ELSEVIER

CrossMark

Alzheimer's & Dementia 14 (2018) 215-229

Alzheimer's & Dementia

Featured Article

A systematic integrated analysis of brain expression profiles reveals *YAP1* and other prioritized hub genes as important upstream regulators in Alzheimer's disease

Min Xu[a,b,1], Deng-Feng Zhang[a,**,1], Rongcan Luo[a,b], Yong Wu[a,b], Hejiang Zhou[a], Li-Li Kong[a,c], Rui Bi[a], Yong-Gang Yao[a,b,d,*]

[a]*Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan, China*

[b]*Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China*

[c]*Institute of Health Science, Anhui University, Hefei, Anhui, China*

[d]*CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China*

Abstract

Introduction: Profiling the spatial-temporal expression pattern and characterizing the regulatory networks of brain tissues are vital for understanding the pathophysiology of Alzheimer's disease (AD).

Methods: We performed a systematic integrated analysis of expression profiles of AD-affected brain tissues (684 AD and 562 controls). A network-based convergent functional genomic approach was used to prioritize possible regulator genes during AD development, followed by functional characterization.

Results: We generated a complete list of differentially expressed genes and hub genes of the transcriptomic network in AD brain and constructed a Web server (www.alzdata.org) for public access. Seventeen hub genes active at the early stages, especially *YAP1*, were recognized as upstream regulators of the AD network. Cellular assays proved that early alteration of *YAP1* could promote AD by influencing the whole transcriptional network.

Discussion: Early expression disturbance of hub genes is an important feature of AD development, and interfering with this process may reverse the disease progression.

Keywords: Alzheimer's disease; Differential expression; Convergent functional genomic; Upstream regulator; Database

1. Introduction

Alzheimer's disease (AD) is the most prevalent neurodegenerative disease in the elderly and is characterized by progressive memory loss and cognitive impairment. Pathological hallmarks of AD include the presence of extracellular amyloid β (Aβ) plaques and intracellular neurofibrillary tangles, synaptic dysfunction, neuronal loss, and brain atrophy [1]. The occurrence and development of AD is affected by age, genetic, and environmental factors [2]. Previous linkage analyses have identified the Aβ production–related genes *APP*, *PSEN1*, and *PSEN2* as the causal genes for familial AD, whereas genome-wide association studies (GWASs) have identified two dozen of the susceptibility loci responsible for sporadic AD [3,4]. Despite the fact that remarkable advances have been made in the understanding of the genetic basis of AD, the pathophysiology of AD is not well understood. A complete characterization of the

The authors declare no conflict of interest.

[1]These authors contributed equally to this work.

*Corresponding author. Tel./Fax: +86-871-68125440.

**Corresponding author. Tel.: +86-871-68125441; Fax: +86-871-68125440.

E-mail address: zhangdengfeng@mail.kiz.ac.cn (D.-F.Z.), yaoyg@mail.kiz.ac.cn (Y.-G.Y.)

http://dx.doi.org/10.1016/j.jalz.2017.08.012

transcriptomic alterations and the regulatory mechanisms that underpin AD may provide essential evidence to fill the gap.

Gene expression profiling of postmortem brain tissues from AD patients and normal controls has identified numerous dysregulated genes and contributed to the understanding of the biological processes disrupted during AD pathogenesis [5–31]. Among the list of differentially expressed genes (DEGs), dysfunction of mitochondrial pathways, calcium signaling, and neuroinflammation were consistently observed in AD, to name a few [32]. Nevertheless, the statistical power and consistency of previous individual studies were limited, mainly due to relatively small sample sizes [5–31]. A comprehensive, robust, cross-validated list of DEGs based on a large sample size is urgently needed in AD research. It is also important to understand the spatial-temporal expression pattern and regulatory network of these DEGs. Identifying the genes which play central regulatory roles during AD pathogenesis, and finding how these genes regulate the downstream DEGs, is vital for understanding both the pathophysiology of AD and looking for potential targets for drug therapy (Fig. 1A).

In this study, we performed an integrative analysis of available high-throughput brain expression profiling data sets from AD patients and controls using a convergent functional genomic (CFG) method [33,34], in an attempt to answer the aforementioned questions (Fig. 1B). We merged all the available expression data for four brain regions affected by AD (entorhinal cortex [EC], hippocampus [HP], temporal cortex [TC], and frontal cortex [FC]) [5–30] through cross-platform normalization to achieve the largest AD brain expression data set for these brain regions (1246 samples, including 139 HP, 78 EC, 697 FC, and 332 TC). We investigated the regulation pattern of DEGs in AD brain and prioritized hub genes and potential upstream regulator genes by the CFG method [33,34], which integrated various levels of AD-related data including GWAS, protein-protein interaction (PPI), brain expressional quantitative trait loci (eQTL), and expression data of mouse AD models. We have been able to provide a complete and robust list of DEGs and hub regulators and identified several candidate upstream regulators in DEG networks, such as *YAP1* in the glial cell differentiation module. Our functional experiments with *YAP1* have suggested that expression perturbations of upstream regulator genes at the early stage of

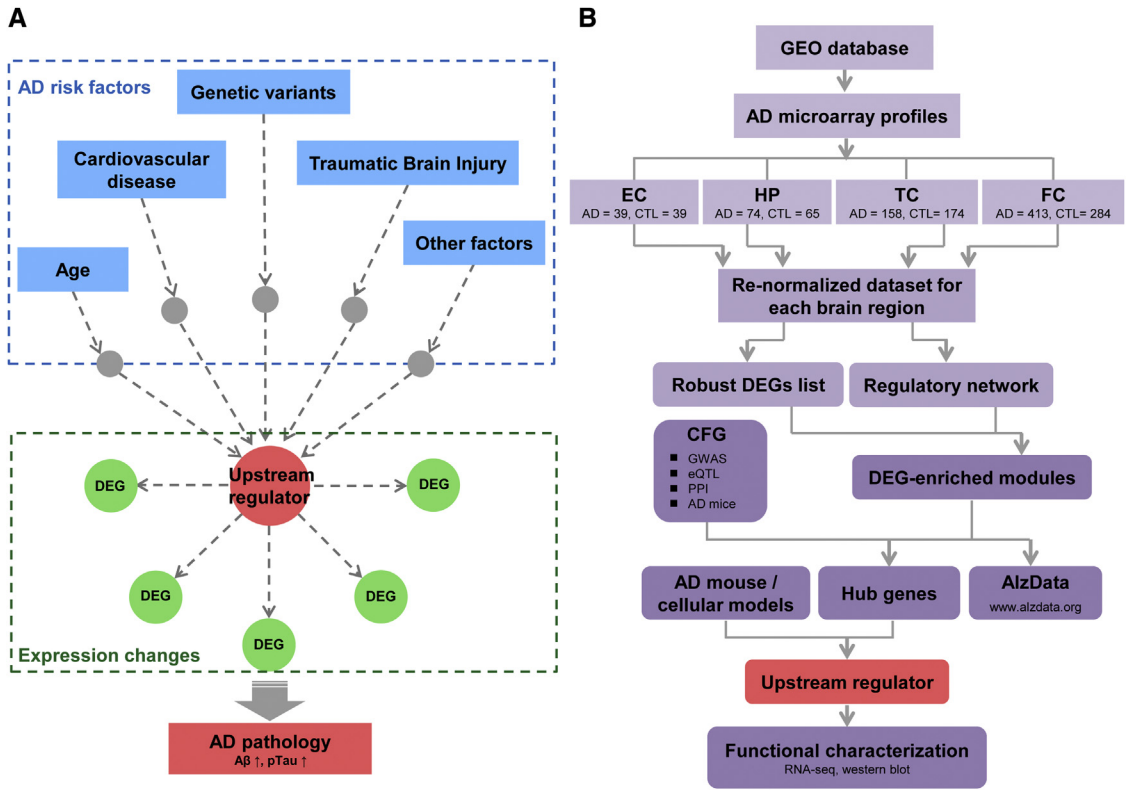


Fig. 1. Rationale and workflow of the present study. (A) Upstream regulator genes would respond to AD risk factors, influence expression of downstream genes, and thus promote AD pathology. Aβ ↑: increase of amyloid β level; pTau ↑: increase of phosphorylated tau. (B) We retrieved and renormalized all relevant expression data sets of samples with and without AD in entorhinal cortex (EC), hippocampus (HP), temporal cortex (TC), and frontal cortex (FC) and then analyzed for DEGs in the compiled data sets. We explored the gene prioritization and regulatory pattern in AD using coexpression network analysis and spatial-temporal expression data of AD mouse and cellular models and identified hub genes in the DEG-enriched coexpression networks. A convergent genomic approach (CFG) integrating multiple lines of evidence (including population genetic association, genetic regulation of expression, protein-protein interaction, early expression alteration, and pathology correlation in AD mice) was used to prioritize potential upstream genes. Abbreviations: eQTL, expressional quantitative trait loci; AD, Alzheimer's disease; CTL, control; DEG, differentially expressed gene; GEO, Gene Expression Omnibus; GWAS, genome-wide association study; PPI, protein-protein interaction.

