medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Multi-transcriptomic analysis points to early organelle dysfunction in human astrocytes in Alzheimer's disease

Elena Galea[1,2,3*], Laura D. Weinstock[5], Raquel Larramona[1,2,4], Alyssa F. Pybus[5], Lydia Giménez-Llort[1,6], Carole Escartin[7], Levi B. Wood[5,8*]

1. Institut de Neurociències, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain.
2. Departament de Bioquímica, Unitat de Bioquímica, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain.
3. ICREA, 08010 Barcelona, Spain.
4. Current address: Celltec-UB, Departament de Biologia Cel.lular, Fisiologia i Immunologia, Universitat de Barcelona, 08028; Institut de Neurociències, Universitat de Barcelona, 08035, Barcelona, Spain.
5. Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, 30332 USA.
6. Departament de Psiquiatria i Medicina Forense, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain.
7. Université Paris-Saclay, CEA, CNRS, MIRCen, Laboratoire des Maladies Neurodégénératives, 92265, Fontenay-aux-Roses, France.
8. George W. Woodruff School of Mechanical Engineering and Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, 30332 USA.

***Corresponding authors**

levi.wood@me.gatech.edu; +1 (617) 388 9950
Elena.Galea@uab.es; +34 935 868 143.

Abstract: 150 words
Main text: 8051 words

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

INTRODUCTION: Our understanding of the impact of astrocytes in Alzheimer's disease (AD) is hindered by limited astrocyte-specific data from AD and mild cognitive impairment (MCI).
METHODS: An astrocytic gene cluster was generated from RNAseq data of isolated healthy human brain cells using a cell-type enrichment score and clustering. The astrocytic cluster was localized in whole-brain transcriptomes from three independent cohorts totaling 766 individuals with MCI or AD, and controls. Changes in astrocytic gene expression were analyzed by gene set and principal component analyses.
RESULTS: Individuals with the same clinical diagnosis, including controls, were molecularly heterogeneous. Astrocytes in MCI and AD showed downregulation of 'mitochondria' and 'endomembrane system', and upregulation of 'stress responses', 'plasticity', and 'perisynaptic astrocyte processes'/'gliotransmission'.
DISCUSSION: Astrocytes undergo a profound transcriptional change in MCI and AD, affecting organelles and astrocyte-neuron interactions. We posit that therapies preventing organelle dysfunction in astrocytes may protect neural circuits in AD.

Key words: Alzheimer's disease, astrocytes, hierarchical clustering, MCI, mitochondria, perisynaptic astrocyte processes, RNA seq, gene enrichment scores, transcriptomics.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1. Background

'Reactive' GFAP-overexpressing astrocytes are found in the vicinity of amyloid-β (Aβ) plaques in *postmortem* brains from patients with Alzheimer's disease (AD), in both the autosomal-dominant Alzheimer's disease (ADAD) and the sporadic late-onset (AD) variants, as well as in mouse models of ADAD. The robust morphological transformation of astrocytes points to major phenotypical and, hence, functional alterations in astrocytes in AD (reviewed in [1]). However, which functions are altered in astrocytes in sporadic AD patients is unknown, in part owing to the scarcity of human astrocyte-specific omics data permitting simultaneous interrogation of numerous pathways. Transcriptomic analyses in laser-microdissected astrocytes [2, 3], systems-biology-based identification of cell-specific clusters in whole-brain AD transcriptomes [8], and recent single-nucleus RNA sequencing (snRNAseq) analyses [4, 5] have been used to generate omics data from human AD astrocytes. Laser-microdissected astrocytes either presented insufficient number of differentially-expressed genes (DEG) for pathway analysis [3], or normalization anomalies, as suggested by the finding that 98% of the DEGs were down-regulated [2]. The problem may lie in the low RNA yields of laser microdissection, exacerbated by the poor quality of RNA after long *post-mortem* intervals. As an alternative to laser microdissection, clustering statistics based on gene co-expression identified function-specific gene modules using 1674 whole-brain AD transcriptomes [6]. Cellular localization of the modules was established using cell-specific markers, and causality between nodes was then inferred with Bayesian statistics. Using this approach, alterations of glutamate and amino acid metabolism in an astrocytic module of 262 genes were discovered to be well-ranked for causal relevance in AD, but no further changes were unraveled. Finally, although snRNAseq improves the cellular resolution of transcriptomics, the numbers of astrocytes being sequenced and of DEG being detected (69 in the astrocytic pool in [5]) are too low to enable a complete understanding of astrocyte pathology in AD. Thus, while large-scale clustering analyses [6] and snRNAseq [4, 5] may serve well to discover global gene networks associated with pathological traits, they might lack statistical power and sensitivity to detect relevant yet under-represented gene modules from low-abundance cells such as astrocytes, which account for 5-10% of the cells in the human neocortex [7]. Moreover, the low total number of subjects (48 in [5] and 12 in [4]) precludes the appraisal of patient heterogeneity [8, 9]. Importantly, the scarcity of human transcriptomic data poses the additional problem that there is no guarantee that mouse models of ADAD based on over-expression of mutant forms of Aβ, presenilin and/or Tau [10-12]), recapitulate the changes of astrocytes in sporadic AD.

Here, we improve on the cellular resolution of large scale analyses [6], as well as on the transcriptomic depth achieved by snRNAseq [4, 5], by taking a reverse approach: instead of determining *a posteriori* which self-organized gene modules might be astrocytic as in these studies, we localized pre-determined astrocytic gene clusters in whole-brain AD transcriptomes (workflow in **Suppl. Fig. 1**). First, we classified brain genes as cell-specific or non-specific using RNAseq data from astrocytes, neurons, microglia, endothelial cells and oligodendrocytes isolated from aged healthy human brains [13], and a combination of two algorithms: hierarchical clustering, and a recently proposed univariate cell-type enrichment τ scoring [14]. Second, astrocytic and, for comparison, neuronal genes were annotated according to function. Third, the gene clusters were localized in three AD whole-brain transcriptomes generated by the Mount Sinai Hospital (MtSINAI) [6], the Religious Order Study and the Rush Memory and Aging project (ROSMAP), which also contains subjects diagnosed with MCI [15], and the Mayo Clinic (MAYO) [16]. These databases present the advantages of including large

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

cohorts (78-219 subjects), which greatly facilitates the use of multivariate systems approaches, and of having been generated by two different approaches, microarrays and RNAseq, thus diminishing technique-associated bias. Finally, alteration of astrocytic and neuronal functional categories in AD and MCI groups versus controls was statistically established using gene set variation analysis (GSVA), and principal component analysis (PCA). We present the most comprehensive transcriptomic analysis of human astrocytes to date using three independent cohorts totaling 766 individuals.

2. Methods

2.1. Transcriptome datasets

The cell-type specific RNAseq datasets generated from healthy human brains [13] were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE73721 pre-aligned to gene symbols by the authors using human genome version 19. The database consisted of transcriptomic data from brain cells isolated from temporal cortical lobes of juvenile (8-18 years old) and adult (21-63 years old) non-demented individuals. Transcriptomes were included from 12 astrocytic, one neuronal, five oligodendrocytic, three microglial, and two endothelial samples from individual patients (**Suppl. data 1, tab 2**). Principal component analysis (PCA, computed in R, R Foundation for Statistical Computing, Vienna, Austria) revealed that transcriptomic data of astrocytes isolated from young (<47 years old) and old (>47 years old) were separated along principal component 1 (PC1, **Suppl. data 1, tab 3**). Since the goal of the study was to identify cell-specific signatures in aged subjects, data from young individuals were discarded, retaining a total of six astrocyte transcriptomes from 47–63-year-old subjects, together with all 11 transcriptomes from the remaining cell types. A threshold was set for gene expression such that a given transcript should present a count > 1 in reads per kilo base per million mapped reads (RPKM) in at least one sample. Genes with no counts > 1 were discarded from the analysis. 11077 genes were above this threshold (**Suppl. data 1, tab 4**).

The MtSINAI database containing microarray data of samples from the prefrontal cortex of 101 controls and 129 AD patients obtained from the Harvard Brain Tissue Resource Center was downloaded from GEO under accession number GSE44770 [6]. The Rosetta gene identifiers were converted to corresponding gene symbols using the Rosetta/Merck Human 44k 1.1 microarray platform table under accession number GPL4372. When multiple Rosetta identifiers mapped to the same gene, the identifier with the greatest variance across samples was selected. The MAYO database, which contains data from the temporal cortex of 78 control samples and 82 AD samples, was downloaded from Synapse.org under Synapse ID syn3163039 [16]. The data were downloaded as MayoRNAseq_RNAseq_TCXCounts_normalized.tsv, pre-aligned to Ensemble Gene Identifier. Ensemble Gene IDs were then converted to HGCN gene symbol using `biomaRt` v. 2.32.1 in R (r-project.org). When multiple Ensemble Gene IDs mapped to the same gene symbol, the one with greatest variance was selected. The ROSMAP database [15] contains data from the dorsolateral prefrontal cortex obtained from autopsied non-demented individuals or patients diagnosed with AD. The data were downloaded from Synapse.org under synapse ID syn8456629 as ROSMAP_DLPFC_netResidualExpression.tsv. The dataset contained 200 control cases, 157 MCI cases, and 219 AD cases. In the MCI group, there was no stratification into MCI and MCI-AD. The dataset as posted by the authors had been adjusted via voom (mean-variance modelling at the observational level) normalization to remove bias associated with batch number, RNA integrity, and

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

other data acquisition and processing co-variates. Missing transcript values in the dataset were imputed using a nearest neighbor averaging method using the `impute` package in R.

2.2. *Univariate cell-type enrichment score.*

To determine the cell-specificity of human genes listed in the brain-cell transcriptomics database [13], we used the τ method, reported to be particularly robust compared with other scoring approaches to identify tissue specificity of a gene among different magnitudes of expression and sizes of datasets [14]. The τ enrichment score was computed as:

$$\tau = \frac{\sum_{i=1}^{n}\left(1-\hat{x}_i\right)}{n-1}; \hat{x}_i = \frac{x_i}{\max_{1\le i\le n}\left(x_i\right)}$$

where x_i is the mean expression of the gene within i^{th} cell type and n is the number of cell types. We then define the cell type enriched in the gene as that with the greatest average expression among all cell types. Top τ scores per gene are in **Suppl. data 1**, tab 5. A score of $\tau \geq 0.8$ indicates that the gene is specific for a given cell type, and a score of $0.8 > \tau \geq 0.6$ that the gene is enriched in that cell type. For simplicity, henceforth we refer to genes with $\tau \geq 0.6$ as 'enriched', although note that the 'enriched' pool also includes specific genes. Finally, $\tau < 0.6$ indicates that the gene is non-specific, although it is assigned to the cell type with the greatest average expression.

2.3. *Hierarchical clustering*

The cell type specific data from [13] (**Suppl. data 1,** tab 4) were z-scored for each gene across samples. Gene and sample hierarchical clustering were conducted using a Euclidian distance measure, and the average agglomeration method in R. Gene clusters associated with neurons and astrocytes were determined by cutting the dendrogram at a height of 3.76 to achieve that at least 70% of the genes in the neuron and astrocyte clusters have the highest τ for that cell type (results from hierarchical clustering are in **Suppl. data 1**, tabs 5-7). Corresponding cell-specific gene clusters in whole tissue data sets (MtSINAI, MAYO, and ROSMAP) were identified by sorting the genes of the whole-tissue data according to the cell type specific dataset. Whole-tissue subject samples (columns) were then clustered within each Control, MCI, or AD cohort (Euclidian distance and Ward.D2 agglomeration) to identify molecularly-defined sub-groups of samples.

