Brain, Behavior, and Immunity 80 (2019) 644–656

Contents lists available at ScienceDirect

Brain, Behavior, and Immunity

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

Transcriptomic analysis of probable asymptomatic and symptomatic alzheimer brains

Hamel Patel[a,b], Angela K. Hodges[c], Charles Curtis[b,d], Sang Hyuck Lee[b,d], Claire Troakes[e,f], Richard J.B. Dobson[a,b,g,h,i,*,1], Stephen J. Newhouse[a,b,g,h,i,*,1]

[a] Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[b] NIHR BioResource Centre Maudsley, NIHR Maudsley Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust (SLaM) & Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, UK
[c] Department of Old Age Psychiatry, Maurice Wohl Clinical Neuroscience Institute Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, London, UK
[d] Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, UK
[e] Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, UK
[f] London Neurodegenerative Diseases Brain Bank, Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK
[g] Health Data Research UK London, University College London, 222 Euston Road, London, UK
[h] Institute of Health Informatics, University College London, 222 Euston Road, London, UK
[i] The National Institute for Health Research University College London Hospitals Biomedical Research Centre, University College London, 222 Euston Road, London, UK

ARTICLE INFO

Keywords:
Alzheimer's disease
Microarray analysis
Gene expression
Human
Neuropathology
Frontal lobe
Gene regulatory networks
Astrocytes
Mitochondria
Brain

ABSTRACT

Individuals with intact cognition and neuropathology consistent with Alzheimer's disease (AD) are referred to as asymptomatic AD (AsymAD). These individuals are highly likely to develop AD, yet transcriptomic changes in the brain which might reveal mechanisms for their AD vulnerability are currently unknown. Entorhinal cortex, frontal cortex, temporal cortex and cerebellum tissue from 27 control, 33 AsymAD and 52 AD human brains were microarray expression profiled. Differential expression analysis identified a significant increase of transcriptomic activity in the frontal cortex of AsymAD subjects, suggesting fundamental changes in AD may initially begin within the frontal cortex region prior to AD diagnosis. Co-expression analysis identified an overactivation of the brain "glutamate-glutamine cycle", and disturbances in the brain energy pathways in both AsymAD and AD subjects, while the connectivity of key hub genes in this network indicates a shift from an already increased cell proliferation in AsymAD subjects to stress response and removal of amyloidogenic proteins in AD subjects. This study provides new insight into the earliest biological changes occurring in the brain prior to the manifestation of clinical AD symptoms and provides new potential therapeutic targets for early disease intervention.

1. Introduction

The increase in life expectancy has profoundly increased the ageing population, which, unfortunately, is also accompanied by a rise in age-related disorders including Alzheimer's disease (AD) (Prince et al., 2016). Alzheimer's disease is a neurodegenerative disorder characterised by progressive accumulation of extracellular amyloid-β (Aβ) protein and intracellular hyperphosphorylated tau filaments in the brain, which form insoluble plaques and tangles respectively. These protein aggregates affect neuronal activity which can lead to progressive loss of neurons associated with deterioration in cognition and development of neuropsychiatric symptoms.

Through longitudinal studies involving autopsy, it has become evident that clinical signs of cognitive impairment are apparent after substantial years of neurodegeneration, which occurs decades after neuropathological changes (Caselli and Reiman, 2012). As the disease is progressively slow and as everyone is expected to experience cognitive change during normal ageing, differentiating AD symptoms from normal ageing at an early stage of the disease can be difficult. Up to 20–30% of the ageing population with intact cognition have amyloid deposition, with these individuals at higher risk of progressing to AD than those without amyloid (Yvette et al., 2010). These individuals are often referred to as asymptomatic AD (AsymAD) (Driscoll and Troncoso, 2011) and have been shown to be distinguishable from

* Corresponding authors.
E-mail addresses: richard.j.dobson@kcl.ac.uk (R.J.B. Dobson), stephen.newhouse@kcl.ac.uk (S.J. Newhouse).
[1] Prof Richard J.B Dobson and Dr Stephen J Newhouse are joint last authors.

https://doi.org/10.1016/j.bbi.2019.05.009
Received 1 January 2019; Received in revised form 2 May 2019; Accepted 3 May 2019
Available online 04 May 2019
0889-1591/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

normal ageing based on neuropathology, brain imaging and cerebrospinal fluid biomarkers (Caselli and Reiman, 2012). While some of these individuals progress to developing symptoms related to cognition, which deviate from normal Mild Cognitive Impairment (MCI), and then to AD, not all do. They are therefore a heterogeneous group, representing those with prodromal AD and those impervious to AD despite having the pathological hallmarks.

Measuring genome-wide expression of transcripts as markers of gene activity has revealed that cognitive decline is accompanied by changes in brain gene expression from normal ageing through to MCI and AD. Studies have suggested that some changes in the pattern of gene expression in normal ageing such as synaptic function and energy metabolism (Saura et al., 2015), are extensively altered in MCI (Nicole et al., 2015) and AD (Berchtold et al., 2013; Miller et al., 2013; Blalock et al., 2011; Blalock et al., 2004; Miller et al., 2008). Additional work has also suggested a number of other biological pathways are more specifically altered in AD, including inflammation (Blalock et al., 2004; Colangelo et al., 2002; Miller et al., 2008), protein misfolding (Blalock et al., 2004; Colangelo et al., 2002), transcription factors (Blalock et al., 2004; Colangelo et al., 2002), cell proliferation (Blalock et al., 2004; Colangelo et al., 2002), immune response (Sekar et al., 2015; Li et al., 2012; Lambert et al., 2010; Chen et al., 2016), protein transcription/translation regulation (Sekar et al., 2015; Li et al., 2012; Liu et al., 2013; Li et al., 2015; Puthiyedth et al., 2016; Godoy et al., 2014), calcium signalling (Sekar et al., 2015; Ramanan et al., 2013; Blalock et al., 2011), MAPK signalling (Puthiyedth et al., 2016; Miller et al., 2013), and various metabolism pathways (Puthiyedth et al., 2016; Di Paolo and Kim, 2012; Liang et al., 2008; Ishii et al., 1997; Godoy et al., 2014; Li et al., 2012; Oshiro et al., 2011, 2011.) which reflect the extent and type of pathology and disruption to cell activity as disease progresses. It is unknown how early the different types of changes occur in the brain, such as in the pre-symptomatic phase or specifically in AsymAD subjects who already have the pathological hallmarks of AD such as amyloid and neurofibrillary tangles (NFTs). Understanding the fundamental changes in this AsymAD group may shed light on specific biological mechanisms that may be involved in early pathological hallmarks of AD, providing new therapeutic targets for early intervention.

In this study, we investigated transcriptomic changes in the human brain of healthy ageing, AsymAD and AD subjects, which have been classified based on the clinical assessment before death and AD neuropathology at autopsy. Typical transcriptomic analysis coupled with a systems-biology approach was used to identify disturbances in the underlying biological mechanisms across the entorhinal cortex, temporal cortex, frontal cortex and cerebellum brain regions. In addition, we provide access of gene-level results to the broader research community through a publicly available R SHINY web-application (https://phidatalab-shiny.rosalind.kcl.ac.uk/ADbrainDE), allowing researchers to quickly query the expression of specific genes through the progression of AD and across multiple brain regions.

2. Materials and methods

2.1. Medical research council london neurodegenerative diseases brain bank

A total of 112 brains were obtained from the Medical Research Council (MRC) London Neurodegenerative Diseases Brain Bank (from now on referred to as MRC-LBB) hosted at the Institute of Psychiatry, Psychology and Neuroscience, KCL. All cases were collected under informed consent, and the bank operates under a licence from the Human Tissue Authority, and ethical approval as a research tissue bank (08/MRE09/38 + 5). Neuropathological evaluation for neurodegenerative diseases was performed in accordance with standard criteria.

2.2. MRC-LBB sample selection

BRAAK staging is a measure of the spread of hallmark AD pathology across the brain and is part of the neuropathological assessment. In general, BRAAK stages I-II, III-IV and V-VI have been suggested to represent low, intermediate, and high likelihood dementia is due to AD respectively (Consensus recommendations for the postmortem diagnosis of Alzheimer's disease, 1997). Twenty-seven control cases were used - classified as showing no clinical sign of any form of dementia and no neuropathological evidence of neurodegeneration. Thirty-three AsymAD cases were also analysed - defined as clinically dementia-free at the time of death, but neuropathological assessment at autopsy showed hallmark AD pathology. Finally, fifty-two AD cases, which had both clinical diagnosis of AD at death and confirmation of this diagnosis through neuropathological evaluation at autopsy, were selected.

2.3. MRC-LBB brain region selection and RNA extraction

Frozen tissues (0.5–1 cm3) from the following brain regions from each case were macrodissected into RNAlater RNA Stabilization Reagent (Qiagen): 1) Frontal Cortex (FC), 2) Temporal Cortex (TC), 3) Entorhinal Cortex (EC) and 4) Cerebellum (CB). Hallmark AD pathology was confirmed in the entorhinal cortex, temporal cortex and frontal cortex but absent from the cerebellum of AsymAD and AD subjects. RNA extraction was performed within 24 h of dissection. Total RNA was extracted using RNeasy Lipid Tissue Mini Kit (Qiagen, 74804) following the manufacturer's protocol. Genomic DNA was removed using gDNA Eliminator Spin Columns (Qiagen). The RNA quality was evaluated with an Agilent 2100 bioanalyzer (Agilent Technologies, Inc., Palo Alto, CA).

2.4. MRC-LBB Illumina beadArray expression profiling

Total RNA (25 ng) was prepared for array expression profiling using the Ovation Pico WTA system (NuGEN Technologies, Inc., San Carlos, CA), as described by the manufacturer's protocol. The Nugen system is optimised for the amplification of degraded RNA, where amplification is initiated at the 3′ end as well as randomly throughout the whole transcriptome. The samples were processed at the NIHR Biomedical Research Centre for Mental Health (BRC-MH), Genomics & Biomarker Core Facility at the Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London (https://www.kcl.ac.uk/ioppn/depts/sgdp-centre/research/The-IoPPN-Genomics–Biomarker-Core-Facility.aspx) in accordance with the manufacturer's protocol using the Illumina HT-12_V4 beadchips (Illumina, USA).

2.5. Microarray expression data processing

Raw gene expression data was exported from Illumina's GenomeStudio (version 2011.1) into RStudio (version 0.99.467) for data processing. Using R (version 3.2.2), raw data was Maximum Likelihood Estimation (MLE) background corrected using R package "MBCB" (version 1.18.0) (Allen et al., 2010), log2 transformed, and underwent Robust Spline Normalisation (RSN) using R package "lumi" (version 2.16.0) (Du et al., 2008).