AD development might be damaging, by affecting expression of downstream genes and so further promoting AD progression. Candidate upstream regulators found in this study might act as potential therapeutic targets for the treatment of AD. The list of all AD-related DEGs and their biological prioritization are available online through our Web server at http://www.AlzData.org.

2. Materials and methods

2.1. Study design and rationale

We initially retrieved and renormalized all relevant expression data of four brain regions (EC, HP, TC, and FC) with and without AD and then pinpointed the DEGs in the compiled data sets. We then explored the regulatory pattern of these DEGs through coexpression network analysis of the renormalized expression data. We hypothesized that the DEGs might be regulated by several hub genes in the DEG-enriched coexpression modules/networks. A CFG approach integrating multiple lines of evidence [33,34], for example, population genetic association, genetic regulation of expression, PPI, gene expression alteration, and pathology correlation in AD mice, was used to prioritize hub genes. In addition, hub genes that met the following criteria were considered likely to be early causal/driver upstream genes: (1) had a response to inducing factors (e.g., Aβ deposits or genetic risk alleles), (2) had expression alterations earlier than the emergence of AD pathology, and (3) were highly correlated with expression of downstream DEGs. Potential upstream regulators were functionally characterized in vitro. The flowchart and rationale of this study are shown in Fig. 1A and 1B.

2.2. AD expression data collection and filtration

All the original microarray data regarding AD were retrieved from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) by searching with the keyword "Alzheimer." An exhaustive, nonredundant data retrieval was achieved for EC, HP, TC, and FC following a series of criteria (see Supplementary Methods). By January 2016, 20 GSE series of expression data sets had been obtained. After manual inspection and filtration, there were four data sets retained for EC (GSE26927 [18,29], GSE48350 [8], GSE26972 [16], GSE5281 [7,9]); five data sets for HP (GSE28146 [14], GSE48350 [8], GSE5281 [7,9], GSE29378 [23], GSE36980 [25]); five data sets for TC (GSE37263 [13], GSE29652 [15], GSE36980 [25], GSE15222 [10], GSE5281 [7,9]); and eight data sets for FC (GSE12685 [11], GSE48350 [8], GSE66333 [30], GSE53890 [27], GSE36980 [25], GSE15222 [10], GSE5281 [7,9], GSE33000 [28]) (Supplementary Table 1). A total of 1246 human postmortem brain samples (684 AD and 562 controls) were compiled for the detailed analysis. Note that among the 20 data sets, GSE15222 for TC and GSE33000 for FC have far more samples than any other studies of the same brain regions, combining these two data sets with others may blur effects of studies with a relatively small sample size. Therefore, we used these two studies as independent validation data sets for the compiled data set of different studies. Thus, we have 269 AD and 271 controls from four brain regions in stage 1 (for EC: 39 vs. 39; for HP: 74 vs. 65; for TC: 52 vs. 39; and for FC: 104 vs. 128) for subsequent analyses (DEG detection and coexpression network construction). Detailed information about data collection, filtration, and sample compositions of each brain region was shown in Supplementary Methods and Supplementary Table 1. The original metadata of each data set including all sample information were accessible at the Alzdata.org Web server (http://www.alzdata.org/download.html).

2.3. Raw data preprocessing, renormalization, and detection of DEGs

Individual expression data set for each brain region was subjected to data normalization, log2 transformation, probe filtration, and probe mapping to entrez gene IDs (see Supplementary Methods). All processed expression data from the same brain region were merged by algorithm ComBat in R package inSilicoMerging [35,36]. The ComBat is an empirical Bayes method that is recommended for removal of batch effects [37,38]. As revealed by our principal variance component analysis [39], ComBat used in this study eliminated the batch effects almost completely in all four combined brain regions (Supplementary Methods and Supplementary Fig. 1).

Cross-platform normalized expression data for each brain region was used to detect DEGs between AD patients and normal controls. The expression profiling was adjusted for age and gender by using a linear regression model during the identification of DEGs. Differential expression analysis was conducted by R package limma and the Benjamini-Hochberg's method was used to correct for multiple comparisons [40].

2.4. Coexpression network construction and identification of hub genes

Coexpression networks for individual brain region (region-specific network) and combined expression profiles of all four brain regions (multitissue network) were constructed by using package weighted gene coexpression network analysis in R (https://www.r-project.org) [41–46] for cases and controls, respectively. Expression residuals after adjustment for age and sex using a linear model were used as input for network construction (Supplementary Methods). Coexpression modules were detected in cases and controls from stage-1 samples, and each module was assigned to a unique color. Pairwise preservation tests (indicated by Z-score and module membership (kME) correlations computed by weighted gene coexpression network analysis) were applied to assess module conservation across brain regions (in cases and controls, respectively), and between cases and controls in multitissue networks. Intramodule connectivity (k) value for each gene in cases (k_{AD}) or controls (k_{CTL}) was calculated independently based on the AD-only or control-only multitissue networks, respectively, and was scaled relative to the

maximally connected gene in each module. As most of the coexpression modules were strongly preserved across brain regions (Supplementary Figs. 2–4), multitissue networks constructed with the combined expression profiles of all four brain regions in AD and controls were used for subsequent analyses. Hub genes were defined as genes with $k > 0.8$ in the AD or control network and were subjected to subsequent analyses [41–46]. Of note, there were some hub genes with a significantly lower (loss-of-connectivity) or higher (gain-of-connectivity) connectivity in the AD network compared with the control network, albeit the overall pattern of the connectivity distribution was preserved in the AD and control networks ($r = 0.7$, $P < 2.2 \times 10^{-16}$). We presented the summary statistics of the connectivity for all genes at our AlzData.org Web server. Anyone who is interested in loss-of-connectivity or gain-of-connectivity hub genes can access this Web resource for reference.

2.5. Convergent functional genomics

The CFG approach integrated multiple lines of AD-related evidence [33,34] and scored each gene based on this evidence (CFG score). A gene was defined as AD-related if it (1) had at least one locus being significantly associated with AD ($P < .001$, from the International Genomics of Alzheimer's Project [4]); (2) was associated with eQTLs that showing an AD risk in International Genomics of Alzheimer's Project ($P_{GWAS} < .001$ and $P_{eQTL} < .001$) [4,47]; (3) physically interacted with any AD core genes (*APP*, *PSEN1*, *PSEN2*, *APOE*, or *MAPT*) ($P < .05$) according to the Human Gene Connectome. [48,49]; (4) were correlated with AD pathology in AD mice ($P < .05$) at the expression level [50]; or (5) showed an early differential expression in AD mice ($P < .05$) [50]. One point was assigned if any of the aforementioned evidence was observed; otherwise zero point, leading to a CFG score ranging from 0 to 5 points (Supplementary Methods).

2.6. Cellular characterization of an upstream hub gene

A U251 glia cell line (U251-APP cell) with a stable expression of mutant APP (APP-K670N/M671L), constructed in our previous studies [51,52], was used in the cellular assay. We harvested cells with knockdown or overexpression of the hub gene at 72 hours after transfection for RNA sequencing (RNA-seq) and western blotting (WB). RNA-seq of each treatment was performed in triplicate (data are accessible at AlzData.org, http://www.alzdata.org/download.html; or GSE100891, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100891). WB was repeated three times (Supplementary Methods). DEGs were detected in cells with *YAP1* knockdown or overexpression versus scramble cells by using RNA-seq data. WB for GAPDH was used as a loading control to measure the densitometry of YAP1, BACE1, PSEN1, PSEN2, nicastrin, amyloid β_{42} ($A\beta_{1-42}$), CDK5, and GSK3α/β. The densitometric signal of phosphorylated tau at threonine 181 (pTau181) or serine 396 (pTau396) was determined by the ratio of the phosphorylated protein to total tau. ImageJ 1.50i (National Institutes of Health, Bethesda, Maryland, USA) was used to evaluate the densitometry.

2.7. Statistical analysis

Genes with log2 fold change greater than 0.1 ($|logFC| > 0.1$) and FDR smaller than 0.05 ($FDR < 0.05$) were defined as DEGs in AD patients in the combined data set. For data sets of AD mouse or cellular models, and data set of cells with *YAP1* knockdown or overexpression, genes with $|logFC| > 0.1$ and $P < .05$ were regarded as DEGs. Enrichment of biological process in Gene Ontology of target gene sets was analyzed by DAVID 6.8 (https://david.ncifcrf.gov) [53]. Fisher's exact test was used to test whether DEGs or cell type–specific genes (Supplementary Methods) were significantly enriched in a target gene set, and the Benjamini-Hochberg's method in R package was used to adjust for multiple comparisons. Comparisons of relative protein levels between two groups from the WB experiment were conducted by the Student's *t* test using the PRISM software (GraphPad Software, Inc., La Jolla, CA, USA). Network was visualized by using GeneMANIA plugin in the Cytoscape software [54,55].