2.4. *Data normalization within each cell type*

A recognized caveat of bulk transcriptomic analyses is that DEGs may reflect changes in cell-type composition rather than real changes in cell-specific transcriptomes. This is relevant to AD because there is neuronal death [17], and microglia proliferation [18]. Since neurons are the most predominant cell type in the brain (40% in the human neocortex [7]), a decrease of the neuronal pool, with no change in the astrocyte pool [18], arguably results in the artificial up-regulation of astrocyte genes. Several approaches have been employed to correct for cellular composition, such as use of several of the so-called "high fidelity" genes primarily expressed by one cell type [19]. Here, we reasoned that reliance on a subset of highly cell-specific genes could lead to artifacts should their expression change

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in the course of the disease, because there is no assurance that they are stably expressed as house-keeping genes. Therefore, we elected to remove gross variance by normalizing gene expression within the neuronal and astrocyte clusters according to the shifts in distribution of differential expressions of all genes. This was accomplished by variance stabilizing normalization using `limma` in R, which simultaneously corrected each dataset for dependence of variance on the mean. Removal of gross variance is widely-used in normalizations of transcriptomic data, based on the premise that, since only a negligible proportion of genes have their expression changed following perturbations [20], a bulk shift in the gene-expression distribution in a given cell population reflects changes in the proportion of that cell type among brain cell populations. Thus, correcting gene expression according to the shift would unmask the real changes in gene expression. We applied this correction separately to the neuronal and astrocytic gene clusters.

2.5. Cell function annotation and gene set variation analysis

Neuronal and astrocytic gene clusters were generated by combining τ [14] and hierarchical clustering of transcriptomes from cells from healthy brains [13] were organized in functional categories and subcategories (**Suppl. data 2**). Categories were manually curated by one expert and cross-validated by another expert using information from GeneCards, perusal of Medline and PubMed, and open-source platforms such as GO, KEGG, and Reactome. Manual curation was indispensable for two reasons. First, over 60% of the genes in the astrocytic cluster were not annotated in open-source platforms. This means a substantial loss of usable gene data if we only relied on genes annotated in open-source platforms. Second, annotation in such platforms represents 'canonical' pathways with little consideration for the fact that different CNS cell types are molecularly and functionally distinct [13]. This limitation is critical for astrocytes. For example, unlike neuronal compartments such as 'synapses', 'dendrites' and 'spines', a highly specialized astrocytic compartment termed 'perisynaptic astrocyte processes' (PAP), concentrating RNA transcripts related to glutamate and GABA metabolism, energy metabolism, as well as the ribosomal machinery to perform local RNA translation [21], is not a category in current KEGG and GO databases. The genes were organized in general categories and subcategories using a mixed set of criteria including general pathways (e.g. carbohydrate metabolism or amino acid metabolism) and subcellular compartments with a clear functional specialization (e.g. mitochondria, peroxisome or lysosome). Such flexibility allowed us to optimize gene inclusion, while generating gene sets with sufficient numbers of genes for statistical purposes. Although most genes were assigned to only one category, around 100 genes were assigned to two categories (in blue, **Supp. Table 2**, tab 'astrocyte cluster 4'). The majority of those genes (96) were PAP genes, because we reasoned that their presence in PAP does not exclude relevance in other subcellular compartments. For example, ALDOC and ACO2 are, plausibly, ubiquitous enzymes relevant for general glycolysis and mitochondrial TCA throughout the cell. Likewise, genes related to 'TNFalpha signaling' are in 'stress responses'/'cytokines' as well as in 'gliotransmission', for TNFalpha has been described to modulate glutamate exocytosis [22]. Overlap between groups was less than 5% with the exception of 'PAP and gliotransmission': 17.8% genes in 'gliotransmission' were in 'PAP', and 10.3% of PAP genes were in 'gliotransmission'. For neurons, each gene was assigned to only one category.

Using our novel annotation database (**Suppl. data 2**), the `GSVA` package in R was used to identify enrichment of each gene set across all samples. GSVA is a generalized gene set enrichment method that detects variations of pathway activity over a sample population in an unsupervised manner

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[23]. Statistical differences in enrichment scores for each gene set between subject groups were computed by comparing the true differences in means against a null distribution obtained by permuting the gene labels and re-computing the GSVA 1000 times. False discovery rate adjust p-values (qFDR) were computed using the method of Benjamini & Hochberg [23]. Gene sets with FDR adjusted p value qFDR < 0.05 were considered significant.

2.6. Relationships between clinical stages and GSVA scores

The ROSMAP dataset includes Braak and CERAD (Consortium to Establish a Registry for Alzheimer's disease) pathological scores, APOE genotype, and sex for each sample. To determine the relationship between these phenotypes and the GSVA enrichment scores for each sample, we used the `stats` R package to fit linear models between enrichment scores and either Braak or CERAD for each gene set. The MAYO and MtSINAI databases do not include additional information per subject other than the clinical classification.

3. Results

3.1. Hierarchical clustering identifies groups of cell-type specific genes.

To identify astrocyte-specific genes, we calculated the metric τ [14] for every gene in the transcriptomes from isolated brain cell types (**Suppl. data 1,** tab 5 'τ scores'). Cell type-specific ($\tau \geq 0.8$) and cell type-enriched ($\tau \geq 0.6$) genes represented 30% and 60% of the 11,077 genes, respectively. Astrocytes (8%), microglia (7.6%) and neurons (7.2%) present the largest percentage of cell-specific genes, whereas neurons (21%) contained the largest pool of cell-enriched genes, followed by microglia (13.4%) and astrocytes (9.4%). In contrast, endothelial cells and oligodendrocytes presented the lowest percentages of cell-enriched (6%) and specific (3%) genes.

In parallel, because cell-specific functions are arguably performed by co-expressed genes, we reasoned that cell-type specific genes would be identified by clustering analysis. Thus, we applied hierarchical clustering (**Methods**) to the cell-type specific dataset, generating 194 clusters that contained between 1-1651 genes (**Fig. 1A**; **Suppl. data 1,** tab 6 'Hierarchical clustering' and tab 7 'Cluster description'). We found that genes from the same cell type segregated together, defining clusters of highly co-expressed genes, plausibly representing cell-specific functions. The dendrogram height of hierarchical clustering was set to identify gene clusters that were highly enriched for individual brain cell types. The criteria for 'enrichment' was that the cluster contained more than 100 genes, of which over 70% corresponded to a unique cell type with $\tau \geq 0.6$ (**Suppl. data 1**, tab 8 'cluster selection'). Five clusters fulfilled these criteria, corresponding to astrocytes (cluster # 4), endothelial cells (cluster # 5), neurons (cluster # 13), microglia (cluster # 17) and oligodendrocytes (cluster # 164). There were other clusters containing over 77% of total of neuronal (e.g., clusters # 10, 116) or microglial (e.g., clusters # 30, 15) genes (**Suppl. data 1**, tab 8 'cluster selection'), but they were excluded from the analysis because less than 54% of their genes was enriched ($\tau \geq 0.6$) for that cell type, suggesting lower cell type specificity.

We then focused on astrocyte cluster 4 (1651 genes) and neuronal cluster 13 (1258 genes) for two reasons: (i) a cluster size of over 1000 genes is sufficient for functional categorization and gene set

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

analysis and (ii) since molecular changes in neurons are better characterized in AD than changes in astrocytes, neurons served as a positive control to assure correct data mining,

In the neuronal cluster # 13, 99.9 % of the genes were neuronal according to τ score, of which 54.1% of genes were neuron-specific (τ ≥0.8), meaning that they encode for proteins that perform functions highly specific to neurons. Indeed, the top 10 ranked genes based on τ (τ > 0.98) were related to neurotransmission (*SYNPR, GABRA1, CNR1, SYT1, GABRG2* and *GAD1),* regulation of membrane potential (*KCN2*), microtubules (*INA),* and cell adhesion *(RELN*). The remaining 45.9% (τ < 0.8) were neuron-enriched or bulk genes plausibly supporting neuron-specific functions, since they were co-expressed with neuron specific genes (**Suppl. data 1**, tab 8 'cluster selection'; **Suppl. data 2**, tab 2 'Neurons cluster 13').

In astrocyte cluster # 4, 80% of the genes were astrocytic according to τ (**Suppl. data 1**, tabs 7 and 8 'Cluster description' and 'Cluster selection'; **Suppl. data 2**, tab 'Astrocytes cluster 4'). Further, 30% had τ ≥ 0.8, including genes that encode typical astrocyte proteins such as *GFAP* (τ = 0.95), glutamine synthase (*GLUL*, τ = 0.95), glutamate decarboxylase (*GLUD2*, τ = 0.91), glutamate transporters (*EAAT1/SLC1A3*, τ = 0.83), lactate dehydrogenase (*LDHB*, τ = 0.91), enzymes involved in mitochondrial fatty-acid oxidation (ACADVL, τ = 0.92). Enriched genes with τ between 0.6-0.8, or bulk genes below 0.6 necessary for astrocyte functions are the aquaporins A*QP1* (τ = 0.739) and *AQP4* (τ = 0.744), and the chaperone clusterin *CLU* (τ =0.373), which was identified by GWAS as a risk factor in AD [24]. These findings indicated that the group of genes gathered in cluster #4 are relevant to astrocyte biology.

3.2. Identification of cell-type specific gene clusters in AD databases

We next searched for the astrocytic and neuronal gene clusters in gene expression data from three AD whole-tissue transcriptomic databases (**Fig. 1B**). Unexpectedly, hierarchical clustering of subjects (columns) of each cohort revealed heterogeneity in groups with the same clinical diagnosis. Subjects were thus divided in at least two molecular groups within the Control, MCI, and AD cohorts across the three databases, as shown by the division into two large groups in each cohort in the dendrograms. We refer to *type 2* subjects as those showing blue z-scores within the neuron gene cluster, suggesting massive down-regulation of genes, as compared to *type 1* subjects, in which genes with red z-scores predominated. Type 1 sub-cohorts included Control1, MCI1 and AD1, while type 2 included Control2, MCI2 and AD2. Importantly, the most profound down-regulation of genes within the neuron gene cluster was observed in AD2. Because down-regulation of neuronal genes due to loss of synapses and neuronal demise is a hallmark of AD [17], we reasoned that AD2 was advanced AD, and Control1 a *bona fide* control composed of subjects that were neither demented, nor preclinical or pre-symptomatic at the time of death. Control1 and AD2 thus appeared to be the extreme groups across the continuum of AD phenotypes. We note that heterogeneity within an AD cohort has been recently reported [25].

Since our goal was to ascertain changes in astrocytic functions, it was critical to eliminate gross variability among samples within the astrocytic and neuronal gene clusters due to global changes in cell composition. To do so, we applied variance stabilizing normalization (**Methods**), which shifted the peaks of the distributions of differences in gene expression between AD2a and Control1 to zero (**Fig. 2A**). Prior to normalization, the distributions of gene-expression differences shifted to the left

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in the neuronal cluster, suggesting depletion of neurons in AD2, while the distribution of astrocytic genes shifted to the right, suggesting a relatively higher content of astrocytes in AD2. After normalization, the global downregulation of all neuronal genes in cluster 13 in type 2 groups was attenuated, but intra-cohort heterogeneity nevertheless persisted (**Fig. 2B** and **Suppl. Fig. 2**). In addition, new gene clusters that were previously masked were unraveled by hierarchical clustering. Thus, there were three neuronal (**Fig. 2B,** N-a, b, c), and four astrocytic gene subgroups (**Fig. 2B,** A-a, b, c, d) with opposite expressions patterns in type 1 versus type 2 subjects. For example, N-b genes were globally downregulated in AD2 and Control 2, as compared to Control 1 and AD1. Importantly, N-b contained genes related to synaptic function, including glutamatergic and GABAergic neurotransmission (e.g., *GABRB3; GABRB2; GAD1; PDYN; SYN2; GABRG2; SYN1 NAPB; GRIA2; SLC17A7; GRIK2*), supporting that this cohort is *bona fide* AD.

Because the ROSMAP database included clinical covariate data, such as CERAD and Braak scores, APOE genotype and sex, we examined whether type 1 or type 2 sub-cohorts were associated with specific covariates to gain insight into why they were molecularly heterogeneous. We found that Braak and CERAD scores were significantly higher in AD groups than in Controls (**Fig. 2C**), confirming the clinical diagnosis. Importantly, type 1 and type 2 subjects with the same clinical diagnoses did not significantly differ in their pathological stages according to CERAD and Braak scoring; nor were type 2 groups more enriched in *APOE4* subjects or females (**Fig. 2D**).

3.3. Gene set analysis reveals altered functions in AD and MCI versus Control1

We next identified changes in astrocytic and neuronal functions using the normalized expression data. Because the majority of the astrocytic genes were not annotated in existing pathways in open-source databases, we manually curated the astrocytic cluster in 17 functional categories (gene sets) and 135 subcategories, and the neuronal cluster in 15 functional categories and 79 subcategories (**Methods**, **Suppl. data 2**).

