subtype, but not all, is involved in disease progression.

Unfortunately, we could not find statistical differences in AD compared with control groups, which is contrary to previous studies [11,56]. We hypothesized that pathological stimulus does not always lead to quantitative changes in astrocyte subtype compositions but could induce functional changes in the form of interference of gene modules. In addition, we have found that some diseases are associated with multiple astrocyte subtypes, and it remains unclear whether the common or specific functions of these astrocytes are responsible for the development of the disease. Therefore, to overcome these problems, we set out to uncover the prevalent regulatory modules across all subsets of astrocytes.

### 3.5. Identification of Seven Common Expression Modules of Astrocytes across Control and Disease Tissues

We used non-negative matrix factorization (NMF) for 294 human brain scRNA-seq datasets to uncover coherent sets of genes that were preferentially co-expressed by subsets of astrocytes. We identified seven expression modules (M1–M7) with different functions and cell statuses (Figure 4A,B; Table S3) and defined cells expressing ≥70% of genes in each module as module cells (Figure 4C). We selected the most activated pathways in the module cells by comparison with non-module cells to analyze their functions using GSVA analysis (Figure 4D). Despite the fact that the number of modules and previously identified clusters are similar, they are not directly corresponded, as many modules tend

to span across multiple clusters. Specifically, M1, M2, M6, and M4 modules are widely expressed with different strengths, suggesting that they are necessary for most astrocyte subtypes to perform basic functions (Figure 4C). The M1 ion homeostasis module was characterized by the expression of genes associated with ion transport (e.g., *GPC5*, *RYR3*, *SLC4A4*, and *GABRB1*). M1 module cells were involved in calcium ion transport, inorganic ion transmembrane transport, and regulation of cellular PH pathways (Figure 4D). The M6 metabolism module consisted of metabolism-associated genes (e.g., *CLU*, *GLUG*, *GJA1*, *APOE*, and *MT3*), which comprise two AD-risk genes (*APOE*, and *CLU*). M6 module cells also displayed small molecule catabolic process, lipid catabolic process, and amyloid beta clearance (Figure 4D). These results suggest M6 plays a critical role in AD pathogenesis by metabolically interacting with neurons [57]. The M2 ECM module had increased expression of ECM-associated genes (e.g., *CD44*, *TNC*, and *VCAN*) and protein kinases (*DCLK1/2*) that were involved in radial migration and axon growth of cortical neurons and associated with neurodevelopmental and neuropsychiatric disorders [58,59]. Additionally, *GLIS3* in the M2 module has been studied to influence the glioma cells' invasion, migration, and proliferation and upregulate the NF-κB signaling pathway [60]. The M4 stress module is a known signature of reactive astrocytes highly expressed in stress/inflammatory genes such as heat shock proteins (*HSPB1*, *HSP1A1*, and *HSPBB*) and cellular stress (e.g., *GFAP*, *UBC*, *FOS*, and *JUN*). M4 stress cells showed activation of the response to oxidative stress, P53 pathway, apoptosis, and TNFα signaling (Figure 4D).

In contrast, other modules were expressed in different astrocyte clusters in a special-cluster manner, such as M3 in cluster 3, M5 in cluster 5, and M7 in cluster 4 (Figure 4C). The M3 synapse module was characterized by high expression of genes involved in synapse organization (e.g., *IL1RAPL1*, *CTNNA3*, *PTPRD*, and *DLG2*) and ensheathment of neurons (e.g., *MBP*, *PLP1*, and *ST18*), in line with displaying regulation of axonogenesis and presynaptic organization pathways (Figure 4B,D). The M5 immunity module consisted of immune response genes (*LRMDA*, *ELMO1*, *MEF2A/C*, *SLC8A1*, and *APBB1IP*) and performed immune activation functions, including positive regulation of phagocytosis and the antigen receptor-mediated signaling pathway [61,62]. The M7 cilium module had overexpressed genes related to the cilium and microtubule, such as cilium proteins (e.g., *CFAP43/54*, *DNAH7/9*, *HYDIN*, and *RFX3*) and microtubule proteins (e.g., *AGBL4*, *DCDC1*, and *SPAG17*). Together, we identified seven co-expression modules of astrocytes from multiple samples, which perform specific functions.

### 3.6. Identification of Dependency among the Expression Modules

We next explored how these modules are co-regulated in pairwise using Spearman correlation analysis (Figure 4E). We identified one significant co-occurring module pair as well as three mutually exclusive module pairs (r ≥ 0.3 or ≤ −0.3, *p* < 0.05, Figure 4E). The M2 ECM showed the strongest positive correlation with the M4 stress/inflammation (r = 0.36). We found an intense interaction network among these two modular genes, probably reflecting their cooperation and coregulation in multiple processes (Figure 4F). Astrocytes as a source for ECM and cytokines are well known, and there is growing evidence that astrocytes communicate through direct cell–cell contacts or via the extracellular matrix to transmit the activated state to individual ECM-connected cells, which is associated with various neurological diseases [63–66]. Another study finds that matrix stiffness affects astrocytic activation and phenotypes in an in vitro injury model [67]. Notably, M1 ion homeostasis showed the strongest negative correlation with the M4 stress/inflammation (r = −0.51), which likely reflects that astrocyte reactivity could disrupt ion homeostasis of the brain microenvironment. Together, these results indicate that astrocytes have function-specific co-expression modules, and these modules are highly heterogeneous among astrocyte subtypes.
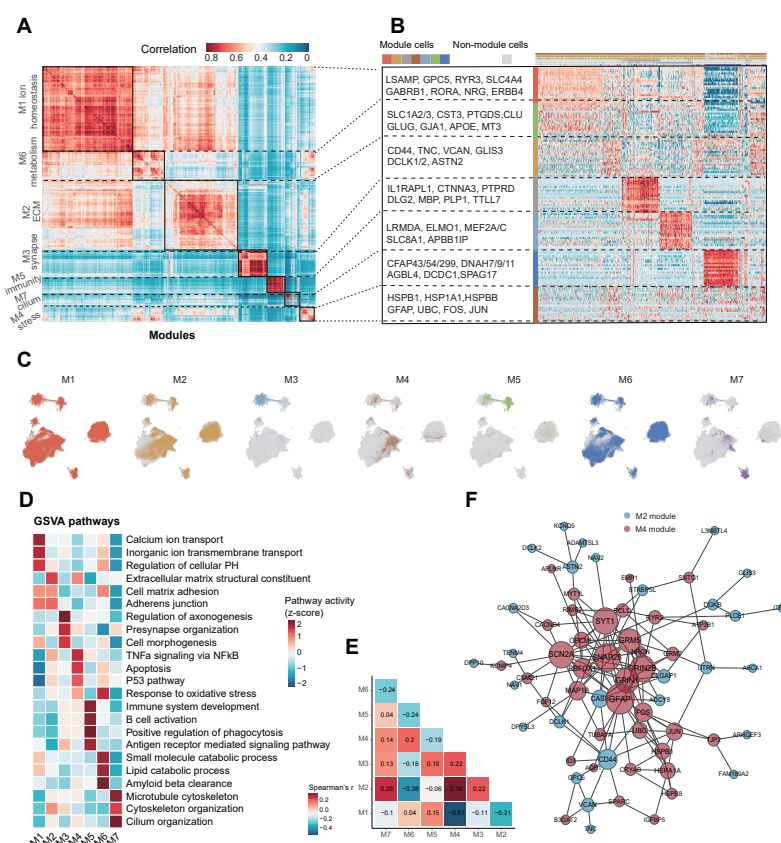
**Figure 4.** Identification of astrocyte transcriptomic gene modules. (**A**) Heatmap showing pairwise correlations of non-negative matrix factorization (NMF) submodules derived from 294 samples. The resembled submodules across samples are aggregated into 7 modules (M1−M7). (**B**) Heatmap showing expression of marker genes within each module across single cells. Colors above columns correspond to cell state. (**C**) The distributions of each expression module. Cells expressing ≥70% of genes in each module are defined as module cells. (**D**) Heatmap showing pathway activity differences between module cells and non-module cells for each module scored by GSVA. Each column is normalized by z−score to indicate the relative pathway activities. (**E**) Heatmap showing Spearman correlation coefficients assessing for each pair of modules (rows, columns) if they are co-occurrent (≥0.3, red, positive correlation) or exclusive (≤−0.3, blue, negative correlation). (**F**) The interaction networks for the M2 ECM and M4 stress modular genes (blue, M2 modular genes; red, M4 modular genes). Node size is the number of links; edges represent interactions between nodes.

### 3.7. Brain Region Is a Key Factor for Transcriptome Differences in Astrocytes

To study whether astrocytes have contextual changes across different CNS disorders, we next analyzed differentially expressed genes (DEGs) and identified both common and disorder-specific transcriptomic changes across different CNS disorders including HD, PD, AD, Epi, CTE, and MS. To avoid spatial bias, the DEG analysis was conducted between disease and controls from the same brain region. In total, we identified 2294 upregulated DEGs and 3407 down-regulated DEGs across all disorders that were presumably driven by disorders (Table S4). Principal component analysis (PCA) of reactivity DEGs was performed to infer the correlation among different diseases (Figure 5A). It was expected that these samples were separated according to disease types. However, the PCA showed near-orthogonal differences, and samples from the same brain region had the closest proximity, indicating that the astrocyte heterogeneity among different brain regions could be stronger than that of disease states. Several recent studies have demonstrated the existence of astrocyte spatial heterogeneity, and astrocytes could produce spatial-specific and subtype-specific responses to inflammatory or pathological stimuli by in situ hybridization and spatial transcriptomics [10,32,68]. Meanwhile, we also found that samples of hippocampus and

cortex tissues are closely grouped, suggesting that hippocampus and cortex astrocyte transcriptional levels are similar. However, epilepsy samples of the hippocampus broke away from clusters of the same tissue (Figure 5A), probably reflecting that epilepsy pathology causes greater transcriptional alterations in astrocytes compared to regional factors.
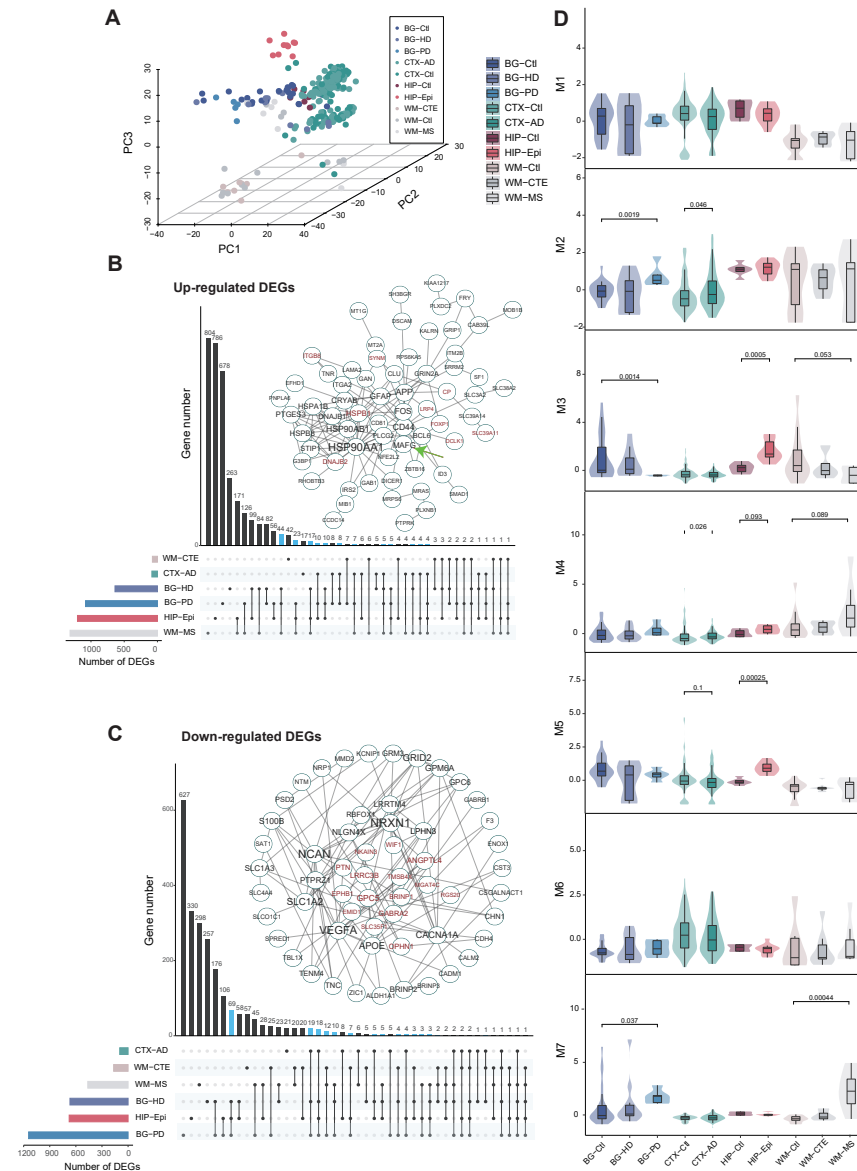


**Figure 5.** Gene and modular transcriptomic changes across multiple CNS disorders and brain regions. (**A**) Three-dimensional PCA plot for different CNS disorder samples based on disease-associated genes. Samples are grouped by both brain region and disease condition: basal ganglion control (BG-Ctl, N = 24), basal ganglion Huntington's disease (BG-HD, N = 11), basal ganglion Parkinson's disease (BG-PD, N = 5), cortex Alzheimer's disease (CTX-AD, N = 120), cortex control (CTX-Ctl, N = 89), hippocampus control (HIP-Ctl, N = 6), hippocampus epilepsy (HIP-Epi, N = 10), white matter control (WM-Ctl, N = 13), white matter chronic traumatic encephalopathy (WM-CTE, N = 8); white matter multiple sclerosis (WM-MS, N = 8). (**B**) UpSet plot showing unique and overlapping up-regulated DEGs for each CNS disorder. The inset panel demonstrates interaction networks among collectively up-regulated genes in at least three out of six disorders. DEGs in at least four out of six disorders are colored in red. (**C**) A similar analysis for down-regulated DEGs across CNS disorders. (**D**) Boxplots showing the distribution of module scores in all control and disorder groups as described in (**A**). *p*-values are derived from a two-sided Wilcoxon test.

### 3.8. Collectively Up- and Down-Regulated Transcriptomic Changes across Multiple CNS Disorders

We further explored transcriptomic differences across different CNS disorders and found that transcriptional changes associated with astrocyte reactivity are highly heterogeneous in a disorder-specific manner. The number of up-regulated DEGs varies greatly across disorders: MS (1315), epilepsy (1203), PD (1082), HD (639), AD (84), and CTE (76) (Figure 5B). Their functions are annotated using GO analysis (Supplementary Figure S5A and Table S4). Among them, 157 genes were up-regulated in at least 3 out of 6 disorders (Figure 5B and Table S4). They formed an interaction network, indicating their highly connected regulations, which implied strong codependency of pathogenic processes in multiple CNS disorders through intercellular communications [69,70]. Reactive and stress signature genes formed the core of this network, including *GFAP*, *FOS*, and heat-shock genes. This network also suggested that many astrocytic genes are associated with multiple pathological processes including *MAFG, CLU, APP*, and *ITM2B/C*. Specifically, a recent study identified that *MAFG*-driven astrocytes promote CNS inflammation and repress antioxidant and anti-inflammatory transcriptional modules in MS and its preclinical model, experimental autoimmune encephalomyelitis (EAE) (Wheeler et al., 2020). Here, we further revealed that in addition to MS, HD, and PD, astrocytes also highly express *MAFG*. Given that comparatively less is known about the role of *MAFG* in other CNS diseases, we hypothesize that the same pathological mechanisms may exist in the other two diseases.

For the down-regulated genes, PD, Epi, HD, and MS had more DEGs compared to AD and CTE disorders (Figure 5C). A total of 178 genes were down-regulated in at least 3 out of 6 disorders. Some of these shared genes include *EPHB1*, *PTN*, *NRP1*, *NRXN1*, and *GABRA2*, which are involved in synapse organization, axonogenesis, and positive regulation of glial cell differentiation (Supplementary Figure S5C and Table S4). The lowered expression levels of some of these transcripts, which have significant potential protective functions in neurodegenerative diseases, might be worrisome. For example, ephrin type-B receptor 1 (*EphB1*) in a mouse ALS model and pleiotrophin (*PTN*) in MS could reduce pro-inflammatory signaling in astrocytes and promote neuronal survival following inflammatory challenge [71,72]. We found that astrocytes in multiple disorders also downregulated pathways associated with calcium channel regulator activity and blood vessel endothelial cell migration, suggesting that astrocytes may be losing important homeostatic and protective functions in disease conditions. Together, we identified both disorder-specific and collective transcriptomic changes across different CNS disorders in astrocytes and further revealed the similarities and differences among CNS disorders. These findings provide a meaningful approach and results for studying associations among diseases.

### 3.9. Gene Module Alterations across Different CNS Disorders

To explore whether there are some associations between the previously constructed modules and the diseases, we firstly calculated co-expressed module scores for the top 20 genes of each module by case status. Then, we compared the module scores between disease and control groups in the same brain region (Figure 5D). Notably, we identified both shared and disorder-specific module changes in different CNS disorders. For instance, the M2 ECM module was significantly increased in PD and AD, suggesting that the ECM performs a general function in both disease processes. Similarly, we found that the M3 synapse module of PD and Epi and the M7 cilium module of PD and MS were also elevated (Figure 5D). In addition to shared modules in multiple disorders, we identified disease-specific modules as well, such as the M4 stress module up-regulated in AD and the M5 immunity module up-regulated in epilepsy (Figure 5D). These results agree with and promote the awareness that glial cells play an important role in epilepsy, which involves inflammation, oxidative stress, and maladaptive myelination [73,74]. Compared to other diseases, we found that AD is most susceptible to the M4 stress/inflammation module. PD in the basal ganglion and AD in the cortex are both more highly expressed in the M2 ECM module, indicating that ECM is significantly associated with pathological processes of

them. In summary, we further assessed the association between diseases and co-expressed modules in astrocytes and found some functional modules played an important role in the development of the disease.

### 3.10. Validation of Astrocyte Modules in AD Proteomic Dataset

To confirm our transcriptome-based results at the protein level, we re-analyzed the tandem mass tag mass spectrometry (TMT-MS) dataset in AD [20]. We included a total of 488 dorsolateral prefrontal cortex tissues including control (N = 106), asymptomatic AD (AsymAD, N = 220), and AD (N = 182) groups (Figure 6A). We then calculated the corresponding protein abundance scores for each module and assessed correlations within modules (Figure 6B). Surprisingly, there was a stronger correlation between some modules at the protein level compared to the RNA level, such as M2/M4 (r = 0.44) and M1/M6 modules pairs (r = 0.4). These results again demonstrated that ECM and inflammation indeed are functionally linked. Interestingly, we found some modules showed an opposite relationship, such as M2/M5, M2/M7, and M5/M6, suggesting that significant differences between protein and RNA co-expression exist [20,75].
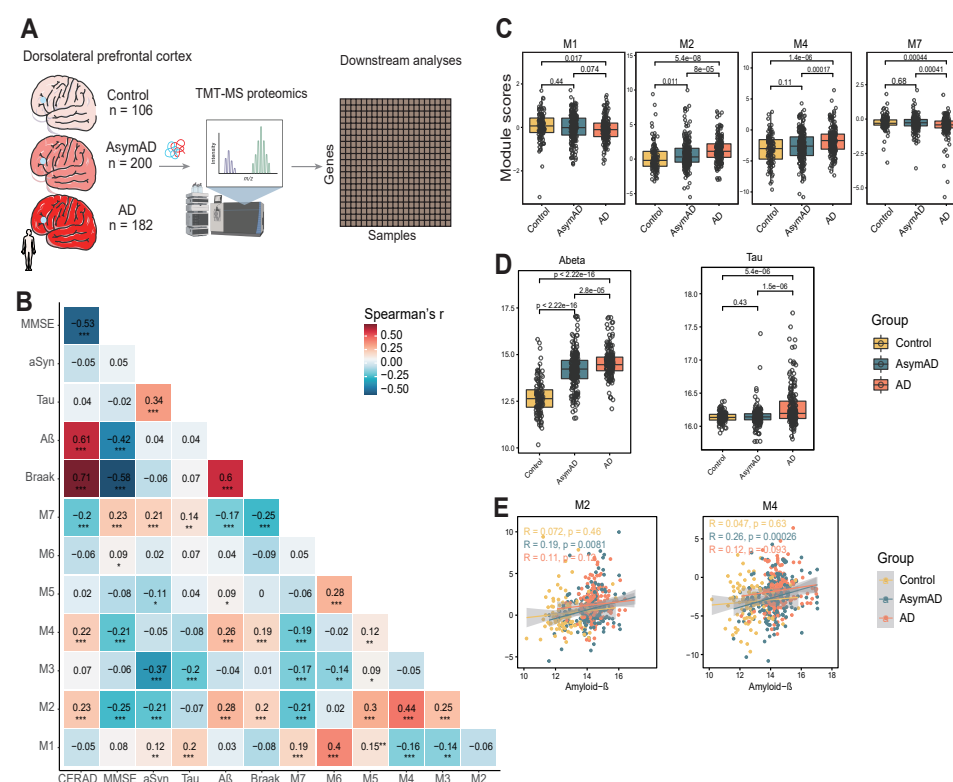


**Figure 6.** Astrocyte transcriptomic modules were validated at the protein level. (**A**) Workflow of TMT-MS-based proteomics on dorsolateral prefrontal cortex (DLPFC) tissues from control (N = 106), asymptomatic AD (AsymAD, N = 200), and AD human brains (N = 182). (**B**) Heatmap shows module relatedness, and modules were correlated with neuropathological and cognitive traits. Correlation coefficients were asserted based on Spearman correlation (red, positive correlation; blue, negative correlation). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Amyloid plaque load (CERAD score), tau neurofibrillary tangle burden (Braak stage), cognitive function (Mini-Mental Status Examination, MMSE), and mass spectrometry molecular measurements (Aβ; tau, the tau microtubule-binding region; αSyn). (**C**) Boxplots showing module score levels by case status. *p*-values are derived from a two-sided Wilcoxon test. (**D**) Boxplots showing Aβ and tau levels by case status. *p*-values are derived from a two-sided Wilcoxon test. (**E**) Aβ levels were correlated with M2 and M4 module scores by case status. Correlation coefficients were asserted based on Spearman correlation, colored by case status.

The module scores were correlated with amyloid plaque load (CERAD score), tau neurofibrillary tangle burden (Braak stage), cognitive function (Mini-Mental Status Examination, MMSE), and mass spectrometry molecular measurements (Aβ; tau, the tau microtubule-binding region; αSyn) (Figure 6B). We observed that the M2 and M4 modules showed the strongest AD trait correlations (CERAD score and Aβ r > 0.2, MMSE r < −0.2). We also found some pathology-specific modules. For example, the M1 ion homeostasis and M7 cilium modules showed a passive correlation with tau and αSyn proteins. However, the correlation between M7 and Aβ was opposed.

We associated our astrocyte modules with diagnostic classifications of AD (Figure 6C). In general, most modules were increased or decreased consistently with AD progressions in the AsymAD group, indicating that these modules could reflect pathophysiologic processes that begin early in the preclinical phase of AD. Notably, the M2 ECM module was more significantly increased in AsymAD compared to control. The accumulation of β-amyloid occurred in early stages of AD (Figure 6D), which was consistent with previous reports [76]. To further investigate susceptibility to the early stages of AD, we correlated each module score to neuropathological traits by case status (Figure 6E). We found that M2 and M4 showed the positive correlation with Aβ only in AsymAD, which suggested that the two network modules could sensitively perceive the pathological changes of the disease in the early stages of AD (Figure 6E). Overall, we again validated these co-expression modules here at the protein level and screened for key modules, such as the astrocytic M2 ECM module, that may have early diagnostic value for AD.

### 3.11. Integrated scRNA-Seq Data Show Alterations in Cell–Cell Communications in Different Disorders

Astrocytes intensively interact with surrounding cell types through signaling crosstalks, which are significantly perturbed in diseases. We explored the changes in intercellular signaling in different diseases with CellChat [37]. We first calculated the number of inferred interactions and interaction strength among all cell subpopulations and found significant differences in cellular communication among different disorders (Figure 7A). We used eight common cell populations among different disorders to perform functional similarity analysis. As the result, the signaling pathways associated with inferred networks from different disorders were mapped onto a shared two-dimensional manifold and clustered into groups. We identified four pathway groups (Figure 7B). Groups 1, 2, and 4 were dominated by growth factor pathways such as PDGF, NEGT, NRG, VEGF, and ANGPTL, while group 3 dominantly contained immune-related pathways such as ncWNT, CD45, ICAM, and MHC-II. We then compared relative information flows for each signaling pathway among different disorders and found significant disease-specific cellular communications (Figure 7C). For example, three signaling pathways were highly active in HD, which were involved in inflammatory and immune responses, such as OPIOID, CD226, and TAC. Ten pathways were specifically active in PD, including known immune signals such as TRAIL, ITGB2, ncWNT, MHC-II, ICAM, and COMPLEMENT. Interestingly, the neurotrophin (NT) signaling pathway was especially enriched in epilepsy, which includes signal molecules such as brain-derived neurotrophic factor (BDNF) and neurotrophin-3 (NT3). Previous studies have reported that BDNF is increased in animal models and humans with epilepsy and involved in epileptogenesis [77,78]. Additionally, some pathways were shared in controls and many disorders, such as NGL, CNTN, and CADM, suggesting that these pathways are essential for the maintenance of normal physiological functions and likely do not critically contribute to disease pathogenesis.

To further investigate the critical disease-related cellular communication in which astrocytes are involved, we compared the significant changed ligand-receptor pairs among different disorders and target cell types (Figure 7D). Specific to NRG and NRXN signaling, we identified two ligand-receptor pairs, NRG3-ERBB4 and NRXN1-NLGN1, as the most significant signaling in epilepsy, contributing to the communication from astrocytes to OPC, Oligo, InN, and ExN. Previous studies reported that NRG3 by non-neuronal cells

interacts with ERBB4 in PV interneurons and can promote the formation and maturation of excitatory synapses in the hippocampus [79,80]. We found that NRG3-ERBB4 signaling from astrocytes to OPC, Oligo, and InN widely participated in epilepsy (Figure 7D). In contrast, this pathway had a relatively weak interaction with other diseases, suggesting significant differences in cellular communication among diseases. Together, we found most alterations in cell–cell communications are specific to disorders, identified key signaling changes in epilepsy, and uncovered that astrocytes might actively participate in disease pathogenesis through the perturbation of cellular interactions.



**Figure 7.** Identification of CNS disorder-specific cellular signaling changes. (**A**) Barplot of the number of inferred interactions and interaction strength among all cell subpopulations in different disorders. (**B**) Jointly projecting and clustering signaling pathways from different disorders into a shared two-dimensional manifold according to their functional similarity. Different symbols represent the signaling networks from different disorders. Each symbol represents the communication network of one signaling pathway. The symbol size is proportional to the total communication probability. Different colors represent different groups of signaling pathways. (**C**) Barplot of all significant signaling pathways showing their relative information flow across disorders. (**D**) Dotplot comparing the significant ligand-receptor pairs among different CNS disorders, which contribute to the signaling from astrocytes to other cell types. Dot color reflects communication probabilities and dot size represents computed *p*-values. Empty space means the communication probability is zero. *p*-values are computed from a one-sided permutation test.

## 4. Discussion

Widespread molecular and functional diversity are recognized to exist in astrocytes across different brain regions and CNS disorders. The development of scRNA-seq enables a high-resolution view of astrocyte heterogeneity, and many studies demonstrated the molecular features of astrocytes in various diseases [1,81–83]. For example, Burda et al. and Endo et al. mainly explored astrocyte reactivity in multiple disorders and astrocyte regional diversity based on a large number of mouse models, respectively [82,83]. Especially, Endo et al. identified genes and pathways related to astrocyte morphological complexity that have a significant association with common CNS disorders [83]. Those studies have dramatically changed our previous superficial perception of astrocyte diversity. However, the comprehensive landscape of context-dependent astrocyte reactivity on the human brain is lacking, mainly due to specimen issues, which limit our general understanding of astrocyte transcriptional changes [11]. Here, we integrate the largest human brain scRNA-seq dataset to date, which includes six main neurological diseases. We profiled the astrocytes from three levels, subtype compositions, regulatory modules, and cell–cell communications, which comprehensively reveals the heterogeneity of pathological astrocytes.

We leverage the integrated dataset to better detect novel subpopulations and regulatory modules that could originally not be discovered in small datasets. For example, we identified cluster 3, OLIGO2$^+$ astrocyte subpopulations located mainly in white matter. The new subtype highly expressed oligodendrocyte-associated genes. Although the existence of this subtype is controversial in many studies, we have confirmed the possibility of its expression of oligodendrocyte-associated genes from our aspect [5,11]. Through joint analysis with the spatial transcriptomics datasets, we hypothesized that this astrocyte subtype may play an important role in the formation of myelin [47,48]. We also identified an astrocyte subtype, cluster 4, specifically located around the ventricles. This astrocyte subtype may have complex interactions with molecules in the cerebrospinal fluid [10,51]. Single-cell transcriptomics combined with spatial transcriptomics can better complement the spatial information of astrocytes, which can help us further understand the function of astrocytes. We have revealed the spatial specificity of astrocytes, therefore, incorporating the brain region as a key intervention factor of astrocytes is expected to be a promising research direction for treating CNS disorders in the future.

Since we found that the conventional astrocyte composition analysis did not fully explain the occurrence of some diseases, we believe that explaining disease occurrence solely from the perspective of cell subtype abundance or functional modules is insufficient, and a combination of both is necessary. Some diseases may be related to multiple subpopulations, and it is unclear whether these subpopulations collectively contribute to disease occurrence. These factors cannot be explained by a single subpopulation alone, and we must examine them from the perspective of multi-subpopulation functional co-clustering modules. Therefore, we applied NMF analysis to construct seven functional specific co-expression gene modules. The M2 ECM and M4 stress modules showed evident signs of co-regulation, which generated a complex network of interactions. Furthermore, we validated that these two modules were significantly elevated in AD at both the transcriptome and protein levels. Notably, the M2 module was significantly elevated in the early stages of AD onset, which could furnish potential markers for AD early diagnosis. For example, M2 hub protein CD44 has been reported to have the potential as a fluid biomarker for the early diagnosis of AD by a recent proteomic study [84].

These astrocyte transcriptomic modules recapitulate the core astrocyte features, which facilitate elucidating diseases mechanisms from a biological perspective. We identified aberration of the M3 synapse and M5 immunity in epilepsy, which are consistent with recent studies [85,86]. The M3 synapse module, which represents the main function of astrocytes, maintains the basic activity of neurons and is also involved in the development of various diseases. A deeper understanding of this pathological process would be very helpful for future research and development of treatments. We also identified the M4 module, which was differentially expressed only in AD. Most importantly, we found, in

addition to AD, the M2 ECM module was also elevated in PD. Moreover, our distillation of collectively up- and down-regulated genes among multiple disorders suggests that some critical genes are widely involved in many diseases, such as *MAFG*, *CD44*, and *DCLK1*.

Finally, we performed cellular communication analysis using integrated data and identified disease-specific signaling pathways that could contribute to disease pathogenesis. HD, PD, and Epi disorders correspond to highly disease-specific pathways. Particularly, the NRG3-ERBB4 signaling pathway from astrocytes to InN was significantly elevated in epilepsy. Therefore, we speculate that this signaling pathway of astrocytes may act as a protective function in epileptic pathology and NRG3 may be a new target for anticonvulsant drugs in epilepsy [79,80]. Future studies on the NRG3-ERBB4 pathway in astrocytes and InN could provide richer information on the mechanism of epileptogenesis, which would be an interesting line of research.

In summary, we integrated published datasets and identified novel and putative functional populations of astrocytes. Future functional studies are needed to evaluate these subpopulations and their potential effectiveness for modulation by therapies. How heterogeneity of astrocytes arises when subjected to different pathogenic stimuli is still an open question. We also hope that continued integration of our current data with future scRNA-seq and spatial datasets will add greater insights into solving these problems and making new discoveries.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/biom13040692/s1, Supplementary Figure S1: Single-cell RNA-seq dataset quality assessment. (A) Violin plots of gene counts (left panel) and UMI counts (right panel) after quality control filtering. (B) UMAP of integrated data, by dataset of origin. (C) Bar chart summarizing the fraction of different cell types in each dataset source. Supplementary Figure S2: Single-cell RNA-seq dataset quality assessment of astrocytes. (A) Violin plots of gene counts (left panel) and UMI counts (right panel) of astrocyte subtypes. (B) The expression of oligodendrocyte-associated gene (*MBP* and *PLP1*) and tanycyte markers (*COL23A1* and *SLC16A2*) in astrocytes. (C) The correlation analysis between the total MBP expression of each sample's astrocytes (left panel) and the total *MBP* expression of astrocyte cluster 3 (right panel) and microglia. Each point represents a sample. Correlation coefficients were asserted based on Pearson correlation. Supplementary Figure S3: The unique GO terms of each astrocyte subpopulation. Barplot of the unique GO terms of each astrocyte subtype with the total number of genes involved in each pathway labeled. Supplementary Figure S4: Relative abundance of astrocyte subpopulations across different CNS disorders. Boxplots showing the relative abundance of astrocyte subpopulations in all control and disorder groups as described in (Figure 5A). *p*-values are derived from a two-sided Wilcoxon test. Supplementary Figure S5: Collective and disorder-specific Go terms across different CNS disorders. (A) Boxplots showing the top 10 up-regulated GO terms per disorder. (B) Boxplots showing the top 10 down-regulated GO terms per disorder. (C) GO terms of genes shared by at least three out of six diseases. Red, up-regulated pathways; blue, down-regulated pathways; The dotted line represents *p* < 0.05. Numbers on top of the barplots represent the total number of genes in each pathway. Table S1: Metadata of published datasets used in this study. Table S2: Astrocyte cluster-specific DEGs and GO annotations. Table S3: Gene lists for seven astrocyte expression modules. Table S4: Common and disorder-specific DEGs and Go terms across different CNS disorders.

**Author Contributions:** Conceptualization, C.Z. and X.Z.; methodology, X.Z. and Z.Q.; software, X.Z. and Z.Q. and J.Q.; validation, J.Q. and Y.L.; formal analysis, X.Z. and Z.Q.; investigation, X.Z. and Z.Q.; resources, C.Z. and X.Z.; data curation, X.Z. and Z.Q.; writing—original draft preparation, Z.Q.; writing—review and editing, C.Z. and X.Z.; visualization, Z.Q.; supervision, C.Z.; project administration, C.Z.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The links and accession ID for public datasets used in this study are listed in Supplementary Table S1.

**Conflicts of Interest:** The authors declare no conflict of interests.

## References

1. Lee, H.-G.; Wheeler, M.A.; Quintana, F.J. Function and Therapeutic Value of Astrocytes in Neurological Diseases. *Nat. Rev. Drug Discov.* **2022**, *21*, 339–358. [CrossRef] [PubMed]
2. Liddelow, S.A.; Barres, B.A. Reactive Astrocytes: Production, Function, and Therapeutic Potential. *Immunity* **2017**, *46*, 957–967. [CrossRef] [PubMed]
3. Escartin, C.; Galea, E.; Lakatos, A.; O'Callaghan, J.P.; Petzold, G.C.; Serrano-Pozo, A.; Steinhäuser, C.; Volterra, A.; Carmignoto, G.; Agarwal, A.; et al. Reactive Astrocyte Nomenclature, Definitions, and Future Directions. *Nat. Neurosci.* **2021**, *24*, 312–325. [CrossRef] [PubMed]
4. Linnerbauer, M.; Wheeler, M.A.; Quintana, F.J. Astrocyte Crosstalk in CNS Inflammation. *Neuron* **2020**, *108*, 608–622. [CrossRef]
5. Lee, S.-H.; Rezzonico, M.G.; Friedman, B.A.; Huntley, M.H.; Meilandt, W.J.; Pandey, S.; Chen, Y.-J.J.; Easton, A.; Modrusan, Z.; Hansen, D.V.; et al. TREM2-Independent Oligodendrocyte, Astrocyte, and T Cell Responses to Tau and Amyloid Pathology in Mouse Models of Alzheimer Disease. *Cell Rep.* **2021**, *37*, 110158. [CrossRef]
6. Araque, A.; Parpura, V.; Sanzgiri, R.P.; Haydon, P.G. Tripartite Synapses: Glia, the Unacknowledged Partner. *Trends Neurosci.* **1999**, *22*, 208–215. [CrossRef]
7. Brandebura, A.N.; Paumier, A.; Onur, T.S.; Allen, N.J. Astrocyte Contribution to Dysfunction, Risk and Progression in Neurodegenerative Disorders. *Nat. Rev. Neurosci.* **2023**, *24*, 23–39. [CrossRef]
8. Terry, R.D.; Masliah, E.; Salmon, D.P.; Butters, N.; DeTeresa, R.; Hill, R.; Hansen, L.A.; Katzman, R. Physical Basis of Cognitive Alterations in Alzheimer's Disease: Synapse Loss Is the Major Correlate of Cognitive Impairment. *Ann. Neurol.* **1991**, *30*, 572–580. [CrossRef]
9. Wheeler, M.A.; Clark, I.C.; Tjon, E.C.; Li, Z.; Zandee, S.E.J.; Couturier, C.P.; Watson, B.R.; Scalisi, G.; Alkwai, S.; Rothhammer, V.; et al. MAFG-Driven Astrocytes Promote CNS Inflammation. *Nature* **2020**, *578*, 593–599. [CrossRef]
10. Hasel, P.; Rose, I.V.L.; Sadick, J.S.; Kim, R.D.; Liddelow, S.A. Neuroinflammatory Astrocyte Subtypes in the Mouse Brain. *Nat. Neurosci.* **2021**, *24*, 1475–1487. [CrossRef]
11. Sadick, J.S.; O'Dea, M.R.; Hasel, P.; Dykstra, T.; Faustin, A.; Liddelow, S.A. Astrocytes and Oligodendrocytes Undergo Subtype-Specific Transcriptional Changes in Alzheimer's Disease. *Neuron* **2022**, *110*, 1788–1805.e10. [CrossRef]
12. Smith, A.M.; Davey, K.; Tsartsalis, S.; Khozoie, C.; Fancy, N.; Tang, S.S.; Liaptsi, E.; Weinert, M.; McGarry, A.; Muirhead, R.C.J.; et al. Diverse Human Astrocyte and Microglial Transcriptional Responses to Alzheimer's Pathology. *Acta Neuropathol.* **2022**, *143*, 75–91. [CrossRef]
13. Mathys, H.; Davila-Velderrain, J.; Peng, Z.; Gao, F.; Mohammadi, S.; Young, J.Z.; Menon, M.; He, L.; Abdurrob, F.; Jiang, X.; et al. Single-Cell Transcriptomic Analysis of Alzheimer's Disease. *Nature* **2019**, *570*, 332–337. [CrossRef]
14. Chancellor, K.B.; Chancellor, S.E.; Duke-Cohan, J.E.; Huber, B.R.; Stein, T.D.; Alvarez, V.E.; Okaty, B.W.; Dymecki, S.M.; McKee, A.C. Altered Oligodendroglia and Astroglia in Chronic Traumatic Encephalopathy. *Acta Neuropathol.* **2021**, *142*, 295–321. [CrossRef]
15. McKellar, D.W.; Walter, L.D.; Song, L.T.; Mantri, M.; Wang, M.F.Z.; De Vlaminck, I.; Cosgrove, B.D. Large-Scale Integration of Single-Cell Transcriptomic Data Captures Transitional Progenitor States in Mouse Skeletal Muscle Regeneration. *Commun. Biol.* **2021**, *4*, 1280. [CrossRef]
16. Ruiz-Moreno, C.; Salas, S.M.; Samuelsson, E.; Brandner, S.; Kranendonk, M.E.G.; Nilsson, M.; Stunnenberg, H.G. Harmonized Single-Cell Landscape, Intercellular Crosstalk and Tumor Architecture of Glioblastoma. *bioRxiv* **2022**. [CrossRef]
17. Argelaguet, R.; Cuomo, A.S.E.; Stegle, O.; Marioni, J.C. Computational Principles and Challenges in Single-Cell Data Integration. *Nat. Biotechnol.* **2021**, *39*, 1202–1215. [CrossRef]
18. Luecken, M.D.; Büttner, M.; Chaichoompu, K.; Danese, A.; Interlandi, M.; Mueller, M.F.; Strobl, D.C.; Zappia, L.; Dugas, M.; Colomé-Tatché, M.; et al. Benchmarking Atlas-Level Data Integration in Single-Cell Genomics. *Nat. Methods* **2022**, *19*, 41–50. [CrossRef]
19. Korsunsky, I.; Millard, N.; Fan, J.; Slowikowski, K.; Zhang, F.; Wei, K.; Baglaenko, Y.; Brenner, M.; Loh, P.-R.; Raychaudhuri, S. Fast, Sensitive and Accurate Integration of Single-Cell Data with Harmony. *Nat. Methods* **2019**, *16*, 1289–1296. [CrossRef]
20. Johnson, E.C.B.; Carter, E.K.; Dammer, E.B.; Duong, D.M.; Gerasimov, E.S.; Liu, Y.; Liu, J.; Betarbet, R.; Ping, L.; Yin, L.; et al. Large-Scale Deep Multi-Layer Analysis of Alzheimer's Disease Brain Reveals Strong Proteomic Disease-Related Changes Not Observed at the RNA Level. *Nat. Neurosci.* **2022**, *25*, 213–225. [CrossRef]
21. Butler, A.; Hoffman, P.; Smibert, P.; Papalexi, E.; Satija, R. Integrating Single-Cell Transcriptomic Data across Different Conditions, Technologies, and Species. *Nat. Biotechnol.* **2018**, *36*, 411–420. [CrossRef] [PubMed]
22. Lau, S.-F.; Cao, H.; Fu, A.K.Y.; Ip, N.Y. Single-Nucleus Transcriptome Analysis Reveals Dysregulation of Angiogenic Endothelial Cells and Neuroprotective Glia in Alzheimer's Disease. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 25800–25809. [CrossRef] [PubMed]

23. Franjic, D.; Skarica, M.; Ma, S.; Arellano, J.I.; Tebbenkamp, A.T.N.; Choi, J.; Xu, C.; Li, Q.; Morozov, Y.M.; Andrijevic, D.; et al. Transcriptomic Taxonomy and Neurogenic Trajectories of Adult Human, Macaque, and Pig Hippocampal and Entorhinal Cells. *Neuron* **2022**, *110*, 452–469.e14. [CrossRef] [PubMed]

24. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters. *Omics J. Integr. Biol.* **2012**, *16*, 284–287. [CrossRef]

25. Zamanian, J.L.; Xu, L.; Foo, L.C.; Nouri, N.; Zhou, L.; Giffard, R.G.; Barres, B.A. Genomic Analysis of Reactive Astrogliosis. *J. Neurosci. Off. J. Soc. Neurosci.* **2012**, *32*, 6391–6410. [CrossRef]

26. Liddelow, S.A.; Guttenplan, K.A.; Clarke, L.E.; Bennett, F.C.; Bohlen, C.J.; Schirmer, L.; Bennett, M.L.; Münch, A.E.; Chung, W.-S.; Peterson, T.C.; et al. Neurotoxic Reactive Astrocytes Are Induced by Activated Microglia. *Nature* **2017**, *541*, 481–487. [CrossRef]

27. Diaz-Castro, B.; Bernstein, A.M.; Coppola, G.; Sofroniew, M.V.; Khakh, B.S. Molecular and Functional Properties of Cortical Astrocytes during Peripherally Induced Neuroinflammation. *Cell Rep.* **2021**, *36*, 109508. [CrossRef]

28. Leifer, D.; Krainc, D.; Yu, Y.T.; Mcdermott, J.; Breitbart, R.E.; Heng, J.; Neve, R.L.; Kosofsky, B.; Nadal-Ginard, B.; Lipton, S.A. MEF2C, a MADS/MEF2-Family Transcription Factor Expressed in a Laminar Distribution in Cerebral Cortex. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1546–1550. [CrossRef]

29. Sousa, J.C.; Cardoso, I.; Marques, F.; Saraiva, M.J.; Palha, J.A. Transthyretin and Alzheimer's Disease: Where in the Brain? *Neurobiol. Aging* **2007**, *28*, 713–718. [CrossRef]

30. Liberzon, A.; Birger, C.; Thorvaldsdóttir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) Hallmark Gene Set Collection. *Cell Syst.* **2015**, *1*, 417–425. [CrossRef]

31. Harrington, A.J.; Raissi, A.; Rajkovich, K.; Berto, S.; Kumar, J.; Molinaro, G.; Raduazzo, J.; Guo, Y.; Loerwald, K.; Konopka, G.; et al. MEF2C Regulates Cortical Inhibitory and Excitatory Synapses and Behaviors Relevant to Neurodevelopmental Disorders. *eLife* **2016**, *5*, e20059. [CrossRef] [PubMed]

32. Morel, L.; Chiang, M.S.R.; Higashimori, H.; Shoneye, T.; Iyer, L.K.; Yelick, J.; Tai, A.; Yang, Y. Molecular and Functional Properties of Regional Astrocytes in the Adult Brain. *J. Neurosci.* **2017**, *37*, 8706–8717. [CrossRef] [PubMed]

33. Durinck, S.; Spellman, P.T.; Birney, E.; Huber, W. Mapping Identifiers for the Integration of Genomic Datasets with the R/Bioconductor Package BiomaRt. *Nat. Protoc.* **2009**, *4*, 1184–1191. [CrossRef] [PubMed]

34. Hänzelmann, S.; Castelo, R.; Guinney, J. GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinform.* **2013**, *14*, 7. [CrossRef]

35. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]

36. Conway, J.R.; Lex, A.; Gehlenborg, N. UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties. *Bioinforma. Oxf. Engl.* **2017**, *33*, 2938–2940. [CrossRef]

37. Jin, S.; Guerrero-Juarez, C.F.; Zhang, L.; Chang, I.; Ramos, R.; Kuan, C.-H.; Myung, P.; Plikus, M.V.; Nie, Q. Inference and Analysis of Cell-Cell Communication Using CellChat. *Nat. Commun.* **2021**, *12*, 1088. [CrossRef]

38. Kanekiyo, T.; Xu, H.; Bu, G. ApoE and Aβ in Alzheimer's Disease: Accidental Encounters or Partners? *Neuron* **2014**, *81*, 740–754. [CrossRef]

39. Roussotte, F.F.; Gutman, B.A.; Madsen, S.K.; Colby, J.B.; Thompson, P.M.; for the Alzheimer's Disease Neuroimaging Initiative. Combined Effects of Alzheimer Risk Variants in the CLU and ApoE Genes on Ventricular Expansion Patterns in the Elderly. *J. Neurosci.* **2014**, *34*, 6537–6545. [CrossRef]

40. Wojtas, A.M.; Sens, J.P.; Kang, S.S.; Baker, K.E.; Berry, T.J.; Kurti, A.; Daughrity, L.; Jansen-West, K.R.; Dickson, D.W.; Petrucelli, L.; et al. Astrocyte-Derived Clusterin Suppresses Amyloid Formation in Vivo. *Mol. Neurodegener.* **2020**, *15*, 71. [CrossRef]

41. Duan, R.; Wang, S.-Y.; Wei, B.; Deng, Y.; Fu, X.-X.; Gong, P.-Y.; Yan, E.; Sun, X.-J.; Cao, H.-M.; Shi, J.-Q.; et al. Angiotensin-(1–7) Analogue AVE0991 Modulates Astrocyte-Mediated Neuroinflammation via LncRNA SNHG14/MiR-223-3p/NLRP3 Pathway and Offers Neuroprotection in a Transgenic Mouse Model of Alzheimer's Disease. *J. Inflamm. Res.* **2021**, *14*, 7007–7019. [CrossRef]

42. Chung, W.-S.; Clarke, L.E.; Wang, G.X.; Stafford, B.K.; Sher, A.; Chakraborty, C.; Joung, J.; Foo, L.C.; Thompson, A.; Chen, C.; et al. Astrocytes Mediate Synapse Elimination through MEGF10 and MERTK Pathways. *Nature* **2013**, *504*, 394–400. [CrossRef]

43. Lee, J.-H.; Kim, J.; Noh, S.; Lee, H.; Lee, S.Y.; Mun, J.Y.; Park, H.; Chung, W.-S. Astrocytes Phagocytose Adult Hippocampal Synapses for Circuit Homeostasis. *Nature* **2021**, *590*, 612–617. [CrossRef]

44. Matsumoto, T.; Imagama, S.; Hirano, K.; Ohgomori, T.; Natori, T.; Kobayashi, K.; Muramoto, A.; Ishiguro, N.; Kadomatsu, K. CD44 Expression in Astrocytes and Microglia Is Associated with ALS Progression in a Mouse Model. *Neurosci. Lett.* **2012**, *520*, 115–120. [CrossRef]

45. Neal, M.L.; Boyle, A.M.; Budge, K.M.; Safadi, F.F.; Richardson, J.R. The Glycoprotein GPNMB Attenuates Astrocyte Inflammatory Responses through the CD44 Receptor. *J. Neuroinflamm.* **2018**, *15*, 73. [CrossRef]

46. Diaz-Castro, B.; Gangwani, M.R.; Yu, X.; Coppola, G.; Khakh, B.S. Astrocyte Molecular Signatures in Huntington's Disease. *Sci. Transl. Med.* **2019**, *11*, eaaw8546. [CrossRef]

47. Cai, J.; Chen, Y.; Cai, W.-H.; Hurlock, E.C.; Wu, H.; Kernie, S.G.; Parada, L.F.; Lu, Q.R. A Crucial Role for Olig2 in White Matter Astrocyte Development. *Development* **2007**, *134*, 1887–1899. [CrossRef]

48. Tatsumi, K.; Isonishi, A.; Yamasaki, M.; Kawabe, Y.; Morita-Takemura, S.; Nakahara, K.; Terada, Y.; Shinjo, T.; Okuda, H.; Tanaka, T.; et al. Olig2-Lineage Astrocytes: A Distinct Subtype of Astrocytes That Differs from GFAP Astrocytes. *Front. Neuroanat.* **2018**, *12*, 8. [CrossRef]

49. Bayraktar, O.A.; Bartels, T.; Holmqvist, S.; Kleshchevnikov, V.; Martirosyan, A.; Polioudakis, D.; Ben Haim, L.; Young, A.M.H.; Batiuk, M.Y.; Prakash, K.; et al. Astrocyte Layers in the Mammalian Cerebral Cortex Revealed by a Single-Cell in Situ Transcriptomic Map. *Nat. Neurosci.* **2020**, *23*, 500–509. [CrossRef]

50. Clark, E.A.; Rutlin, M.; Capano, L.; Aviles, S.; Saadon, J.R.; Taneja, P.; Zhang, Q.; Bullis, J.B.; Lauer, T.; Myers, E.; et al. Cortical RORβ Is Required for Layer 4 Transcriptional Identity and Barrel Integrity. Available online: https://elifesciences.org/articles/52370 (accessed on 20 August 2022).

51. Mizrak, D.; Levitin, H.M.; Delgado, A.C.; Crotet, V.; Yuan, J.; Chaker, Z.; Silva-Vargas, V.; Sims, P.A.; Doetsch, F. Single-Cell Analysis of Regional Differences in Adult V-SVZ Neural Stem Cell Lineages. *Cell Rep.* **2019**, *26*, 394–406.e5. [CrossRef]

52. Chen, R.; Wu, X.; Jiang, L.; Zhang, Y. Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. *Cell Rep.* **2017**, *18*, 3227–3241. [CrossRef] [PubMed]

53. Wu, Y.E.; Pan, L.; Zuo, Y.; Li, X.; Hong, W. Detecting Activated Cell Populations Using Single-Cell RNA-Seq. *Neuron* **2017**, *96*, 313–329.e6. [CrossRef] [PubMed]

54. Falcone, C.; McBride, E.L.; Hopkins, W.D.; Hof, P.R.; Manger, P.R.; Sherwood, C.C.; Noctor, S.C.; Martínez-Cerdeño, V. Redefining Varicose Projection Astrocytes in Primates. *Glia* **2022**, *70*, 145–154. [CrossRef] [PubMed]

55. Oberheim, N.A.; Takano, T.; Han, X.; He, W.; Lin, J.H.C.; Wang, F.; Xu, Q.; Wyatt, J.D.; Pilcher, W.; Ojemann, J.G.; et al. Uniquely Hominid Features of Adult Human Astrocytes. *J. Neurosci. Off. J. Soc. Neurosci.* **2009**, *29*, 3276–3287. [CrossRef]

56. Habib, N.; McCabe, C.; Medina, S.; Varshavsky, M.; Kitsberg, D.; Dvir-Szternfeld, R.; Green, G.; Dionne, D.; Nguyen, L.; Marshall, J.L.; et al. Disease-Associated Astrocytes in Alzheimer's Disease and Aging. *Nat. Neurosci.* **2020**, *23*, 701–706. [CrossRef]

57. Li, X.; Zhang, J.; Li, D.; He, C.; He, K.; Xue, T.; Wan, L.; Zhang, C.; Liu, Q. Astrocytic ApoE Reprograms Neuronal Cholesterol Metabolism and Histone-Acetylation-Mediated Memory. *Neuron* **2021**, *109*, 957–970.e8. [CrossRef]

58. Koizumi, H.; Fujioka, H.; Togashi, K.; Thompson, J.; Yates, J.R.; Gleeson, J.G.; Emoto, K. DCLK1 Phosphorylates the Microtubule-Associated Protein MAP7D1 to Promote Axon Elongation in Cortical Neurons. *Dev. Neurobiol.* **2017**, *77*, 493–510. [CrossRef]

59. Bergoglio, E.; Suzuki, I.K.; Togashi, K.; Tsuji, M.; Takeuchi, S.; Koizumi, H.; Emoto, K. Spatial and Temporal Diversity of DCLK1 Isoforms in Developing Mouse Brain. *Neurosci. Res.* **2021**, *170*, 154–165. [CrossRef]

60. Liu, Z.; Liu, L.; Qi, Y.; Li, H.; Pan, S. GLIS Family Zinc Finger 3 Promoting Cell Malignant Behaviors and NF-KB Signaling in Glioma. *Brain Res.* **2021**, *1770*, 147623. [CrossRef]

61. Elliott, M.R.; Zheng, S.; Park, D.; Woodson, R.I.; Reardon, M.A.; Juncadella, I.J.; Kinchen, J.M.; Zhang, J.; Lysiak, J.J.; Ravichandran, K.S. Unexpected Requirement for ELMO1 in Clearance of Apoptotic Germ Cells in Vivo. *Nature* **2010**, *467*, 333–337. [CrossRef]

62. Xiao, Q.; Gan, Y.; Li, Y.; Fan, L.; Liu, J.; Lu, P.; Liu, J.; Chen, A.; Shu, G.; Yin, G. MEF2A Transcriptionally Upregulates the Expression of ZEB2 and CTNNB1 in Colorectal Cancer to Promote Tumor Progression. *Oncogene* **2021**, *40*, 3364–3377. [CrossRef] [PubMed]

63. Wiese, S.; Karus, M.; Faissner, A. Astrocytes as a Source for Extracellular Matrix Molecules and Cytokines. *Front. Pharmacol.* **2012**, *3*, 120. [CrossRef] [PubMed]

64. Edeleva, E.V.; Shcherbata, H.R. Stress-Induced ECM Alteration Modulates Cellular MicroRNAs That Feedback to Readjust the Extracellular Environment and Cell Behavior. *Front. Genet.* **2013**, *4*, 305. [CrossRef] [PubMed]

65. Gaudet, A.D.; Popovich, P.G. Extracellular Matrix Regulation of Inflammation in the Healthy and Injured Spinal Cord. *Exp. Neurol.* **2014**, *258*, 24–34. [CrossRef]

66. Song, I.; Dityatev, A. Crosstalk between Glia, Extracellular Matrix and Neurons. *Brain Res. Bull.* **2018**, *136*, 101–108. [CrossRef]

67. Hu, Y.; Huang, G.; Tian, J.; Qiu, J.; Jia, Y.; Feng, D.; Wei, Z.; Li, S.; Xu, F. Matrix Stiffness Changes Affect Astrocyte Phenotype in an in Vitro Injury Model. *NPG Asia Mater.* **2021**, *13*, 35. [CrossRef]

68. Zeng, H.; Huang, J.; Zhou, H.; Meilandt, W.J.; Dejanovic, B.; Zhou, Y.; Bohlen, C.J.; Lee, S.-H.; Ren, J.; Liu, A.; et al. Integrative in Situ Mapping of Single-Cell Transcriptional States and Tissue Histopathology in a Mouse Model of Alzheimer's Disease. *Nat. Neurosci.* **2023**, *26*, 430–446. [CrossRef]

69. Kim, J.; Miller, V.M.; Levites, Y.; West, K.J.; Zwizinski, C.W.; Moore, B.D.; Troendle, F.J.; Bann, M.; Verbeeck, C.; Price, R.W.; et al. BRI2 (ITM2b) Inhibits Abeta Deposition in Vivo. *J. Neurosci. Off. J. Soc. Neurosci.* **2008**, *28*, 6030–6036. [CrossRef]

70. Yao, W.; Yin, T.; Tambini, M.D.; D'Adamio, L. The Familial Dementia Gene ITM2b/BRI2 Facilitates Glutamate Transmission via Both Presynaptic and Postsynaptic Mechanisms. *Sci. Rep.* **2019**, *9*, 4862. [CrossRef]

71. Tyzack, G.E.; Hall, C.E.; Sibley, C.R.; Cymes, T.; Forostyak, S.; Carlino, G.; Meyer, I.F.; Schiavo, G.; Zhang, S.-C.; Gibbons, G.M.; et al. A Neuroprotective Astrocyte State Is Induced by Neuronal Signal EphB1 but Fails in ALS Models. *Nat. Commun.* **2017**, *8*, 1164. [CrossRef]

72. Linnerbauer, M.; Lößlein, L.; Farrenkopf, D.; Vandrey, O.; Tsaktanis, T.; Naumann, U.; Rothhammer, V. Astrocyte-Derived Pleiotrophin Mitigates Late-Stage Autoimmune CNS Inflammation. *Front. Immunol.* **2021**, *12*, 800128. [CrossRef]

73. Binder, D.K.; Steinhäuser, C. Astrocytes and Epilepsy. *Neurochem. Res.* **2021**, *46*, 2687–2695. [CrossRef]

74. Chen, R.; Xue, G.; Hölscher, C. The Role of the TNFα-Mediated Astrocyte Signaling Pathway in Epilepsy. *Acta Epileptol.* **2021**, *3*, 24. [CrossRef]

75. Kaufmann, M.; Schaupp, A.-L.; Sun, R.; Coscia, F.; Dendrou, C.A.; Cortes, A.; Kaur, G.; Evans, H.G.; Mollbrink, A.; Navarro, J.F.; et al. Identification of Early Neurodegenerative Pathways in Progressive Multiple Sclerosis. *Nat. Neurosci.* **2022**, *25*, 944–955. [CrossRef]

76. Jack, C.R.; Knopman, D.S.; Jagust, W.J.; Shaw, L.M.; Aisen, P.S.; Weiner, M.W.; Petersen, R.C.; Trojanowski, J.Q. Hypothetical Model of Dynamic Biomarkers of the Alzheimer's Pathological Cascade. *Lancet Neurol.* **2010**, *9*, 119–128. [CrossRef]

77. Li, K.-X.; Lu, Y.-M.; Xu, Z.-H.; Zhang, J.; Zhu, J.-M.; Zhang, J.-M.; Cao, S.-X.; Chen, X.-J.; Chen, Z.; Luo, J.-H.; et al. Neuregulin 1 Regulates Excitability of Fast-Spiking Neurons through Kv1.1 and Acts in Epilepsy. *Nat. Neurosci.* **2012**, *15*, 267–273. [CrossRef]

78. Iughetti, L.; Lucaccioni, L.; Fugetto, F.; Predieri, B.; Berardi, A.; Ferrari, F. Brain-Derived Neurotrophic Factor and Epilepsy: A Systematic Review. *Neuropeptides* **2018**, *72*, 23–29. [CrossRef]

79. Müller, T.; Braud, S.; Jüttner, R.; Voigt, B.C.; Paulick, K.; Sheean, M.E.; Klisch, C.; Gueneykaya, D.; Rathjen, F.G.; Geiger, J.R.; et al. Neuregulin 3 Promotes Excitatory Synapse Formation on Hippocampal Interneurons. *EMBO J.* **2018**, *37*, e98858. [CrossRef]

80. Zhou, L.; Fisher, M.L.; Cole, R.D.; Gould, T.J.; Parikh, V.; Ortinski, P.I.; Turner, J.R. Neuregulin 3 Signaling Mediates Nicotine-Dependent Synaptic Plasticity in the Orbitofrontal Cortex and Cognition. *Neuropsychopharmacology* **2018**, *43*, 1343–1354. [CrossRef]

81. Giovannoni, F.; Quintana, F.J. The Role of Astrocytes in CNS Inflammation. *Trends Immunol.* **2020**, *41*, 805–819. [CrossRef] [PubMed]

82. Burda, J.E.; O'Shea, T.M.; Ao, Y.; Suresh, K.B.; Wang, S.; Bernstein, A.M.; Chandra, A.; Deverasetty, S.; Kawaguchi, R.; Kim, J.H.; et al. Divergent Transcriptional Regulation of Astrocyte Reactivity across Disorders. *Nature* **2022**, *606*, 557–564. [CrossRef] [PubMed]

83. Endo, F.; Kasai, A.; Soto, J.S.; Yu, X.; Qu, Z.; Hashimoto, H.; Gradinaru, V.; Kawaguchi, R.; Khakh, B.S. Molecular Basis of Astrocyte Diversity and Morphology across the CNS in Health and Disease. *Science* **2022**, *378*, eadc9020. [CrossRef] [PubMed]

84. Johnson, E.C.B.; Dammer, E.B.; Duong, D.M.; Ping, L.; Zhou, M.; Yin, L.; Higginbotham, L.A.; Guajardo, A.; White, B.; Troncoso, J.C.; et al. Large-Scale Proteomic Analysis of Alzheimer's Disease Brain and Cerebrospinal Fluid Reveals Early Changes in Energy Metabolism Associated with Microglia and Astrocyte Activation. *Nat. Med.* **2020**, *26*, 769–780. [CrossRef]

85. Knowles, J.K.; Xu, H.; Soane, C.; Batra, A.; Saucedo, T.; Frost, E.; Tam, L.T.; Fraga, D.; Ni, L.; Villar, K.; et al. Maladaptive Myelination Promotes Generalized Epilepsy Progression. *Nat. Neurosci.* **2022**, *25*, 596–606. [CrossRef]

86. Kumar, P.; Lim, A.; Hazirah, S.N.; Chua, C.J.H.; Ngoh, A.; Poh, S.L.; Yeo, T.H.; Lim, J.; Ling, S.; Sutamam, N.B.; et al. Single-Cell Transcriptomics and Surface Epitope Detection in Human Brain Epileptic Lesions Identifies pro-Inflammatory Signaling. *Nat. Neurosci.* **2022**, *25*, 956–966. [CrossRef]