3. Results

3.1. DEGs list and AlzData.org Web server of compiled AD brain expression profiling

In total, 1246 postmortem brain samples containing highly curated and focused expression data from four brain regions were obtained (Supplementary Table 1). As the fold change was affected by the sample size that was used for identifying DEGs and the analyzed sample sizes of the four brain regions were different in this study, an increase of the cutoff of fold change might cause a bias for scoring DEGs in these brain regions (Supplementary Table 2). Therefore, we defined the DEGs using a relatively low threshold ($|logFC| > 0.1$, $FDR < 0.05$) for differential expression. Around 9% to 20% of the total genes in the merged data set could be identified as DEGs in the four brain regions (Supplementary Table 3). Our final DEG list of the merged data sets captured a high proportion of the DEGs in the original studies, ranging from 30% to 70% (except for GSE12685 [13%]; Supplementary Table 4), indicating a reasonably high confidence of the DEGs identified by the merged data sets. With the enlarged sample size, we were able to identify DEGs that were missed in previous individual studies (from 15% to 70%, Supplementary Table 3), especially when the sample size of individual studies was relatively small (e.g., in EC and HP). This quite variable range might be partially caused by limited coverage or power of the originally individual data sets and a relatively low threshold in our DEG definition. Among the list, 139 genes had a consistently differential expression in all four brain regions, with 35 genes being upregulated and 104 genes being downregulated (Fig. 2A

and Supplementary Table 5). Consistent with previous studies [7,9,10,28], the expression pattern of well-known AD-risk genes, such as *APP*, *PSEN1*, and *PSEN2*, across EC, HP, TC, and FC was only slightly altered or unchanged in AD patients. Only one GWAS risk gene *MEF2C* [4] showed consistent downregulation in all four brain regions (Supplementary Table 6).

For easy access to the complete list of DEGs for each brain region (top DEGs for each region were listed in Supplementary Table 7), we established a Web server at www.AlzData.org. At this Web server, it is possible to search, browse, or download the differential expression pattern of genes of interest in either an individual GSE series or the normalized data sets along with graphic views and statistic results (Supplementary Figs. 5–8).

3.2. Dysregulated pathways in four regions of AD brain

We investigated the biological pathways dysregulated in different brain regions based on the identified DEGs. As shown in Fig. 2B, DEGs related to the metabolic processes were significantly enriched in all four brain regions, which indicated a global dysregulation of energy metabolism across all brain regions during AD development. Processes involved in the synaptic functions were enriched in EC, HP, and TC, but not FC (which is the last brain region affected by AD [1]). The cortex had an enrichment of DEGs involved in the neurotrophin tropomyosin-related kinase receptor signaling pathway and mitogen-activated protein kinase (MAPK) pathway. In contrast to early affected brain regions (EC, HP, and TC), DEGs in FC were mainly involved in toll-like receptors signaling pathways, which have been reported to mediate microglia activation and promote clearance of Aβ [56]. The overall pattern of the DEG-enriched pathways (Supplementary Table 8) based on our normalized data sets was consistent with previous reports [5,7–9,11,13–15,20,23,24,57,58]. Intriguingly, DEGs related to the Hippo signaling pathway were enriched in EC (the first brain region affected by AD [1]), which has not been reported in individual expression studies.

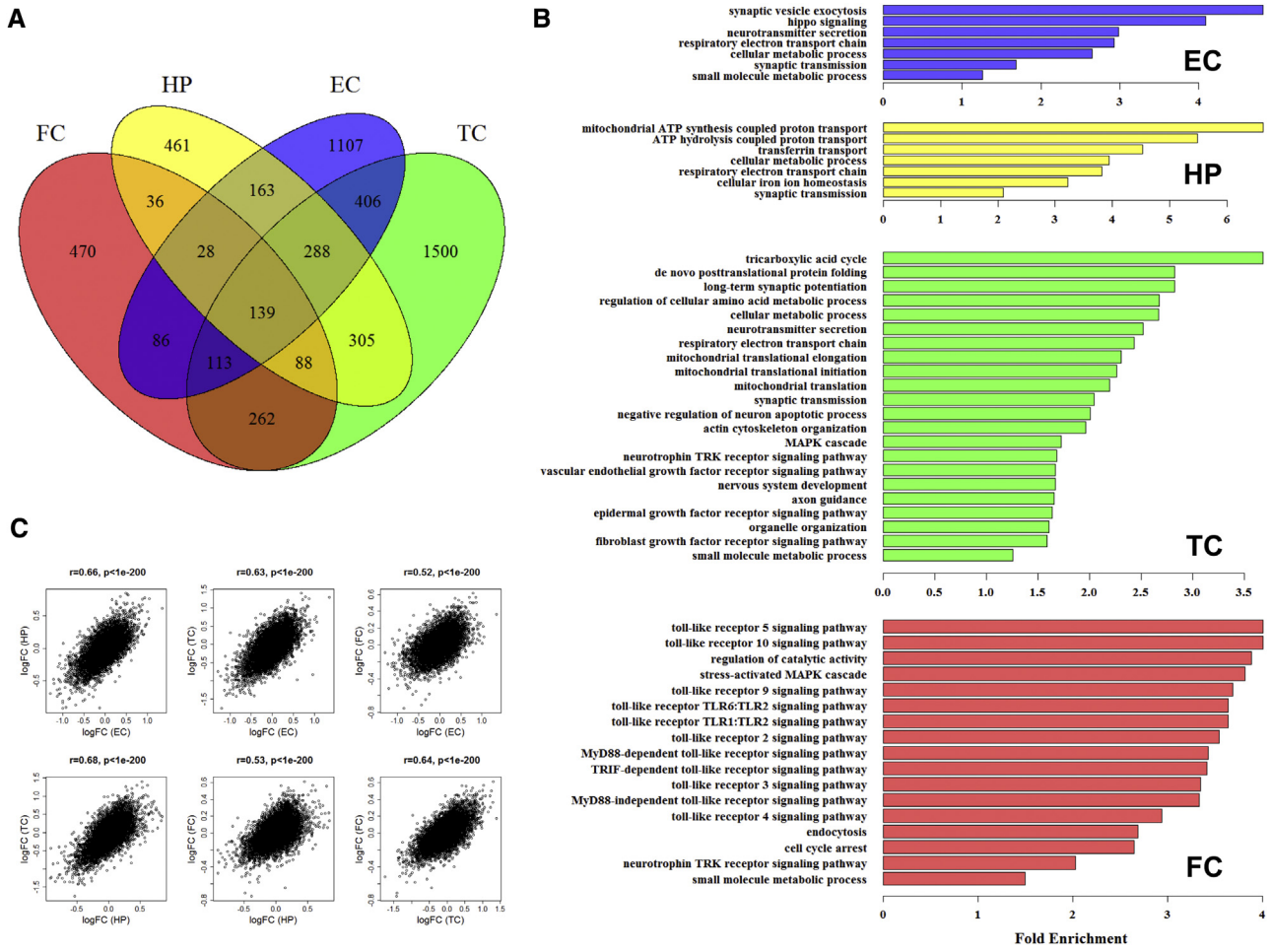


Fig. 2. Summary of differential expression. Distribution of (A) differential expressed genes and (B) disrupted pathways in entorhinal cortex (EC), hippocampus (HP), temporal cortex (TC), and frontal cortex (FC). Enrichment of biological processes in Gene Ontology (GO) of the target gene sets was analyzed by DAVID 6.8 (https://david.ncifcrf.gov) [53]. Shown are enriched terms with FDR <0.05; x-axis, fold enrichment; y-axis, names of the enriched GO terms. (C) Correlation of gene expression changes between each pair of brain regions. Abbreviations: logFC, log2 fold change of expression of a gene in AD patients compared with controls; *p*, *P*-value calculated by Pearson's correlation; *r*, Pearson's correlation; AD, Alzheimer's disease.

3.3. Coexpression network and DEG-enriched modules

In addition to the identification of dysregulated genes and pathways, we attempted to explore the disturbed networks and their regulatory pattern in AD brain. A coexpression network, which served as a high-order functional module, was constructed and subjected to subsequent analysis. Expression alterations and coexpression network structures were highly correlated (Fig. 2C, $P < 1 \times 10^{-200}$) and preserved among EC, HP, TC, and FC (Supplementary Figs. 2–3), indicating a similar expression network organization in brain regions. Therefore, to increase the statistical power and achieve a more refined network structure, we combined the expression data of all four brain regions and built a multitissue coexpression network for AD patients (AD network) and controls (control network) using the stage 1 samples (269 AD patients and 271 controls), respectively. We noted that different brain tissues might have distinct expression and network patterns during AD progression. Our current strategy of using a multitissue coexpression network does not rule out a region-specific network or regulation pattern.

In the AD network, 13 transcriptional modules were found, with a modular size ranged from 155 to 3626 genes. The observed variable size of the modules might reflect different levels of organization and complexity of gene regulation in AD brain. Among these modules, five modules were significantly enriched with DEGs (Fig. 3A and Supplementary Table 9), and the genes in these five modules were mainly involved in synaptic transmission, myelination, transcription regulation, glial cell differentiation, and regulation of inflammation response, respectively (Fig. 3B). Single-cell RNA-Seq data (GSE67835 [59]) showed that genes in each module had a cell-type specificity and were consistent with their predicted modular biological functions (Fig. 3B). AD core genes, such as *APP*, *PSEN1*, *PSEN2*, *MAPT*, and *APOE*, all appeared in the DEG-enriched modules, supporting the biological relevance of these modules with AD. These five DEG-enriched modules were defined as disturbed networks in AD and were used for identifying hub regulators.