GSVA was then used to compare the gene sets corresponding to different functional categories among cohorts, yielding an enrichment score (ES) for each subject and gene set. We started by comparing the extreme cases, AD2 and Control1, using two criteria. First, probability values were computed using a permutation analysis across genes (cut-off for statistical significance was set at FDR-adjusted $q < 0.05$, **Methods**). Second, we considered that a function was altered in AD2 as compared to Control1 if such alteration was detected in at least two out of the three databases (consensus criterion).

For neurons, seven functions were altered in AD2 versus Control1 **(Fig. 3A, Suppl. Fig. 3,** statistics in **Suppl. data 3**): 'neurotransmission' and 'gene expression' were altered in the three databases, and 'synaptic plasticity', 'regulation of membrane potential', 'neural development', 'mitochondria' and 'stress response' in MAYO and ROSMAP, but not MtSINAI. 'Synaptic plasticity', 'neurotransmission', 'neural development' and 'membrane potential' were down-regulated, and 'gene expression', 'mitochondria' and 'stress response' up-regulated. Among significantly different gene sets in MAYO and ROSMAP (**Fig. 3A**), Control 2 and MCI2 resembled AD2, and AD1 and MCI1 resembled Control1 (**Fig. 3B**). In MtSINAI, we found no statistically significant difference between Control 1 and Control 2, and AD groups were similar, although changes in AD2 were more dramatic, as compared to Control1 (**Fig. 3B**, statistics in **Suppl. data 3**).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

For astrocytes, five functions were altered in AD2 versus Control 1 (**Fig. 4A**): 'PAP', 'plasticity' 'stress response' were up-regulated in three/three databases, 'mitochondria' was down-regulated in three/three databases, while 'endomembrane system' was down-regulated in two/three databases. Because 'PAP' are specialized astrocyte-neuron contacts and, hence, this category includes many genes related to gliotransmission, and since 'gliotransmission' was highly significantly changed in ROSMAP (qFDR<0.001), and trending in MAYO (qFDR=0.07), henceforth 'PAP' and 'gliotransmission' were considered together. As with neurons, Control2 and MCI2 showed the same direction of change as AD2 compared to Control1 (**Fig. 4B**, **Suppl. data 4**). Likewise, enrichment of these gene sets in AD1 resembled Control1, particularly in MAYO and ROSMAP, while AD1 appeared to be a milder version of AD2 in MtSINAI.

3.4. *Co-variance of gene sets with one another and with pathological covariates*

Correlations between altered functions and pathological stages. Since CERAD and Braak scores were available for the ROSMAP data, we examined relationships between the enrichment of each gene set and pathological stages by regressing enrichment scores for each gene set and sample against the corresponding CERAD and Braak scores, regardless of clinical diagnoses (**Methods**). Data from all the individuals in the three groups (control, MCI and AD) were used in the analysis. For neurons, we found no correlation between Braak and CERAD stages for any of the seven top dysregulated categories described in Section 3.3 (in parenthesis p values for regressions with Braak/CERAD): 'neurotransmission' (0.536/0.919), 'synaptic plasticity' (0.342/0.703), 'neural development' (0.85/0.552), 'membrane potential' (0.794/0.479), 'gene expression' (0.55/0.563), 'mitochondria' (0.357/0.254) and 'stress response' (0.548/0.755) (**Suppl. Fig. 4**). The lack of correlation can be interpreted to indicate similar gene dysregulation in all the stages, suggesting that neuronal dysfunction is an early event in AD pathology such that it appears in subjects with no (i.e., Control2) or incipient (i.e., MCI2) cognitive deficits, and low CERAD and Braak scores.

For astrocytes, of the five functions that were significantly altered in AD2 vs Control1 (**Fig. 3A**), the following ones significantly correlated with Break/CERAD: 'mitochondria' (0.0000203/0.000742), 'stress response' and 'plasticity' (0.0768/0.00967) (**Suppl. Fig. 5**). In contrast, 'PAP' (0.563/0.838), 'gliotransmission' (0.811/0.805), and 'endomembrane system' (0.709/0.942) did not correlate with pathological stages, suggesting that, as we reasoned with neurons, they represent early manifestations of alterations in astrocytic functions. For example, there is a progressive decline in the expression of mitochondrial genes from CERAD 1 to 3 and from Braak 1 to 6, while the expression of genes involved in 'endomembrane system' is altered even before the detection of cognitive deficits in patients.

Correlations among altered functions. We examined correlations among all neuronal and astrocytic gene sets by regressing enrichment scores for each gene set against each other gene set. Hierarchical clustering of correlation coefficients and adjusted p-values are in **Suppl. Fig. 6** for ROSMAP, MAYO, and MtSINAI. In neurons, the strongest direct correlations were among highly related functions such as 'membrane potential', 'neurotransmission' and 'synaptic plasticity' in neurons. Likewise, in astrocytes, the strongest correlations were between 'PAP' and 'gliotransmission' (orange squares). The strongest anti-correlations point to intertwined changes between: (i) morphological remodeling of PAP and loss of synaptic functions, and (ii) survival pathways included in 'plasticity' in astrocytes (see below) and organelle dysfunction.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3.5. Alteration of functional subcategories in astrocytes

We next aimed to gain further insight into the functional alterations in AD astrocytes by searching for functional subcategories with statistically significant differences in enrichment scores between AD2 and Control1 (qFDR< 0.05) in at least two out of three of the databases (**Suppl. data 4**).

The consensus subcategory significantly upregulated was 'PAP morphology' (qFDR <0.001, AD2 vs Control1, in MAYO and ROSMAP, **Suppl. data 4**). Because PAP are specialized in mediating astrocyte-to-neuron communication, the rearrangement of PAP may enhance local plasticity by modifying the localization and efficacy of transmitter uptake and release, thereby contributing to changes in synaptic strength [17].

Subcategories in 'Plasticity' include annotations associated with morphological remodeling (e.g., 'adherens junction', 'ECM', 'cytoskeleton', 'ciliogenesis', 'TGFbeta signaling', 'metalloproteases'), and pathways involved in brain development whose role in adult astrocytes is not well understood (e.g., 'axon outgrowth', 'synaptogenesis, 'WNT signaling', 'hippo signaling', 'smoothened signaling', 'hedgehog signaling', 'homeobox signaling', 'notch signaling', 'FGF signaling', 'pluripotency', and 'glial cell fate commitment') (**Suppl. data 4**). Significantly dysregulated pathways were 'ECM', upregulated in the three databases, and 'hippo signaling' and 'ciliogenesis' (i.e., the process of generation of a microtubule-based and centriole-derived cilium whose function in astrocytes is unknown), respectively upregulated and down-regulated in two databases—note that the fact that the general trend of a given general category is towards up-regulation does not preclude that some of its subcategories are down-regulated.

'Stress responses' includes 'AMPK signaling', 'antioxidant response', 'cell death', 'chaperone', 'DNA repair', 'hypoxia', 'complement system' and 'cytokines'. No subcategory was dysregulated according to qFDR in at least two databases. However, it is worth noting that 'TNFalpha', which is a subcategory of 'cytokines' as well as of 'gliotransmission' was significantly dysregulated in MAYO and ROSMAP, and trending in MtSINAI (qFDR=0.093). In 'endomembrane system', encompassing genes related to 'endoplasmic reticulum', 'Golgi', 'lysosome' and 'vesicles', no single subcategories were significantly dysregulated, while in 'mitochondria', 'ETC' and 'TCA' were significantly downregulated in AD2 versus Control1 in three and two databases, respectively. For 'ETC', qFDR=<0.001, <0.001, 0.012, <0.001 in MAYO, ROSMAP and MtSINAI. For 'TCA', qFDR=<0.001, 0.39, 0.01 in MAYO, ROSMAP and MtSINAI (**Suppl. data 4**). All in all, the dysregulated subcategories in astrocytes point to alterations in mitochondrial respiration concomitant to morphological remodeling.

3.6. Principal component analysis confirms main functional changes in AD astrocytes

GSVA was dependent on gene sets organized into pre-determined functions. Thus, we also used PCA as an alternative, unsupervised approach to gain independent insight into the predominant functional changes of AD astrocytes. First, we examined whether expression of genes contained in the astrocytic cluster 4 segregated cohorts in ROSMAP. We found that PCA separated the Control1 and AD1 groups to the left, and the AD2 group to the right, according to the principal component 1 (PC1, horizontal axis), while the other groups appeared in between (**Fig. 5A**). Violin plots confirmed this tendency, and further showed that the three type 1 groups segregated from the three type 2 groups. Second, we examined the functions performed by the proteins encoded by the top 50 dysregulated genes in PC1, representing the maximally co-varying astrocytic genes. Mirroring our results from GSVA, genes related to 'plasticity' and 'PAP/gliotransmission' were over-represented in the up-regulated portion,

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

whereas genes related to 'endomembrane system', 'gene expression', and 'signaling' were predominant in the down-regulated portion (**Fig. 5B, Suppl. data 6**). Genes related to 'mitochondria' and 'stress responses' appeared in both sections, indicating mixed direction of dysregulation in these functions. Unlike GSVA, PCA also detected changes in 'signaling', as suggested by the presence of genes related to this function in the up and down segments of the top PC1 genes. Third, we asked if the changes detected in AD astrocytes in ROSMAP with PCA were reproducible. Thus, we determined the capacity of the PC1 genes to segregate cohorts in MAYO and MtSINAI. As in ROSMAP, Control1 subjects were mostly negative, and AD2 cases were mostly positive in MAYO and MtSINAI, thus validating the discriminating capacity of the maximally co-varying genes in PC1 (**Fig. 5C**). AD1 segregated with Control 1 in MAYO, but not in MtSINAI, where AD1 aligned with AD2, confirming less intra-cohort heterogeneity in MtSINAI than in the other two databases, as shown with GSVA. Altogether, these results point to multi-factorial changes in AD astrocytes, encompassing changes in astrocyte-neuron interactions and organelle dysfunction (model in **Fig. 6A** and description in **Discussion**).

Finally, comparison of the top 50 UP and 50 DOWN PC1 genes with the DEG astrocytic genes identified with snRNAseq studies in AD brains [4, 5] reveals little concordance (**Suppl. data 8**). Specifically, only 0.3% and 1.3% of PC1 genes are in the DEG reported in [5] and [4], respectively, while [5] and [4] share 32 genes, representing 47% of the DEG in [5] and 4.5 % of those in [4]. The remarkable differences among the databases may be due to methodologies, classifications of disease (e.g., clinical in our study and in [4], but associated to Aβ pathology in [5] regardless of clinical diagnosis), brain regions sampled (prefrontal or temporal cortex in [5] and in our case, but entorhinal cortex in [4]), and/or the total number of subjects analyzed (48 in [5], 12 in [4], and 766 herein).

3.7. Cortical APP/PS1 astrocytes partially mimic cortical AD astrocytes

We next asked whether changes identified in cortical AD astrocytes are recapitulated in mouse models. To answer this question, we examined the expression of human astrocytic genes from cluster 4 in the transcriptome of cortical astrocytes isolated from 15-18 month-old APP/PS1 versus WT mice (N=4 per group) [10], using the same functional categorization and statistical tools utilized for human AD astrocytes. Comparison with recent snRNAseq mouse data [12], is not appropriate since these studies were carried in hippocampi. We found robust down-regulation of all human astrocytic functions in APP/PS1 cortical astrocytes (**Fig. 6B**). Thus, APP/PS1 astrocytes mimic the down-regulation in 'endomembrane system' and 'mitochondria' detected in AD astrocytes, but not the up-regulation of 'PAP', 'stress responses' and 'plasticity'. We also compared the top 50 and bottom 50 DEG from APP/PS1 astrocytes [5] with the top 50 and bottom 50 genes from PC1 in human AD astrocytes (**Fig. 5B**). GSVA confirmed that both the top and bottom human AD astrocyte gene sets were down-regulated in APP/PS1 mice (**Fig. 6C**, left). Importantly, the pool of immune-related genes reported to be up-regulated in mouse APP/PS1 astrocytes [10] are predominantly microglial according to τ scores (**Suppl. data 6**), and, with the exception of *C1S,* do not coincide with functionally related human astrocyte-enriched genes in cluster 4 subcategories such as 'cytokines' and 'complement' (**Suppl. data 7**). This divergence points to the existence of two distinct gene pools in cortical APP/PS1 astrocytes that change in opposite directions in advance disease stages: a down-regulated pool of typical astrocyte genes, also noted in [10], and an up-regulated pool of typical microglia genes. The loss of astrocytic functions and the adoption of a microglia-like phenotype suggests that cortical APP/PS1 astrocytes may undergo a phenotypical involution with time.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Finally, regardless of the direction of change in mice, DEG in APP/PS1 vs WT cortical astrocytes were largely up-regulated in AD2 the three human databases (**Fig. 6C**, right panels). This suggests predominance of typical microglial genes in advanced AD, plausibly upregulated in both microglia and astrocytes.

4. Discussion

Because of the challenges associated with isolation of astrocytes or their nuclei, and to profit from large publicly available human datasets, the goal of this study was to gain insight into what happens to astrocytes in AD using whole brain transcriptomes and bioinformatics tools. The approach consisted of re-compartmentalizing whole-brain transcriptomes from Control, MCI and AD subjects into contributions from individual cell types by using pre-clustered cell-specific gene sets. The study has four main findings. First, GSVA and PCA show that AD astrocytes—so defined because they appear in patients with the down-regulation of neuronal synaptic functions typical of AD—undergo a profound phenotypical transformation encompassing down-regulation of genes encoding for organelle proteins, and up-regulation of genes related to PAP, gliotransmission, plasticity and stress responses. Second, unbiased hierarchical clustering of patient cohorts reveals major intra-cohort heterogeneity, i.e., relationships are not robust between gene profiles, clinical diagnosis, and pathological stage according to CERAD and Braak scoring, as also reported in [25]. In other words, some non-demented subjects present AD-like astrocytes and neurons, while some demented subjects present healthy ones, particularly in MAYO and ROSMAP. Third, according to regression analyses between functions and CERAD or Braak scores, some astrocytic functions may be altered early in the AD continuum (e.g., PAP, gliotransmission and endomembrane system), while other functions may be increasingly dysregulated in parallel with the deposition of neuritic plaques and neurofibrillary tangles (e.g., mitochondrial functions, plasticity-related pathways and stress responses). Fourth, multivariate statistics and hierarchical clustering show that cortical astrocytes from human AD and from a 15–18-month-old mouse model of ADAD [5] are distinct transcriptional entities.

4.1. *Limitations*

A limitation of our study is that pathway dysregulation during AD may result in the remodeling of cell-type specific gene clusters such that gene clusters identified in healthy brain transcriptomes may no longer exist in AD. However, quantification of network reorganization in AD subjects using a metric called modular differential connectivity (MDC) showed equal or enhanced connectivity in 95.5% of the modules [6], suggesting that clusters identified by co-expression analyses in healthy cells are, for the most part, preserved in AD. Nevertheless, we acknowledge that changes in transcripts ubiquitous in all cell types, or typical of other cell types, namely microglia, may go undetected in our analysis relying on astrocyte-specific and enriched genes. That is, our results may be just the tip of the iceberg in terms of defining the totality of changes in human AD astrocytes. Pathways plausibly overlooked include the inflammation-like response common to astrocytes and microglia, as described in 15-18-month-old APP/PS1 transgenic mice [10], and ubiquitous metabolic pathways, such as glycolysis. Further, because transcriptomic data are cross-sectional and descriptive, they do not demonstrate cause-effect relationships between pathway dysregulation and pathology, or unequivocally prove that changes are adaptive (meant to compensate for dysregulated functions and maintain or restore homeostasis), maladaptive (irreversibly contributing to dyshomeostasis and neurodegeneration), or

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

epiphenomena with scarce bearing on disease onset and progression. Finally, genetic data does not directly inform about changes in proteins and metabolites. Hence, predictions based upon these data (below) need validation in appropriate models.

4.2. *Key implications and predictions*

Use of molecular data for subject stratification in clinical trials. Neurocognitive measurements, atrophy patterns assessed with neuroimaging, neurofibrillary tangles assessed *postmortem*, CSF levels of amyloid-β1-42 and Tau, as well as Tau neuroimaging, have revealed subtypes of AD [8, 9]. This clinical and biological heterogeneity, which is increasingly recognized to be a major obstacle to establishing drug efficacy in clinical trials, starts to be addressed with molecular tools, as shown by the recent detection of three molecular classes in a large cohort of AD patients using RNAseq transcriptomics of the parahippocampal gyrus [25]. Although the areas under study in [25] and in the present analysis are different, we reason that our AD2 group might correspond to their C typical class, and AD1 to their A atypical class, because both AD2 and C present downregulation of neuronal networks as compared to AD1 and A. We speculate that AD1 subjects might have been resilient to AD pathology, which they had, according to high CERAD and Braak scores, but that they developed dementia due to comorbidities (e.g., Lewy bodies, vascular disease, TDP43 deposits [26]). However, the possibility exists that the tissues used for transcriptomic analysis and neuropathological assessment of neuritic plaques and neurofibrillary tangles were not the same. This means that the transcriptomic analysis in demented AD1 subjects with a healthy-like molecular profile might have been performed in cortical tissue spared from ongoing pathology, such that AD1 subjects were not truly resilient. By contrast, regardless of how omics and neuropathological assessments were performed, Control2 subjects with an AD-like molecular profiles in ROSMAP and MAYO were cognitively normal. Thus, Control2 subjects might high-risk, but resistant subjects [24], in whom the accumulation of amyloid-β plaques and Tau tangles, as well as cognitive impairment, were somehow delayed despite the decreased expression of synaptic genes due soluble oligomeric amyloid-β [27]. Whatever the case, our study and [25] support the use of omics-based molecular phenotyping and clustering statistics as unbiased tools to stratify patients in clinical trials in the spirit of personalized medicine.

Early dysregulation of the endolysosomal system in astrocytes. The category 'endomembrane system' and genes related to endoplasmic reticulum (e.g., *FAF2, PIGX*), endocytosis (e.g., *ATXN2, MEGF8, RAB11FIP2*), and autophagy (*ATG9A*) were found to be down-regulated in AD2 astrocytes by PCA and GSVA (in parentheses, consensus genes dysregulated in at least 2 out of three databases, **Supp. Data 5**). These findings support cumulative evidence in mouse and *in vitro* models of AD pointing to inefficient phagocytosis and degradation of amyloid-β protofibrils [28-30] and dystrophic neurites [31] in astrocytes, as well as decreased exosome production [32]. Thus, our study and others do not support the notion that astrocytes efficiently clear amyloid-β plaques, as initially suggested by evidence in non-primary cultured astrocytes (discussed in [1]), which, arguably, are endowed with greater motility and morphological plasticity than primary astrocytes. Moreover, targeted enhancement of lysosomal biogenesis by overexpression of the transcription factor EB (TFEB) in astrocytes enhances Tau clearance, and reduces pathological hallmarks in the hippocampus of PS19 tauopathy mice [33], suggesting that the improvement of lysosomal function in astrocytes may also reverse Tau pathology

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in AD. Along these lines, TFEB has been reported to regulate AD risk genes in an astrocyte subpopulation detected in AD brains by snRNAseq [4].

Early and concomitant alteration of neuronal-astrocyte contacts. This prediction is based on two findings: the inverse correlation of changes in synaptic functions and PAP, and the lack of correlation between these functional categories with pathological stages, suggesting that they change early in AD. Accordingly, both PAP and synaptic genes are dysregulated in non-demented subjects with a type 2 molecular profile (Control2), which, as noted earlier, may represent preclinical AD. Specifically, GSVA showed up-regulation of 'PAP morphology', whose leading gene *EZR* encodes ezrin, a protein involved in PAP motility and in the anchoring of the astrocyte membrane to neurons or the extracellular matrix [34]. Also, genes related to integrin and cell motility (e.g., *FERMT2*, *PREX2*) were part of the top genes detected by PCA as defining AD astrocytes (**Fig. 5B**). We posit that upregulation of PAP genes is an adaptive change to improve the coverage of neurons to counteract the down-regulation of synaptic elements directly caused by oligomeric amyloid-β [27]. This scenario is supported by the upregulation of genes encoding for synaptogenic factors such as thrombospondin and glypicans (*THSD1, GPC4*) in the three databases (**Suppl. data 5**).

Mitochondrion dysfunction in AD astrocytes may impact gliotransmission. It is generally assumed that the well-established mitochondrion dysfunction [35] and reduced glucose metabolism (often referred to as 'hypometabolism' [36]) in AD are neuronal phenomena. However, our analysis reveals down-regulation of nuclear-encoded mitochondrial genes encoding TCA and ETC components in astrocytes, suggesting deficiencies in energy-generating mitochondrial pathways. For example, one of the consensus downregulated genes, *ATPAF1*, is a component of the ATP synthase (complex V), and it is known that impairment of the ETC beyond complexes I-III increases reactive-oxygen species (ROS) formation while decreasing ATP formation [37]. Moreover, i*n vitro* and *ex vivo* studies link mitochondrial dysfunction in astrocytes to amyloid-β. Thus, in co-cultures of rat hippocampal neurons and astrocytes, amyloid-β causes loss of mitochondrial potential and transient mitochondrial depolarization, concomitant to activation of NADPH oxidase and decrease glutathione production, in astrocytes, but not neurons [38]. Loss of membrane potential and depolarization of astrocytic mitochondria is prevented by antioxidants, and by the use of glutamate by mitochondrial complex I as a substrate, suggesting that amyloid-β elicited mitochondrial dysfunction is caused by oxidative stress and deficits in substrate supply [38]. Interestingly, the mitochondrial glutamate transporter *SLC25A18*, as well as the plasma membrane glutamate transporters *SLC1A3, SLC6A11* and *SLC7A11,* are up-regulated in the three databases, while glutamine synthase (*GLUL*), which encodes the enzyme that produces glutamine from glutamate, is downregulated in ROSMAP and MtSINAI. Taken together, the data suggest that AD astrocytes recycle glutamate not to produce glutamine, but to counteract the damage caused by amyloid-β to astrocytic respiration by increasing the supply of fuels amenable for oxidation. Along these lines, *HADH8,* which encodes for an enzyme involved in mitochondrial beta oxidation, was found upregulated in AD astrocytes by PCA.

Likewise, the upregulation in AD astrocytes in MAYO and ROSMAP of the monoamine oxidase *MAO-B*, a dopamine-degrading enzyme located in the outer mitochondrial membrane, may be interpreted as an attempt to reverse respiratory deficits. This is supported by the recent discovery that MAO-B increases the polarization of the inner mitochondrial membrane and ATP production by shuttling electrons through the inter-membrane space [39]. MAO-B upregulation in AD is relevant for two

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

reasons. First, the MAO-B ligand (11)C-deuterium-L-deprenyl may serve as a PET-based biomarker of early astrocyte dysfunction in AD, as documented in ADAD [40]. Second, MAO-B may affect the excitatory/inhibitory balance of neural circuits by producing GABA from putrescine, as shown in APP/PS1 mice [41]. In support of this scenario, the genes encoding the GABA transporter *SLC6A11/GAT3*, which may extrude GABA coupled to glutamate uptake, and *SLC7A2*, which imports the putrescine precursor L-arginine, are consensus dysregulated genes (**Suppl. data 5).** Taken together, the data support that impaired mitochondrial respiration in astrocytes switches gliotransmission to a GABA-dominant mode. A detrimental effect of astrocytic GABA has been proposed based on the beneficial effects of the MAO-B inhibitor selegiline on electrophysiological readouts in APP/PS1 mouse hippocampi [41]; however, there is no benefit of selegiline in patients [42], thus raising doubts as to whether APP/PS1 mice appropriately model the impact of astrocytic GABA in AD. Alternatively, astrocytic GABA might counteract the loss of GABAergic tone that has been causally related with neuronal hyperactivity, desynchronization of neural circuits, and amyloid-β production [43]. Since our regression analyses reveal that mitochondrial dysregulation in astrocytes increases in advanced pathological stages, and since the upregulation of MAO-B detected by PET in the prodromal stages in ADAD is transient [40], we reason that the two strategies used by mitochondria to preserve membrane potential and ETC (enhanced supply of substrates such as aromatic amino acids like glutamate, and of electrons *via* MAO-B) eventually fail in astrocytes. As a consequence, glutathione production might be decreased and NADPH oxidase activity increased [37], perhaps exacerbating oxidative stress and overall astrocyte malfunction in advanced AD.

Pathological-stage dependent upregulation of 'plasticity' and 'stress response' pathways might explain why, unlike neurons, astrocytes survive in AD. The progressive upregulation of the subcategory 'ECM', which includes integrins and proteoglycans, is not surprising considering the indisputable adoption of a 'reactive' morphology by astrocytes in AD, consisting of engrossment of primary and secondary processes due to over-expression of the intermediate filament protein GFAP, and increased production of ECM proteins [44]. More intriguing is the consensus up-regulation of hippo signaling, an evolutionarily conserved network with a central role in the regulation of cell proliferation, cell fate, organ growth and regeneration [45], and no reported role in healthy adult astrocytes, despite the fact that some of the genes in this pathway are highly specific to human adult astrocytes according to τ (e.g., *WWC1* τ=0.91; *WWOX* τ =0.86, and *YAP1* τ=0.88). *YAP1* was, moreover, detected by PCA as a relevant gene in the astrocytic phenotype in AD. The upregulation of hippo signaling in AD astrocytes, as well as the global upregulation of genes involved in brain development may be a manifestation of the well documented phenomenon of re-activation of developmental pathways in reactive astrocytes that some authors interpret as a failed attempt at reprogramming into neural stem cells [46], and we interpret to be a protective response to facilitate survival of defective astrocytes. In addition, the recent observation that *YAP1* regulates scar formation in spinal cord injury in mice [47], suggests that hippo signaling is coordinated with the increased production of ECM in AD astrocytes. Finally, the consensus dysregulation of 'TNF-alpha' (**Suppl. data 4**), and the consensus upregulation of *TNFRSF11B, TNFSF13 and TRAF3IP2*, members of the IL17 family (*Il17D, IL17RB, and IL17RD*), and complement factors (*CD59, C1S*) (**Suppl. data 5**) might be interpreted as detrimental 'neuroinflammation', but evidence supports modulatory or protective actions of these factors. First, TNFα may facilitate the release of glutamate by improving vesicle docking at the astrocyte plasma membrane prior to exocytosis [22], such that TNFα signaling may help adjust gliotransmission in AD astrocytes. Second, the finding that Nrf2/IL17D axis acts as an

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

antioxidant protective pathway in stress responses induced by tumorigenic stimuli and viral infections [48], supports an astroprotective role of members of the IL17 family in AD. Also, members of the complement system have been shown to facilitate amyloid-β phagocytosis [49].

4.3. Model of the phenotypical transformation of cortical astrocytes in AD

The key predictions from our analysis are summarized in **Fig. 6A.** We posit that amyloid-β drives the phenotypical change of astrocytes in AD in several stages by causing dysfunction of the endolysosomal/mitochondrial axis, which, in turn, prompts a change in gliotransmission. Control2 subjects (cognitively normal/AD-like molecular profile) would be at stage a, and AD2 patients (demented, AD molecular profile) at stage b (**Fig. 6A**). In both a and b, the phenotypical change is mixed with regards to possible clinical impact, because it includes changes that disrupt or preserve homeostasis in astrocytes themselves, and changes that disrupt or preserve homeostasis in astrocyte-neuronal communication. The model thus represents a departure from simplistic neuroprotective/neurotoxic classifications of reactive astrocytes from the point of view of neurons. As discussed in a recent consensus article [50], an 'astrocyte-centric' stance focused on established whether pathways changes are adaptive or maladaptive in the complex astrocyte biology should be adopted when studying reactive astrocytes. What is then the *net* impact of the mixed functional changes of reactive astrocytes on AD? In short, we posit that, despite organelle dysfunction, astrocytes manage in early disease stages to partially preserve their functions, including the modulation of neural circuits, through adaptive changes, while they become globally malfunctional in advanced stages. Thus, the early changes in astrocyte-neuron interplay may delay the onset of clinical symptoms, while detrimental phenomena would prevail in stage b. If these predictions are correct, therapies aimed at protecting and restoring the functions of the endolysosomal system and mitochondria to prevent astrocytes from progressing to stage b might help arrest the progression of AD. Even if GABA were detrimental, as reported in [41], preservation of mitochondrial respiration is therapeutically indicated, because increased GABA production, whatever its effect, is the result of mitochondrial impairment. In summary, the present study points to prevention of organelle dysfunction in astrocytes as a therapeutic strategy in AD, to be tested in transgenic mice or human cellular models that appropriately recapitulate the complex transformation of astrocytes in the AD continuum.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figures

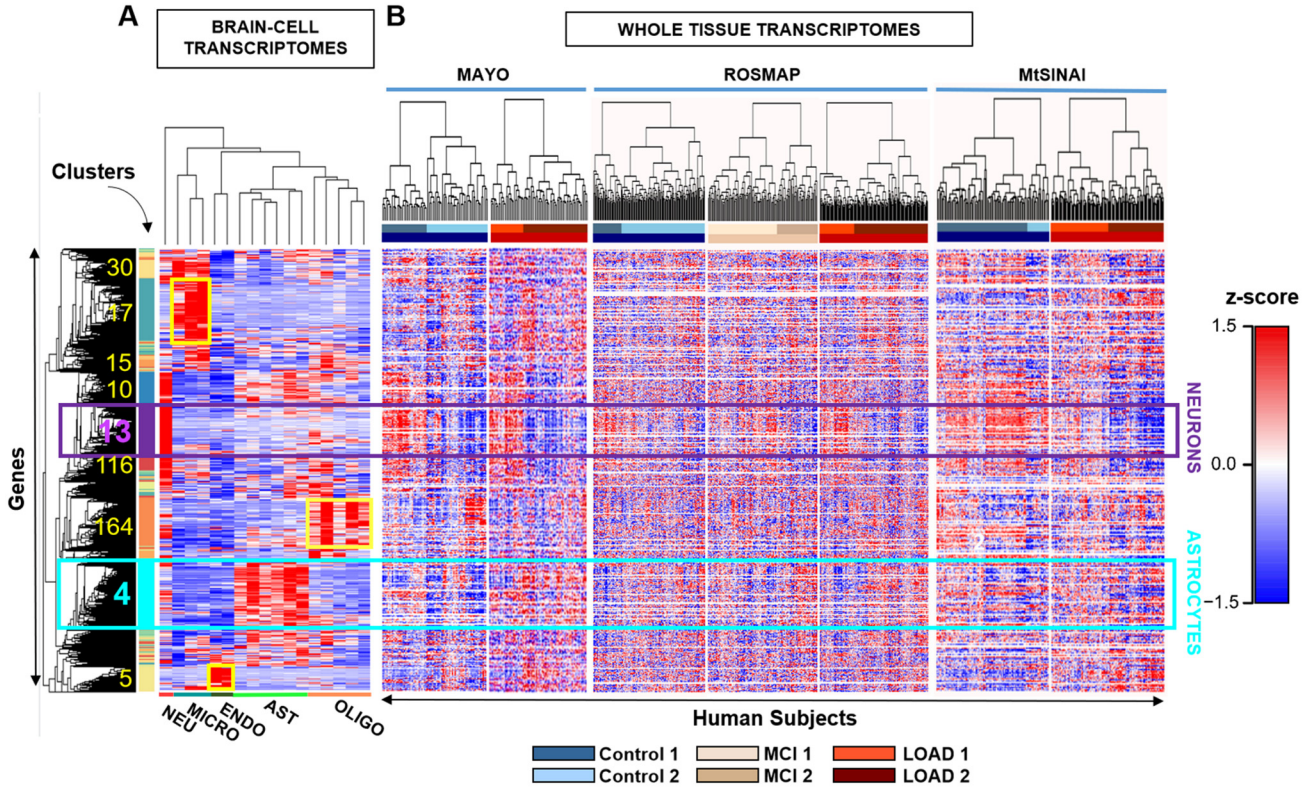


Figure 1. Identification of cell-specific gene clusters in whole-brain transcriptomes by two-step hierarchical clustering.

A. Hierarchical clustering was performed on transcriptomes of brain cell populations isolated from individual healthy subjects. Each column is a sample from a single subject/patient and each row is a z-scored gene. In yellow, the number and location of clusters with highest z scores, coinciding with highest τ scores for a given cell type (cluster information in **Suppl. data 1**, tabs 6-8). The turquoise and purple squares are the main astrocytic and neuronal clusters (cluster # 4 and cluster # 13), respectively. Microglia, oligodendrocyte and endothelial-cell clusters are 17, 164 and 5, respectively (shown in yellow). Clusters 30, 15, 10 and 116 are highly enriched in microglial or neuronal genes, but fall below the criteria established for cell-specificity (**Methods**).

B. Whole-brain transcriptomes from individual subjects of control (blue), MCI (beige) and AD (red) groups from MAYO, ROSMAP, MtSINAI were hierarchically clustered. The rows from the cell-type clustering in **A.** were kept fixed to identify cell-specific clusters. The figure was generated with raw data before correction for cell composition. Clustering of subjects segregated all cohorts into at least 2 subgroups (defined as 1 and 2) showing opposite changes in gene-expression.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

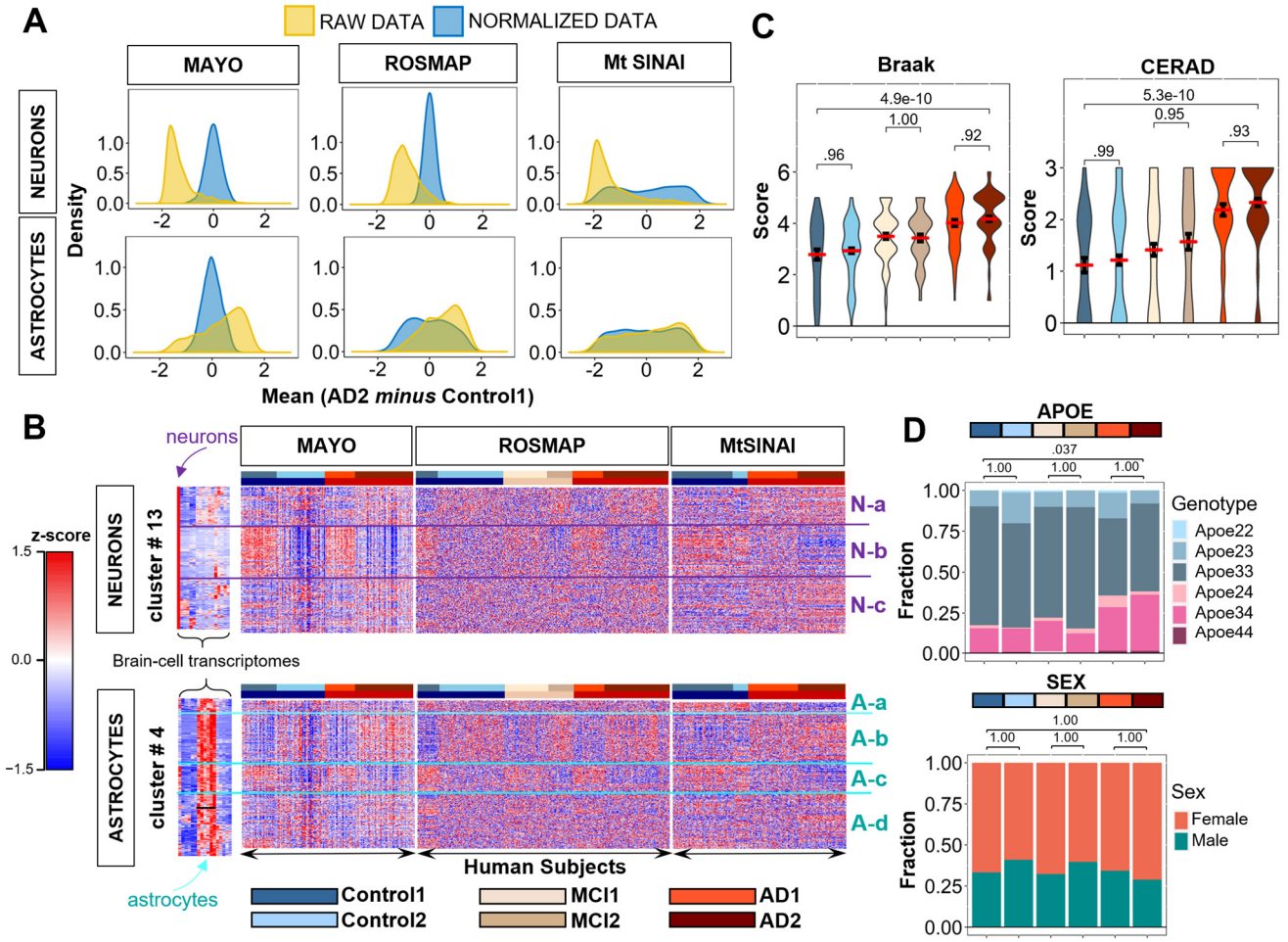


Figure 2. Subject and gene subclusters in the neuronal and astrocytic gene sets.
A. Distributions of gene expression differences between AD2 and Control1 cohorts before (yellow) and after (blue) variance stabilizing normalization, yielding zero-centered distributions (**Methods**).
B. Hierarchical clustering of Control (blue), MCI (beige), and AD (red) subjects from Fig. 1B using normalized data from Fig. 2A in rows. For both astrocytic and neuronal genes, subjects self-organized into at least two distinct subgroups (defined as 1 and 2) within each clinical group, while genes self-organized into up to four groups (Na-c and Aa-d). In MAYO and ROSMAP, clusters showed opposite directions of change in type 1 and 2 subjects. In MtSINAI, Control2 was like Control1, and AD1 like AD2, although with less pronounced changes.
C. Violin plots of pathological stages according to CERAD and Braak scoring shows no significant difference in pathological stages between sub-groups within each cohort in ROSMAP (means ± SEM, one-way ANOVA with Tukey post-hoc).
D. Distribution of *APOE* genotypes and sex in the six groups. While 3/3 was the predominant genotype across all groups, and 3/4 was more abundant in AD, there were no differences in frequency of *APOE4* alleles between type 1 and 2 subgroups within each clinical group. Likewise, females were more abundant than males in all groups, but not in type 2 subjects as compared to type 1 (p values determined by pairwise Fisher's Exact test with Bonferroni adjustment).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

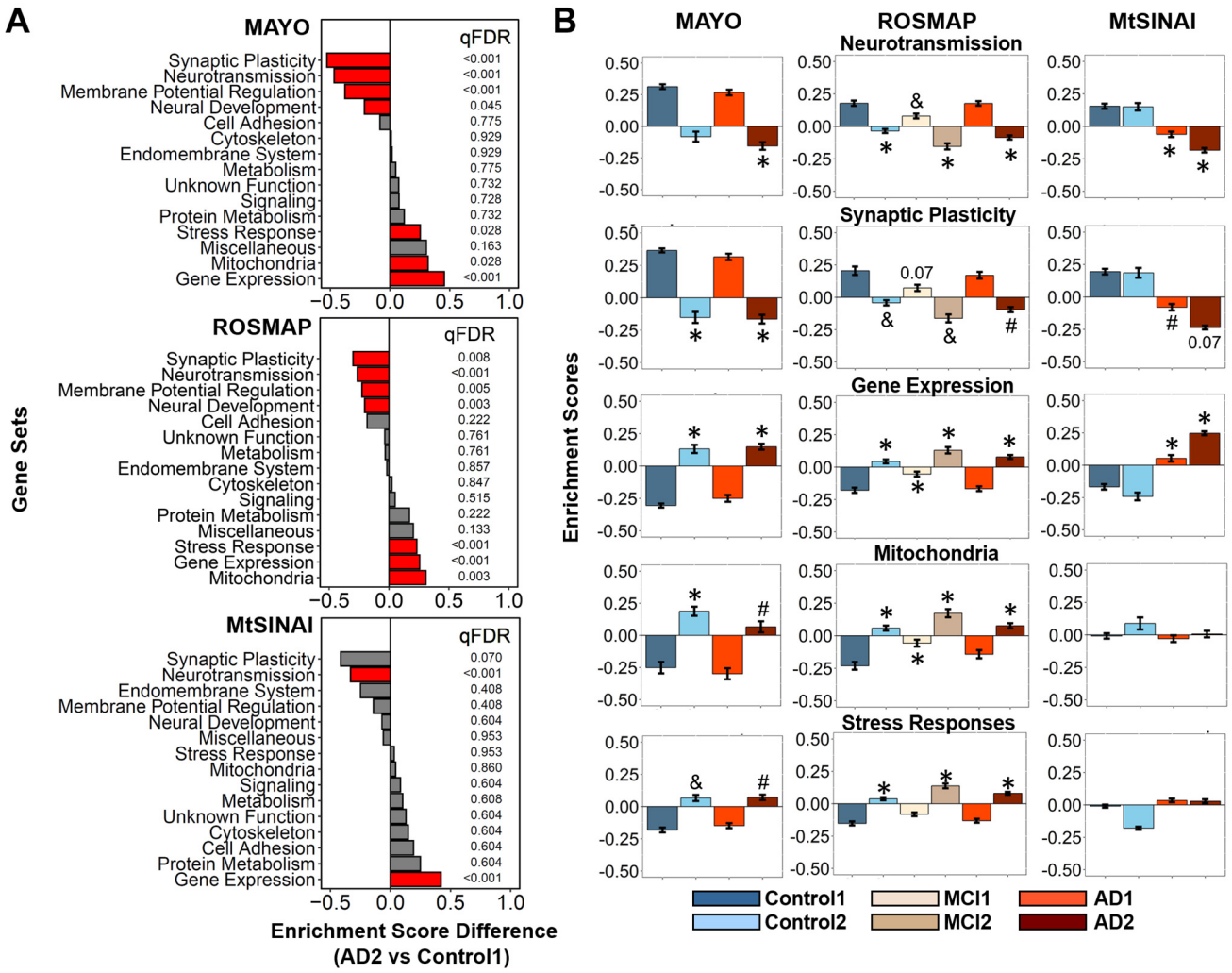
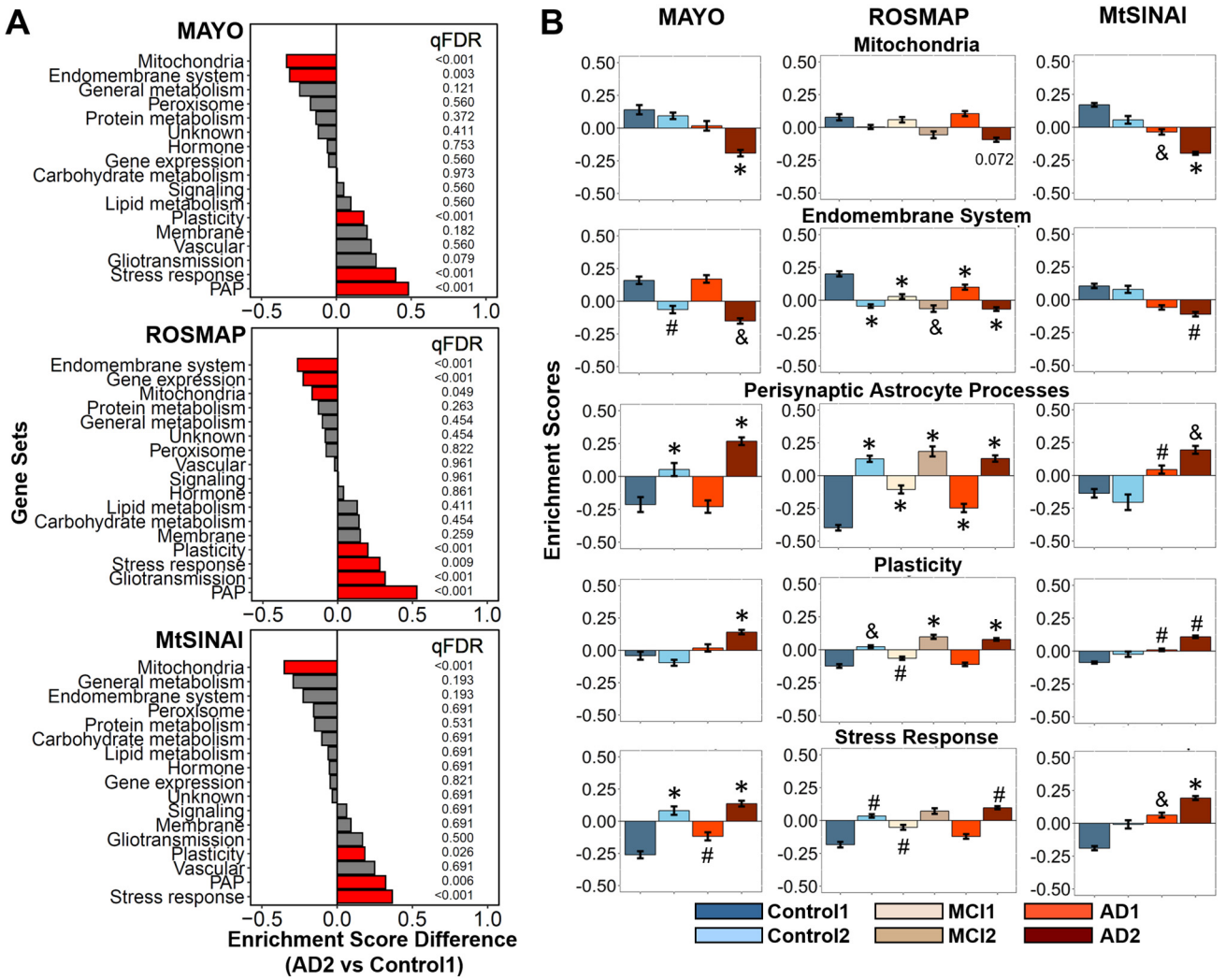


Figure 3. GSVA unravels neuron-enriched functions altered in AD. Differences in the general functional categories among cohorts in MAYO, ROSMAP and MtSINAI databases were examined by GSVA.

A. Comparison between AD2 and Control1. Functions are ranked according to the difference in enrichment scores (ES). Significantly changed functions (qFDR < 0.05) are labeled in red.

B. Enrichment scores of functional categories significantly changed in at least 2 out of 3 databases. (#) qFDR< 0.05 (&) qFDR <0.01, and (*) qFDR < 0.001, versus Control1.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



743

744 **Figure 4. GSVA unravels astrocyte-enriched functions altered in AD.** Differences in the
745 functional categories among cohorts in MAYO, ROSMAP and MtSINAI databases were examined by
746 GSVA.
747 **A.** Comparison between AD2 and Control1. Functions are ranked according to the difference in
748 enrichment scores (ES). Significantly changed functions (qFDR < 0.05) are labeled in red.
749 **B.** Enrichment scores of functional categories significantly changed in at least 2 out of 3 databases.
750 (#) qFDR < 0.05, (&) qFDR <0.01, and (*) qFDR < 0.001, versus Control1.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

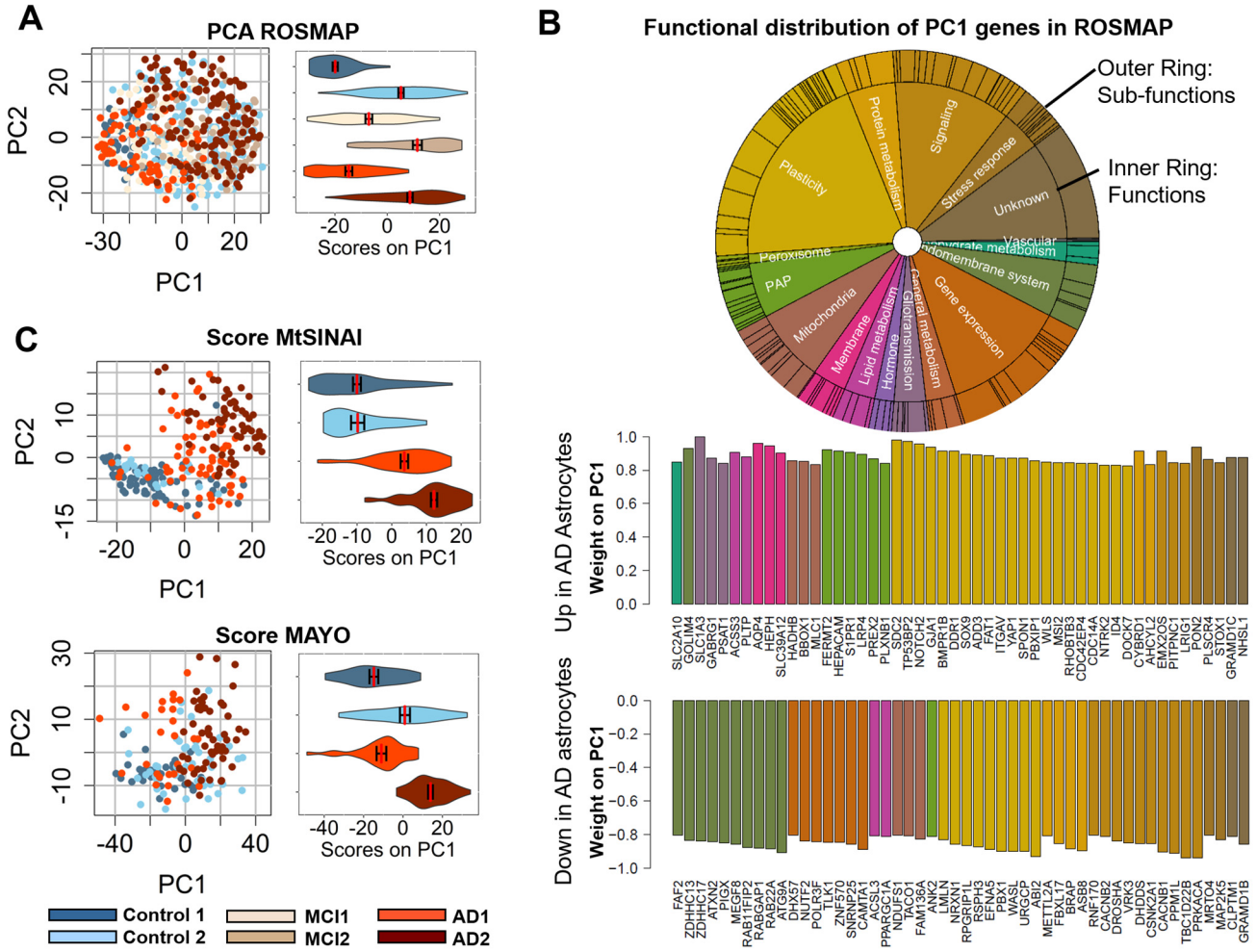


Figure 5. PCA identifies top altered functions in AD astrocytes.

A. PCA analysis of astrocyte-enriched genes in ROSMAP. Dot plots and violin plots along PC1 (mean ± SEM) show differential expression of astrocyte genes in type 2 versus type 1 subjects.

B. Top varying genes in PC1 (up-regulated and down-regulated), color-coded for function according to the circos plot, inform about functions altered in AD astrocytes. The vertical axis represents the relative contribution (importance/weight) of each gene for the scoring of individual samples in panel **A** along PC1.

C. Scoring of MtSINAI and MAYO databases using genes overlapping with ROSMAP PCA in **A.** discriminates Control and AD cohorts along PC1, as shown by the dot and violin plots (means ± SEM).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

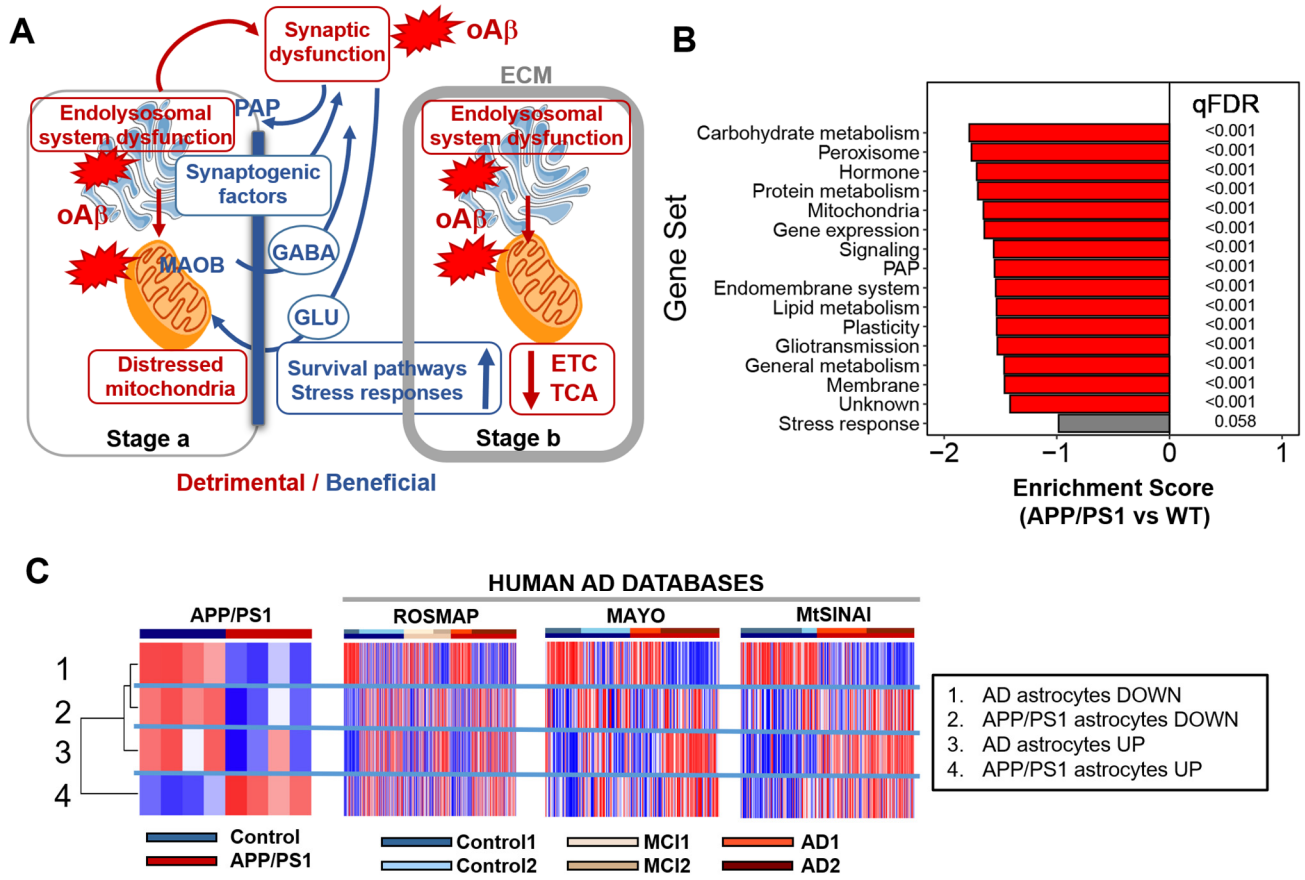


Figure 6. Model of cortical AD astrocytes and comparison with mice. **A**. The model revolves around two predictions. The first one is that amyloid-β drives the phenotypical change of cortical astrocytes in AD by two parallel mechanisms: direct and indirect. On the one hand, amyloid–β directly causes endolysosomal and mitochondrial dysfunction in astrocytes, which results in an adaptive shift of gliotransmission to a GABA-predominant mode as a strategy to preserve mitochondrial respiration. Impairment of mitophagy due to endolysosomal malfunction may further exacerbate mitochondrial dysregulation. On the other hand, synaptic dysfunction caused by amyloid-β in neurons triggers increased production of the PAP machinery, including synaptogenic factors—a phenomenon of enhanced protective local astrocytic plasticity. The second prediction is that the phenotypical transformation of astrocytes in AD happens in early and late stages. Control2 subjects, which are cognitively normal would be at stage a, and AD2 patients, diagnosed with dementia, at stage b. In stage a, we posit that the change in glutamate/GABA fluxes is an astroprotective response that, at the same time, preserves neural-circuit homeostasis, because the GABA produced by MAO-B may locally compensate for the deficits in inhibitory tone that disrupts circuitry synchronization [43]. By contrast, deficits in the endolysosomal system would result in partial phagocytosis of amyloid-β [28] and dystrophic neurites [31], thus exacerbating amyloid-β accumulation, and hindering neuronal repair due to lack of debris elimination. In stage b, we posit that the mechanisms triggered to protect mitochondria fail, and cortical astrocytes become deeply malfunctional, as suggested by the reduction in TCA and ETC transcripts, despite the increase in protective stress responses and pro-survival pathways aimed at preventing astrocyte demise. For simplicity, only predominantly altered functions are depicted in stages a and b. ECM, extracellular matrix; ETC: electron transport chain, GLU: glutamate; oAβ: oligomeric amyloid-β; PAP: perisynaptic astrocyte processes; TCA:

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

tricarboxylic acid cycle. This figure was created using cell art adapted from Servier Medical Art (https://smart.servier.com/).

B. GSVA of the transcriptome of astrocytes isolated from APP/PS1 mouse cortices using our manually curated functional categorization shows downregulation of astrocyte-enriched functions.

C. Left, GSVA of transcriptome of APP/PS1 astrocytes according to the top dysregulated astrocyte genes in PC1 from the ROSMAP database (**Fig. 5B**). Astrocyte-enriched genes dysregulated in AD (up and down) are down-regulated in APP/PS1 astrocytes, as shown in B. Right, GSVA of PC1 AD and mouse APP/PS1 DEG gene sets in ROSMAP, MAYO and MtSINAI whole-brain transcriptomes shows that both the up and down-regulated mouse genes pools tend to be up-regulated in AD databases.

Acknowledgements

This work was supported by Fundació Marato-TV3 (grant number 201414-30), and Grup de Recerca en demències: Sant Pau SGR2017547, Generalitat de Catalunya (E.G.), by funds from the Coins for Alzheimer's Research Trust (L.B.W.) and startup funds from the Woodruff School of Mechanical Engineering at Georgia Tech (L.B.W.). L.D.W. and A.F.P were supported in part by the National Institutes of Health Cell and Tissue Engineering Biotechnology Training Grant (T32-GM008433). CNRS, CEA and France Alzheimer (C.E.).

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.synapse.org). Study data were provided by the following sources: The Mayo Clinic Alzheimers Disease Genetic Studies, led by Dr. Nilufer Ertekin-Taner and Dr. Steven G. Younkin, Mayo Clinic, Jacksonville, FL using samples from the Mayo Clinic Study of Aging, the Mayo Clinic Alzheimers Disease Research Center, and the Mayo Clinic Brain Bank. Data collection was supported through funding by NIA grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949, NINDS grant R01 NS080820, CurePSP Foundation, and support from Mayo Foundation. Study data includes samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The Brain and Body Donation Program is supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium) and the Michael J. Fox Foundation for Parkinson's Research. Study data were also provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161 (ROS), R01AG15819 (ROSMAP; genomics and RNAseq), R01AG17917 (MAP), R01AG30146, R01AG36042 (5hC methylation, ATACseq), RC2AG036547 (H3K9Ac), R01AG36836 (RNAseq), R01AG48015 (monocyte RNAseq) RF1AG57473 (single nucleus RNAseq), U01AG32984 (genomic and whole exome sequencing), U01AG46152 (ROSMAP AMP-AD, targeted proteomics), U01AG46161(TMT proteomics), U01AG61356 (whole genome sequencing, targeted proteomics, ROSMAP AMP-AD), the Illinois Department of Public Health (ROSMAP), and the Translational Genomics Research Institute (genomic). ROSMAP data can be requested at www.radc.rush.edu.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supplementary data

Supplementary data 1. Generation of cell-specific gene modules. Tabs:

1. Tab description.
2. Transcriptomic databases of human-brain cells obtained from [13].
3. PCA of transcriptomic data of astrocytes isolated from young (<47 years old) and old (>47 years old) shows that they were significantly different.
4. Specific data used for τ scoring and hierarchical clustering.
5. Top τ scores per gene, indicating which cell type presents that score; and pie charts of the distribution of specific and enriched genes among brain cells.
6. Raw results of hierarchical clustering.
7. Cluster description.
8. Cluster selection.

Supplementary data 2. Functional categories of astrocytic and neuronal modules. Tabs:

1. Circo plots showing functional categories (inner ring) and subcategories (outer ring) of astrocytic gene cluster 4 (1650 genes) and neuronal gene cluster 13 (1258 genes). For astrocytes, there are 17 groups and 133 subgroups. For neurons, 15 groups and 76 subgroups.
2. Functional categorization of neuronal cluster 13.
3. Functional categorization of astrocytic cluster 4.

Supplementary data 3. GSVA statistics of inter-cohort comparisons of changes in neuronal and astrocytic general functions in MAYO, ROSMAP and MtSINAI databases. Comparisons: ROSMAP: Control2 vs Control1 (pFDR1), MCI1 vs Control1 (pFDR2), MCI2 vs Control1 (pFDR3), AD1 vs Control1 (pFDR4), AD2 vs Control1 (pFDR4). MAYO and MtSINAI: Control2 vs Control1 (pBootFDR1), AD1 vs Control1 (pFDR2), AD2 vs Control1 (pFDR3). Tabs:

1. Neurons.
2. Astrocytes.

Supplementary data 4. GSVA statistics of inter-cohort comparisons of changes in astrocytic subcategories in MAYO, ROSMAP and MtSINAI databases. Comparisons: ROSMAP: Control2 vs Control1 (pFDR1), MCI1 vs Control1 (pFDR2), MCI2 vs Control1 (pFDR3), AD1 vs Control1 (pFDR4), AD2 vs Control1 (pFDR4). MAYO and MtSINAI: Control2 vs Control1 (pFDR1), AD1 vs Control1 (pFDR2), AD2 vs Control1 (pFDR3). Tabs:

1. All functional subcategories/MAYO.
2. All functional subcategories /ROSMAP.
3. All functional subcategories /MtSINAI.
4. PAP in all databases.
5. Gliotransmission / all databases.
6. Plasticity /all databases.
7. Stress response /all databases.
8. Mitochondria in all data /all databases.
9. Endolysosomal system /all databases.

Supplementary data 5. Expression of astrocytic genes from cluster 4 in the clinical databases. The statistical comparison between AD2 and Control1 was performed by the Wilcoxon rank sum test. Tabs:

1. MAYO.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2. ROSMAP.
3. MtSINAI.
4. Consensus genes dysregulated in at least two of the databases used in Discussion to generate the model of phenotypical transformation of astrocytes.

Supplementary data 6. Astrocyte cluster 4 ranked according to PC1.

Supplementary data 7. Comparison of immune-like genet sets in AD (our human astrocyte annotations) and APP/PS1 cortical astrocytes [10]. The functional category of cluster # 4 of human astrocytes containing immunity-related genes were compared with the immunity-related GO groups found significantly dysregulated in APP/PS1 mouse astrocytes. Only C1S appears in both gene sets.

Supplementary data 8. Comparison of PC1 genes with DEG identified by snRNAseq in human AD astrocytes [5] [4]. Tabs:

1. Venn diagrams.
2. Comparison with [5].
3. Comparison with [4].

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5. References

1. Perez-Nievas, B.G. and A. Serrano-Pozo, *Deciphering the Astrocyte Reaction in Alzheimer's Disease*. Front Aging Neurosci, 2018. **10**: p. 114.
2. Simpson, J.E., et al., *Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype*. Neurobiology of Aging, 2011. **32**(10): p. 1795-1807.
3. Sekar, S., et al., *Alzheimer's disease is associated with altered expression of genes involved in immune response and mitochondrial processes in astrocytes*. Neurobiology of Aging, 2015. **36**(2): p. 583-591.
4. Grubman, A., et al., *A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation*. Nature Neuroscience, 2019. **22**(12): p. 2087-+.
5. Mathys, H., et al., *Single-cell transcriptomic analysis of Alzheimer's disease (vol 570, pg 332, 2019)*. Nature, 2019. **571**(7763): p. E1-E1.
6. Zhang, B., et al., *Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease*. Cell, 2013. **153**(3): p. 707-720.
7. Pelvig, D.P., et al., *Neocortical glial cell numbers in Alzheimer's disease. A stereological study*. Dement Geriatr Cogn Disord, 2003. **16**(4): p. 212-9.
8. Devi, G. and P. Scheltens, *Heterogeneity of Alzheimer's disease: consequence for drug trials?* Alzheimers Res Ther, 2018. **10**(1): p. 122.
9. Dujardin, S., et al., *Tau molecular diversity contributes to clinical heterogeneity in Alzheimer's disease*. Nature Medicine, 2020. **26**(8): p. 1256-1263.
10. Orre, M., et al., *Isolation of glia from Alzheimer's mice reveals inflammation and dysfunction*. Neurobiol Aging, 2014. **35**(12): p. 2746-2760.
11. Ceyzeriat, K., et al., *Modulation of astrocyte reactivity improves functional deficits in mouse models of Alzheimer's disease*. Acta Neuropathol Commun, 2018. **6**(1): p. 104.
12. Habib, N., et al., *Disease-associated astrocytes in Alzheimer's disease and aging*. Nature Neuroscience, 2020. **23**(6): p. 701-706.
13. Zhang, Y., et al., *Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse*. Neuron, 2016. **89**(1): p. 37-53.
14. Kryuchkova-Mostacci, N. and M. Robinson-Rechavi, *A benchmark of gene expression tissue-specificity metrics*. Brief Bioinform, 2017. **18**(2): p. 205-214.
15. Bennett, D.A., et al., *Religious Orders Study and Rush Memory and Aging Project*. J Alzheimers Dis, 2018. **64**(s1): p. S161-S189.
16. Allen, M., et al., *Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases*. Sci Data, 2016. **3**: p. 160089.
17. Gomez-Isla, T., et al., *Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease*. Ann Neurol, 1997. **41**(1): p. 17-24.
18. Marlatt, M.W., et al., *Proliferation in the Alzheimer hippocampus is due to microglia, not astroglia, and occurs at sites of amyloid deposition*. Neural Plast, 2014. **2014**: p. 693851.
19. Kelley, K.W., et al., *Variation among intact tissue samples reveals the core transcriptional features of human CNS cell classes*. Nat Neurosci, 2018. **21**(9): p. 1171-1184.
20. Lin, S.M., et al., *Model-based variance-stabilizing transformation for Illumina microarray data*. Nucleic Acids Research, 2008. **36**(2): p. e11-e11.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

21. Sakers, K., et al., *Astrocytes locally translate transcripts in their peripheral processes*. Proc Natl Acad Sci U S A, 2017. **114**(19): p. E3830-E3838.
22. Santello, M., P. Bezzi, and A. Volterra, *TNFα controls glutamatergic gliotransmission in the hippocampal dentate gyrus*. Neuron, 2011. **69**(5): p. 988-1001.
23. Hanzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data*. BMC Bioinformatics, 2013. **14**: p. 7.
24. Han, Z., et al., *Analyzing 74,248 Samples Confirms the Association Between CLU rs11136000 Polymorphism and Alzheimer's Disease in Caucasian But Not Chinese population*. Sci Rep, 2018. **8**(1): p. 11062.
25. Neff, R.A., et al., *Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets*. Science Advances, 2021. **7**.
26. Arenaza-Urquijo, E.M. and P. Vemuri, *Resistance vs resilience to Alzheimer disease: Clarifying terminology for preclinical studies*. Neurology, 2018. **90**(15): p. 695-703.
27. Marsh, J. and P. Alifragis, *Synaptic dysfunction in Alzheimer's disease: the effects of amyloid beta on synaptic vesicle dynamics as a novel target for therapeutic intervention*. Neural Regen Res, 2018. **13**(4): p. 616-623.
28. Sollvander, S., et al., *Accumulation of amyloid-beta by astrocytes result in enlarged endosomes and microvesicle-induced apoptosis of neurons*. Mol Neurodegener, 2016. **11**(1): p. 38.
29. Funato, H., et al., *Astrocytes containing amyloid beta-protein (Abeta)-positive granules are associated with Abeta40-positive diffuse plaques in the aged human brain*. Am J Pathol, 1998. **152**(4): p. 983-92.
30. Sanchez-Mico, M.V., et al., *Amyloid-beta impairs the phagocytosis of dystrophic synapses by astrocytes in Alzheimer's disease*. Glia, 2020.
31. Gomez-Arboledas, A., et al., *Phagocytic clearance of presynaptic dystrophies by reactive astrocytes in Alzheimer's disease*. Glia, 2018. **66**(3): p. 637-653.
32. Abdullah, M., et al., *Amyloid-beta Reduces Exosome Release from Astrocytes by Enhancing JNK Phosphorylation*. J Alzheimers Dis, 2016. **53**(4): p. 1433-41.
33. Martini-Stoica, H., et al., *TFEB enhances astroglial uptake of extracellular tau species and reduces tau spreading*. J Exp Med, 2018. **215**(9): p. 2355-2377.
34. Derouiche, A. and K.D. Geiger, *Perspectives for Ezrin and Radixin in Astrocytes: Kinases, Functions and Pathology*. Int J Mol Sci, 2019. **20**(15).
35. Swerdlow, R.H., *Mitochondria and Mitochondrial Cascades in Alzheimer's Disease*. Journal of Alzheimers Disease, 2018. **62**(3): p. 1403-1416.
36. Hoffman, J.M., et al., *FDG PET imaging in patients with pathologically verified dementia*. J Nucl Med, 2000. **41**(11): p. 1920-8.
37. Shi, Q. and G.E. Gibson, *Oxidative stress and transcriptional regulation in Alzheimer disease*. Alzheimer Dis Assoc Disord, 2007. **21**(4): p. 276-91.
38. Abramov, A.Y., L. Canevari, and M.R. Duchen, *β-Amyloid Peptides Induce Mitochondrial Dysfunction and Oxidative Stress in Astrocytes and Death of Neurons through Activation of NADPH Oxidase*. The Journal of Neuroscience, 2004. **24**(2): p. 565-575.
39. Graves, S.M., et al., *Dopamine metabolism by a monoamine oxidase mitochondrial shuttle activates the electron transport chain*. Nat Neurosci, 2020. **23**(1): p. 15-20.
40. Rodriguez-Vieitez, E., et al., *Diverging longitudinal changes in astrocytosis and amyloid PET in autosomal dominant Alzheimer's disease*. Brain, 2016. **139**(Pt 3): p. 922-36.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.25.21252422; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

41. Jo, S., et al., *GABA from reactive astrocytes impairs memory in mouse models of Alzheimer's disease*. Nat Med, 2014. **20**(8): p. 886-96.
42. Birks, J. and L. Flicker, *Selegiline for Alzheimer's disease*. Cochrane Database Syst Rev, 2003(1): p. CD000442.
43. Lee, Y.F., et al., *Slow Wave Sleep Is a Promising Intervention Target for Alzheimer's Disease*. Frontiers in Neuroscience, 2020. **14**(705).
44. Escartin, C., O. Guillemaud, and M.A. Carrillo-de Sauvage, *Questions and (some) answers on reactive astrocytes*. Glia, 2019. **67**(12): p. 2221-2247.
45. Misra, J.R. and K.D. Irvine, *The Hippo Signaling Network and Its Biological Functions*. Annu Rev Genet, 2018. **52**: p. 65-87.
46. Torper, O. and M. Gotz, *Brain repair from intrinsic cell sources: Turning reactive glia into neurons*. Functional Neural Transplantation Iv: Translation to Clinical Application, Pt A, 2017. **230**: p. 69-97.
47. Xie, C., et al., *Astrocytic YAP Promotes the Formation of Glia Scars and Neural Regeneration after Spinal Cord Injury*. J Neurosci, 2020. **40**(13): p. 2644-2662.
48. Saddawi-Konefka, R., et al., *Nrf2 Induces IL-17D to Mediate Tumor and Virus Surveillance*. Cell Rep, 2016. **16**(9): p. 2348-58.
49. Iram, T., et al., *Astrocytes from old Alzheimer's disease mice are impaired in Abeta uptake and in neuroprotection*. Neurobiol Dis, 2016. **96**: p. 84-94.
50. Escartin, C., et al., *Reactive astrocyte nomenclature, definitions, and future directions*. Nat Neurosci, 2021.