A series of quality control steps were carried out before data analysis. Duplicate samples were removed based on lowest RIN score. Sex was predicted for each sample using the R package "massiR" (version 1.0.1) (Buckberry et al., 2014), with any discrepancies in predicted and clinically recorded sex from the same individual across all tissues removed from further analysis. For each sample, probesets "not reliably detected" or "unexpressed" were removed to eliminate noise (Lazar et al., 2013) and increase power (Hackstadt and Hess, 2009). If the expression of a probe was below the 90th percentile of the log2 expression scale in over 80% of samples across all groups (based on

disease status, brain region and sex), the probe was deemed "unexpressed" and was removed from further analysis.

Batch effects were then explored using Principal Component Analysis (PCA) and Surrogate Variable Analysis (SVA) using the R package "sva" (version 3.10.0) (Leek et al., 2012). Sex and diagnosis information was used as covariates in sva when correcting for unknown batch effects. To ensure homogeneity among the biological groups, outlying samples per tissue and disease group were iteratively identified and removed following the fundamental network concepts described in (Oldham et al., 2012). Finally, Illumina-specific probe ID's were converted to the universal Entrez Gene ID using the R package "illuminaHumanv4.db" (version 1.22.1).

2.6. Differential expression and gene set enrichment analysis

Differential Expression (DE) analysis was performed using the R package "limma" (version 3.20.9) (Smyth, 2005). As we had theoretically corrected for unwanted batch effects in our data using sva, we only used sex in the DE model as a covariate. A gene was regarded as significantly differentially expressed if the false discovery rate (FDR) adjusted p-value ≤ 0.05.

Gene set enrichment analysis (GSEA) was performed using an Over-Representation Analysis (ORA) implemented through the ConsensusPathDB (http://cpdb.molgen.mpg.de) web-based platform (version 32) (Kamburov et al., 2009) in October 2017. ConsensusPathDB incorporates numerous well-known biological pathway databases including BioCarta, KEGG, Reactome and Wikipathways. It performs a hypergeometric test while combining a background gene list, compiles results from each database and corrects for multiple testing using FDR (Kamburov et al., 2009). During GSEA analysis, a minimum overlap of the query signature and database was set as 2.

2.7. Weighted gene co-expression network analysis

Weighted gene co-expression analysis (WGCNA) was performed using R package "WGCNA" (version 1.51) to identify clusters (modules) of highly correlated genes, with the underlying hypothesis that such modules could possess a common function. The WGCNA analysis was performed as described in (Langfelder and Horvath, 2008). In brief, a co-expression network based on "signed" adjacency was independently created for all three phenotypes (control, AsymAD and AD group), topological overlap calculated, and hierarchical clustering used to group genes into modules. The control group module was assigned default colours based on module size, and the AsymAD and AD module colours determined based on the control module gene overlap. Module cross-tabulations were generated across the three phenotypes and Fisher's exact test used to test for enrichment between modules-gene assignments between the control, AsymAD and AD groups. To aid in identifying significant changes in the co-expression network within the same modules in the three phenotypes, additional statistics known as "Module preservation Zsummary" and "median rank" were calculated as described in (Langfelder et al., 2011).

2.8. Protein–Protein interaction network analysis

Protein-protein interaction (PPI) networks were generated by uploading gene lists (referred to as seeds in network analysis) to NetworkAnalyst's (http://www.networkanalyst.ca/faces/home.xhtml) web-based platform in December 2017. The "zero-order network" option was incorporated to allow only seed proteins directly interacting with each other, preventing the well-known "hairball effect" and allowing for better visualisation and interpretation (Xia et al., 2014). Submodules with a p-value ≤ 0.05 based on the "InfoMap" algorithm (Zaki and Mora, 2015) were deemed significant "hubs" and the gene(s) with the most connections within this network as the "key hub gene(s)".

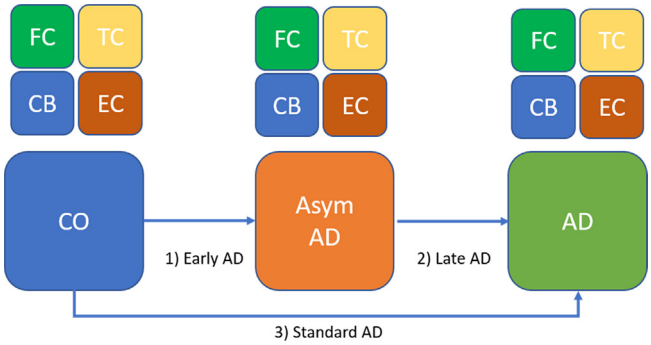

Fig. 1. Overview of study design. Four brain regions; frontal cortex (FC), temporal cortex (TC), entorhinal cortex (EC) and cerebellum (CB) from the three subject groups; control (CO), Asymptomatic AD (AsymAD) and Alzheimer's disease (AD) were expression profiled. The typical comparison between the CO and AD group is referred to as the "Standard AD" analysis, the comparison between the CO and AsymAD group is referred to as the "Early AD" analysis and the comparison between the AsymAD and AD group is referred to as the "Late AD" analysis.

2.9. Study design

Differential and co-expression analysis was performed between the three disease groups and for each of the four brain regions. First, the control and AsymAD groups were compared, and from this point onwards is referred to as the "**Early AD**" analysis. Second, the AsymAD and AD groups were compared, and from this point onwards is referred to as the "**Late AD**" analysis. Finally, the control and AD groups were compared, and from this point onwards is referred to as the "**Standard AD**" analysis. An overview of the study design and analyses is shown in Fig. 1.

2.10. Data availability

The microarray data have been deposited in NCBI's GEO database under the accession number GSE118553. Additionally, a shiny application was written in R using the "shiny" framework (version 0.14) to allow quick visualisation of specific gene expression in the control, AsymAD and AD subjects, and across the EC, TC, FC and CB brain regions. The application also displays DE results of each gene and can be accessed at https://phidatalab-shiny.rosalind.kcl.ac.uk/ADbrainDE. All data analysis scripts used in this study are available at https://doi.org/10.5281/zenodo.1400644

3. Results

3.1. Data processing

Of the 401 tissue samples assessed (extracted from the 112 brains) 48 samples were removed due to duplication, 4 samples due to outlier detection analysis and 2 samples due to sex discrepancies between recorded and actual sex, leaving 347 tissue samples from 111 brains for DE and co-expression analysis. As a result of samples not being microarray profiled due to sample quality, and samples being removed during the Quality Control (QC) process, not all subjects had tissue samples extracted from all four brain regions. The demographics for datasets by brain region and sample group is provided in Table 1.

After further QC and annotation to determine Entrez gene identifiers, the final data represented 3518 "reliably detected" genes across all samples. Chi-squared tests revealed no significant difference in the proportion of males to females across the three disease groups or brain regions. Mann-Whitney U test revealed no significant difference between post-mortem (PM) delay or disease duration across analyses; however, age was significantly ($p \leq 0.01$) lower in the control groups

Table 1
Summary of MRC-LBB sample characteristics.

Brain Region	Phenotype	No. Samples	Sex (M/F)	Age (± SD)	BRAAK (± SD)	PM Delay (h)	Disease duration (yrs)
Entorhinal Cortex	Control	16	9/7	71.9 (15.6)	0	33.8 (17.8)	0
	AsymAD	28	8/20	85.4 (9.5)	2.2 (± 1.2)	52.5 (15.9)	0
	AD	34	13/21	83.9 (9.7)	4.9 (± 1)	39.5 (21.2)	11.8 (5.2)
Temporal Cortex	Control	24	14/10	71.5 (16.9)	0	37.2 (19.8)	0
	AsymAD	28	9/19	86.3 (8.6)	2.5 (± 1.1)	54.2 (16.6)	0
	AD	45	20/25	82.7 (9.8)	4.9 (± 0.9)	40.4 (21.4)	9.7 (5.4)
Frontal Cortex	Control	21	12/9	69.8 (15.4)	0	40.4 (24.6)	0
	AsymAD	32	10/22	86 (8.9)	2.3 (± 1.2)	54.1 (16.2)	0
	AD	38	13/25	82.5 (4.7)	4.9 (± 1)	39.4 (20.5)	10.5 (5.7)
Cerebellum	Control	18	10/8	69.4 (16)	0	37.9 (20.7)	0
	AsymAD	27	8/19	86.3 (9.2)	2.4 (± 1.2)	56 (16.5)	0
	AD	36	17/19	82.6 (10.6)	5.1 (± 0.3)	40.2 (22.3)	9.4 (5.6)

when compared to the AsymAD and AD group in each tissue (see Supplementary Table 1). It is important to note, although the control group are collectively younger than both the AsymAD and AD groups, some controls exist in this study that are older than the average age of both the AsymAD and AD groups. Phenotype information per sample is provided in Supplementary Table 2.

The table provides a summary of sample characteristics used in this study. From the Initial 401 samples expression profiled, 48 samples were removed due to duplication, 2 samples removed due to sex discrepancies and 4 samples removed due to being identified as outliers. The total number of samples available after quality control was 347. BRAAK staging is a measure of the spread of hallmark AD pathology across the brain and does not reflect pathology within a distinct brain region. In general, BRAAK stages I-II, III-IV and V-VI have been suggested to represent prodromal, early-moderate AD, and moderate-late AD respectively. BRAAK scores deviate between brain regions as not all four brain regions were available from all donors. Hallmark AD pathology was confirmed in the entorhinal cortex, temporal cortex and frontal cortex but absent from the cerebellum of AsymAD and AD subjects. The values provided in Age, BRAAK and PM Delay represent the mean ± standard deviation. Demographics per sample is provided in Supplementary Table 2. Abbreviation: M/F: the ratio of male and female samples, PM: Post-Mortem, h: Hours, yrs: Years, SD: Standard deviation.

3.2. Summary of differentially expressed genes across disease groups and tissues

A summary of DEG's identified in each brain tissue and analyses are illustrated in Fig. 2, and a full list of DEG's is provided in Supplementary Table 3. The general trend of DEG's in subjects with AD ("Late AD" and "Standard AD" analysis) decreases across brain regions in the order of EC (n = 1904 and n = 1690 respectively) > TC (n = 1546 and n = 1517 respectively) > FC (n = 52 and n = 299 respectively) > CB (n = 13 and n = 176 respectively). This expression pattern corresponds to the route AD pathology is seen to spread through the brain. By contrast, the pattern differs in the AsymAD group ("Early AD" analysis), where most DEGs are detected in the FC (n = 398) followed by the TC (n = 253), EC (n = 19) and CB (n = 1), suggesting initial molecular changes may begin in the FC brain region prior to AD symptoms.

3.3. AD tau pathology marker suggests AsymAD subjects are an Intermediate state between normal ageing and AD

A previous study identified eight genes highly correlated with AD tau pathology (Miyashita et al., 2014), of which two genes (**RELN, TRIL**) are present in our data. The RELN gene is not significantly DE in any of the analysis in this study; however, DE analysis results indicate the **TRIL** gene expression gradually increases through the control, AsymAD and then the AD group. In addition, the expression increase is only observed in brain regions known to be affected by tau pathology (EC, TC and FC), and the extent of expression change within these affected brain regions shadows the route of disease manifestation through the brain (Fig. 3a). The EC exhibits the most significant increase of **TRIL** expression (logFC = 0.99, FDR adjusted p-value = 2.77e−8), followed by the TC (logFC = 0.48, FDR adjusted p-value = 1.41e−3) and then FC brain region (logFC = 0.44, FDR adjusted p-value = 2.21e−2). This expression pattern further suggests the **TRIL** gene is a reliable brain marker for tau pathology, and our AsymAD samples are a good representation of early-intermediate state between normal ageing and AD.

3.4. The most significant differentially expressed genes per analysis

The most DEG's from each analysis is 1) **MOSPD3** (downregulated in the TC brain region in "Early AD", FDR adjusted p-value = 1.18e−10, 2) **NPC2** (upregulated in the EC brain region in the "Late AD" analysis, FDR adjusted p-value = 2.39e−20,) and 3) **NOTCH2NL** (upregulated in the EC brain region in the "Standard AD" analysis, FDR adjusted p-value = 1.29e−15. Expression boxplots of these genes are available to view in the SHINY web-app.

3.5. Common differentially expressed genes across all brain regions

The overlap of DE genes across brain regions is shown in Fig. 4. **MOSPD3** is the only gene significantly differentially expressed across all four brain regions in the "Early AD" analysis. No gene was significantly differentially expressed in the "Late AD" analysis across all four brain regions; however, six genes (**NPC2, DUSP1, GPM6B, SLC38A2, ANKEF1, MOSPD3**) were identified in "Standard AD" analysis. Three of these genes (**DUSP1, SLC38A2** and **MOSPD3**) are consistently expressed in the same direction across all four brain regions. **DUSP1** and **SLC38A2** gene expression are upregulated during disease progression (Control to AsymAD to AD). **MOSPD3**, however, is downregulated in the disease in both the "Early AD" and "Standard AD" analyses, with no significant difference between the AsymAD and AD subjects. The remaining three genes (**NPC2, GPM6B, ANKEF1**) are DE in the same direction across all brain regions but reversed in the CB; a brain region suggested to be spared by hallmark AD pathology.

3.6. Differentially expressed genes in brain regions with hallmark AD pathology

The EC, TL and FC are all affected by hallmark AD pathology (amyloid and NFT's), while the CB is known to be partially spared. Gene's DE in the EC, TC and FC brain regions and not the CB, may identify hallmark AD pathology specific genes. Three (**ALDH2, FBLN2** and **METTL7A**) and nine (**FLCN, ASPHD1, ARL5A, GPR162, HBA2, PCID2, NDRG2, BEND3, RAP1Gap**) genes were significantly

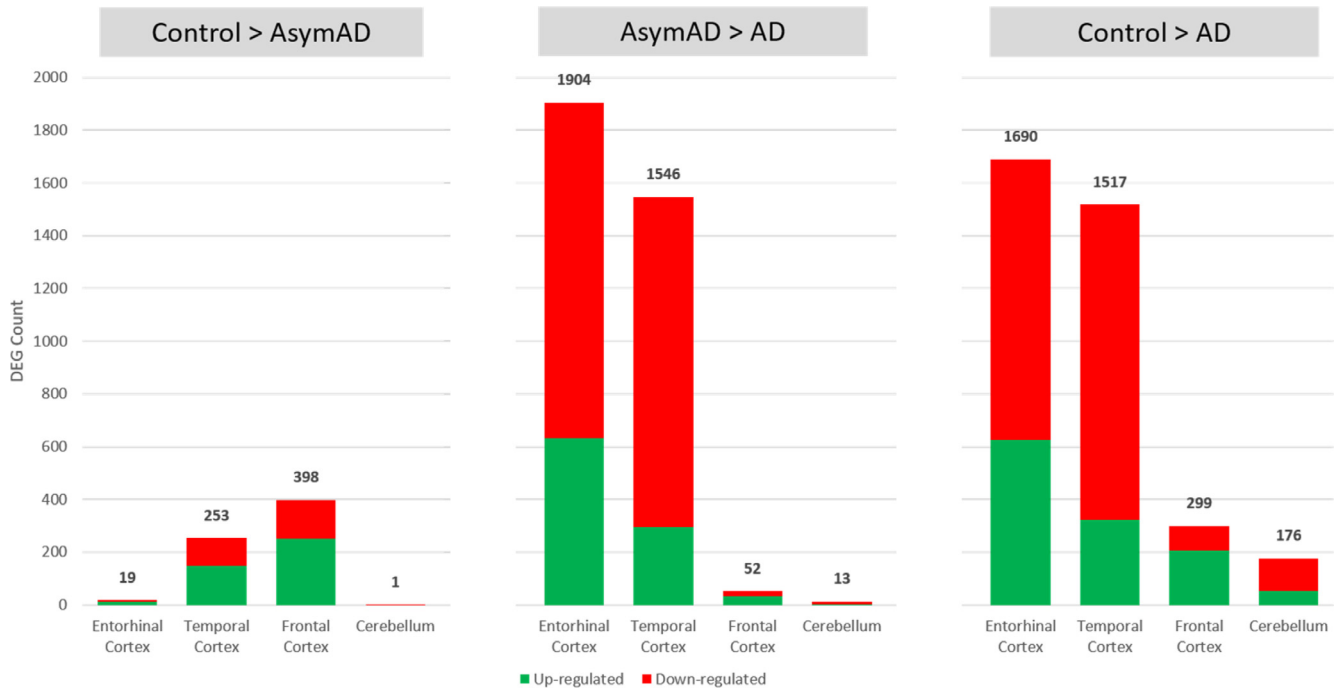

Fig. 2. Distribution of Significant DEG (FDR adjusted p-value ≤ 0.05) across brain regions and analyses. "Control > AsymAD" summarises the number of DEGs between the control and AsymAD group. "AsymAD > AD" summarises the number of DEGs between the AsymAD and AD group. "Control > AD" summarises the number of DEGs between the control and AD group. The proportion of up-regulated genes is represented in green while the down-regulated genes are represented in red. The total number of significantly differentially expressed genes in each brain region and analysis is provided on top of each bar. More genes are observed to be generally perturbed when comparing the AD group to the AsymAD or healthy ageing group, with the general pattern of more genes perturbed in the entorhinal cortex, followed by the temporal cortex, frontal cortex and then the cerebellum, a pattern generally representing the spread of hallmark AD pathology. In contrast, comparing the AsymAD group to the healthy ageing group reveals more genes are perturbed in the frontal cortex, followed by the temporal cortex, entorhinal cortex and then the cerebellum, suggesting initial molecular changes in AD may begin in the frontal cortex before the manifestation of clinical AD symptoms.

differentially expressed across the EC, TC and FC brain regions and not the CB brain region in the "Early AD" and "Late AD" analysis respectively.

3.7. Gene set enrichment analysis of differentially expressed genes

To understand the functional implications of DEG's, GSEA was performed using the significant DEG list from all three analyses ("Early AD", "Late AD" and "Standard AD") and across all four brain regions, resulting in 12 enrichment result tables (provided in Supplementary Table 4). No biological pathway is significantly enriched across all four brain regions in the "Early AD", "Late AD" or "Standard AD" analysis. However, when excluding the brain region spared by hallmark AD pathology (CB), the "**glutamate glutamine metabolism**" and "**gluconeogenesis and glycolysis**" pathways are the only pathways significantly enriched in the "Early AD" and "Late AD" analysis respectively. For the "Standard AD" analysis, excluding the CB brain region additionally identified "**mRNA processing**", "**synaptic vesicle pathway**" and "**TNF-alpha**" pathways as significantly enriched in the remaining three brain regions.

3.8. Summary of weighted co-expression network analysis

Weighted gene co-expression analysis was performed on the FC and EC brain regions. We focused on these two brain regions as differential expression analysis identified an increased number of significant DEG's in the FC brain region prior to AD symptoms and the EC is widely regarded as one of the first areas of the brain to be affected in AD. Network preservation and cross-tabulation statistics were calculated to identify co-expression networks that may be preserved or disrupted between the Control, AsymAD and AD subjects. Fig. 5 illustrates the WGCNA module assignments and module preservation statistics, and Fig. 6 shows the cross-tabulation statistics across phenotypes.

Co-expression analysis in the FC brain region identified 13, 7, and 12 modules within the control, AsymAD and AD groups respectively, while the analysis in the EC identified 8, 8 and 11 modules within the control, AsymAD and AD groups respectively. GSEA analysis was performed for all fifty-nine modules to identify potential biological pathways the co-expressed genes may be involved with. A summary of the GSEA results on the co-expression module in the FC and EC is provided in Tables 2 and 3 respectively, with complete GSEA results for the Control, AsymAD and AD groups in the FC and EC brain regions provided in Supplementary Tables 5–10 respectively.

3.9. Co-expression modules are weakly preserved in AsymAD and AD entorhinal cortex

Module preservation statistics were calculated for each brain region to identify co-expression networks that are weakly preserved through the course of the disease. Modules below a "preservation Zsummary" statistic of 10 and "preservation median rank" higher than the gold module (random 100 genes) are suggested to be weakly preserved. Module colours for the AsymAD and AD groups were mapped to the control module colours, allowing for changes and preservation in the co-expression networks to be observed as the disease progresses. The module colours assigned in the EC brain region are independently assigned to modules colours assigned in the FC brain region and therefore similar module colours across these two tissues bare no relation.

The FC "preservation Zsummary" statistics (Fig. 5b and c) suggests all modules from the control group are relatively well-preserved in the AsymAD and AD groups. In contrast, the EC "preservation median rank" statistics suggest the green control module is weakly preserved in AsymAD group (Fig. 5e), and both the green and brown control modules are weakly preserved in the AD group (Fig. 5f). In addition, the

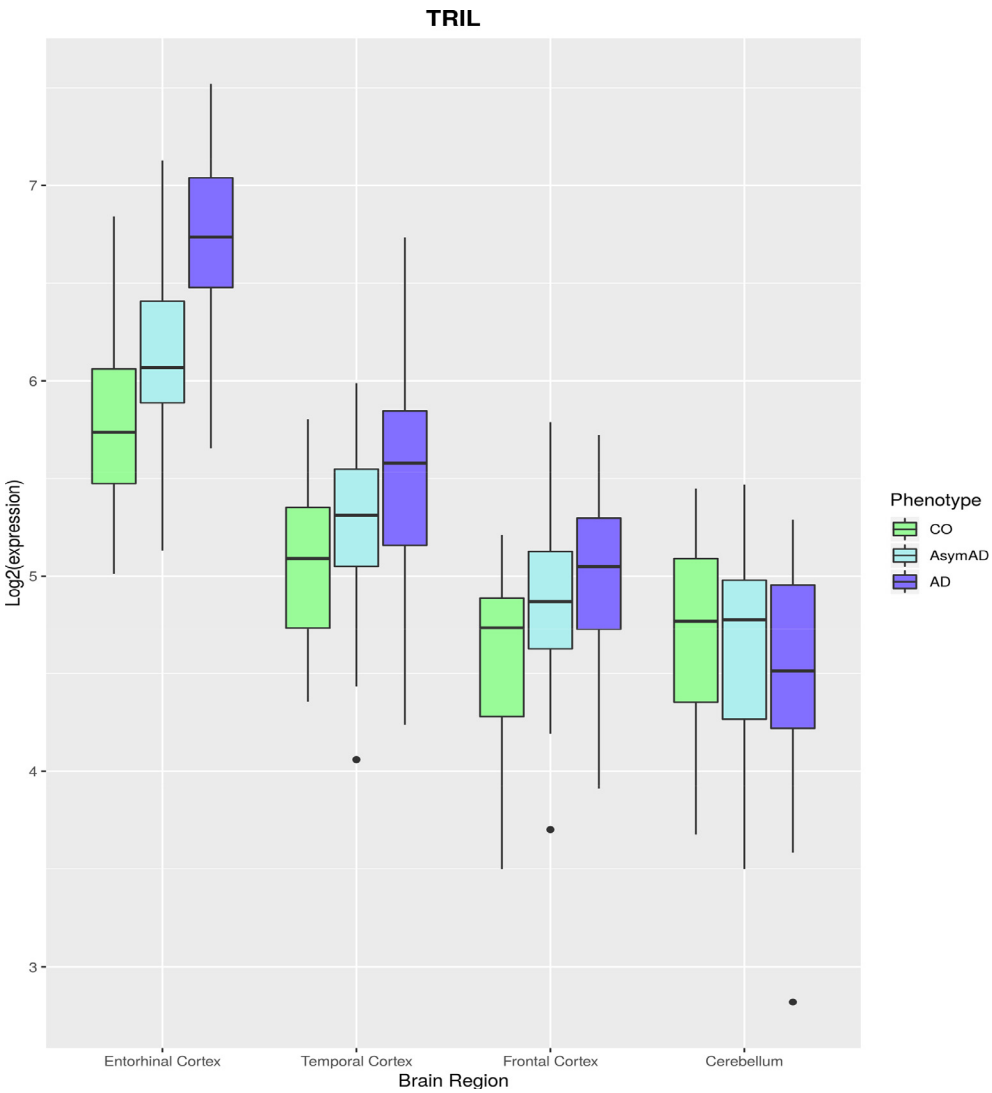

Fig. 3. Expression boxplots of the TRIL gene. Previous research identified the TRIL gene as a marker for Tau pathology. The TRIL gene is significantly up-regulated (before multiple correction) from control to AsymAD (EC: logFC = 0.39 & p-value = 0.01, TC: logFC = 0.24 & p-value = 0.04, FC: logFC = 0.24 & p-value = 0.04) and then further to AD (EC: logFC = 0.6 & p-value = 6.57e-6, TC: logFC = 0.29 & p-value = 4.19e-3, FC: logFC = 0.18 & p-value = 0.05), but not in the cerebellum (control to AsymAD: logFC = −0.01 & p-value = 1, AsymAD to AD: logFC = −0.19 & p-value = 0.2), a region spared by hallmark AD pathology. The expression pattern of the TRIL gene further supports the assignment of AsymAD samples, which were based on clinical records and neuropathological assessment, as an early intermediate state between healthy ageing and AD.

cross-tabulation statistics are also indicative of disruption to the EC green control module (Fig. 6d). GSEA reveals the EC brown module in control, AsymAD and AD group is most significantly enriched for "**selenocysteine synthesis**" (control q-value = 4.71e54, AsymAD q-value = 5.89e−90, AD = 1.35E−96), suggesting this process is not significantly disrupted in AsymAD or AD subjects. In contrast, the EC control green module is significantly enriched (before multiple corrections) for **"neutrophil degranulation"** (p-value = 0.5e−4), **"TYROBP casual network"** (p-value = 2.5e−3) and the **"innate immune system"** (p-value = 2.7e−3), none of which are present in the green module of the AsymAD, suggesting these pathways may be disrupted in AsymAD subjects.

Clusters of co-expressed genes in both the FC and EC brain regions were enriched for specific cell types including neurons, astrocytes, oligodendrocytes and microglia (results not shown); however, we did not detect a disturbance in any cell type in AsymAD subjects.

3.10. Frontal cortex co-expression network re-wired in AsymAD

Co-expression analysis identified 13 and 12 co-expressed modules in the control and AD subjects respectively. However, the AsymAD group exhibits 7 larger modules of highly co-expressed genes, suggesting the co-expression network is re-wired in the FC brain region in this intermediate stage of AD. The module preservation analysis suggested all modules within the control group are relatively preserved through the course of the disease, however, through cross-tabulation of the modules we observe subtle changes leading to a much larger magenta module in the AsymAD group. The biological processes associated with the magenta module changes from being enriched for "**glucose metabolism**" (q-value 6.26e−02) in the control group to "**oxidative phosphorylation**" (q-value = 2.26e−11), **Parkinson's disease** (q-value = 5.12e−9), **electron transport chain** (q-value = 5.83e−9) and **Alzheimer's disease** (q-value = 8.21e−9) in the AsymAD group. Then the large magenta module in the AsymAD group, branches into four new AD modules (blue, turquoise, midnightblue, and yellow), which are most enriched for **Parkinson's disease** (q-value = 3.09e−4), **neurotransmitter receptors and postsynaptic signal transmission** (q-value = 0.01), **the citric acid (TCA) cycle and respiratory electron transport** (q-value = 0.01), and **fas signalling** (q-value = 0.0030) respectively.

3.11. Entorhinal cortex yellow module enriched for all "early AD" analysis DEG's

The yellow module contained all genes identified as significantly DE in the "Early AD" analysis (**ALDH2**, **FBLN1** and **METTL7A**) and contained a large number of genes disrupted from the green module, which was the least preserved module through disease progression. Overall, this made the yellow module a prime candidate for further investigation. Gene set enrichment analysis of the yellow module in the AsymAD

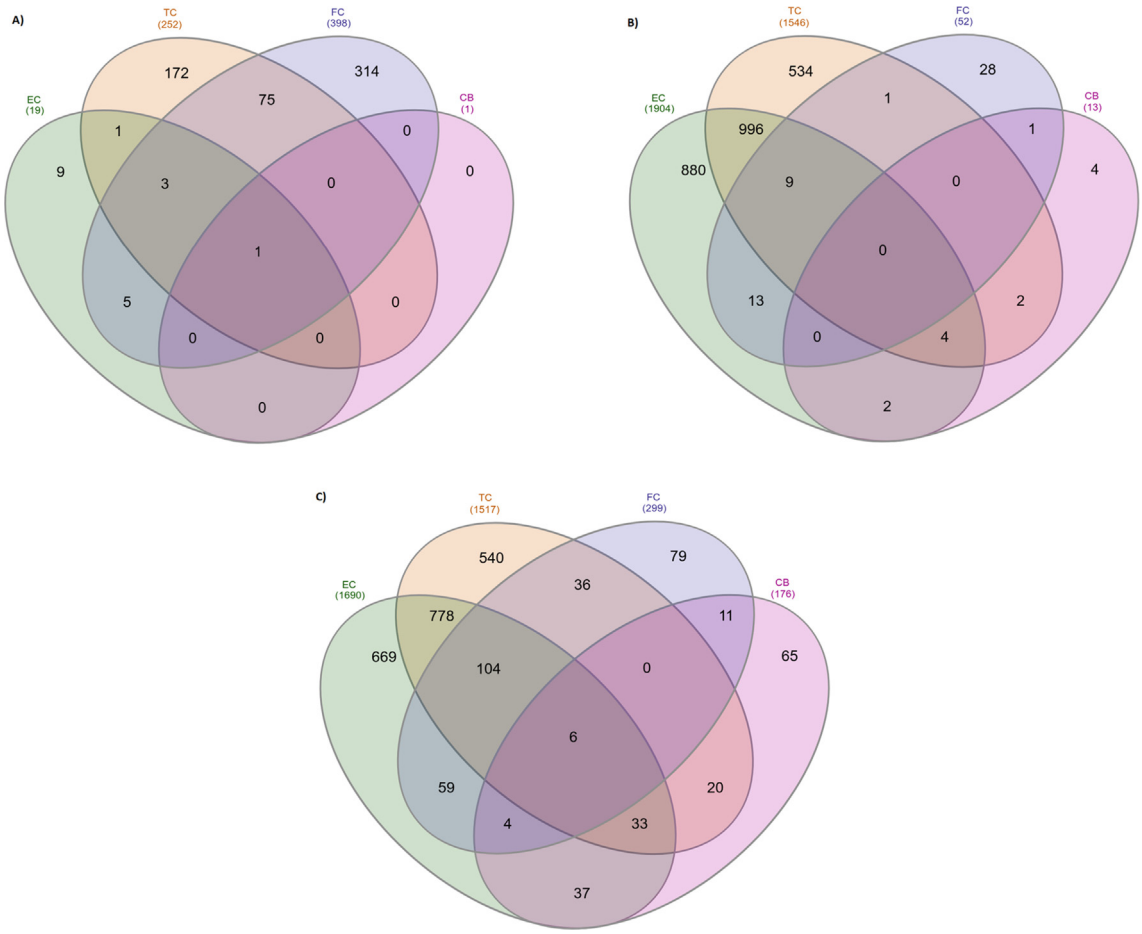

Fig. 4. Overlap of significant DEG across brain regions in A)" Early AD" analysis, B) "Late AD" analysis and C) "Standard AD" analysis. All brain regions in this study are affected in AD, specifically by atrophy and neuronal loss, while only three brain regions in this study (EC, TC and FC) are affected by the additional accumulation of hallmark AD pathology (Aβ and NFT). Genes perturbed across all brain regions may be markers of cell dysfunction in AD, while genes consistently perturbed in the EC, TC and FC but not in the CB may be associated with AD pathology. **MOSPD3** gene is the only gene DE in all brain regions of the "Early AD" analysis. No gene is DE across all brain regions in the "Late AD" analysis. Three (**ALDH2, FBLN2**, **METTL7A**) and nine genes (**FLCN, ASPHD1, ARL5A, GPR162, HBA2, PCID2, NDRG2, BEND3, RAP1Gap**) are consistently DE across all brain regions affected by hallmark AD pathology in the "Early AD" and "Late AD" analyses respectively.

group reveals enrichment in "**fatty acid degradation**" (q-value = 0.03), "**glycerophospholipid metabolism**" (q-value = 0.008), "**urea cycle and metabolism of arginine, proline, glutamate, aspartate and asparagine**" (q-value = 0.05), "**astrocytic glutamate-glutamine uptake and metabolism**" (q-value = 0.05) and "**neurotransmitter uptake and metabolism in glial cells**" (q-value = 0.05), all of which were not previously enriched in the matched yellow module in the control group.

Protein-protein interaction analysis in the yellow control module generated six networks, with the largest containing 28 nodes and 30 edges, and identified **EGFR** gene as the only significant key hub gene (p-value = 0.01). **APOE** was not a member of this network. Further PPI analysis in the AsymAD yellow module generated five networks, with the largest containing 71 nodes and 81 edges, and **EGRF** was still the key hub gene (p-value = 0.007). In the equivalent AD yellow module, PPI analysis identified a single network generated with 284 nodes and 420 edges. This network contained far higher numbers of genes and now integrated the **APOE** gene as part of the network with **UBC** as the key hub (p-value = 4.12e63). This suggests protein interactions in this yellow module increases gradually through the course of the disease, with up-regulated **EGRF** interacting with more genes in the AsymAD group when compared to controls, followed by significant changes occurring in the AD group where up-regulated **UBC** gene takes more of a central role.

4. Discussion

4.1. Transcriptomic perturbations suggest AsymAD subjects could be an intermediate stage between control and MCI/AD

This study hypothesises the samples we have labelled as "AsymAD" subjects are an intermediate state between healthy ageing and MCI/AD. The assignment of these samples to the AsymAD group was based on the fact that these individuals had no reported clinical record of dementia prior to death as indicated in the MRC-LBB database; however, upon autopsy, these samples were found to have low levels of hallmark AD pathology, i.e. BRAAK Staging ≥ 2. Furthermore, an independent expression study identified the **TRIL** gene as being highly correlated with AD neuropathology, specifically tau pathology (Miyashita et al., 2014). Our study shows that the **TRIL** gene expression gradually increases from the Control to AsymAD, and then further increases in AD subjects (Fig. 2a), and this expression pattern is only observed in brain regions known to be affected by hallmark AD pathology (amyloid and NFT's), i.e. the EC, TC and FC, and not in the CB brain region. This observation suggests the phenotype assignments (controls, AsymAD, AD) are a suitable representation of three points in AD progression (assuming the AsymAD subjects are all prodromal AD), and as suggested by the **TRIL** gene expression pattern across brain regions and the fact the CB has been consistently reported to be partially spared from hallmark AD pathology (amyloid and NFT's), even those with severe AD pathology

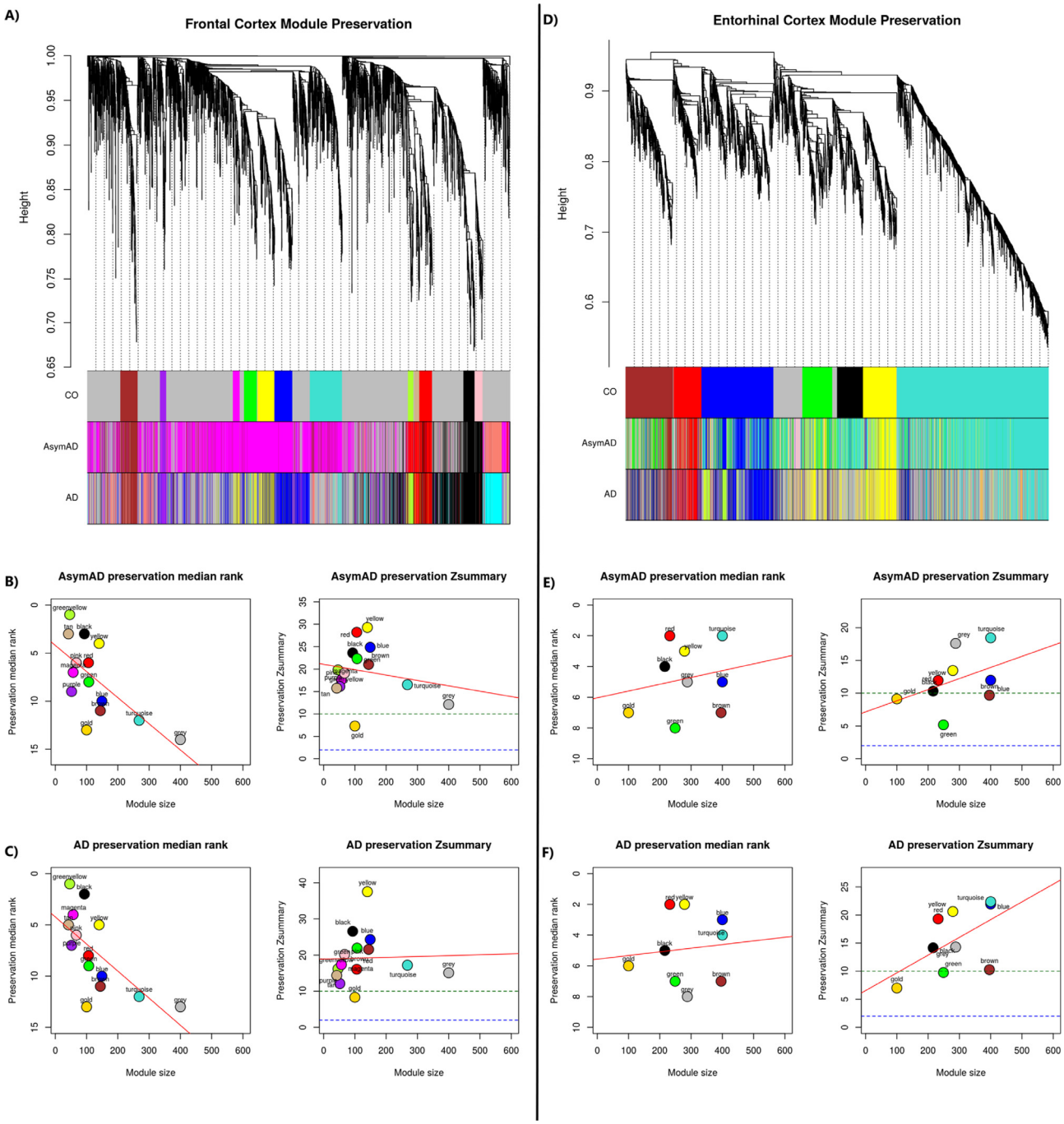

Fig. 5. Hierarchical clustering of genes and module preservations statistics for the frontal cortex is illustrated in A–C) and entorhinal cortex in D–F). In brief, a co-expression network based on "signed" adjacency was independently created for all three phenotypes (control, AsymAD and AD group), topological overlap calculated, and hierarchical clustering used to group genes into modules. For the Hierarchical clustering plots, the y-axis represents the network distance with values closer to 0 indicating greater similarity of probe expression across the control group. The x-axis represents the modules in the control, AsymAD and AD group. The AsymAD and AD module colours are mapped to the control group, with the AsymAD and AD colour panel representing how well the control modules are preserved through the disease. The red line in the module preservation statistics (B, C, E, F) represents the correlation between module size and preservation statistics. The gold module represents 100 random genes, and the grey module represents uncharacterised genes. The FC preservation plots (B and C) suggest all modules in the control group are relatively preserved in the AsymAD and AD group. In contrast, the EC preservation plots (E and F) suggest the green module is not well preserved in the AsymAD and AD group and requires further investigation.

(Convit et al., 2000); genes whose expression pattern differs significantly in the CB from that consistently seen in the EC, TC and FC tissues may be associated with hallmark AD pathology.

4.2. MOSPD3 gene is perturbed in the brains of AsymAD and blood of AD subjects

We identify **MOSPD3** as the only significant DE gene which is consistently down-regulated across all four brain regions in the AsymAD subjects suggesting this may be an early marker of cell dysfunction in AD. The **MOSPD3** gene encodes for a Motile Sperm Domain

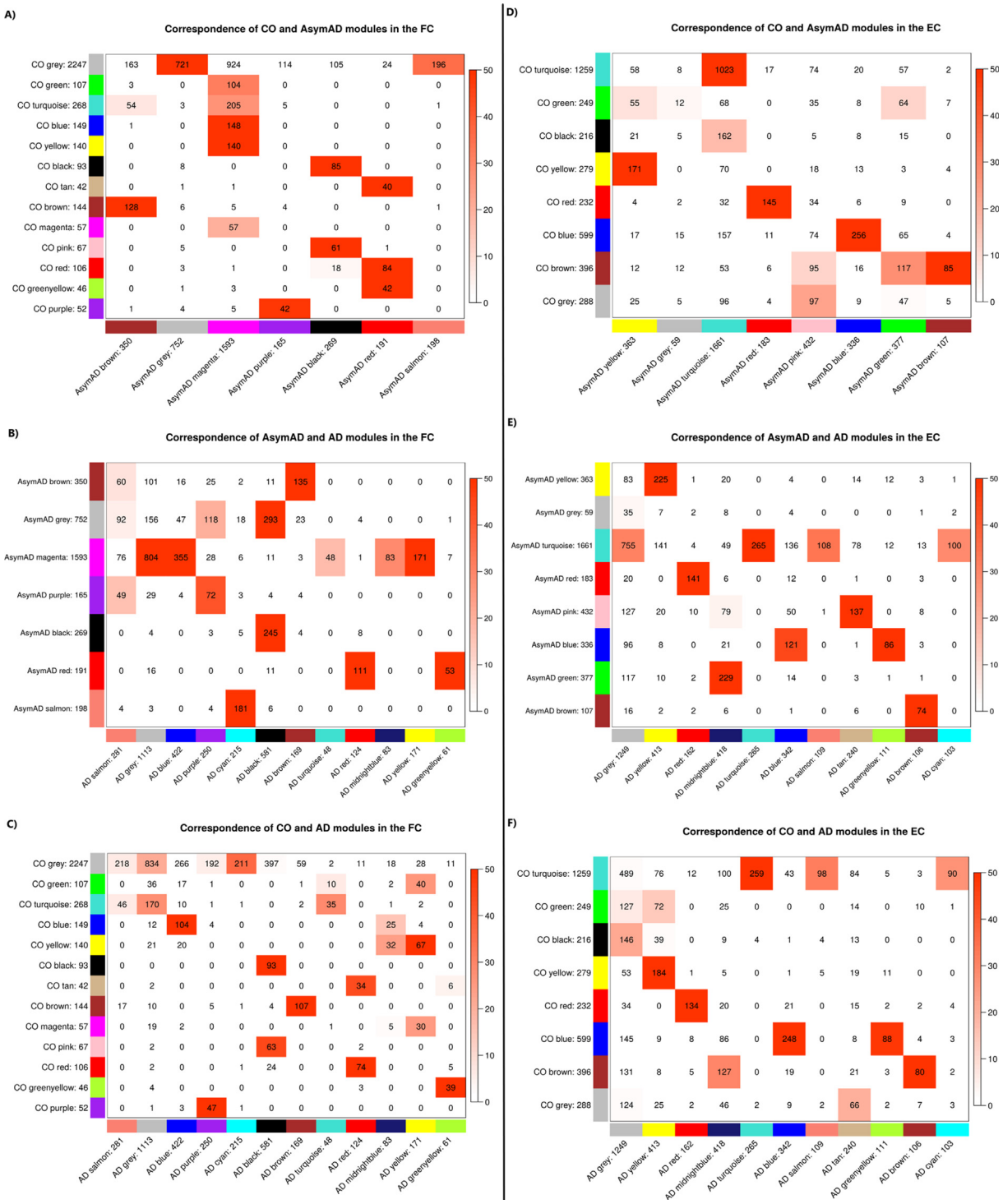

Fig. 6. Illustrates the "module correspondence" between A) FC control and AsymAD group, B) FC AsymAD and AD group, C) FC control and AD group, D) EC control and AsymAD, E) EC AsymAD and AD group, and F) EC control and AD groups. The modules represent clusters of highly correlated genes which were calculated independently in each brain region and diagnosis group. The module colours in the AsymAD and AD group were assigned based on the gene overlap of the control module. The total number of genes within each module is indicated next to the module colour. The numbers in each cell represent the overlap of genes between modules, with increased red intensity cells indicating increased significant overlap based on Fisher's exact test. This "module correspondence" plot provides a visual overview of how modules of highly correlated genes are preserved or disrupted between, control, AsymAD and AD groups. Module preservation statistics suggested the green module in the EC control group is not well preserved in the AsymAD and AD groups, indicating possible disruption to the co-expression network in this module. This "module correspondence" plot identifies the disrupted genes in the control green module synchronises with the genes of the AsymAD yellow module, identifying the yellow module for further investigation.

Table 2
Summary of frontal cortex co-expression module GSEA results.

Phenotype	Module	Module size	Most significant GSEA result	Pathway source	FDR adjusted q-value
CO	Black	93	Pentose phosphate pathway	HumanCyc	1.90E−02
	Blue	149	Parkinson's disease	KEGG	8.32E−10
	Brown	144	Differentiation Pathway	Wikipathways	2.26E−02
	Green	107	Oxidative phosphorylation	KEGG	1.43E−02
	GreenYellow	46	TNFs bind their physiological receptors	Reactome	1.60E−02
	Grey	2247	Processing of Capped Intron-Containing PrE−mRNA	Reactome	1.47E−02
	Magenta	57	Glucose metabolism	Reactome	6.26E−02
	Pink	67	Tight junction interactions	Reactome	1.49E−01
	Purple	52	Selenocysteine synthesis	Reactome	8.66E−53
	Red	106	TNF receptor superfamily (TNFSF) members mediating non-canonical NF-kB pathway	Reactome	3.88E−02
	Tan	42	Ovarian steroidogenesis	KEGG	4.43E−02
	Turquoise	268	Attenuation of gpcr signaling	BioCarta	6.13E−02
	Yellow	140	FCERI mediated MAPK activation	Reactome	5.62E−02
AsymAD	Black	269	RNA Polymerase II Transcription	Reactome	5.89E−04
	Brown	350	Differentiation Pathway	Wikipathways	3.17E−01
	Grey	752	mRNA Processing	Wikipathways	2.03E−02
	Magenta	1593	Oxidative phosphorylation - Homo sapiens (human)	KEGG	2.26E−11
	Purple	165	Peptide chain elongation	Reactome	5.03E−58
	Red	191	TNFR2 non-canonical NF-kB pathway	Reactome	2.71E−02
	Salmon	198	Hematopoietic cell lineage - Homo sapiens (human)	KEGG	2.82E−01
AD	Black	581	RNA Polymerase II Transcription	Reactome	6.61E−06
	Blue	422	Parkinson's disease	KEGG	3.09E−04
	Brown	169	Differentiation Pathway	Wikipathways	3.23E−02
	Cyan	215	Hematopoietic cell lineage - Homo sapiens (human)	KEGG	3.54E−01
	GreenYellow	61	Steroid hormone biosynthesis - Homo sapiens (human)	KEGG	2.82E−02
	Grey	1113	Neuronal System	Reactome	2.89E−01
	MidnightBlue	83	The citric acid (TCA) cycle and respiratory electron transport	Reactome	1.31E−02
	Purple	250	Eukaryotic Translation Elongation	Reactome	9.36E−66
	Red	124	TNF receptor superfamily (TNFSF) members mediating non-canonical NF-kB pathway	Reactome	9.31E−02
	Salmon	281	Histidine metabolism	EHMN	4.87E−04
	Turquoise	215	Neurotransmitter receptors and postsynaptic signal transmission	Reactome	1.15E−02
	Yellow	171	Fas	INOH	2.94E−03

Co-expression analysis in the frontal cortex brain region identified 13, 7, and 12 modules within the control, AsymAD and AD groups respectively. Gee set enrichment analysis was performed on each module, and the most significant result from each module is provided above.

Table 3
Summary of entorhinal cortex module co-expression results.

Phenotype	Module	Module size	Most significant GSEA result	Pathway source	FDR adjusted q-value
CO	Black	216	Hedgehog	INOH	7.69E−03
	Blue	599	Generic Transcription Pathway	Reactome	3.32E−05
	Brown	396	Ribosome	KEGG	4.71E−54
	Green	249	Neutrophil degranulation	Reactome	1.35E−01
	Grey	288	Pink/Parkin Mediated Mitophagy	Reactome	2.97E−01
	Red	232	Leptin Insulin Overlap	Wikipathways	6.01E−02
	Turquoise	1259	Neuronal System	Reactome	1.68E−08
	Yellow	279	Histidine metabolism	EHMN	5.24E−03
AsymAD	Blue	336	Generic Transcription Pathway	Reactome	2.04E−04
	Brown	107	Eukaryotic Translation Elongation	Reactome	1.13E−93
	Green	337	Processing of Capped Intron-Containing PrE−mRNA	Reactome	2.08E−06
	Grey	1661	Antigen processing and presentation	KEGG	2.10E−04
	Pink	432	Neural Crest Differentiation	Wikipathways	1.44E−01
	Red	183	Hematopoietic cell lineage	KEGG	2.21E−01
	Turquoise	1661	Parkinson's disease	KEGG	2.82E−10
	Yellow	363	Metallothioneins bind metals	Reactome	8.08E−03
AD	Blue	342	Generic Transcription Pathway	Reactome	6.31E−04
	Brown	106	Eukaryotic Translation Elongation	Reactome	4.46E−98
	Cyan	103	Cardiac conduction	Reactome	4.69E−04
	GreenYellow	111	Steroid hormone biosynthesis - Homo sapiens (human)	KEGG	2.04E−01
	Grey	1249	Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins.	Reactome	3.55E−06
	MidnightBlue	418	Processing of Capped Intron-Containing PrE−mRNA	Reactome	7.92E−07
	Red	162	Fat digestion and absorption	KEGG	2.80E−01
	Salmon	109	How progesterone initiates the oocyte maturation	BioCarta	2.08E−02
	Tan	240	Differentiation Pathway	Wikipathways	1.95E−01
	Turquoise	265	Neuronal System	Reactome	5.10E−11
	Yellow	413	Propanoate metabolism	EHMN	3.67E−03

Co-expression analysis in the entorhinal cortex brain region identified 8, 8, and 11 modules within the control, AsymAD and AD groups respectively. Gee set enrichment analysis was performed on each module, and the most significant result from each module is provided above.

Containing 3 protein [provided by RefSeq, Jul 2008] and has been reported to be significantly down-regulated (p-value = 6.47E-05) in the blood of AD subjects when compared to MCI subjects (Lunnon et al., 2012). This suggests MOSPD3 gene expression is significantly decreased in the brain before clinical signs of AD are apparent; however, blood gene expression levels are only significantly decreased after clinical signs of AD are apparent. It is difficult to interpret the biological relevance of this gene in AD, and further investigation is required.

4.3. Genes perturbed in brain regions affected explicitly by hallmark AD pathology may be associated with plaques and tangles, providing new therapeutic targets

Many molecular and cellular changes occur in AD brains including nerve cell death, atrophy, loss of neurons and accumulation hallmark AD pathology, specifically plaques and tangles. However, not all brain regions are affected to the same degree. The CB, which only accounts for 10% of the brain but contains over 50% of the brains total neurons, is often regarded as being partially spared from AD as plaques and tangles are generally not reported (Convit et al., 2000; Jacobs et al., 2017), and in this study are free from hallmark AD pathology in both AsymAD and AD subjects. For subjects with hallmark AD pathology (BRAAK ≥ 2, AsymAD and AD), genes significantly and consistently perturbed across the EC, TC and FC tissues that are not or are significantly reversed in the CB, may be associated with hallmark AD pathology, although, it remains unclear if these genes are causative or a response to the pathology itself.

We identified a total of 15 genes (**ALDH2, FBLN2**, **METTL7A**, **FLCN**, **ASPHD1, ARL5A, GPR162, HBA2, PCID2, NDRG2, BEND3, RAP1Gap, GPM6B, ANKEF1** and **NPC2**) with expression patterns suggestive of an association with hallmark AD pathology. Previous studies have already demonstrated an increased expression of **ALDH2** accelerated neurodegeneration and increased the accumulation of hyperphosphorylated tau protein (Ohsawa et al., 2008) in mice, while another demonstrated **NDRG2** might play a role in generating Aβ (Rong et al., 2017). Collectively, the 15 genes are not significantly enriched to be involved with any biological pathway; however, individually, these genes may play an essential role in the pathological aspect of AD and may provide new therapeutic targets for disease intervention.

4.4. Individuals with milder disease (early BRAAK pathology) show increased changes in the frontal cortex compared to the entorhinal cortex

The molecular changes in AD may initially begin in the FC, a region involved in working memory, as there were relatively more changes in the FC of mild pathology AD cases (AsymAD) than the EC region. This mirrors changes described in a longitudinal study involving ageing controls, where positron emission tomography (PET) scans were used to detect increased activity in the medial frontal cortex and decrease activity in the temporal lobe brain region in subjects who subsequently acquired cognitive impairment (Beason-Held et al., 2013). In addition, a higher degree of atrophy has also been detected in the FC than the temporal lobe brain region in MCI when compared to AD (Tabatabaei-Jafari et al., 2015). Our observations provide further evidence to suggest that brain perturbations at the molecular/transcriptomic level may initially occur in the FC before the presentation of more severe clinical symptoms consistent with a diagnosis of probable AD.

At the later point of the disease when clinical signs of AD are present, we find that the most substantial number of transcriptomic changes occur in the EC, followed by the TC, FC and only minor changes in the CB. This observation matches the common route AD neuropathology is seen to spread through the brain. Furthermore, we detect more DEG in the "Late AD" analysis compared to "Early AD" analysis, signifying more genes are disrupted in the later stage of the disease when the clinical symptoms of cognitive impairment are apparent.

4.5. Neutrophil, TYROBP network and the innate immune system disrupted in asymptomatic AD

Co-expression analysis of the EC brain region identified a green module of highly co-expressed genes which is disrupted in the AsymAD and AD subjects according to both module preservation statistics and cross-tabulation analysis. This green module is significantly enriched for "**neutrophil degranulation**", "**TYROBP casual network**" and the "**innate immune system**" processes in the control subjects, but not in the AsymAD or AD subjects, suggesting these pathways are most likely disrupted during the disease. Disturbance in TYROBP and Immune system pathways have been widely accepted in AD (Ma et al., 2015; Lambert et al., 2010), and a previous mouse study demonstrated disruptions in neutrophil levels impact memory loss and neurological features of AD (Pietronigro et al., 2017). We now suggest these pathways are specifically perturbed in the EC brain region early in the disease when hallmark AD pathology exists but clinical symptoms of AD are absent.

4.6. Disruption in brain energy pathways is detectable early in the disease

Co-expression analysis of the FC identifies disruptions in the "**glucose metabolism**", "**glucogenesis**" and"**oxidative phosphorylation**" processes in the AsymAD group, while DE analysis identified disruption in the "**gluconeogenesis and glycolysis**" pathway in the AD subjects. The brain critically relies on a constant supply of energy which is known to be generated by glycolysis followed by oxidative phosphorylation. Changes in the brain energy pathways have been widely accepted in AD (Shoffner, 1997; Cunnane et al., Jan. 2011), with a general decrease in glycolysis suggested to be a result of decreased brain functionality. Here we demonstrate disruptions in the energy pathway are detectable early in the disease, in subjects with low levels of AD pathology.

4.7. The glutamate-glutamine cycle is disturbed in AsymAD and AD subjects

Gene set enrichment analysis on DEGs identified the "**glutamate-glutamine cycle**" as the only biological pathway significantly perturbed across all brain regions in the AsymAD subjects. Furthermore, co-expression analysis of the EC brain regions was indicative of disruptions to the "**urea cycle and metabolism of arginine, proline, glutamate, aspartate and asparagine**" and "**astrocytic glutamate-glutamine uptake and metabolism**" in AsymAD and AD subjects, further confirming a possible disruption in glutamate-related activities in the brain.

Astrocytes are the most common form of neuroglial cells in the brain, and its primary function is to protect neurons against excitotoxicity by converting excess ammonia and glutamate to glutamine through the glutamate-glutamine cycle. Glutamate is the principal excitatory neurotransmitter in the brain and plays a vital role in linking carbohydrate and amino acid metabolism via the tricarboxylic acid (TCA) cycle. Glutamate is also a precursor of γ-aminobutyric acid (GABA) which binds and inhibits neuron activity; hence, an accumulation of glutamate can cause failures in synaptic connectivity, leading to deficient cognition and memory (Schousboe et al., 2014). A disruption in the glutamate-glutamine cycle would have a severe knock-on effect on many other biological pathways, including a disruption in amino acid metabolism which could explain the enrichment of "**urea cycle and metabolism of arginine, proline, glutamate, aspartate and asparagine**" in our results as well. In addition, glutamate stimulates astrocytes to derive energy from oxidative and glycolytic pathways, both of which have been identified as disrupted in AsymAD subjects.

The genes enriched in this pathway were all significantly up-regulated, indicating an overactive cycle. This could be part of the brain defence mechanism in preventing the accumulation of brain glutamate

levels or a broken cycle which is consistently being overactive, leading to decreased levels of brain glutamate, a phenomenon observed in AD subjects (Fayed et al., 2011). Targetting this pathway for AD treatment is extraordinarily complex and challenging as over inhibition or excitation may lead to increased levels of glutamate and glutamine respectively, both of which can be neurotoxic at high levels. Therapeutic compounds affecting the "**glutamate-glutamine cycle**" have already been identified, such as memantine, which is already a clinically established therapeutic drug used to for the symptomatic treatment of AD, which blocks N-methyl-d-aspartate (NMDA) receptors (Johnson and Kotermanski, 2006), essentially preventing excitotoxicity caused by neurotransmitters such as glutamate and ultimately increasing cognition temporarily.

The glutamate-glutamine cycle has been previously suggested to be disrupted in AD (Walton and Dodd, 2007), along with many other central nervous system disorders including Huntington's disease and Amyotrophic Lateral Sclerosis (ALS) (Matthews et al., 2012). Through this study, we now demonstrate this is one of the earliest biological pathways perturbed across all brain regions in AD, before clinical symptoms of AD are apparent, which can have a knock-on effect on other biological pathways also observed to be disrupted in the disease. Clinically established drugs to relieve AD symptoms already interact with this pathway and could also be effective in the asymptomatic period to prolong cognitive impairment, although clinical identification and measuring effectiveness in AsymAD subjects would be a challenge in itself.

4.8. Co-expression network changes indicate a shift from "cell proliferation" in AsymAD subjects to "removal of amyloidogenic proteins" in AD subjects

Protein-protein interactions identified **EGFR** as a key hub gene in both the control and AsymAD groups; however, it achieves more connections with neighbouring proteins in the AsymAD group, suggesting a possible increase in the **EGFR** activity. The **EGFR** gene is up-regulated in the AsymAD group and encodes for a transmembrane glycoprotein that binds to epidermal growth factor, leading to cell proliferation. In contrast, **EGFR** is replaced by **UBC** as the key hub gene in AD subjects, indicating it may play a more central role in the disease once the accumulation of hallmark AD pathology is at a level where clinical symptoms are apparent. The **UBC** gene is significantly up-regulated in the EC of AD subjects and is considered a stress gene which encodes for polyubiquitin precursor protein, a member of the ubiquitin-proteasome system (UPS) which removes toxic proteins and impacts on the amyloidogenic pathway of amyloid precursor protein (**APP**) processing that generates Abeta (Hong et al., 2014). A previous AD study had also observed **UBC** as a novel key hub gene and demonstrated **UBC** knockout models in C. elegans accelerated age-related AB toxicity (Mukherjee et al., 2017). Effectively, a portion of the co-expression network may have a central role involved in cell proliferation in control subjects, with increased activity in AsymAD subjects, followed by a shift towards the removal of toxic proteins such as amyloid beta in AD subjects.

5. Limitations

We cannot exclude the fact AsymAD group may represent a heterogeneous group consisting of cognitively normal, MCI, mixed dementia and AD subjects. It remains unclear these AsymAD subjects would remain free from clinical symptoms of dementia with longer survival and can be argued to be a possible extension to general ageing. However, the extent of BRAAK staging in AsymAD subjects was at a level consistently found with early cognitive impairment, and therefore, we make the strong assumption that these subjects are more likely to be prodromal AD rather than an extension of natural ageing. As AsymAD subjects are extremely rare, hence the low sample numbers in this study, larger AsymAD cohorts are required for better discovery and to validate our findings.

6. Conclusion

We believe this is the first study to explore the emergence of transcriptomic changes in the human brain from normal ageing through to mild AD pathology and diagnosis of AD. Using DE analysis, coupled with a "systems-biology" approach, we were able to detect disturbances in the energy pathways and the "**glutamate-glutamine cycle**" in the brains of subjects with mild and severe AD pathology. We found that changes in the FC brain region dominate in mild pathology, but are greater in the EC in subjects with more severe pathology, thus mirroring the changes in an aggregate spread in AD. The work undertaken in this study is an exploration of microarray-based transcriptomic changes, with future studies encouraged to explore changes based on RNA-Seq technologies, which are not limited by pre-defined probes allowing for unbiased insight into all transcripts.

Acknowledgements

This study presents independent research supported by the NIHR BioResource Centre Maudsley at South London and Maudsley NHS Foundation Trust (SLaM) & Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR, Department of Health or King's College London. We gratefully acknowledge capital equipment funding from the Maudsley Charity (Grant Ref. 980) and Guy's and St Thomas's Charity (Grant Ref. STR130505. Tissue samples were supplied by the London Neurodegenerative Diseases Brain Bank, which receives funding from the MRC and as part of the Brains for Dementia Research programme, jointly funded by Alzheimer's Research UK and Alzheimer's Society.

RJBD and SJN are supported by 1. Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust. 2. The National Institute for Health Research University College London Hospitals Biomedical Research Centre.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbi.2019.05.009.

References

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., Karagiannidou, M., 2016. World Alzheimer report 2016 improving healthcare for people living with dementia. Coverage, quality and costs now and in the future. Alzheimer's Dis. Int. 1–140.

Caselli, R.J., Reiman, E.M., 2012. Characterizing the preclinical stages of Alzheimer's disease and the prospect of presymptomatic intervention. Adv. Alzheimer's Dis. 3 (1), 405–416.

Yvette, G., Rodrigue, K.M., Kennedy, K.M., Park, D.C., Deng, W., Aimone, J.B., Gage, F.H., 2010. Beta-amyloid deposition and the aging brain karen. North 11 (4), 436–450.

Driscoll, I., Troncoso, J., 2011. Asymptomatic Alzheimer's disease: a prodrome or a state of resilience? Curr. Alzheimer Res. 8 (4), 330–335.

Saura, C.A., Parra-Damas, A., Enriquez-Barreto, L., 2015. Gene expression parallels synaptic excitability and plasticity changes in Alzheimer's disease. Front. Cell. Neurosci. 9, 1–14.

Nicole, P.D., Berchtold, C., Sabbagh, M.N., Beach, T.G., Kim, R.C., Cribbs, D.H., Cotman, C.W., Berchtold, N.C., Sabbagh, M.N., Beach, T.G., Kim, R.C., Cribbs, D.H., Cotman, C.W., 2015. Brain gene expression patterns differentiate Mild Cognitive Impairment from normal Aged and Alzheimer Disease. Neurobiol. Ageing 27 (3), 320–331.

Berchtold, N.C., Coleman, P.D., Cribbs, D.H., Rogers, J., Gillen, D.L., Cotman, C.W., 2013. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. Neurobiol. Aging 34 (6), 1653–1661.

Miller, J.A., Woltjer, R.L., Goodenbour, J.M., Horvath, S., Geschwind, D.H., 2013. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. Genome Med. 5 (5), 48.

Blalock, E.M., Buechel, H.M., Popovic, J., Geddes, J.W., Landfield, P.W., 2011. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. J. Chem. Neuroanat. 42 (2), 118–126.
Blalock, E.M., Geddes, J.W., Chen, K.C., Porter, N.M., Markesbery, W.R., Landfield, P.W., 2004. Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proc. Natl. Acad. Sci. U.S.A. 101 (7), 2173–2178.
Miller, J.A., Oldham, M.C., Geschwind, D.H., 2008. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J. Neurosci. 28 (6), 1410–1420.
Colangelo, V., Schurr, J., Ball, M.J., Pelaez, R.P., Bazan, N.G., Lukiw, W.J., 2002. Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling. J. Neurosci. Res. 70 (3), 462–473.
Sekar, S., McDonald, J., Cuyugan, L., Aldrich, J., Kurdoglu, A., Adkins, J., Serrano, G., Beach, T.G., Craig, D.W., Valla, J., Reiman, E.M., Liang, W.S., 2015. Alzheimer's disease is associated with altered expression of genes involved in immune response and mitochondrial processes in astrocytes. Neurobiol. Aging 36 (2), 583–591.
Li, Y., Wu, Z., Jin, Y., Wu, A., Cao, M., Sun, K., Jia, X., Chen, M., 2012. Analysis of hippocampal gene expression profile of Alzheimer 's disease model rats using genome chip. Bioinformatics 7 (5), 332–340.
Lambert, J.C., Grenier-Boley, B., Chouraki, V., Heath, S., Zelenika, D., Fievet, N., Hannequin, D., Pasquier, F., Hanon, O., Brice, A., Epelbaum, J., Berr, C., Dartigues, J.F., Tzourio, C., Campion, D., Lathrop, M., Amouyel, P., 2010. Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis. J. Alzheimer's Dis. 20 (4), 1107–1118.
Chen, J., Xie, C., Zhao, Y., Li, Z., Xu, P., Yao, L., 2016. Gene expression analysis reveals the dysregulation of immune and metabolic pathways in Alzheimer's disease. Oncotarget 7 (45), 1–6.
Liu, T., Ren, D., Zhu, X., Yin, Z., Jin, G., Zhao, Z., Robinson, D., Li, X., Wong, K., Cui, K., Zhao, H., Wong, S.T.C., 2013. Transcriptional signaling pathways inversely regulated in Alzheimer's disease and glioblastoma multiform. Sci. Rep. 3 (1), 3467.
Li, X., Long, J., He, T., Belshaw, R., Scott, J., 2015. Integrated genomic approaches identify major pathways and upstream regulators in late onset Alzheimer's disease. Sci. Rep. 5 (1), 12393.
Puthiyedth, N., Riveros, C., Berretta, R., Moscato, P., 2016. Identification of differentially expressed genes through integrated study of Alzheimer's disease affected brain regions. PLoS One 11 (4), 1–29.
Godoy, J.A., Rios, J.A., Zolezzi, J.M., Braidy, N., Inestrosa, N.C., 2014. Signaling pathway cross talk in Alzheimer's disease. Cell Commun. Signal. 12 (1), 23.
Ramanan, V.K., Kim, S., Holohan, K., Shen, L., Nho, K., Risacher, S.L., Foroud, T.M., Mukherjee, S., Crane, P.K., Aisen, S., Petersen, R.C., Weiner, M.W., Saykin, A.J., 2013. Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. Brain Imag. Behav. 6 (4), 634–648.
Di Paolo, G., Kim, T., 2012. Linking lipids to Alzheimer ' s disease: cholesterol and beyond. Aging (Albany. NY) 125, 284–296.
Liang, W.S., Reiman, E.M., Valla, J., Dunckley, T., Beach, T.G., Grover, A., Niedzielko, T.L., Schneider, L.E., Mastroeni, D., Caselli, R., Kukull, W., Morris, J.C., Hulette, C.M., Schmechel, D., Rogers, J., Stephan, D.A., 2008. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc. Natl. Acad. Sci. U.S.A. 105 (11), 4441–4446.
Ishii, K., Sasaki, M., Kitagaki, H., Yamaji, S., Sakamoto, S., Matsuda, K., Mori, E., 1997. Reduction of cerebellar glucose metabolism in advanced Alzheimer's disease. J. Nucl. Med. 38 (6), 925–928.
Oshiro, S., Morioka, M.S., Kikuchi, M., 2011, 2011... Dysregulation of iron metabolism in Alzheimer's disease, parkinson's disease, and amyotrophic lateral sclerosis. Adv. Pharmacol. Sci., vol.
Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. Neurobiol. Aging 18(4)1997 pp. S1–S2.
Allen, J.D., Chen, M., Xie, Y., 2010. Model-based background correction (MBCB): R methods and gui for illumina bead-array data. J. Cancer Sci. Ther. 1 (1), 25–27.
Du, P., Kibbe, W.A., Lin, S.M., 2008. lumi: a pipeline for processing Illumina microarray. Bioinformatics 24 (13), 1547–1548.
Buckberry, S., Bent, S.J., Bianco-Miotto, T., Roberts, C.T., 2014. MassiR: a method for predicting the sex of samples in gene expression microarray datasets. Bioinformatics 30 (14), 2084–2085.
Lazar, C., Meganck, S., Taminau, J., Steenhoff, D., Coletta, A., Molter, C., Weiss-Solís, D.Y., Duque, R., Bersini, H., Nowé, A., 2013. Batch effect removal methods for microarray gene expression data integration: a survey. Brief. Bioinform. 14, 469–490.
Hackstadt, A.J., Hess, A.M., 2009. Filtering for increased power for microarray data analysis. BMC Bioinformatics 10 (1), 11.
Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., Storey, J.D., 2012. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28 (6), 882–883.
Oldham, M.C., Langfelder, P., Horvath, S., Jan. 2012. Network methods for describing sample relationships in genomic datasets: application to Huntington's disease. BMC Syst. Biol. 6 (1), 63.
Smyth, G., 2005. limma: linear models for microarray data. Bioinform Comput. Biol. Solut. Using R Bioconductor 2005, 397–420.
Kamburov, A., Wierling, C., Lehrach, H., Herwig, R., 2009. ConsensusPathDB – A database for integrating human functional interaction networks. Nucleic Acids Res. 37 (Suppl. 1), 623–628.
Langfelder, P., Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559.
Langfelder, P., Luo, R., Oldham, M.C., Horvath, S., 2011. Is my network module preserved and reproducible? PLoS Comput. Biol. 7 (1).
Xia, J., Benner, M.J., Hancock, R.E.W., 2014. NetworkAnalyst - Integrative approaches for protein-protein interaction network analysis and visual exploration. Nucleic Acids Res. 42 (W1), 167–174.
Zaki, N., Mora, A., 2015. A comparative analysis of computational approaches and algorithms for protein subcomplex identification. Sci. Rep. 4 (1), 4262.
Miyashita, A., Hatsuta, H., Kikuchi, M., Nakaya, A., Saito, Y., Tsukie, T., Hara, N., Ogishima, S., Kitamura, N., Akazawa, K., Kakita, A., Takahashi, H., Murayama, S., Ihara, Y., Ikeuchi, T., Kuwano, R., 2014. Genes associated with the progression of neurofibrillary tangles in alzheimer's disease. Transl. Psychiatry 4, 1–8.
Convit, A., de Asis, J., de Leon, M., Tarshish, C., De Santi, S., Rusinek, H., 2000. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease☆. Neurobiol. Aging 21 (1), 19–26.
Lunnon, K., Ibrahima, Z., Proitsi, P., Lourdusamy, A., Newhouse, S., Sattlecker, M., Furney, S., Saleem, M., Soininen, H., Kłoszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Coppola, G., Geschwind, D., Simmons, A., Lovestone, S., Dobson, R., Hodges, A., 2012. Mitochondrial dysfunction and immune activation are detectable in early alzheimer's disease blood. J. Alzheimer's Dis. 30 (3), 685–710.
Jacobs, H.I.L., Hopkins, D.A., Mayrhofer, H.C., Bruner, E., van Leeuwen, F.W., Raaijmakers, W., Schmahmann, J.D., 2017. The cerebellum in Alzheimer's disease: evaluating its role in cognitive decline. Brain.
Ohsawa, I., Nishimaki, K., Murakami, Y., Suzuki, Y., Ishikawa, M., Ohta, S., 2008. Age-dependent neurodegeneration accompanying memory loss in transgenic mice defective in mitochondrial aldehyde dehydrogenase 2 activity. J. Neurosci. 28 (24), 6239–6249.
Rong, X.-F., Sun, Y.-N., Liu, D.-M., Yin, H.-J., Peng, Y., Xu, S.-F., Wang, L., Wang, X.-L., 2017. The pathological roles of NDRG2 in Alzheimer's disease, a study using animal models and APPwt-overexpressed cells. CNS Neurosci. Ther. 23 (8), 667–679.
Beason-Held, L.L., Goh, J.O., An, Y., Kraut, M.A., O'Brien, R.J., Ferrucci, L., Resnick, S.M., 2013. Changes in brain function occur years before the onset of cognitive impairment. J. Neurosci. 33 (46), 18008–18014.
Tabatabaei-Jafari, H., Shaw, M.E., Cherbuin, N., 2015. Cerebral atrophy in mild cognitive impairment: a systematic review with meta-analysis. Alzheimer's Dement Diagnosis, Assess. Dis. Monit. 1 (4), 487–504.
Ma, J., Jiang, T., Tan, L., Yu, J.T., 2015. TYROBP in Alzheimer's disease. Mol. Neurobiol. 51 (2), 820–826.
Pietronigro, E.C., Della Bianca, V., Zenaro, E., Constantin, G., 2017. NETosis in Alzheimer's disease. Front. Immunol. 8, 1–12.
Shoffner, J.M., 1997. Oxidative phosphorylation defects and Alzheimer's disease. Neurogenetics 1 (1), 13–19.
Cunnane, S., Nugent, S., Roy, M., Courchesne-Loyer, A., Croteau, E., Tremblay, S., Castellano, A., Pifferi, F., Bocti, C., Paquet, N., Begdouri, H., Bentourkia, M., Turcotte, E., Allard, M., Barberger-Gateau, P., Fulop, T., Rapoport, S.I., Jan. 2011. Brain fuel metabolism, aging, and Alzheimer's disease. Nutrition 27 (1), 3–20.
Schousboe, A., Scafidi, S., Bak, L.K., Waagepetersen, H.S., Mary, C., Sciences, M., 2014. Glutamate and ATP at the Interface of Metabolism and Signaling in the Brain. Adv. Neurobiol. 11, 13–30.
Fayed, N., Modrego, P.J., Rojas-Salinas, G., Aguilar, K., 2011. Brain glutamate levels are decreased in alzheimer's disease: a magnetic resonance spectroscopy study. Am. J. Alzheimers Dis. Other Demen. 26 (6), 450–456.
Johnson, J.W., Kotermanski, S.E., 2006. Mechanism of action of memantine. Curr. Opin. Pharmacol. 6 (1), 61–67.
Walton, H.S., Dodd, P.R., 2007. Glutamate-glutamine cycling in Alzheimer's disease. Neurochem. Int. 50 (7–8), 1052–1066.
Matthews, D.C., Henter, I.D., Zarate, C.A., 2012. Targeting the glutamatergic system to treat Major Depressive Disorder. Drugs 72 (10), 1313–1333.
Hong, L., Huang, H.-C., Jiang, Z.-F., 2014. Relationship between amyloid-beta and the ubiquitin–proteasome system in Alzheimer's disease. Neurol. Res. 36 (3), 276–282.
Mukherjee, S., Russell, J.C., Carr, D.T., Burgess, J.D., Younkin, M.A., Serie, D.J., Boehme, K.L., Kauwe, J.S.K., Naj, A.C., Fardo, D.W., Dickson, D.W., Montine, T.J., Ertekin-Taner, N., Kaeberlein, M.R., Crane, P.K., Allen, M., Serie, D.J., Boehme, K.L., Kauwe, J.S.K., Naj, A.C., Fardo, D.W., Dickson, D.W., Montine, T.J., Ertekin-Taner, N., Kaeberlein, M.R., Crane, P.K., 2017. Systems biology approach to late-onset Alzheimer's disease genome-wide association study identifies novel candidate genes validated using brain expression data and Caenorhabditis elegans experiments. Alzheimer's Dement 1–10.