3.4. Identification and prioritization of hub genes

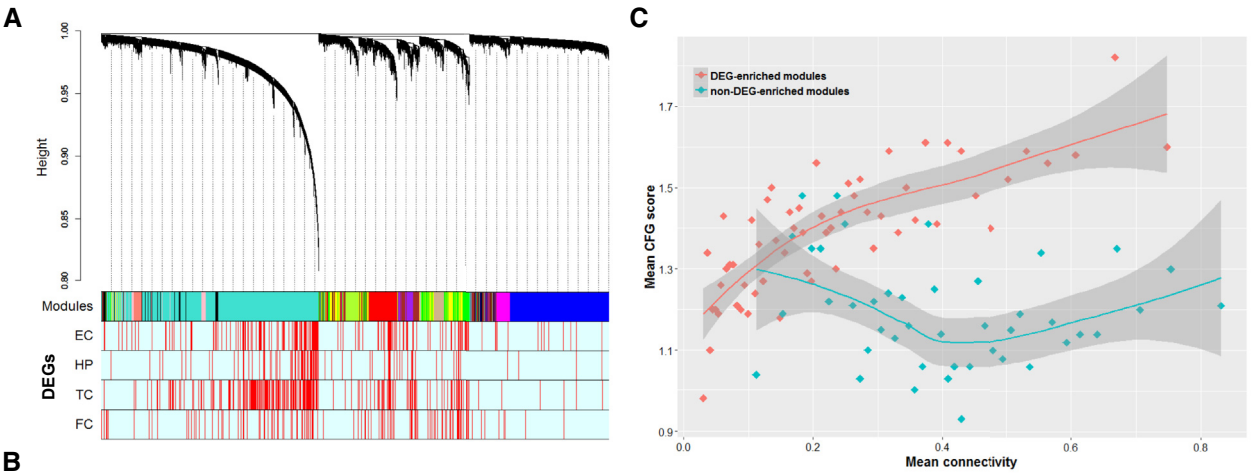
A hub gene holds nodes together and is important for the integrity and proper functioning of the whole network [60]. In the five modules, we identified 156 hub genes with $k > 0.8$ (AD network, $k_{AD} > 0.8$; control network, $k_{CT} > 0.8$). To prioritize these hub genes, we introduced a CFG approach integrating data from GWAS, brain eQTL, PPI, and AD mouse models as supporting evidence [4,47–50]. Genes in DEG-enriched modules had an overall higher CFG score than genes in non–DEG-enriched modules (*t* test, $P = 5.53 \times 10^{-11}$; Fig. 3C), and the score showed significantly positive correlation with connectivity (Pearson's correlation $r = 0.766$, $P = 2.39 \times 10^{-12}$) in DEG-enriched modules, whereas no such tendency was observed in non–DEG-enriched modules (Pearson's correlation $r = -0.145$, $P = .348$) (Fig. 3C). Therefore, hub genes in DEG-enriched modules were more likely to be AD relevant compared with nonhubs or hubs of non–DEG-enriched modules and thus deserved further investigation. Indeed, the well-known AD genes *APP*, *MAPT*, *APOE*, and *BIN1* showed top CFG signals (CFG ≥ 4), indicating that the CFG system could assign proper weight to genes that were involved in AD. Of note, most of these well-recognized AD-risk genes presented low connectivity ($k < 0.8$) (Supplementary Table 6), whereas two GWAS reported top genes *FERMT2* and *MS4A6A* that were implicated in tau metabolism and inflammatory response, respectively [61], were also hub genes (Supplementary Table 6). For these 156 hub genes observed in the five DEG-enriched modules, 89 received at least two lines of AD-related evidence (CFG score > 1; Supplementary Table 10) and were prioritized as highly AD-relevant hub genes. The full genome-wide CFG prioritization results can be retrieved at the AlzData.org Web server.

3.5. Identification of 17 potential upstream regulators and highlighting the transcriptional coactivator YAP1

To investigate the potentially causal role of the hub genes, we examined the early expression patterns of the 89 hub genes by using AD mouse or cellular models (Aβ treatment). Fifty-eight genes showed early expression alterations in the hippocampus of 2-month-old AD mice (a model of preclinical stage, Mouseac [50]) (Supplementary Table 10). Among these genes, 17 genes had a consistent trend of differential expression in AD mouse or cellular model data sets (Mouseac [50], GSE29317 [62], and GSE31372 [63]). These genes had central positions (hub) in the network and were differentially expressed at the early stage and therefore might act as potential upstream regulators in AD. We then considered whether there were transcriptional factors, which might be more likely to be upstream driver regulators, in those early-altered hub genes. Among the upstream regulators, we found that *YAP1* (the main effector of the Hippo signaling pathway [64]) in module "red" (Fig. 3D; Supplementary Table 10) was the only gene with transcriptional activity and therefore was of particular interest.

3.6. Functional validation of regulatory and upstream roles of YAP1

The aforementioned prioritization had indicated that the transcriptional cofactor *YAP1* was a hub node ($k_{AD} = 0.84$, module "red") in the AD network and showed a high level of AD relevance (CFG score = 3). Notably, the connectivity of *YAP1* increased from 0.50 in the control network to 0.84 (in the corresponding module) in the AD network, supporting a gain-of-connectivity position of *YAP1* in AD pathology. Moreover, we observed that *YAP1* was downregulated in hippocampus of AD mice at the early stage (2 months old, before the onset of Aβ deposits or tau tangle, Fig. 4A) [50]. Importantly, the early declined



B

Module	Module size	DEGs		Module functions	Enriched cell type	AD core genes in this module
		Up	Down			
red	599	344	8	glial cell differentiation (Fold = 10.7, FDR = 3.48×10^{-104})	astrocyte (Fold = 13.6, FDR = 2.49×10^{-68})	*APOE*, *CLU*, *FERMT2*, *CD2AP*
turquoise	3626	56	1862	synaptic transmission (Fold = 4.6, FDR = 5.62×10^{-60})	neuron (Fold = 2.6, FDR = 1.75×10^{-18})	*APP*, *PSEN2*, *MAPT*, *PTK2B*, *SORL1*, *MEF2C*
green	671	261	11	myelination (Fold = 18.7, FDR = 1.07×10^{-58})	oligodendrocyte (Fold = 13.0, FDR = 1.32×10^{-13})	*PSEN1*, *BIN1*, *PICALM*, *ZCWPW1*
purple	273	147	0	regulation of inflammatory response (Fold = 27.6, FDR = 1.79×10^{-42})	endothelial cell (Fold = 20.8, FDR = 1.61×10^{-33})	*BACE2*
yellow	722	291	10	transcription regulation (Fold = 2.6, FDR = 5.02×10^{-3})	OPC (Fold = 4.8, FDR = 0.02)	-

D

Modules	Gene	Connectivity			AD-related evidence							Expression change in AD patients
					CFG	eQTL	GWAS	PPI	Pathology		Early change	
		k_{AD}	k_{CTL}	k_{AD}-k_{CTL}					Aβ	Tau		EC/HP/TC/FC
red	*AGT*	0.93	0.9	0.03	4	1	0	APP, PSEN1, APOE	-0.359,*	0.002,ns	↑↑--	--↑-
	GJA1	0.8	0.81	-0.02	4	2	0	PSEN1, APOE, MAPT	0.388,**	0.131,ns	--↓-	-↑↑↑
	HDAC1	0.56	0.98	-0.43	4	1	0	APP, PSEN2, MAPT	0.449,**	0.643,**	↓-↓-	--↑↑
	YAP1	0.84	0.5	0.34	3	5	0	-	0.207,ns	0.563,*	↓--↓	↑↑↑-
	ATP1A2	0.82	0.71	0.11	3	1	0	-	0.460,**	-0.110,ns	-↓--	--↑-
	PON2	0.74	0.85	-0.11	3	2	0	-	0.697,***	0.760,**	-↓↓-	↑---
	MSN	0.63	0.96	-0.33	3	0	NA	APOE	0.791,***	0.549,*	↓↓--	↑-↑↑
	HSPB1	0.18	0.82	-0.64	3	0	0	APP, PSEN1, MAPT	0.520,***	0.711,**	↓↓--	---↑
turquoise	*SH3GL2*	0.85	0.68	0.17	4	1	0	PSEN2	-0.365,*	-0.393,ns	↓--↓	↓↓↓-
	UCHL1	0.83	0.49	0.34	3	1	0	-	-0.338,*	-0.209,ns	↑↑--	↓↓↓↓
	STMN2	1	0.83	0.17	2	0	0	-	-0.662,***	-0.713,**	↑↑↑-	↓↓↓-
	REEP1	0.93	0.75	0.18	2	0	0	-	-0.621,***	-0.858,***	--↑↑	↓↓↓↓
green	*NDE1*	0.81	0.79	0.01	4	2	0	PSEN2, MAPT	0.626,***	0.297,ns	↓-↓-	---↑
	TJAP1	0.82	0.95	-0.13	3	1	0	-	0.481,***	0.008,ns	↓-↓↑	↑-↑↑
	KLK6	0.79	0.91	-0.12	3	2	0	-	0.795,***	0.458,ns	-↓--	---↑
	SLC44A1	0.87	0.86	0.01	2	0	0	-	0.666,***	0.246,ns	-↓↓↑	↑--↑
	HSPA2	0.74	0.89	-0.15	2	1	0	-	0.132,ns	0.472,ns	↓---	↑-↑↑

Fig. 3. Coexpression modules, DEG-enriched modules, hub genes, and upstream regulator genes. Hub genes in DEG-enriched modules were important candidate genes in AD. (A) DEG-enriched modules in AD coexpression network. Tree branches were colored by module membership (row 1) and expression changes in entorhinal cortex (EC, row 2), hippocampus (HP, row 3), temporal cortex (TC, row 4), and frontal cortex (FC, row 5). DEGs and non-DEGs were marked by red and light cyan, respectively. (B) Putative functions and related AD core genes in each DEG-enriched modules. (C) Hub genes in DEG-enriched modules had a higher level of AD-related evidence than those in non–DEG-enriched modules ($P = 5.53 \times 10^{-11}$). The connectivity was divided into 100 intervals, ranging from 0 to 1. Genes with a connectivity falling in the same interval were combined as one set (dot) and were featured by mean connectivity (x-axis) and mean CFG score (y-axis). Correlation between mean connectivity and mean CFG score was measured by Pearson's correlation. Difference of mean CFG score

expression of *YAP1* could be verified in incipient AD patients [14] and in Aβ-treated neonatal astrocyte (from APP/PSEN1 double mutant transgenic mice [62]) (Fig. 4B and 4C). The opposite expression tendency was observed in old AD mice [50] (Fig. 4A) and AD patients of severe stage [14] (Fig. 4B and 4D). Expression alterations of hub genes at the early stage might be a response to upstream trigger factors, for example, Aβ deposit or risk SNPs, and further influence downstream genes and disease development (Fig. 1A). Indeed, our analysis revealed that the mRNA expression level of *YAP1* can respond to several GWAS reported AD-risk SNPs (Supplementary Table 11) [4,47] and Aβ treatment (GSE29317 [62], Fig. 4C), thus supporting the upstream role of *YAP1* in AD.

Downregulation of *YAP1* at the early stage of AD development was hypothesized to be a damaging event, and therefore, any direct interference with *YAP1* could increase AD pathology, and vice versa. We detected the changes of key proteins involved in Aβ production and tau phosphorylation in response to *YAP1* deficiency and redundancy. As expected, knockdown of *YAP1* significantly increased the $A\beta_{1-42}$ level and tau phosphorylation (pTau181 and pTau396), and the opposite tendency was observed in cells overexpressing *YAP1* (Fig. 4E). Moreover, proteins responsible for Aβ generation (i.e., β-cleavage enzyme BACE1 and γ-cleavage enzyme components, including PSEN1, PSEN2, and Nicastrin), and proteins phosphorylate tau (i.e., CDK5 and GSK3α/β) all showed a consistent tendency to alter with changing levels of Aβ production and tau phosphorylation (Fig. 4E), supporting a damaging role of downregulated *YAP1*.

To test our hypothesis that expression perturbation of *YAP1* as an upstream hub regulator could promote AD pathogenesis by disrupting expression of downstream genes (Fig. 1A), we performed RNA sequencing of U251-APP cells with *YAP1* knockdown or overexpression (Supplementary Fig. 9). In total, 3154 and 1467 DEGs were found for *YAP1* overexpression and knockdown, respectively. Among these DEGs, 1008 genes were affected by both *YAP1* reduction and overexpression with opposite direction of differential expression and were considered to be real downstream genes regulated by *YAP1*. Notably, these *YAP1*-regulated genes ($N = 455$, Supplementary Table 12) were significantly enriched in DEG-enriched modules (Fisher's exact test, $P = 2.33 \times 10^{-9}$) but not in non–DEG-enriched modules (Fisher's exact test, $P = 1.00$), adding a robust support to the hub role and global impact of *YAP1* on the whole AD transcriptomic network (Fig. 5). Of note, the experimentally confirmed target genes (e.g., *YES1* and *TEAD1*) of *YAP1* were captured in both coexpression modules and our RNA-seq data, albeit we observed no enrichment of *YAP1*-regulated genes in the *YAP1*-centered module. All these results suggested that the reduction of *YAP1*, and the downstream effects consequent on this, might play a key role at the early stages of AD development.

According to the single-cell RNA-seq data of mouse (http://web.stanford.edu/group/barres_lab/brainseqMariko/brainseq2.html) and human brain (http://www.alzdata.org/single_RNAseq.php), *YAP1* was primarily expressed in astrocytes. Given that genes in the *YAP1*-centered module were also enriched in astrocytes (FDR $= 2.49 \times 10^{-68}$, Fig. 3B), dysregulation of astrocyte-expressed genes such as *YAP1* and *APOE* in AD subjects may be important factors in the onset of AD.

4. Discussion

AD is characterized by the presence in the brain of Aβ plaque, tau tangles, and neuron loss [1], but the molecular changes underpinning these pathological features have not been fully elucidated. The characterization of transcriptional alterations of the brain during disease development might offer some insights into the pathogenesis of AD. Dozens of studies looking at the expression profiling of brain tissues with and without AD have been reported [5–31], and a large number of AD-related DEGs were identified [5–31], albeit the conclusions were often affected by small sample size and poor levels of statistical robustness. A comprehensive list of DEGs and a systematic analysis of the priority of genes and regulatory networks, based on a considerable large sample size, are vital for understanding the pathophysiology of AD and for identifying potential targets for drug therapy.

Access to brain tissues has long been and will continue to be a challenge in Alzheimer's research. Previous studies

between DEG-enriched modules and non–DEG-enriched modules was measured by unpaired Student's *t* test. A strong positive correlation between mean connectivity and mean CFG score was observed for DEG-enriched modules ($r = 0.766$, $P = 2.39 \times 10^{-12}$), whereas no correlation was observed in non–DEG-enriched modules ($r = -0.145$, $P = .348$). (D) Candidate upstream regulator genes. Connectivity of a gene in AD network (k_{AD}) or in control network (k_{CTL}) was estimated for each module; $k_{AD} - k_{CTL}$, an index to show the gain-of-connectivity (positive value) or loss-of-connectivity (negative value) in the AD network compared to the control network. Expression correlation (r) of the target gene and AD pathology in AD mice (Pearson's correlation, $P < .05$) was performed for the Aβ line AD mice in Mouseac (marked as Aβ) [50] and the tau line AD mice in Mouseac (marked as tau) [50]; Early change: expression alterations in hippocampus of 2-month-old AD mice (in an order of HO-TASTPM: homozygous APP/PSEN1 double mutant mice, HET-TASTPM: heterozygous APP/PSEN1 double mutant mice, TAS10: human mutant APP mice, TAU: mutant human MAPT mice) [50]; Expression change in AD patients: expression changes of the target gene in AD patients in the merged data sets of entorhinal cortex (EC), hippocampus (HP), temporal cortex (TC), and frontal cortex (FC); "↑", upregulated; "↓", downregulated; "-", no reported AD core genes, or no significant cell-type enrichment, no significant PPI interaction, expression correlation or expression change; "NA", not applicable due to missing related data for the target gene. Abbreviations: AD, Alzheimer's disease; CFG, convergent functional genomics score based on the total number of lines of AD-related evidence; DEG, differentially expressed gene; eQTL, the total number of risk SNPs based on the IGAP data set [4] ($P < .001$) that were able to regulate expression of the target gene (Braincloud [47], $P < .001$); GWAS, the total number of risk SNPs within the target gene based on the IGAP data set [4] ($P < .001$); PPI, AD core genes (*APP*, *PSEN1*, *PSEN2*, *MAPT*, and *APOE*) that had a significant protein-protein interaction ($P < .05$) with the target genes [48,49]. ns, $P > .05$; *$P < .05$; **$P < .01$; ***$P < .001$.

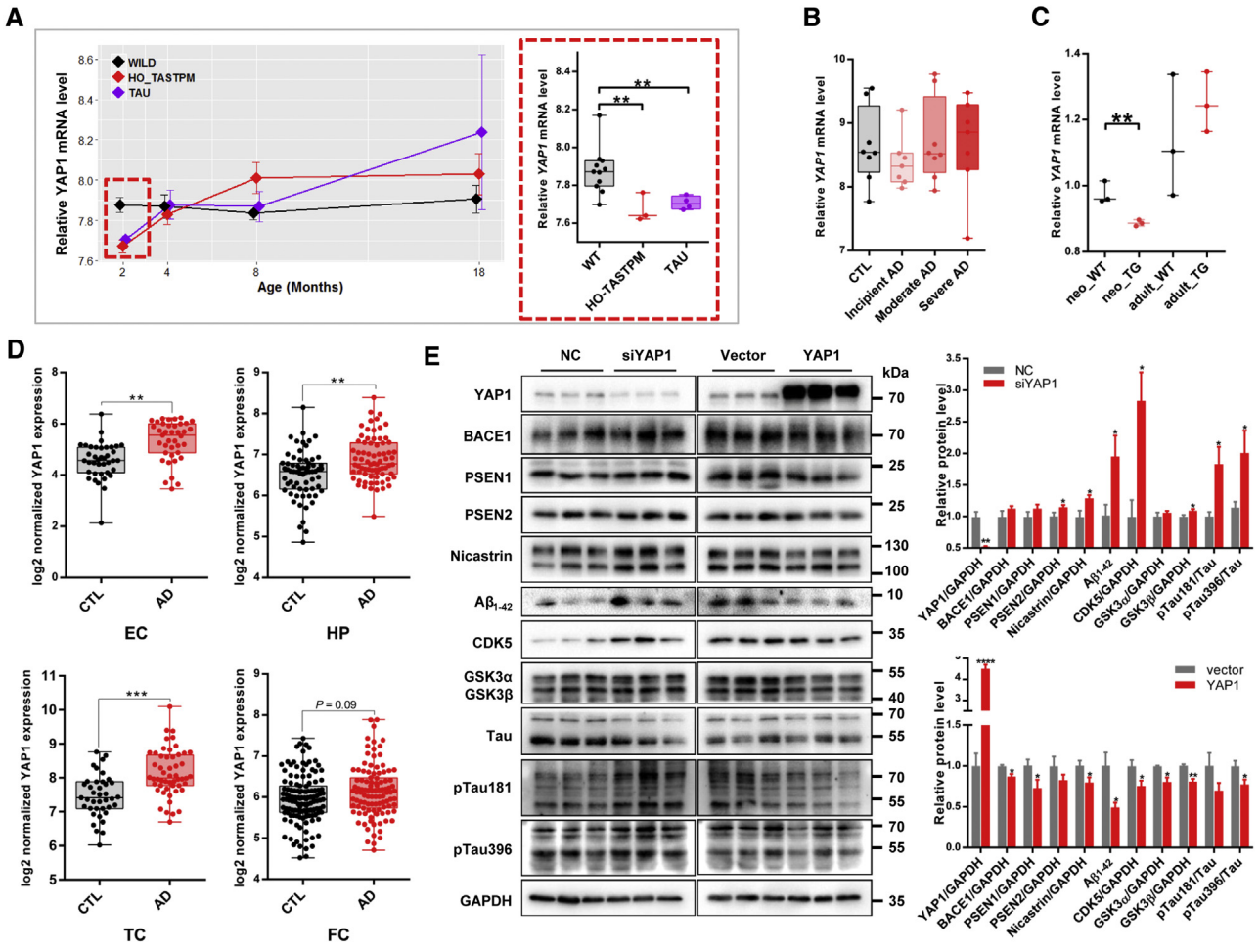


Fig. 4. *YAP1* was an upstream regulator in Alzheimer's disease (AD) network. (A) mRNA expression of *YAP1* was downregulated at the early stage and upregulated at the late stage in AD mice (Mouseac [50], WILD: wild-type mice; HO_TASTPM: homozygous APP/PSEN1 double mutant mice; TAU: mutant human MAPT mice). (B) The same tendency was observed in incipient to severe AD patients (GSE28146 [14], AD: AD patients, CTL: controls). (C) mRNA expression of *YAP1* was downregulated in Aβ-treated neonatal astrocyte (GSE29317 [62], neo_WT/adult_WT: astrocytes cultured on brain sections from neonatal/adult wild-type mice; neo_TG/adult_TG: cultured astrocytes from brain sections of neonatal/adult APdE9 mice [APP/PSEN1 double mutant]). (D) mRNA expression level of *YAP1* was upregulated in AD patients in merged data sets (EC: entorhinal cortex, HP: hippocampus; TC: temporal cortex, FC: frontal cortex). (E) Knockdown of *YAP1* expression led to increased Aβ production and tau phosphorylation in U251-APP cells, whereas overexpression of *YAP1* had opposite effects. Western blotting was performed for cell lysate (cytosolic proteins) and culture supernatant ($A\beta_{42}$). GAPDH was used as the loading control to measure the densitometries of YAP1, BACE1, PSEN1, PSEN2, nicastrin, amyloid β42 (Aβ1-42), CDK5, and GSK3α/β. The densitometric signal of phosphorylated tau at threonine 181 (pTau181) or serine 396 (pTau396) was determined by the ratio of the phosphorylated protein to total tau. Quantitative data were represented as mean ± SEM of three independent experiments. Statistical differences were calculated by the Student's *t* test. $^{*}P < .05$, $^{**}P < .01$, $^{***}P < .001$, $^{****}P < .0001$.

have taken the advantage of accumulated sample collection in reported microarray data sets through meta-analysis [7,9,65]. However, the meta-analysis had some limitations as it mixed samples at the statistical level rather than starting with the original expression data. In this study, we have used the cross-platform normalization method, which has better performance in the detection of robust DEGs than meta-analysis, and provides a complete expression profiling at the individual level for downstream analyses [33,34,66]. Based on the enlarged, highly curated sample collection (1246 brain tissues), we were able to get a complete list of DEGs in AD brain. The finalized DEG list covered not only a high proportion of the DEGs reported in individual studies but also DEGs that were missed in individual studies. To make our results a usable resource, we constructed a publicly available, user-friendly Web server (AlzData: www.alzdata.org). This list of DEGs and the open-access Web server might benefit future hypothesis-driven researches. It should be mentioned that the neuronal expressed genes might appear to be lower in AD samples, as AD-affected brain regions had fewer neurons and greater numbers of microglia and astrocyte among AD subjects compared with controls [67]. It is difficult to differentiate, therefore, the true expression changes at the cellular level from those that reflect cell population changes.

In addition to producing the DEG lists, we performed the gene coexpression network analysis, to identify functionally related gene modules and hub genes in the regulatory network. The coexpression module could serve as a

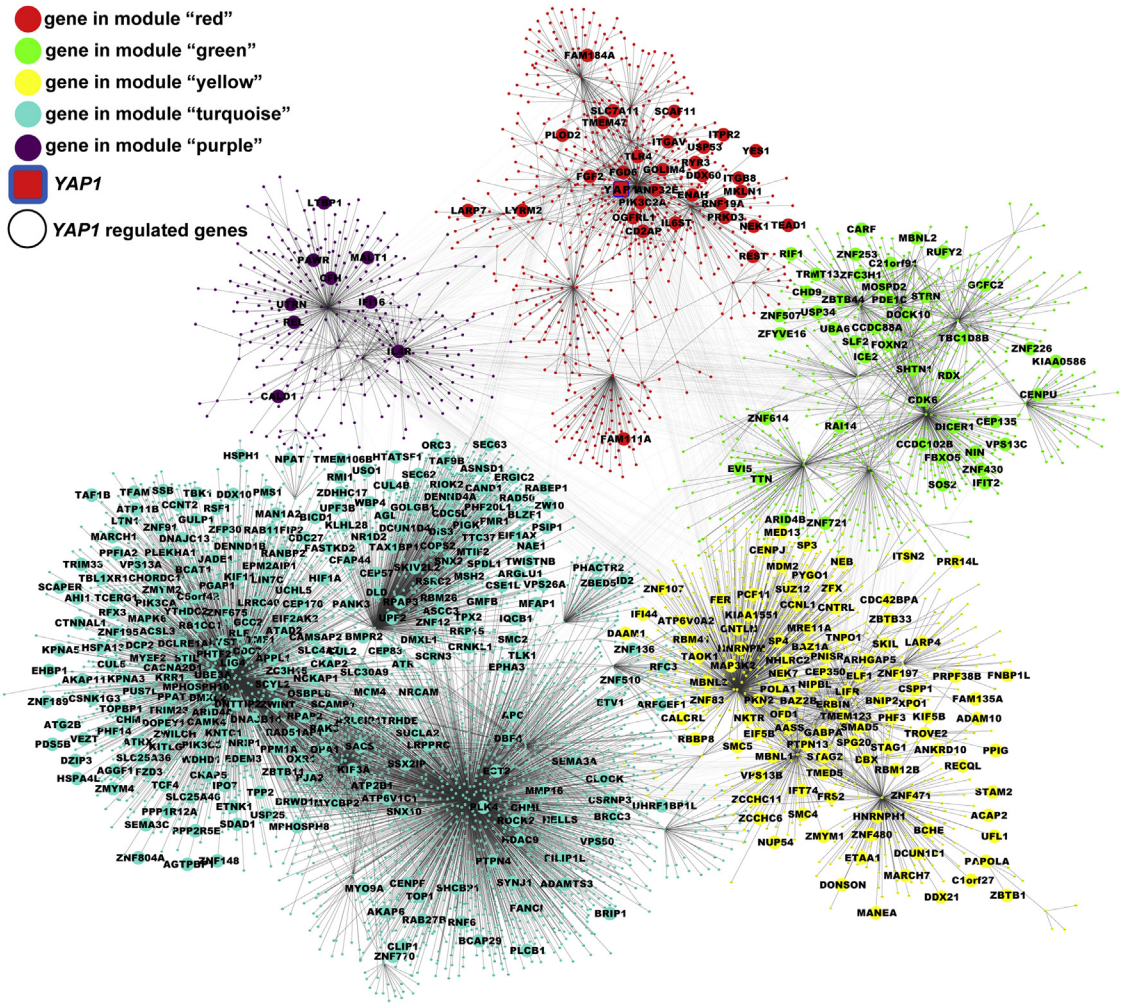


Fig. 5. Global impact of *YAP1* alteration in the DEG-enriched coexpression network. Coexpression network of the DEG-enriched modules (module "red," "turquoise," "green," "yellow," and "purple") was constructed for normalized expression data sets of all brain regions. *YAP1*-regulated genes (marked in circle with gene name) were identified by RNA-seq of the U251-APP cells with *YAP1* overexpression or knockdown. Abbreviation: DEG, differentially expressed gene.

robust, high-order, functional unit. Identification of the hub genes, which are essential for the integrity of the functional module [60], offers a possibility to explore the disrupted regulatory pattern in AD brain. The coexpression network method had been widely recognized and was used to identify central genes in gene expression networks [24,68]. For instance, Zhang et al. combined massive microarray (376 AD patients and 173 controls) and eQTL data and identified a hub gene *TYROBP* as the key causal regulator of the microglia module in AD brain [24]. The hub gene, *TYROBP*, was featured as a central node in our analysis for AD cases (n = 269, k_{AD} = 1 in module "greenyellow") and controls (n = 271, k_{CTL} = 0.83 in module "greenyellow"). Moreover, we found that *TYROBP* was an early-altered gene and was positively correlated with pathology of AD mouse, supporting its robust role in AD. *YWHAZ*, which was identified to be a hub gene in AD and aging [68,69], together with its family members *YWHAB/YWHAE/YWHAH*, were also featured in our results. The *SYT1* gene in the synaptic transmission module, reported as an important molecule for neurotransmitter release at the synapse [70], showed the highest connectivity (k_{AD} = 1 in module "turquoise") in our results. The consistency of these identified hub genes between our study and previous reports [24,68–70] indicates the repeatability and importance of the focus on hub genes. Nevertheless, there might be hundreds of hub genes and it is necessary to reduce this list of genes. In addition, a systems biology approach using other high-throughput data is warranted to fully characterize the genetic architecture of AD. Recently, Mukherjee et al. integrated massive PPI with genetic association data and indicated novel AD-risk genes [71], and this acts as a good example of this method.

Different from those previous network analyses that were mainly based on eQTL [24], or PPI data [71], in this study, we used a CFG approach integrating multiple lines of biological evidence (GWAS, eQTL, PPI, AD pathology correlation, and early alteration) to identify the hub genes which are highly relevant in AD. We found that the CFG score was positively correlated with connectivity (Fig. 3C) with

remarkable significance, and hub genes in the DEG-enriched modules showed a higher level of AD relevance. This observation suggested that it is desirable to focus on several prioritized genes selected from the numerous AD-related hub genes. The prioritized hub genes might act as key targets for the characterization of the regulatory pattern in AD brain.

As it is important to distinguish whether these hub genes are upstream drivers of AD or just downstream effectors, we analyzed the temporal expression of the hub genes with the help of AD mouse and cellular models, based on the hypothesis that gene expression changes before the emergence of AD pathology are more likely to be causal [50]. A dozen hub genes with early expression change were identified, such as *AGT*, *HDAC1*, *SH3GL2*, *STMN2*, *NDE1*, and *TJAP1* (Fig. 3D). *YAP1* is the only gene with transcription factor activity among these genes. The recognition of *YAP1* to be active in AD was not unexpected, as this gene had an active role in regulating neural precursor proliferation [72], neuronal specification [73], and neocortical astrocytic differentiation [74]. In addition, *YAP1* deletion could hyperactivate the inflammatory pathway and reactive astrogliosis [75]. The YAP/TAZ has also been proposed to be a downstream mediator of the APP signaling through a transcriptionally active protein complex containing APP and amyloid β precursor protein–binding family A member 3, also known as Mint3 and thus might mediate gene transcription induced by APP [76–78]. In our study, *YAP1* was downregulated in brain of preclinical AD mice and incipient patients and was decreased in response to Aβ deposit and genetic risk alleles (Fig. 4 and Supplementary Table 11), showing a promising potential as a link of downstream dysregulated events with upstream inducing factors. Indeed, our RNA-seq data of cellular assay showed that *YAP1* alterations affected the expression profiles of downstream genes involved in the DEG-enriched modules. *YAP1* depletion promoted the pathological change of AD, whereas overexpression of *YAP1* had opposite effects on AD-related process (Figs. 4 and 5). All these results suggested that the reduction of *YAP1*, and the downstream effects consequent on this, might play a key role in the early stage of AD development. Moreover, we observed a gain-of-connectivity of *YAP1* in the AD network compared with control network in this study. This implicates that the gain of hub role of *YAP1* might be disease specific. The differential correlations between *YAP1* and downstream genes might indicate AD-specific changes in functional interactions and coordinated activities under specific conditions or perturbations, rather than a normal physiology function.

Intriguingly, we observed a positive correlation of the gene expression levels between *YAP1* and *REST* (the only functionally recognized upstream regulator in AD and aging [27]) in our compiled data sets (Supplementary Fig. 10). The damaging role of *YAP1* depletion was consistent with the observation with reduced *REST* [27]. *REST* was in the *YAP1*-centered module (module "red"), with a moderate connectivity in the AD network ($k_{AD} = 0.29$) and control network ($k_{CT} = 0.42$). In contrast to the original report [27], we observed an increase of *REST* mRNA expression level in AD brains compared with controls. *YAP1* was also increased in AD brain compared with controls in all four brain regions in our combined data sets, which was different from its expression pattern at the early stage. We speculated that this was caused by a potentially compensatory effect. In our RNA-seq data, we observed an increase of *REST* expression in response to *YAP1* knockdown (logFC = 0.292, $P = .0039$) and a decrease of *REST* expression in response to *YAP1* overexpression (logFC = −0.305, $P = .0102$) (Supplementary Fig. 10). To determine if *YAP1* regulates *REST*, or if there is a compensatory feedback loop linking the expression of these two genes, we reanalyzed publically available expression data of neuronal cells with *REST* perturbation. We found no significant expression change of *YAP1* in response to knockdown (GSE28289 [79], $P = .59$, logFC = −0.09) or knockout (GSE27341 [80], $P = .7$, logFC = −0.05) of *REST*. Taken together, these results indicate that *YAP1* might be an upstream regulator of *REST*. These observations further supported the reliability of the upstream role of *YAP1* in AD development.

The identification of *YAP1*, as an upstream regulator, raised an important question whether modulation of the *YAP1* pathway, for example, use of *YAP1* activator at the early stage of cognitive impairment, would be a good approach for AD intervention. Note that *YAP1* and *TAZ* are major downstream effectors of the Hippo pathway that regulates key processes related to cell growth, tissue homeostasis, organ size, and regeneration [81–83], especially in ocular development [84,85] and liver regeneration [86]. *YAP1* acts as a central signal–responsive regulator of multipotent pancreatic progenitor in the embryonic development of pancreas [87]. Systemic effects should be taken into account when modulating the *YAP1* pathway. Although our finding indicates a protective role of increased *YAP1* level in AD, elevated YAP1 activity has been found in several types of human cancers [81,83,88]. Therefore, overactivation of *YAP1* in the elderly with cognitive impairment might increase the risk of cancer and this would hinder future efforts using YAP1 as a valid therapeutic target of AD.

Although we have provided a comprehesive and highly curated analysis of AD expression data sets, this study had several limitations. First, the cross-platform normalization only retained the common genes from different studies, thereby decreasing the total number of genes under consideration. Second, we did not include the most recently released expression data sets during our analysis, and we will include these data sets later for the AlzData.org Web server. Third, the sample size in respect of hippocampus and entorhinal cortex available for study was still limited, and the larger sample size of the frontal cortex might contribute more weight to

the results. In addition, because of the lack of detailed information regarding RNA quality and potentially inconsistent diagnosis criteria for AD in some data sets, the results should be interpreted with caution.

In summary, we obtained a robust DEG list for four AD brain regions based on a large sample size. We identified several hub genes with multiple lines of AD-related supporting evidence and established a publicly accessible Web server for these results.

In particular, we found a previous unknown *YAP1*-initiated regulatory network active during AD development. Our results offer a place of reference for gene expression alterations in AD brain available to future hypothesis-driven, gene-central studies, as well as providing information about potential therapeutic targets. Further functional experiments, such as Chip-seq and animal experiments, are required to validate our results and to uncover the functions of the genes identified in this study.

Acknowledgments

The authors thank Ian Logan for language editing and helpful comments and the three anonymous reviewers for their constructive comments and suggestions. This study was supported by the Key Program of National Natural Science Foundation of China (NSFC, 31730037), the Bureau of Frontier Sciences and Education (QYZDJ-SSW-SMC005), the Strategic Priority Research Program (B) (XDB02020003), and the West Light Foundation of the Chinese Academy of Sciences. The sponsor had no role in the study design, in the collection, analysis, and interpretation of data; in the writing of the article; or in the decision to submit the article for publication. The authors appreciate the colleagues for their sharing of expression data: John Hardy and colleagues for AD mouse models; Webster J., Gibbs R., Myers A., and colleagues for GSE15222; Nakabeppu Y., Hokama M., and colleagues for GSE36980; Liang and colleagues for GSE5281. The authors thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i–Select chips were funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2, and the Lille University Hospital. GERAD was supported by the Medical Research Council (grant no. 503480), Alzheimer's Research UK (grant no. 503176), the Wellcome Trust (grant no. 082604/2/07/Z), and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant no. 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01–AG–12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, the Erasmus Medical Center, and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the grant ADGC–10–196728.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2017.08.012.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed the related literature studies using PubMed and retrieved relevant data from Gene Expression Omnibus. There were dozens of studies focusing on expression alterations of Alzheimer's disease (AD) brain, and a few of them investigated the transcriptomic regulatory network. These relevant citations were appropriately cited.
2. Interpretation: Our data indicated potentially causal upstream regulators in the transcriptional network of AD brain. The findings were validated by functional assay and were consistent with previous observations.
3. Future directions: The article generated a robust differentially expressed gene list and a comprehensive Web server based on a large number of AD brain tissues, leading to a framework for continued studies of AD brain transcriptomics. The prioritized hub genes and upstream regulators might benefit further gene-centered, hypothesis-driven research.

References

[1] Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. Cold Spring Harb Perspect Med 2011;1:a006189.
[2] Alzheimer's association. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 2016;12:459–509.
[3] Tanzi RE. The genetics of Alzheimer disease. Cold Spring Harb Perspect Med 2012;2:a006296.
[4] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45:1452–8.
[5] Blalock EM, Geddes JW, Chen KC, Porter NM, Markesbery WR, Landfield PW. Incipient Alzheimer's disease: microarray correlation

analyses reveal major transcriptional and tumor suppressor responses. Proc Natl Acad Sci U S A 2004;101:2173–8.
[6] Dunckley T, Beach TG, Ramsey KE, Grover A, Mastroeni D, Walker DG, et al. Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. Neurobiol Aging 2006;27:1359–71.
[7] Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Walker DG, et al. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol Genomics 2007;28:311–22.
[8] Berchtold NC, Cribbs DH, Coleman PD, Rogers J, Head E, Kim R, et al. Gene expression changes in the course of normal brain aging are sexually dimorphic. Proc Natl Acad Sci U S A 2008; 105:15605–10.
[9] Liang WS, Reiman EM, Valla J, Dunckley T, Beach TG, Grover A, et al. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci U S A 2008;105:4441–6.
[10] Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P, et al. Genetic control of human brain transcript expression in Alzheimer disease. Am J Hum Genet 2009;84:445–58.
[11] Williams C, Mehrian Shai R, Wu Y, Hsu YH, Sitzer T, Spann B, et al. Transcriptome analysis of synaptoneurosomes identifies neuroplasticity genes overexpressed in incipient Alzheimer's disease. PLoS One 2009;4:e4936.
[12] Astarita G, Jung KM, Berchtold NC, Nguyen VQ, Gillen DL, Head E, et al. Deficient liver biosynthesis of docosahexaenoic acid correlates with cognitive impairment in Alzheimer's disease. PLoS One 2010; 5:e12538.
[13] Tan MG, Chua WT, Esiri MM, Smith AD, Vinters HV, Lai MK. Genome wide profiling of altered gene expression in the neocortex of Alzheimer's disease. J Neurosci Res 2010;88:1157–69.
[14] Blalock EM, Buechel HM, Popovic J, Geddes JW, Landfield PW. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. J Chem Neuroanat 2011;42:118–26.
[15] Simpson JE, Ince PG, Shaw PJ, Heath PR, Raman R, Garwood CJ, et al. Microarray analysis of the astrocyte transcriptome in the aging brain: relationship to Alzheimer's pathology and APOE genotype. Neurobiol Aging 2011;32:1795–807.
[16] Berson A, Barbash S, Shaltiel G, Goll Y, Hanin G, Greenberg DS, et al. Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. EMBO Mol Med 2012;4:730–42.
[17] Cribbs DH, Berchtold NC, Perreau V, Coleman PD, Rogers J, Tenner AJ, et al. Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study. J Neuroinflammation 2012; 9:179.
[18] Durrenberger PF, Fernando FS, Magliozzi R, Kashefi SN, Bonnert TP, Ferrer I, et al. Selection of novel reference genes for use in the human central nervous system: a BrainNet Europe Study. Acta Neuropathol 2012;124:893–903.
[19] Sarvari M, Hrabovszky E, Kallo I, Solymosi N, Liko I, Berchtold N, et al. Menopause leads to elevated expression of macrophage-associated genes in the aging frontal cortex: rat and human studies identify strikingly similar changes. J Neuroinflammation 2012; 9:264.
[20] Silva AR, Grinberg LT, Farfel JM, Diniz BS, Lima LA, Silva PJ, et al. Transcriptional alterations related to neuropathology and clinical manifestation of Alzheimer's disease. PLoS One 2012; 7:e48751.
[21] Berchtold NC, Coleman PD, Cribbs DH, Rogers J, Gillen DL, Cotman CW. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. Neurobiol Aging 2013;34:1653–61.
[22] Blair LJ, Nordhues BA, Hill SE, Scaglione KM, O'Leary JC 3rd, Fontaine SN, et al. Accelerated neurodegeneration through chaperone-mediated oligomerization of tau. J Clin Invest 2013; 123:4158–69.
[23] Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. Genome Med 2013;5:48.
[24] Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 2013;153:707–20.
[25] Hokama M, Oka S, Leon J, Ninomiya T, Honda H, Sasaki K, et al. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. Cereb Cortex 2014;24:2476–88.
[26] Lai MK, Esiri MM, Tan MG. Genome-wide profiling of alternative splicing in Alzheimer's disease. Genom Data 2014;2:290–2.
[27] Lu T, Aron L, Zullo J, Pan Y, Kim H, Chen Y, et al. REST and stress resistance in ageing and Alzheimer's disease. Nature 2014; 507:448–54.
[28] Narayanan M, Huynh JL, Wang K, Yang X, Yoo S, McElwee J, et al. Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases. Mol Syst Biol 2014;10:743.
[29] Durrenberger PF, Fernando FS, Kashefi SN, Bonnert TP, Seilhean D, Nait-Oumesmar B, et al. Common mechanisms in neurodegeneration and neuroinflammation: a BrainNet Europe gene expression microarray study. J Neural Transm (Vienna) 2015;122:1055–68.
[30] Simpson JE, Ince PG, Minett T, Matthews FE, Heath PR, Shaw PJ, et al. Neuronal DNA damage response-associated dysregulation of signalling pathways and cholesterol metabolism at the earliest stages of Alzheimer-type pathology. Neuropathol Appl Neurobiol 2016; 42:167–79.
[31] Bossers K, Wirz KT, Meerhoff GF, Essing AH, van Dongen JW, Houba P, et al. Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. Brain 2010; 133:3699–723.
[32] Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ. Gene expression profiling in human neurodegenerative disease. Nat Rev Neurol 2012;8:518–30.
[33] Niculescu AB, Le-Niculescu H. Convergent functional genomics: what we have learned and can learn about genes, pathways, and mechanisms. Neuropsychopharmacology 2010;35:355–6.
[34] Niculescu AB 3rd, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR. Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. Physiol Genomics 2000;4:83–91.
[35] Taminau J, Meganck S, Lazar C, Steenhoff D, Coletta A, Molter C, et al. Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. BMC Bioinformatics 2012;13:335.
[36] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007; 8:118–27.
[37] Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, et al. Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PLoS One 2011; 6:e17238.
[38] Luo J, Schumacher M, Scherer A, Sanoudou D, Megherbi D, Davison T, et al. A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. Pharmacogenomics J 2010;10:278–91.
[39] Bushel P. pvca: Principal Variance Component Analysis (PVCA). R package version 1.10.0 2013. Available at: https://bioconductor.org/packages/release/bioc/html/pvca.html. Accessed September 27, 2017.
[40] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[41] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.

[42] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.
[43] Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics 2008;24:719–20.
[44] Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. J Stat Softw 2012;46:i11.
[45] Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? PLoS Comput Biol 2011;7:e1001057.
[46] Horvath S, Dong J. Geometric interpretation of gene co-expression network analysis. PLoS Comput Biol 2008;4:e1000117.
[47] Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011;478:519–23.
[48] Itan Y, Mazel M, