Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology

Amy M. Smith[1,2], Karen Davey[1,2], Stergios Tsartsalis[2,3], Combiz Khozoie[2], Nurun Fancy[1,2], See Swee Tang[2], Eirini Liaptsi[2], Maria Weinert[2], Aisling McGarry[1,2], Robert C. J. Muirhead[1,2], Steve Gentleman[2], David R. Owen[2] and Paul M. Matthews[1,2]

UK Dementia Research Institute, Imperial College London[1]; Department of Brain Sciences, Imperial College London, United Kingdom[2]; Department of Psychiatry, University of Geneva, Switzerland[3]

Address for correspondence:

Prof. Paul M. Matthews

E515, Department of Brain Sciences
Imperial College London
Hammersmith Hospital
DuCane Road, London WC12 0NN
Tel: 0044 207 594 2612

p.matthews@imperial.ac.uk

Abstract

To better define roles that astrocytes and microglia play in Alzheimer's disease (AD), we used single-nuclei RNA sequencing to comprehensively characterize transcriptomes in astrocyte and microglia nuclei isolated *post mortem* from neuropathologically-defined AD and control brains with a range of amyloid-beta and phospho-tau (pTau) pathology. Significant differences in glial gene expression (including AD risk genes expressed in astrocytes [*CLU*, *MEF2C*, *IQCK*] and microglia [*APOE*, *MS4A6A*, *PILRA*]) were correlated with tissue amyloid and pTau expression. Astrocytes were enriched for proteostatic, inflammatory and metal ion homeostasis pathways. Pathways for phagocytosis, proteostasis and autophagy were highly enriched in microglia and perivascular macrophages. Gene co-expression analyses revealed potential functional associations of soluble biomarkers of AD in astrocytes (*CLU*) and microglia (*GPNMB*). Our work highlights responses of both astrocytes and microglia for pathological protein clearance and inflammation, as well as glial transcriptional diversity in AD.

Introduction

Astrocytes and microglia play important and potentially causal roles in the disease[1]. Activated astrocytes and microglia are found around amyloid plaques[2] and genes associated with AD risk are enriched in both cell types, particularly in microglia[3,4]. However, the mechanisms by which microglia and astrocytes contribute to disease genesis, progression and response still are poorly defined; a recent meta-analysis suggested that the diversity of glial responses in late onset sporadic Alzheimer's disease is not well captured by mouse models[5].

Single-nuclei RNA-sequencing (snRNASeq) from *post mortem* brain tissue[6,7] is transforming the potential to characterise the molecular neuropathology of AD at the level of single cells[8-10]. However, because of their lower cellular abundance in the brain, microglia and astrocytes have been poorly represented in most studies published to date (e.g., 449-3982 microglia, representing only 1-3% of the total nuclei annotated [8,9,11]), limiting the depth to which they can be characterised. Astrocytes and microglia also have highly heterogeneous phenotypes[12-14]. To address these limitations, we employed a novel negative-selection approach that enriches for them in nuclei isolated from *post mortem* tissue. This allowed us to characterise snRNASeq transcriptomes from much larger numbers of these nuclei (52,706 astrocytes and 27,592 microglia) efficiently. After quantitative assessments of neuropathological features in each brain region, we generated transcriptional signatures associated with amyloid-beta or pTau pathology in non-diseased control (NDC) and AD brains. These data allowed us to develop comprehensive descriptions of gene co-expression networks that provide both further insights into responses of astrocytes and microglia to AD pathology and evidence for cell-type specific functions of genes associated with risk for AD. Transcription factors potentially responsible for the differential gene expression with pathology were identified from these co-expression modules. We confirmed our major observations by re-analyses of data from four previously reported AD snRNASeq studies. Our work provides new insights into linked glial-specific responses mediating pathological protein clearance and inflammation in AD, highlighting diverse roles of astrocytes and microglia.

Results

Selective astrocyte and microglia transcriptome sequencing

The proportions of microglia and astrocytes defined by snRNASeq of nuclei isolated from the human brain *post mortem* are low and variable[8,9]. To enable comprehensive analyses of differential transcript expression in astrocytes and microglia with AD pathology, we enriched for these glia by selectively removing neuronal (NeuN-positive[7]) and oligodendrocyte (Sox10-positive[15]) nuclei using FACS (Fig. 1 and Extended Data Fig. 1). For this, we isolated nuclei from each of two cortical regions (entorhinal and somatosensory cortex) taken from six brains with low levels of AD neuropathology provided by donors without reported cognitive impairment and from six brains with high levels of AD pathology (Extended Data Fig. 2). Astrocyte nuclei had a mean unique molecular identifier (UMI) count of 8,775 with an average of 3,166 distinct genes and microglial nuclei had a mean UMI count of 4,808 with an average of 2,132 genes. Amongst the 52,706 astrocyte nuclei, we found expression of 90% of astrocyte transcripts previously reported from human brain astrocytes (500 genes)[16]. 16/65 AD risk genes were represented in the astrocyte co-expression network (Extended Data Fig. 3a). The 27,592 total microglial nuclei included expression of 96% of the recently described microglial "core" consensus transcriptome (249 genes)[17]. Microglia also were highly enriched in genes associated previously with genetic risk for AD (27/65)[1,18,19](Extended Data Fig. 4a).

Increased expression of genes related to metal ion homeostasis, proteostasis and inflammation in astrocytes with AD pathology

Gene expression associated with extracellular amyloid plaques or intraneuronal neurofibrillary tangles (pTau) (Fig. 2) were discovered by regressing gene expression against amyloid-beta (expressed as $\log_2$ fold difference/% area stained) or pTau (expressed as $\log_2$ fold difference/% pTau positive cells) densities in sections prepared from homologous regions of the contralateral hemispheres for each of the brains. Half of the significantly positively associated genes expressed were correlated with both amyloid-beta and pTau pathology, but almost three-fold more transcripts were associated uniquely with amyloid-beta (313 genes) expression relative to pTau (106 genes) (Fig. 2). We found significant astroglial functional enrichment for

pathways involved in the 'cellular response to zinc ion', 'cellular response to copper ion' and 'response to metal ions' with both amyloid-beta and pTau expression (Fig. 3); genes encoding proteins involved in metal ion homeostasis (*MT1G*, *MT1F*, *MT1E*, *MT2A*, *MT3* and *FTL*) were amongst the top transcripts most highly *positively* associated with pathology in astrocytes. Transcripts involved in 'chaperone-mediated protein complex assembly' and 'response to unfolded protein' pathways, such as *CRYAB*, *HSPB1*, *HSPH1* and *HSP90AA1* also were positively differentially expressed. Increased expression of the AD risk gene *CLU* was associated with pTau pathology in astrocytes (Extended Data Fig. 5). Expression of the AD risk gene *IQCK* was positively associated with both amyloid-beta and pTau (Extended Data Fig. 5). Pathways involved in inflammatory processes also were significantly enriched ('NLRP3 inflammasome' and 'NFkB is activated and signals survival'). By contrast, "core" or homeostatic astrocyte transcripts, such as those for glutamate transporters *SLC1A3* and *SLC1A2* or for *IL-33* (CSF1R ligand, which promotes microglial synaptic remodelling[20]) were down-regulated. The AD risk-associated *MEF2C* transcription factor, as well as *MAFG*, *JUND*, *CEBPB*, *MAF* and *LHX2* were up-regulated, suggesting roles for these transcription factors in the regulation of responses to AD pathology.

We confirmed with gene set enrichment (AUCell) that transcripts positively differentially expressed with pathology also were significantly enriched in nuclei with human AD pathology reported in previously published snRNASeq studies [8-10,21], albeit with very low log fold changes in one[21] out of the four datasets analysed.

Expression of genes related to autophagy, phagocytosis and proteostasis in microglia with AD pathology

Microglial transcripts most highly *positively* associated with tissue amyloid-beta and tissue pTau density included those for genes associated with AD risk (*APOE*, *MS4A6A* and *PILRA*, Fig. 2 and Extended Data Fig. 5), as well as those for genes associated with risks for other neurodegenerative disorders (*LRRK2*, *SNCA* and *GPNMB*, associated with Parkinson's disease[22], and *GRN*, associated with ceroid lipofuscinosis[23] and frontotemporal dementia[24]). Four-fold more transcripts were associated uniquely with amyloid-beta

(109 genes) expression relative to pTau (27 genes), while 60% of the significantly positively associated genes expressed were correlated with both amyloid-beta and pTau pathology (Fig. 2). Differentially expressed transcripts were functionally enriched in 'selective autophagy' and 'microglia pathogen phagocytosis' pathways (Fig. 4); *ASAH1*, *ATG7*, *STARD13* and *MYO1E* were amongst the most strongly positively associated genes. Perivascular macrophages (PVM) also showed functional enrichment in 'selective autophagy', as well as several CCT/TriC molecular chaperone complex pathways involved in proteostasis and actin/tubulin folding (Extended Data Fig. 6). Up-regulation of the transcription factors *MAFG*, *MITF* and *JUND* with amyloid-beta or pTau pathology suggests their involvement in transcriptional regulation of these microglial and PVM responses to pathology.

We confirmed that transcripts positively differentially expressed in microglia were significantly enriched in nuclei with human AD pathology in previously published snRNASeq studies [8-10,21]. Toll-like receptors (TLR2 and TLR10), HK2 (hexokinase 2), JAK2 (Janus kinase 2) and ITGAM (CD11b) were amongst the smaller number of transcripts in these cells that were significantly *negatively* associated with tissue amyloid-beta or pTau.

Gene co-expression modules suggest glial cell-specific functional roles of AD GWAS genes

Evidence for involvement of CLU in astrocyte metal ion homeostasis and proteostasis pathways with AD

Co-expression network analyses were used to characterise gene expression modules (MEGENA) in astrocytes and microglia, suggesting potential functional relationships. AD GWAS genes *CLU* and *IQCK* were co-expressed in an astrocyte module (module 9) which was amongst the most strongly positively correlated with both amyloid-beta and pTau density. Both *CLU* and *GJA1* (Gap Junction Protein Alpha 1; Connexin-43) are hub genes in this module, which was functionally enriched in transcripts for proteins involved in metal ion homeostasis (e.g., 'metallothioneins bind metals' and 'response to metal ions') and proteostasis ('HSF1 activation', 'response to unfolded protein' and 'chaperone-mediated protein complex assembly'). They also were hub genes in the related child module 30, which includes genes in pathways for 'ceramide transport' and 'gap junction assembly' (Extended Data Fig. 3b). *CLU* was associated with the

astrocyte regulons (for transcription factors MAF, MAFG, JUND and CEBPB) that had the strongest correlations with pTau and amyloid-beta. *CLU*-containing modules and the associated regulons also were enriched in AD nuclei reported in earlier studies.

Evidence for a cell-specific role for APOE in microglia linking phagocytic, complement and inflammatory activation pathways in AD

APOE, the largest genetic risk factor for AD, was up-regulated in microglia with both pTau and amyloid-beta pathology. *APOE* was a hub gene in microglia co-expressed both with *TREM2* and inflammatory activation and response genes (e.g., *C1QB, C1QC, CD74, CTSB*) in a module functionally enriched for pathways including the 'endosomal/vacuolar pathway', 'microglia pathogen phagocytosis', and 'antigen processing – cross presentation' (module 19) (Extended Data Fig. 4b). Regulons inferred to be responsible for microglial *APOE* expression included those for transcription factors MXD4, MITF, PBX3 and JUND. *APOE* expression in astrocytes was not significantly correlated with either amyloid or pTau pathology and co-expression relationships suggested a different functional role for *APOE* in astrocytes as a hub gene in a module functionally enriched for 'dermatan sulphate biosynthesis', 'extracellular matrix organisation' and 'ferroptosis' (module 13).

Microglial and PVM GPNMB are up-regulated with AD pathology in modules related to lipid homeostasis

Glycoprotein nonmetastatic melanoma protein B (GPNMB) is elevated in plasma and CSF with AD and has been proposed as a biomarker of disease[25]. We found that *GPNMB* is up-regulated in microglia with amyloid-beta and pTau pathology and in PVMs with pTau pathology. Consistent with this, *GPNMB* was found in association with the MAFG, JUND, MAFB, CEBPD and CEBPA transcription factor regulons which showed strong correlations to amyloid-beta and pTau in microglia (Fig. 5a). *GPNMB* also is a hub gene in microglial co-expression modules 11 and 34, which are strongly associated with amyloid-beta and pTau expression. As well as *GPNMB*, hub genes for module 11 also include *ASAH1, ATG7, STARD13, IQGAP2, CPVL, TANC2* and *MITF*, all of which were positively differentially expressed with one or both of the AD pathologies (Fig. 5a). We found that module 11 was enriched in AD (relative to control) samples

from three out of four previous human snRNASeq studies analysed[8,10,21]. Pathways involving the differentially expressed genes in module 11 suggest a functional role in cholesterol homeostasis ('regulation of cholesterol storage'). We found that the smaller module 34 was enriched in AD samples from two out of four previous studies[9,21]. Functional pathways enriched in the *GPNMB* hub gene module 34 relate to phospholipid and lipoprotein homeostasis ('phospholipid efflux', 'phospholipid homeostasis', 'lipoprotein metabolism').

Consistent with the expected GPNMB expression in PVM, immunohistochemistry for GPNMB showed prominent staining around blood vessels, as well as in a subset of parenchymal microglia (Fig. 5b). There was a mean ~1.8-fold increase in GPNMB-positive cells/μm^2 in AD (relative to NDC) in the cortical tissue studied here (Fig. 5c). GPNMB staining density in sections of samples was positively correlated both with tissue pTau (R=0.39) and with expression of modules 11 and 34 (R=0.507 and 0.576, respectively) defined by snRNASeq of the same region in the contralateral hemisphere.

Transcriptional heterogeneity associated with AD pathology in sub-sets of human astrocytes and microglia

Our snRNASeq data reduction identified six distinguishable clusters that expressed the core set of astrocyte genes. Each also expressed distinct sets of genes that suggested sub-types of astrocytes (Fig. 1). Astro1 and Astro2 expressed higher levels of genes involved in core astrocyte functions, such as *SLC1A2* (*GLT1*) and glutamine synthetase (*GLUL*) and Astro2 was enriched for pathways including 'neurotransmitter uptake', 'glutamatergic synapses' and 'amino acid import'. In contrast, Astro4 and Astro5 were characterized by relative expression of genes involved in extracellular matrix formation and functions[26] and were enriched for pathways including 'carbohydrate binding' and 'cell-matrix adhesion'. Astro4 was distinguished from Astro5 by relatively high expression of *VEGFA*, while Astro5 was enriched uniquely for immune response pathways, such as the Toll-like receptor cascade and the activated astrocyte marker *GFAP*. Transcripts in Astro6 were enriched for metallothionein genes. Astro5 and Astro6 showed the greatest specificity for regulons most highly up-regulated with pTau (for transcription factors JUND, MAF and CEBPB).

The disease associated astrocyte (DAA) expression signature defined in a mouse model[27] was represented in all astrocyte clusters other than Astro3. 4.6% of DAA genes were up-regulated with amyloid-beta and 9.2% with pTau. Although the A1 (12/15 genes expressed) or A2 (13/13 genes expressed) gene sets[28] associated previously with injury-responsive or homeostatic astrocytes in rodent models, respectively, were represented in the astrocyte clusters, neither A1 nor A2 gene sets were enriched significantly in the total astrocyte nuclei or in any of the clusters.

Clustering parameters selected to distinguish PVM (identified by markers such as *CD163*, *MRC1* (*CD206*) and *MSR1*[29]) also identified three clusters of nuclei (designated Micro1, 2 and 3) expressing different sets of microglial marker genes (Fig. 1). Micro1 was most highly enriched in transcripts for human 'core'[30] homeostatic genes, Micro2 showed a relative functional enrichment for the 'TYROBP causal network' and 'ferroptosis' pathways and Micro3 expressed lower levels of both homeostatic and activation genes, but had higher expression of *C3* and *LPAR6*. Micro2 showed the greatest specificity for the MAFG, CEBPA, JUND, CEBPD and MITF regulons found here to be highly correlated with amyloid-beta and pTau. Plaque-induced (PIG[31]), disease associated (DAM[32]), activated response (ARM[33]) and interferon response microglial (IRM[33.34]) gene sets identified in rodent amyloid models were expressed in all of the clusters, with relative enrichment of DAM, ARM and, to a lesser degree, IRM gene sets in Micro2 (Extended Data Fig. 7). PVMs were also relatively enriched in these gene sets related to microglial activation, as well as in a gene set associated with human microglial aging[35] (Extended Data Fig. 7).

Discussion

Our results describe potentially neuroprotective gene up-regulation for proteostasis, phagocytosis and protein clearance with amyloid-beta and pTau pathology in both astrocytes and microglia. A 3-4 fold greater number of genes were uniquely differentially expressed with amyloid-beta relative to pTau, mirroring observations with microglia in preclinical transgenic models overexpressing the two proteins[36]. Differentially expressed genes in astrocytes were associated with increased expression of genes for the

regulation of metal ion homeostasis that may both protect cells from oxidative injury and facilitate redox-dependent chaperone clearance of abnormal proteins by astrocytes[37]. However, astrocytes also expressed inflammasomes and inflammatory activation pathways commonly found in microglia, both of which may promote neurodegeneration. Transcriptional functional diversity suggests distinguishable sub-types of human astrocytes and microglia enriched in gene sets related to (but distinct from) those previously described in transgenic amyloid or tau mouse models or with aging[27,32,33,35]. Moreover, co-expression networks suggest cell-specific functional roles for genes associated with AD risk, most notably highlighting *APOE* as a hub gene in microglial co-expression modules linking gene expression subserving phagocytic, complement and inflammatory activation pathways.

We extended previous approaches to enrich our *post mortem* brain snRNASeq data for microglia and astrocytes in our study. Several recent publications have described approaches for selective glial nuclei isolation from biopsy and rapid *post mortem* delay tissue, including the use of antibodies to microglial transcription factor PU.1[38], a combination of NeuN and Olig2[39], and IRF8[40]. Our negative-selection approach is distinguished by relying on nuclear markers that we have found to be robust to *post mortem* delay, reliably at least to 24 hr. Our analytical approach for addressing heterogeneity in the samples also differs from most prior studies (although it is not without similar precedents[11,41]). Categorical descriptions of brains by Braak stage do not directly reflect tissue pathology in the local regions studied, so we associated each glial nuclear preparation with quantitative measures of pTau and amyloid-beta density specific to the regions studied in each of the brains. We also used of a mixed regression model to account for uncontrolled sources of variance related to each sample[42].

Astrocytes expressed genes suggesting functional roles for proteostasis with amyloid-beta and pTau-associated functional enrichment for HSF1 activation and chaperone pathways. This was accompanied by expression of genes for pro-inflammatory NF-kB and inflammasome pathways. Greater regional amyloid-beta and pTau also was associated with enrichment of astrocytes for metal ion homeostasis pathways and the expression of metallothioneins [43-45]. Identification of *CLU* as a hub gene in astrocyte co-expression modules

including genes for metal ion homeostasis and proteostasis may reflect dependence of clusterin chaperone functions on the local redox environment[37]. We found increased expression of the astrocyte *CLU* module with greater amyloid or pTau.

Co-expression analysis in microglia identified *APOE* as a hub gene for a module including complement genes (*C1QA, C1QB, C1QC*) and functionally enriched in phagocytosis pathways. Microglia may contribute to astrocyte activation through C1q expression and C1q and other complement cascade proteins co-localise with amyloid plaques in AD[46]. These analyses also highlighted potential functional relationships of *GPNMB* (concentrations of the protein product of which are increased in brain samples and cerebrospinal fluid of sporadic AD patients[25]) with a diverse group of genes connected to transcription factors whose regulons were up-regulated with amyloid and pTau, suggesting that GPNMB provides a biomarker for a human microglial neurodegenerative activation state central to pathological responses in AD.

We tested the relevance of microglial sub-cluster expression signatures defined in mouse transgenic models of AD using gene set enrichment analyses (Extended Data Fig. 7). We found evidence for similar, moderate levels of relative enrichment for inflammatory activation gene sets (DAM and ARM) in Micro2 and PVM[32,33]. The relatively low expression of both homeostatic and activation genes in Micro3 corresponded to patterns associated with 'dystrophic', immuno-senescent microglia[47]. Transcripts from the disease associated astrocyte (DAA) gene set were up-regulated with amyloid-beta and pTau, but not restricted to any particular astrocyte sub-set[27]. Our results highlight distinct features and a greater functional diversity amongst microglia and astrocytes in the human disease relative to related preclinical models.

We recognise limitations of our data, as well as their considerable promise for further exploration. Although restricted to end-stage tissue *post mortem*, we attempted to maximise the dynamic range of pathology sampled by evaluating correlations with quantitative pathological measures across two anatomical regions and brains of different Braak stages. We tried to compensate for the sparse 10X technology sampling of the transcriptome by increasing the number of glial nuclei of interest and provided evidence for generalisability

of our results by demonstration that major co-expression modules, regulons and gene expression analyses discovered here also were represented in previous human *post mortem* snRNASeq datasets[8-10,21]. Even so, the nuclear transcriptome may be biased relative to that from a whole cell, potentially reducing the power to detect some genes reported from studies of related pathologies in mouse models[48]. The need to maximise detection sensitivity motivated us to enrich our sample for microglia and astrocytes for maximising the number of nuclei characterised and to make use of co-expression-based analyses, which rely less on detection of absolute expression levels than do single gene differential expression analyses.

Our work has extended previous studies defining AD-associated molecular pathology of glial cells substantively by describing proteostasis, metal ion homeostasis and inflammatory mechanisms in astrocytes and phagocytotic, proteostatic and autophagic pathways in microglia. We found that gene sets described in transgenic mouse models are variably represented in AD in the context of more complex glial phenotypes. Our data add to evidence that there are functionally distinct sub-types of astrocytes and microglia, with particular diversity amongst the former, in which we distinguished enrichments for gene sets associated with synaptic function, extracellular matrix formation, immune responses and control of metal ion homeostasis/redox state. While this diversity suggests multiple potential targets for therapeutic modulation, the complexity of human astroglia and microglial phenotypes simultaneously expressed in AD also needs to be taken into account.

Methods

Brain tissue

This study was carried out in accordance with the Regional Ethics Committee and Imperial College Use of Human Tissue guidelines. Cases were selected from the London Neurodegeneration (King's College London) and Parkinson's UK (Imperial College London) Brain Banks. Entorhinal and somatosensory cortex from 6 non-diseased control (NDC) cases (Braak stage 0-II) and 6 AD cases (Braak stage III-VI) were used (total of 24 brain samples). Cortical samples from two regions were prepared from each brain in order to characterise transcript expression with both higher (entorhinal cortex) and lower (somatosensory cortex) tissue densities of pTau in neurofibrillary tangles and amyloid-beta plaques. Brains used for this study excluded cases with clinical or pathological evidence for small vessel disease, stroke, cerebral amyloid angiopathy, diabetes, Lewy body pathology (TDP-43), or other neurological diseases. Where the information was available, cases were selected for a brain pH greater than 6 and all but one had a *post mortem* delay of less than 24 hr.

Table 1: Cohort information

	M:F ratio	**Age at death (yrs, mean +/- SD)**	**Post mortem delay (hr, mean +/- SD)**	**RIN (mean +/- SD)**
Non-diseased controls (Braak 0-II)	4:2	79.3 +/- 6.5	18 +/- 6.9	4.9 +/- 2.0
Alzheimer's disease (Braak III-VI)	4:2	81 +/- 6.8	22.1 +/- 15.9	7.1 +/- 0.7

Immunohistochemistry

Immunohistochemical staining was performed on formalin-fixed paraffin-embedded sections from homologous regions of each brain used for snRNASeq. Standard immunohistochemical procedures were followed using the ImmPRESS Polymer (Vector Laboratories) and SuperSensitive Polymer-HRP (Biogenex) kits (Table 2). Briefly, endogenous peroxidase activity and non-specific binding was blocked with 0.3% H_2O_2 and 10% normal horse serum, respectively. Primary antibodies were incubated overnight at 4°C. Species-specific ImmPRESS or SuperSensitive kits and DAB were used for antibody visualization. Counter-staining for nuclei was performed by incubating tissue sections in haematoxylin (TCS Biosciences)

for 2 min. AD pathology was assessed by amyloid plaque (4G8, BioLegend 17-24) and pTau (AT8, NBS Biologics) staining. GPNMB staining in microglia was assessed using R&D antibody AF2550.

Table 2. List of antibodies and immunostaining methods.

Antigen	**Antibody**	**Dilution**	**Antigen Retrieval**	**IHC Staining Kit**
Amyloid-β	4G8 BioLegend (800702)	1:15,000	Citrate Buffer, in Steamer	Supersensitive Kit
pTau	AT8 Invitrogen (MN1020)	1:1,600	Citrate Buffer, in Steamer	Supersensitive Kit
GPNMB	R&D (AF2550)	1:500	Citrate Buffer, in Steamer	ImmPRESS Kit

Quantitative image analysis

Labelled tissue sections were imaged using a Leica Aperio AT2 Brightfield Scanner (Leica Biosystems). Images were analysed using HALO software (Indica Labs, Version 2.3.2989.34). The following image analysis macros were used for the study: area quantification macro (amyloid), multiplex macro (pTau) and microglia macro (GPNMB). GraphPad Prism version 8 for Windows (GraphPad Software, La Jolla, CA, USA) was used for plotting immunohistochemical results and performing statistical analysis. A Mann Whitney test was used to test for the significance of differences between NDC and AD.

We found that densities of pTau staining were greater in brains from the higher Braak stages, and that pTau and amyloid pathology both were found in the regions studied even in brains with lower Braak stages (Extended Data Fig. 2). This variability of local tissue pathology within similar Braak stages emphasises the potential importance of matching regional neuropathology with transcript expression for each brain individually.

Nuclei isolation and selective glial enrichment

Fresh frozen entorhinal and somatosensory cortical tissue blocks were sectioned to 80 μm on a cryostat and grey matter separated by scoring the tissue with sharp forceps to collect ~200 mg grey matter in an RNAse-free Eppendorf tube. Nuclei from NDC and AD samples were isolated in parallel using a protocol based on Krishnaswami *et al.* (2016)[6]. All steps were carried out on ice or at 4°C. Tissue was homogenized in a 2 ml glass douncer containing homogenization buffer (0.1% Triton-X + 0.4 U/μl RNAseIn + 0.2 U/μl SUPERaseIn). The tissue homogenate was centrifuged at 1000 g for 8 min, and the majority of supernatant removed without disturbing the tissue pellet. Homogenised tissue was filtered through a 70 μm filter and centrifuged in an Optiprep (Sigma) density gradient at 13,000 g for 40 min to remove myelin and cellular debris. The nuclei pellet was washed and filtered twice in PBS buffer (PBS + 1% BSA + 0.2 U/ml RNAseIn). Isolated nuclei were labelled in suspension in 1 ml PBS buffer with 1:500 anti-NeuN antibody (Millipore, MAB377, mouse) and 1:250 anti-Sox10 antibody (R&D, AF2864, goat) for 1 hr on ice. Nuclei were washed twice with PBS buffer and centrifuged at 500 g for 5 min. Nuclei were incubated with Alexa-fluor secondary antibodies at 1:1000 (goat-anti-mouse-647 and donkey-anti-goat-488) and Dapi (1:1000) for 30 min on ice, and washed twice. Nuclei were FACS-sorted on a BD Aria II, using BD FACSDiva software, gating first for Dapi +ve nuclei, then singlets and then Sox10- and NeuN-negative nuclei. A minimum of 150,000 double-negative nuclei were collected.

Single nucleus capture and snRNA sequencing

Sorted nuclei were centrifuged at 500 g, resuspended in 50 μl PBS buffer and counted on a LUNA-FL Dual Fluorescence Cell Counter (Logos Biosystems, L20001) using Acridine Orange dye to stain nuclei. Sufficient nuclei were added for a target of 7,000 nuclei for each library prepared. Barcoding, cDNA synthesis and library preparation were performed using 10X Genomics Single Cell 3' Gene Expression kit v3 with 8 cycles of cDNA amplification, after which up to 25 ng of cDNA was taken through to the fragmentation reaction and a final indexing PCR was carried out to 14 cycles. cDNA concentrations were measured using Qubit dsDNA HS Assay Kit (ThermoFisher, Q32851), and cDNA and library preparations were assessed using the Bioanalyzer High-Sensitivity DNA Kit (Agilent, 5067-4627). Samples were pooled to equimolar concentrations and the pool sequenced across 24 lanes of an Illumina HiSeq 4000 according to

the standard 10X Genomics protocol. The snRNAseq data will be made available for download from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE160936.

Processing of FASTQ files, dimensionality reduction and clustering

snRNASeq data were pre-processed and clustered using 10X Genomics Cell Ranger and Seurat analysis tools[10,39,49]. Illumina sequencing files were aligned to the genomic sequence (introns and exons) using GRCh38 annotation in Cell Ranger v3.1. Nuclei were identified above background by the Cell Ranger software. Filtered gene matrices from CellRanger were loaded into R where Seurat v3 single-cell analysis package was used for analysis[50]. Genes that were expressed in three or more nuclei were used for further analysis. Further QC was performed to exclude nuclei with less than 200 genes or greater than 6000 genes or 25,000 UMIs, which likely represent low quality or doublet nuclei, respectively. Nuclei with greater than 5% mitochondrial genes were also excluded. Mitochondrial gene reads were excluded. The 24 samples had an average of 3819 nuclei per sample after passing QC.

Data was normalized and scaled using the *NormalizeData* function with normalisation.method = "LogNormalize" and scale.factor = 10,000. Variable genes then were identified using the *FindVariableFeatures* function with nfeatures (number of variable genes) set to 2000. To integrate the data from all samples, *FindIntegrationAnchors* (dims = 1:20) and *IntegrateData* (dims = 1:20) functions were used. PCA analysis was run, using variable genes, for the top 30 components. Clusters were identified using *FindClusters* (resolution = 0.5) and UMAP was used for 2D visualization of clusters (with the top 15 PCs, based on the ranking of PCs by the variance explained by each).

Cell-type identification of clusters was performed using AUCell (see below) with cell-type specific genes identified by previous human brain snRNASeq studies [51]. Cell type annotation was confirmed by visual inspection of key marker genes (Fig. 1 and Extended Data Fig. 1). Glial cluster specific genes were identified using the *FindMarkers* function. All clusters were composed of nuclei from all samples and did

not represent a single case or disease group. A small number of nuclei that either did not express any major cell-type markers, or expressed a combination of cell-type markers, were categorized as 'unclassified'. We used the thresholding method described above for identification and removal of unclassified clusters (Extended Data Fig. 1).

Differential gene expression analysis

MAST was used to identify genes differentially expressed (associated) with histopathological features (using pTau or amyloid-beta as markers)[42,52] to perform zero-inflated regression analysis by fitting a mixed model. The model specification was zlm(~histopath_marker + (1|sample) + cngeneson + pc_mito + sex + brain_region, sca, method = "glmer", ebayes = F). The fixed effect term pc_mito accounts for the percentage of counts mapping to mitochondrial genes. The term cngeneson is the cellular detection rate. Each nuclei preparation was considered as a distinct sample for the mixed effect. Models were fit with and without the dependent variable and compared using a likelihood ratio test. Units for differential expression are defined as $\log_2$ fold difference/% pTau positive cells (or $\log_2$ fold difference/% amyloid area), i.e., a one unit change in immunohistochemically-defined pTau (or amyloid density) is associated with one $\log_2$ fold change in gene expression. Genes expressed in at least 10% of nuclei from each cell type (either total microglia or total astrocytes) were tested. Genes with a $\log_2$ fold-change of at least 0.25 and adjusted p-value <0.05 were defined as meaningfully differentially expressed. As an additional filter, the percentage of the inter-individual variance in expression between the NDC subjects was calculated for each gene and three genes with unusually high (>2 standard deviations) variance (*LINGO1*, *SLC26A3* and *RASGEF1B*) in one or two samples were excluded.

Gene ontology and pathway enrichment analysis

The gene ontology (GO) enrichment and the pathway enrichments analysis were carried out using the R package enrichR (v 3.0), which uses Fisher's exact test (Benjamini-Hochberg FDR < 0.1)[53,54]. Genesets with minimum and maximum genes of 10 and 500 respectively were considered. To improve biological interpretation of functionally related gene ontology and pathway terms and to reduce the number of

redundant gene sets, we first calculated a pairwise distance matrix using Cohen's kappa statistics based on the overlapping genes between the enriched terms and then performed hierarchical clustering of the enriched terms[55].

Gene Set Enrichment Analyses (GSEA)

AUCell[56] (R package v1.6.1) was used to quantify the expression of published gene set signatures. Mouse genes were converted to human orthologues where applicable using bioDBnet. Normalised data was processed in AUCell using the *AUCell_buildRankings* function. The resulting rankings, along with the gene lists of interest, were then put into the function *AUCell_calcAUC* (aucMaxRank set to 1% of the number of input genes). Resulting AUC scores were scaled across clusters and plotted in a heatmap (Extended Data Fig. 7).

We also used AUCell to test for enrichment of the gene sets that we identified on previously published human single-nuclei data [8-10,21]. Where possible for the AUCell tests for enrichment of gene sets from previously reported data, the cell type annotations from the published data were used. Filtered matrices were processed and cell types identified using the methods described above for the Zhou *et al.*[10] data set. The scFlow single cell analysis pipeline (https://nf-co.re/scflow) was employed for analysis of the Gerrits *et al.* data set. In brief, quality control was performed using the same criteria as described above. Sample integration was performed using Liger[57] and dimension reduction using PCA and UMAP and finally clustering was performed using the Leiden algorithm[58]. Microglia and astrocytes were then identified using the same sets of marker genes used for the primary analysis (Fig. 1). AUCell was then run separately on microglia and astrocyte populations from each study using lists of our significantly up-regulated and down regulated genes with pTau and amyloid pathology, from microglia and astrocytes (thresholded as described above). The aucMaxRank term was set to 200 genes. LogFC values between Control and AD samples were estimated using the limma package in R[59], using the default configuration and the following linear model: ~diagnosis+nFeature, where nFeature is the total number of distinct features expressed in each nucleus (to account for the fact that nuclei that express a higher number of features may have higher AUCell scores).

Over-representation analyses of literature gene sets in our glial sub-clusters and differentially expressed gene lists were performed by Fisher's exact test using the "enrichment" function of the R package "bc3net" (https://github.com/cran/bc3net). The p values associated with the Fisher's exact test correspond to the probability that the overlap between the literature gene sets and the sub-cluster markers/differentially regressed genes from our dataset has occurred by chance.

Gene co-expression (module), regulatory network (regulon) and enrichment analyses

Gene co-expression module and hub-gene identification analysis were performed separately for microglial and astrocyte populations using the MEGENA (v1.3.7) package[59]. The top 15% most variably expressed genes were used as input[13]. We evaluated the mean expression of each of these genes across all the nuclei in the expression matrices: all the genes in both astrocyte and microglia matrices were among the top 25% most highly expressed genes confirming that the choice to filter the expression matrix based on variability did not bias the inclusion towards genes with a particularly low level of expression. In addition, we verified that the filtered expression matrices for both astrocytes and microglia included a substantial proportion of the differentially expressed genes: over 90% for astrocytes and all but two of the microglial differentially expressed genes were included in the respective filtered expression matrices, suggesting that the filtered expression matrices contained a biologically meaningful gene set. The MEGENA pipeline then was applied using default parameters, using Pearson's correlations and a minimum module size of 10 genes. Parent modules were produced from which a sub-set of genes form smaller child modules. For downstream analysis, interpretation and presentation of results, modules with >20 genes were retained. Co-expression modules were represented graphically using Cytoscape software (Mac OS version 3.8.0)[60] with hub genes represented with a triangle and nodes with a circle with a diameter proportional to the node degree[59]. Genes previously associated with AD as defined by Kunkle *et al.* (Table 1[19], 23 genes), Jansen *et al.* (Table 1[18], a further 17 genes) and Andrews *et al.* (Table 1[1], a further 25 genes), for a total of 65 genes, were annotated in the co-expression networks.

Gene regulatory networks were built using pySCENIC (0.10.3) [61 62] package default parameters. The inputs for the pySCENIC gene regulatory network analyses were the same filtered expression matrices as for the MEGENA gene co-expression module analyses. Correlations between a list of 1390 human transcription factors (TFs) curated by Lambert et al (2018)[63] and the genes in the expression matrix were evaluated and co-expression modules with a minimum size of 20 genes were defined. From these, regulons (gene modules sharing a common association with a TF) were built after removing the genes without a recognition motif (based on the hg19-500bp-upstream-7species.mc9nr and hg19-tss-centered-10kb-7species.mc9nr databases provided in the pySCENIC package) for the correlated TF. Only regulons with activator-TFs were retained[61]. Regulons including 50 or more genes were retained for downstream analyses and Cytoscape software (Mac OS version 3.8.0)[60] was used for their graphical representation.

To evaluate module and regulon enrichment with AD pathology, the AUCell scores for each gene co-expression module or regulon in each nucleus were calculated (aucMaxRank set to 5% of the number of input genes). The statistical analysis was performed using the limma package in R[64], using the default configuration and the following linear model: ~pathology+nFeature+pc_mito, where pathology is the average immunohistochemistry quantification value for Aβ or pTau, nFeatures is the total number of distinct features expressed in each nucleus (to account for the fact that nuclei that express a higher number of features may have higher AUCell scores) and pc_mito is the percentage of counts mapping to mitochondrial genes. We also corrected for a potential pseudoreplication bias[65], by using the duplicateCorrelation function of the limma package with the sample as the “blocking” variable.

Sub-cluster-specificity of modules and regulons were estimated using the regulon_specificity_scores function in pySCENIC[66]. Briefly, the module/regulon specificity score employs the Jensen-Shannon Divergence, a metric previously used to assess cell type specificity of transcripts[67] and regulons[66]. Modules and regulons with the highest specificity score may be considered subcluster-specific. A specificity score of

1 indicates a gene set that is only expressed in one sub-cluster while a specificity score of 0 indicates an evenly expressed gene set across all sub-clusters.

Acknowledgements

We thank the donors and their families for the use of human brain tissue in this study and the London Neurodegeneration and Parkinson's UK Brain Bank staff for making it available. Infrastructure, including particularly the LMS/NIHR_Imperial Biomedical Research Centre Flow Cytometry Facility and the Imperial BRC Genomics Facility, were supported by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC). ST was supported by an "Early Postdoc.Mobility" scholarship (P2GEP3_191446) from the Swiss National Science Foundation and a "Clinical Medicine Plus" scholarship from the Prof Dr Max Cloëtta Foundation (Zurich, Switzerland). PMM acknowledges generous personal support from the Edmond J Safra Foundation and Lily Safra and an NIHR Senior Investigator Award. This work was supported by the UK Dementia Research Institute, which receives its funding from UK DRI Ltd., funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK.

Author Contributions

Amy Smith – experimental design, performed experiments, data analysis, interpretation of results, wrote the paper

Karen Davey – performed experiments, data analysis, interpretation of results

Stergios Tsartsalis – co-expression data analysis and visualisation, interpretation of results

Combiz Khozoie – bioinformatic data analysis, differential expression analysis

Nurun Fancy – enrichment analysis and visualisation, interpretation of results

Elijah See – performed immunohistochemistry experiments and image analysis

Eirini Liaptsi – performed immunohistochemistry experiments and image analysis

Maria Weinert – performed optimisation experiments

Aisling McGarry - performed immunohistochemistry experiments

Callum Muirhead – performed immunohistochemistry experiments

Steve Gentleman – cohort selection and neuropathology

David Owen – experimental design, interpretation of results

Paul M. Matthews- experimental design, interpretation of results, wrote the paper

All of the co-authors reviewed and approved the manuscript.

Competing Interests

PMM has received consultancy fees from Roche, Adelphi Communications, Celgene, Neurodiem and Medscape. He has received honoraria or speakers' fees from Novartis and Biogen and has received research or educational funds from Biogen, Novartis and GlaxoSmithKline.

Figures

Figure 1: Analysis of human brain microglia and astrocytes from low and high AD pathology brains by single-nuclei RNASeq.

a) FACS gating method for sorting of human brain nuclei to enrich for microglia and astrocyte double-negative population (pink; mean for 24 samples = 17.9%, standard deviation = 6.2%). Dapi+ve nuclei were selected first, followed by separation based on NeuN and Sox10 staining.

b) UMAP 2D visualisation of clustering of 91,655 single nuclei from the 24 brain samples (average of 3,819 nuclei per sample) including 52,706 (58%) astrocytes and 27,592 (30%) microglia. Smaller numbers of nuclei from neurons, oligodendrocytes and vascular cells also were found (Extended Data Fig. 1), but these cell types formed distinct UMAP clusters that were not analysed further.

c) Heatmap showing cell type-specific marker gene expression in the nuclei clusters.

d) Heatmap of top differentially expressed genes in each microglial cluster.

e) Heatmap of top differentially expressed genes in each astrocyte cluster.

Figure 2: Differential gene expression in astrocytes and microglia with amyloid-beta and pTau pathology.

a,b) Volcano plot of transcripts differentially expressed in astrocyte nuclei (threshold of $\leq$0.05 adjusted p-value, abs logFC$\geq$ 0.25, omitting the top three most variable genes between samples) with immunohistochemically-defined tissue amyloid-beta (a) and pTau (b) density.

c) Venn diagram illustrating the number of genes positively correlated (top) and negatively correlated (bottom) with amyloid-beta (blue) and pTau pathology (pink) in astrocytes.

d,e) Volcano plot of transcripts differentially expressed in microglial nuclei (threshold of $\leq$0.05 adjusted p-value, abs logFC$\geq$ 0.25, omitting the top three most variable genes between samples) with immunohistochemically-defined tissue amyloid-beta (d) and pTau (e) density.

f) Venn diagram illustrating the number of genes positively correlated (top) and negatively correlated (bottom) with amyloid-beta (blue) and pTau (pink) pathology in microglia

Figure 3: Functional enrichment of differential gene expression with amyloid-beta (left) and pTau (right) pathology in astrocytes. The plots describe the significant functionally enriched pathways in astrocytes obtained using enrichR (see Methods) from Gene Ontology (GO), Reactome and Wikipathways (WP) databases

Figure 4: Functional enrichment of differential gene expression with amyloid-beta (left) and pTau (right) pathology in microglia. The plots describe the significant functionally enriched pathways in microglia obtained using enrichR (see Methods) from Gene Ontology (GO), Reactome and Wikipathways (WP) databases

Figure 5: *GPNMB* is a hub gene in microglial gene co-expression modules up-regulated with AD pathology.

a) Graph of microglial gene co-expression module 11 enriched with amyloid-beta and pTau pathology for which *GPNMB* is a hub gene (Triangles = module hub genes). *GPNMB* is expressed in regulons identified by their transcription factors (in green).

b) Immunohistochemical staining of GPNMB in *post mortem* human brain tissue from representative NDC and AD brains with nuclear counterstain. Staining is present in both microglia and perivascular macrophages (insets). Scale = 100μm.

c) Quantification of density of GPNMB-positive cells in cortical tissue by automated image analysis showing an increase in GPNMB-positive cells per μm^2 for AD cases compared to NDC ($p = 0.0023$; Mann Whitney test, two-tailed; N = 12). Each point represents a single sample and the horizontal bar indicates the median.

References

1 Andrews, S. J., Fulton-Howard, B. & Goate, A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. *The Lancet Neurology* **19**, 326-335, doi:https://doi.org/10.1016/S1474-4422(19)30435-1 (2020).

2 Serrano-Pozo, A. *et al.* Differential relationships of reactive astrocytes and microglia to fibrillar amyloid deposits in Alzheimer disease. *J Neuropathol Exp Neurol* (2013).

3 Efthymiou, A. G. & Goate, A. M. Late onset Alzheimer's disease genetics implicates microglial pathways in disease risk. *Molecular neurodegeneration* **12**, 43-43, doi:10.1186/s13024-017-0184-x (2017).

4 Skene, N. G. & Grant, S. G. N. Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. *Frontiers in Neuroscience* **10**, 16 (2016).

5 Wan, Y.-W. *et al.* Meta-Analysis of the Alzheimer's Disease Human Brain Transcriptome and Functional Dissection in Mouse Models. *Cell reports* **32**, 107908-107908, doi:10.1016/j.celrep.2020.107908 (2020).

6 Krishnaswami, S. R. *et al.* Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nature Protocols* **11**, 499-524, doi:10.1038/nprot.2016.015 (2016).

7 Lake, B. B. *et al.* Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586 LP-1590, doi:10.1126/science.aaf1204 (2016).

8 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* **22**, 2087-2097, doi:10.1038/s41593-019-0539-4 (2019).

9 Mathys, H. *et al.* Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332-337, doi:10.1038/s41586-019-1195-2 (2019).

10 Zhou, Y. *et al.* Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nature medicine* **26**, 131-142, doi:10.1038/s41591-019-0695-9 (2020).

11 Nguyen, A. T. *et al.* APOE and TREM2 regulate amyloid-responsive microglia in Alzheimer's disease. *Acta Neuropathol* **140**, 477-493, doi:10.1007/s00401-020-02200-3 (2020).

12 Sankowski, R. *et al.* Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nature Neuroscience* **22**, 2098-2110, doi:10.1038/s41593-019-0532-y (2019).

13 Al-Dalahmah, O. *et al.* Single-nucleus RNA-seq identifies Huntington disease astrocyte states. *Acta Neuropathologica Communications* **8**, 19-19, doi:10.1186/s40478-020-0880-6 (2020).

14 Das, S., Li, Z., Noori, A., Hyman, B. T. & Serrano-Pozo, A. Meta-analysis of mouse transcriptomic studies supports a context-dependent astrocyte reaction in acute CNS injury versus neurodegeneration. *J Neuroinflammation* **17**, 227-227, doi:10.1186/s12974-020-01898-y (2020).

15 Yeung, M. S. Y. *et al.* Dynamics of oligodendrocyte generation in multiple sclerosis. *Nature* **566**, 538-542, doi:10.1038/s41586-018-0842-3 (2019).

16 Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37-53 (2016).

17 Patir, A., Shih, B., McColl, B. W. & Freeman, T. C. A core transcriptional signature of human microglia: Derivation and utility in describing region-dependent alterations associated with Alzheimer's disease. *Glia* **67**, 1240-1253, doi:10.1002/glia.23572 (2019).

18 Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer’s disease risk. *Nature Genetics* **51**, 404-413, doi:10.1038/s41588-018-0311-9 (2019).

19 Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics* **51**, 414-430, doi:10.1038/s41588-019-0358-2 (2019).

20 Vilalta, A. & Brown, G. C. Neurophagy, the phagocytosis of live neurons and synapses by glia, contributes to brain development and disease. *FEBS J* **285**, 3566-3575, doi:10.1111/febs.14323 (2018).

21 Gerrits, E. *et al.* Distinct amyloid-β and tau-associated microglia profiles in Alzheimer's disease. *Acta Neuropathologica* **141**, 681-696, doi:10.1007/s00401-021-02263-w (2021).

22 Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nature Genetics* **49**, 1511-1516, doi:10.1038/ng.3955 (2017).

23 Ward, M. E. *et al.* Individuals with progranulin haploinsufficiency exhibit features of neuronal ceroid lipofuscinosis. *Sci Transl Med* **9**, eaah5642, doi:10.1126/scitranslmed.aah5642 (2017).

24 Baker, M. *et al.* Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature* **442**, 916-919, doi:10.1038/nature05016 (2006).

25 Hüttenrauch, M. *et al.* Glycoprotein NMB: a novel Alzheimer's disease associated marker expressed in a subset of activated microglia. *Acta neuropathologica communications* **6**, 108-108, doi:10.1186/s40478-018-0612-3 (2018).

26 Naba, A. *et al.* The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Molecular & cellular proteomics : MCP* **11**, M111.014647-M014111.014647, doi:10.1074/mcp.M111.014647 (2012).

27 Habib, N. *et al.* Disease-associated astrocytes in Alzheimer's disease and aging. *Nature Neuroscience* **23**, 701-706, doi:10.1038/s41593-020-0624-8 (2020).

28 Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481-487, doi:10.1038/nature21029 (2017).

29 Kim, W.-K. *et al.* CD163 identifies perivascular macrophages in normal and viral encephalitic brains and potential precursors to perivascular macrophages in blood. *Am J Pathol* **168**, 822-834, doi:10.2353/ajpath.2006.050215 (2006).

30 Galatro, T. F. *et al.* Transcriptomic analysis of purified human cortical microglia reveals age-associated changes. *Nature neuroscience* **20**, 1162-1171, doi:10.1038/nn.4597 (2017).

31 Chen, W.-T. *et al.* Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer’s Disease. *Cell* **182**, 976-991.e919, doi:10.1016/j.cell.2020.06.038 (2020).

32 Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169**, 1276-1290.e1217, doi:10.1016/j.cell.2017.05.018 (2017).

33 Sala Frigerio, C. *et al.* The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques. *Cell reports* **27**, 1293-1306.e1296, doi:10.1016/j.celrep.2019.03.099 (2019).

34 Roy, E. R. *et al.* Type I interferon response drives neuroinflammation and synapse loss in Alzheimer disease. *The Journal of Clinical Investigation* **130**, 1912-1930, doi:10.1172/JCI133737 (2020).

35 Olah, M. *et al.* A transcriptomic atlas of aged human microglia. *Nature Communications* **9**, 539-539, doi:10.1038/s41467-018-02926-5 (2018).

36 Sierksma, A. *et al.* Novel Alzheimer risk genes determine the microglia response to amyloid-β but not to TAU pathology. *EMBO molecular medicine* **12**, e10606-e10606, doi:10.15252/emmm.201910606 (2020).

37 Rohne, P., Prochnow, H., Wolf, S., Renner, B. & Koch-Brandt, C. The Chaperone Activity of Clusterin is Dependent on Glycosylation and Redox Environment. *Cellular Physiology and Biochemistry* **34**, 1626-1639, doi:10.1159/000366365 (2014).

38 Nott, A. *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science (New York, N.Y.)* **366**, 1134-1139, doi:10.1126/science.aay0793 (2019).

39 Gerrits, E., Heng, Y., Boddeke, E. W. G. M. & Eggen, B. J. L. Transcriptional profiling of microglia; current state of the art and future perspectives. *Glia* **68**, 740-755, doi:10.1002/glia.23767 (2020).

40 van der Poel, M. *et al.* Transcriptional profiling of human microglia reveals grey-white matter heterogeneity and multiple sclerosis-associated changes. *Nature communications* **10**, 1139-1139, doi:10.1038/s41467-019-08976-7 (2019).

41 Castanho, I. *et al.* Transcriptional Signatures of Tau and Amyloid Neuropathology. *Cell Rep* **30**, 2040-2054 e2045, doi:10.1016/j.celrep.2020.01.063 (2020).

42 Finak, G. *et al.* MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology* **16**, 278-278, doi:10.1186/s13059-015-0844-5 (2015).

43 Sharma, S. & Ebadi, M. Significance of metallothioneins in aging brain. *Neurochem Int* **65**, 40-48, doi:10.1016/j.neuint.2013.12.009 (2014).

44 Hozumi, I. Roles and therapeutic potential of metallothioneins in neurodegenerative diseases. *Curr Pharm Biotechnol* **14**, 408-413, doi:10.2174/1389201011314040004 (2013).

45 Hao, Q. & Maret, W. Imbalance between pro-oxidant and pro-antioxidant functions of zinc in disease. *J Alzheimers Dis* **8**, 161-170; discussion 209-115, doi:10.3233/jad-2005-8209 (2005).

46 Schartz, N. D. & Tenner, A. J. The good, the bad, and the opportunities of the complement system in neurodegenerative disease. *Journal of Neuroinflammation* **17**, 354, doi:10.1186/s12974-020-02024-8 (2020).

47 Streit, W. J., Braak, H., Xue, Q.-S. & Bechmann, I. Dystrophic (senescent) rather than activated microglial cells are associated with tau pathology and likely precede neurodegeneration in Alzheimer's disease. *Acta neuropathologica* **118**, 475-485, doi:10.1007/s00401-009-0556-6 (2009).

48 Thrupp, N. *et al.* Single nucleus sequencing fails to detect microglial activation in human tissue. *bioRxiv*, 2020.2004.2013.035386-032020.035304.035313.035386, doi:10.1101/2020.04.13.035386 (2020).

49 Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e722, doi:https://doi.org/10.1016/j.cell.2019.06.029 (2019).

50 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).

51 Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75-82, doi:10.1038/s41586-019-1404-z (2019).

52 Sorrells, S. F. *et al.* Immature excitatory neurons develop during adolescence in the human amygdala. *Nature Communications* **10**, 2748-2748, doi:10.1038/s41467-019-10765-1 (2019).

53 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, doi:10.1186/1471-2105-14-128 (2013).

54 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

55 Huang, D. W. *et al.* The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology* **8**, R183, doi:10.1186/gb-2007-8-9-r183 (2007).

56 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083-1086, doi:10.1038/nmeth.4463
10.1038/nmeth.4463. Epub 2017 Oct 9. (2017).

57 Welch, J. D. *et al.* Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* **177**, 1873-1887.e1817, doi:10.1016/j.cell.2019.05.006 (2019).

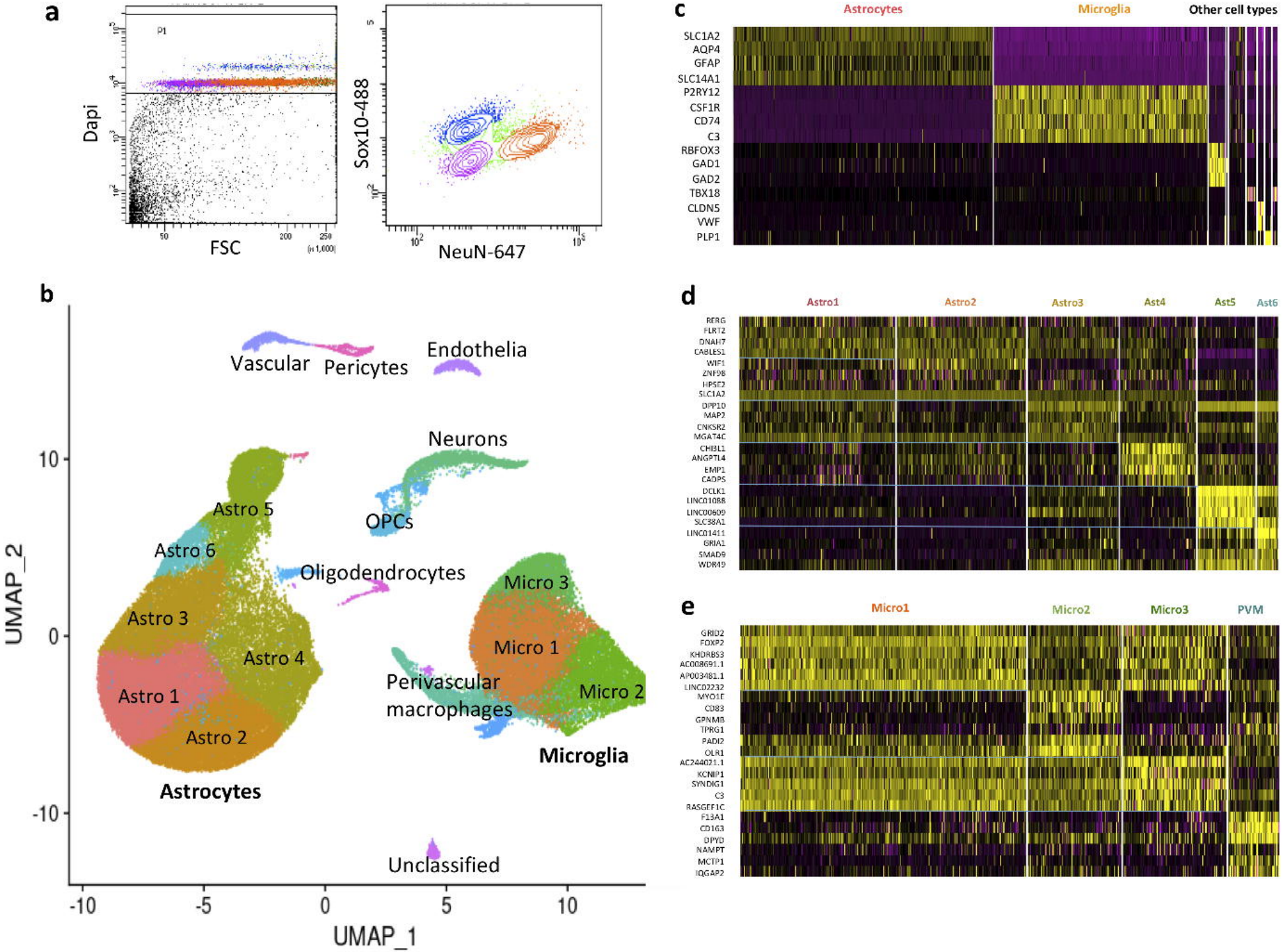
58 Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* **9**, 5233, doi:10.1038/s41598-019-41695-z (2019).

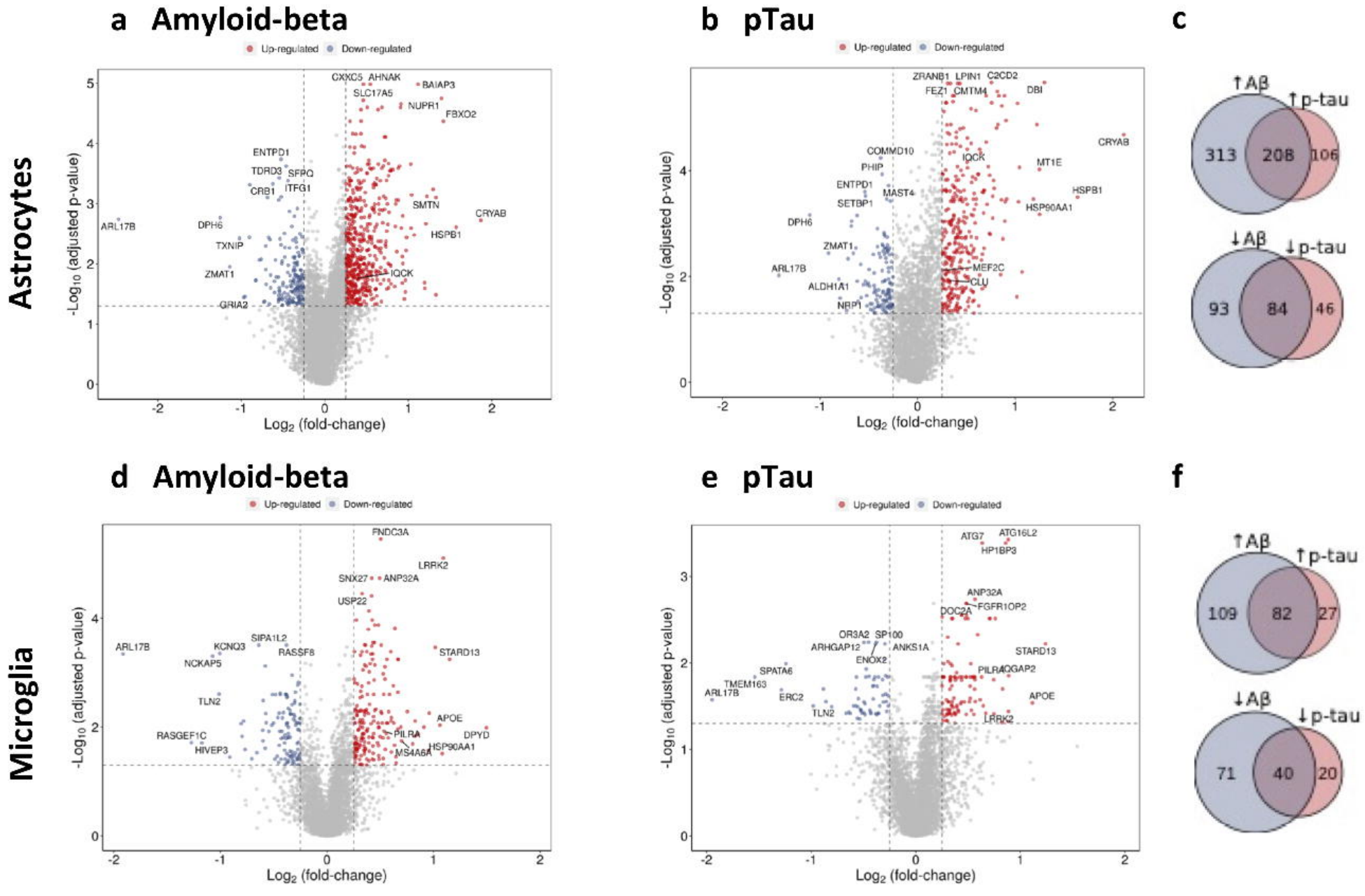
59 Song, W. M. & Zhang, B. Multiscale Embedded Gene Co-expression Network Analysis. *PLoS Comput Biol* **11**, e1004574-e1004574, doi:10.1371/journal.pcbi.1004574
10.1371/journal.pcbi.1004574. eCollection 2015 Nov. (2015).

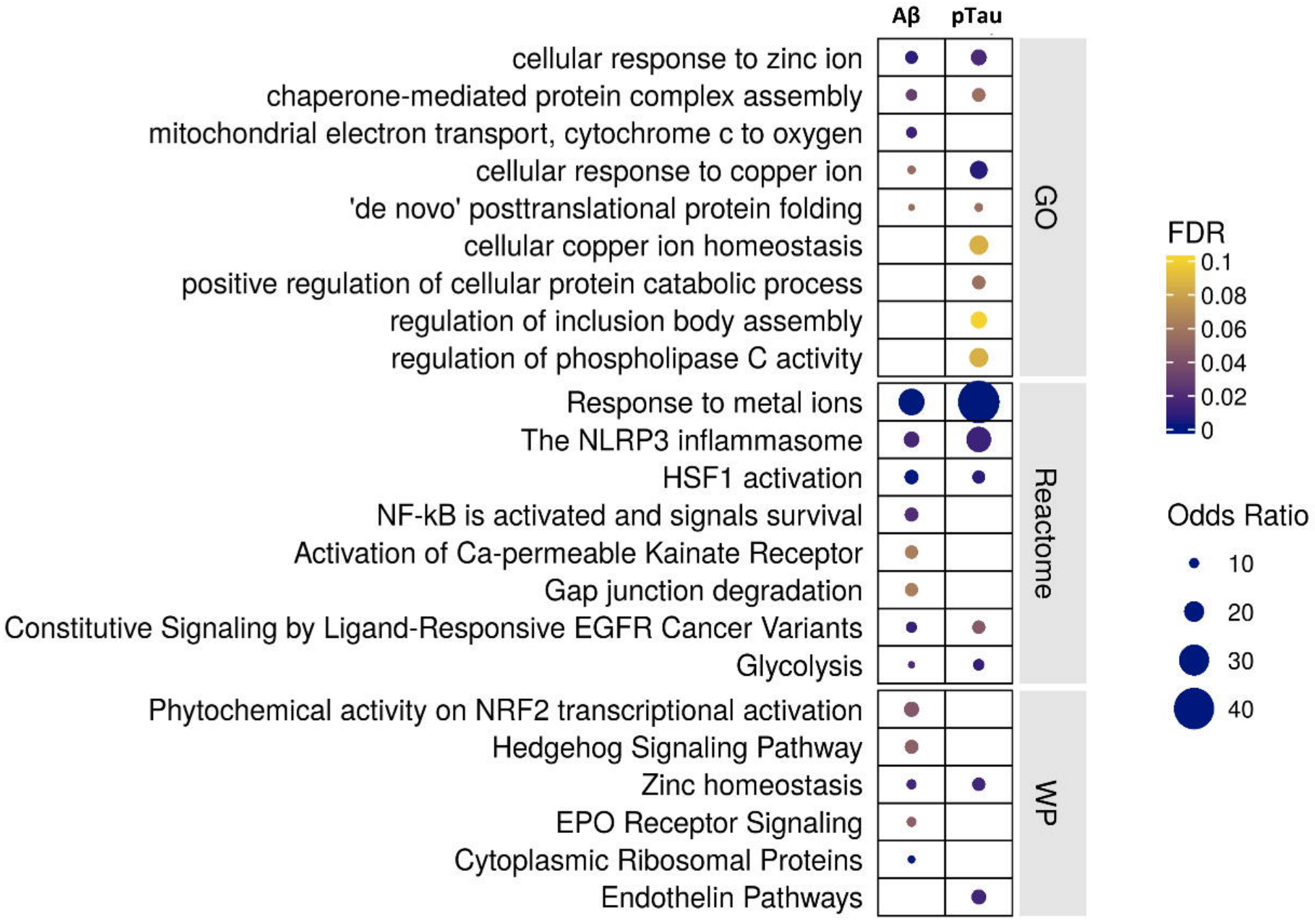
60 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303
10.1101/gr.1239303. (2003).

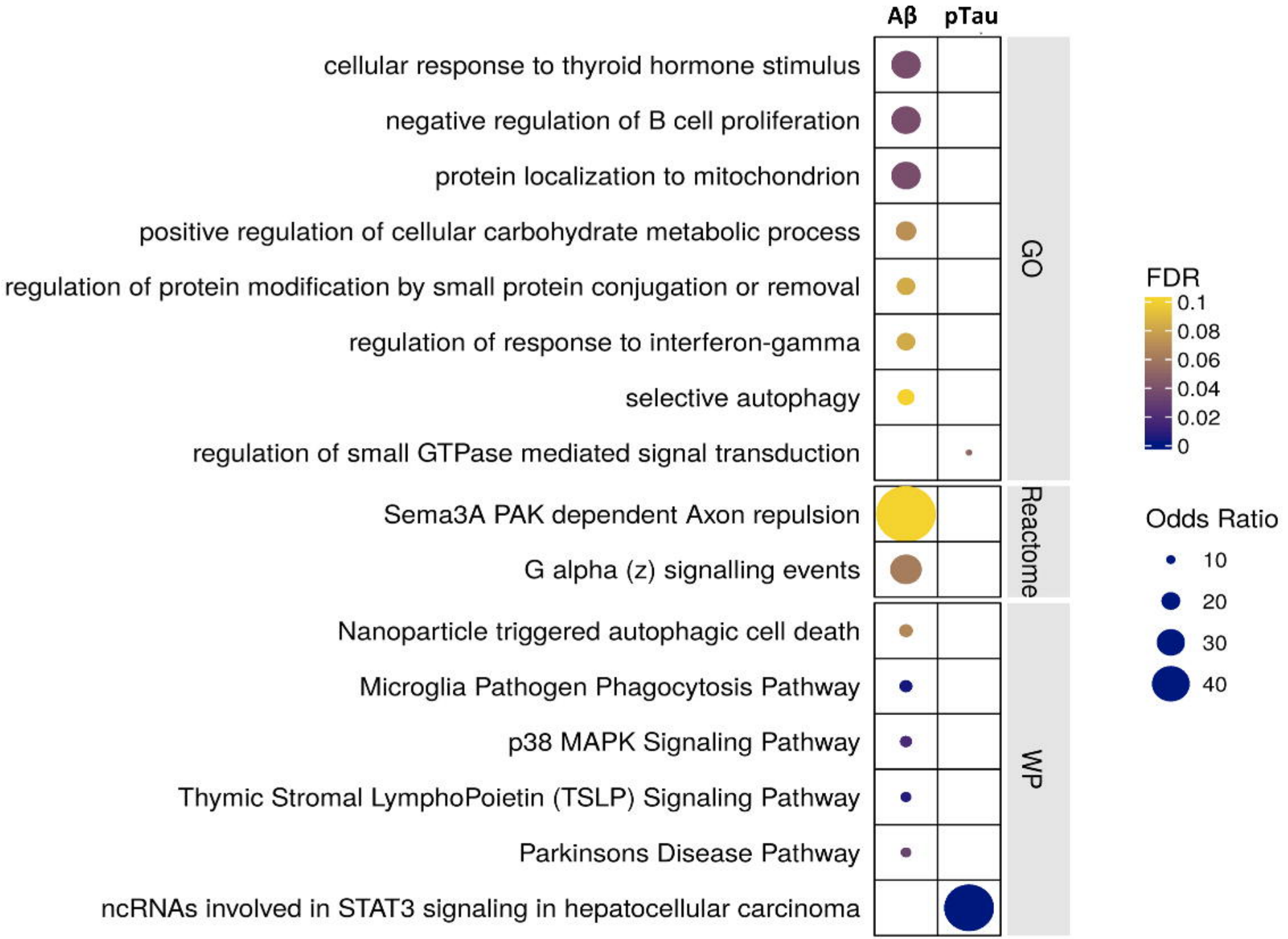
61 Van de Sande, B. *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* **15**, 2247-2276, doi:10.1038/s41596-020-0336-2
10.1038/s41596-020-0336-2. Epub 2020 Jun 19. (2020).

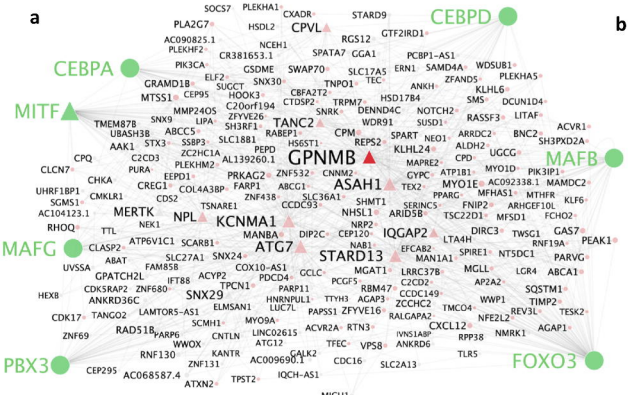
62 Moerman, T. *et al*. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159-2161, doi:10.1093/bioinformatics/bty916
10.1093/bioinformatics/bty916. (2019).
63 Lambert, S. A. *et al*. The Human Transcription Factors. *Cell* **175**, 598-599, doi:10.1016/j.cell.2018.09.045
10.1016/j.cell.2018.09.045. (2018).
64 Ritchie, M. E. *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* (2015).
65 Zimmerman, K. D., Espeland, M. A. & Langefeld, C. D. A practical solution to pseudoreplication bias in single-cell studies. *Nature Communications* **12**, 738, doi:10.1038/s41467-021-21038-1 (2021).
66 Suo, S. *et al*. Revealing the Critical Regulators of Cell Identity in the Mouse Cell Atlas. *Cell Rep* **25**, 1436-1445 e1433, doi:10.1016/j.celrep.2018.10.045
10.1016/j.celrep.2018.10.045. (2018).
67 Cabili, M. N. *et al*. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* **25**, 1915-1927, doi:10.1101/gad.17446611
10.1101/gad.17446611. Epub 2011 Sep 2. (2011).

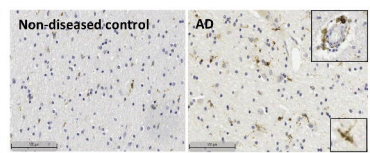

a
P1
Dapi
FSC
Sox10-488
NeuN-647
b
Vascular
Pericytes
Endothelia
Neurons
OPCs
Astro 5
Astro 6
Oligodendrocytes
Astro 3
Astro 4
Astro 1
Astro 2
Micro 3
Micro 1
Micro 2
Perivascular macrophages
Microglia
Astrocytes
Unclassified
UMAP_2
UMAP_1
c
Astrocytes
Microglia
Other cell types
SLC1A2
AQP4
GFAP
SLC14A1
P2RY12
CSF1R
CD74
C3
RBFOX3
GAD1
GAD2
TBX18
CLDN5
VWF
PLP1
d
Astro1
Astro2
Astro3
Ast4
Ast5
Ast6
RERG
FLRT2
DNAH7
CABLES1
WIF1
ZNF98
HPSE2
SLC1A2
DPP10
MAP2
CNKSR2
MGAT4C
CHI3L1
ANGPTL4
EMP1
CADPS
DCLK1
LINC01088
LINC00609
SLC38A1
LINC01411
GRIA1
SMAD9
WDR49
e
Micro1
Micro2
Micro3
PVM
GRID2
FOXP2
KHDRBS3
AC008691.1
AP003481.1
LINC02232
MYO1E
CD83
GPNMB
TPRG1
PADI2
OLR1
AC244021.1
KCNIP1
SYNDIG1
C3
RASGEF1C
F13A1
CD163
DPYD
NAMPT
MCTP1
IQGAP2

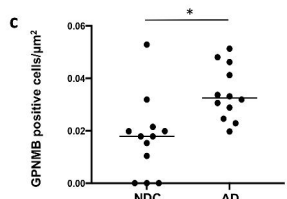

Astrocytes
Microglia
a Amyloid-beta
b pTau
c
d Amyloid-beta
e pTau
f
Up-regulated
Down-regulated
-Log10 (adjusted p-value)
Log2 (fold-change)
CXXC5
AHNAK
SLC17A5
BAIAP3
NUPR1
FBXO2
ENTPD1
TDRD3
SFPQ
CRB1
ITFG1
SMTN
CRYAB
ARL17B
DPH6
HSPB1
TXNIP
ZMAT1
IQCK
GRIA2
ZRANB1
LPIN1
C2CD2
FEZ1
CMTM4
DBI
COMMD10
PHIP
MT1E
MAST4
SETBP1
HSP90AA1
ALDH1A1
NRP1
MEF2C
CLU
FNDC3A
LRRK2
SNX27
ANP32A
USP22
SIPA1L2
KCNQ3
RASSF8
NCKAP5
STARD13
TLN2
APOE
RASGEF1C
PILRA
DPYD
HIVEP3
MS4A6A
ATG7
ATG16L2
HP1BP3
DOC2A
FGFR1OP2
OR3A2
SP100
ARHGAP12
ANKS1A
ENOX2
SPATA6
TMEM163
ERC2
↑Aβ
↑p-tau
313
208
106
↓Aβ
↓p-tau
93
84
46
109
82
27
71
40
20


Aβ
pTau
cellular response to zinc ion
chaperone-mediated protein complex assembly
mitochondrial electron transport, cytochrome c to oxygen
cellular response to copper ion
'de novo' posttranslational protein folding
cellular copper ion homeostasis
positive regulation of cellular protein catabolic process
regulation of inclusion body assembly
regulation of phospholipase C activity
GO
Response to metal ions
The NLRP3 inflammasome
HSF1 activation
NF-kB is activated and signals survival
Activation of Ca-permeable Kainate Receptor
Gap junction degradation
Constitutive Signaling by Ligand-Responsive EGFR Cancer Variants
Glycolysis
Reactome
Phytochemical activity on NRF2 transcriptional activation
Hedgehog Signaling Pathway
Zinc homeostasis
EPO Receptor Signaling
Cytoplasmic Ribosomal Proteins
Endothelin Pathways
WP
FDR
0.1
0.08
0.06
0.04
0.02
0
Odds Ratio
10
20
30
40


Aβ
pTau
cellular response to thyroid hormone stimulus
negative regulation of B cell proliferation
protein localization to mitochondrion
positive regulation of cellular carbohydrate metabolic process
regulation of protein modification by small protein conjugation or removal
regulation of response to interferon-gamma
selective autophagy
regulation of small GTPase mediated signal transduction
Sema3A PAK dependent Axon repulsion
G alpha (z) signalling events
Nanoparticle triggered autophagic cell death
Microglia Pathogen Phagocytosis Pathway
p38 MAPK Signaling Pathway
Thymic Stromal LymphoPoietin (TSLP) Signaling Pathway
Parkinsons Disease Pathway
ncRNAs involved in STAT3 signaling in hepatocellular carcinoma
GO
Reactome
WP
FDR
0.1
0.08
0.06
0.04
0.02
0
Odds Ratio
10
20
30
40


a
CEBPD
CEBPA
MITF
MAFB
MAFG
PBX3
FOXO3
GPNMB
ASAH1
CPVL
TANC2
KCNMA1
IQGAP2
ATG7
STARD13
NPL
MERTK
SNX29
SOCS7
PLEKHA1
CXADR
STARD9
PLA2G7
HSDL2
GTF2IRD1
AC090825.1
RGS12
PLEKHF2
NCEH1
CR381653.1
SPATA7
GGA1
PCBP1-AS1
PIK3CA
GSDME
SWAP70
SLC17A5
ERN1
SAMD4A
WDSUB1
ELF2
SNX30
TNPO1
TEC
ZFAND5
PLEKHA5
GRAMD1B
SUGCT
ANKH
KLHL6
MTSS1
CEP95
HOOK3
CBFA2T2
HSD17B4
SMS
DCUN1D4
C20orf194
CTDSP2
TRPM7
MMP24OS
SNRK
DENND4C
NOTCH2
ZFYVE26
RASSF3
LITAF
TMEM87B
SNX9
LIPA
SH3RF1
WDR91
SUSD1
ACVR1
UBASH3B
ABCC5
RABEP1
CPM
SPART
NEO1
ARRDC2
BNC2
SH3PXD2A
AAK1
STX3
SSBP3
SLC18B1
HS6ST1
REPS2
KLHL24
ALDH2
UGCG
CPQ
C2CD3
ZC2HC1A
PEPD
AL139260.1
CPD
MAPRE2
CLCN7
PURA
PLEKHM2
ZNF532
CNNM2
GYPC
ATP1B1
MYO1D
PIK3IP1
CHKA
EEPD1
PRKAG2
ABCG1
TEX2
MYO1E
AC092338.1
MAMDC2
UHRF1BP1
CREG1
COL4A3BP
FARP1
PPARG
MFHAS1
MTHFR
SGMS1
CMKLR1
CDS2
ZNF438
SLC36A1
SHMT1
SERINC5
FNIP2
ARHGEF10L
KLF6
AC104123.1
TSNARE1
CCDC93
NHSL1
ARID5B
TSC22D1
MFSD1
FCHO2
RHOQ
NEK1
NRP2
DIRC3
TWSG1
GAS7
TTL
MANBA
DIP2C
CEP120
LTA4H
RNF19A
PEAK1
CLASP2
ATP6V1C1
SCARB1
NAB1
EFCAB2
SPIRE1
NT5DC1
ABAT
SLC27A1
SNX24
MAN1A1
PARVG
UVSSA
FAM85B
COX10-AS1
MGAT1
LRRC37B
MGLL
LGR4
ABCA1
GPATCH2L
IFT88
ACYP2
PDCD4
GCLC
PCGF5
C2CD2
AP2A2
CDK5RAP2
ZNF680
TPCN1
PARP11
RBM47
SQSTM1
HEXB
ANKRD36C
HNRNPUL1
TTYH3
AGAP3
CCDC149
WWP1
TIMP2
ELMSAN1
LUC7L
ZCCHC2
TMCO4
REV3L
TESK2
CDK17
TANGO2
LAMTOR5-AS1
PAPSS1
ZFYVE16
RALGAPA2
NFE2L2
SCMH1
MYO9A
CXCL12
RAD51B
ACVR2A
RTN3
AGAP1
ZNF69
PARP6
CNTLN
LINC02615
IVNS1ABP
RPP38
NMRK1
WWOX
ATG12
TFEC
VPS8
ANKRD6
RNF130
KANTR
GALK2
TLR5
ZNF131
AC009690.1
CDC16
SLC2A13
CEP295
AC068587.4
TPST2
IQCH-AS1
ATXN2
MICU1


b
Non-diseased control
AD


c
*
GPNMB positive cells/μm²
0.06
0.04
0.02
0.00
NDC
AD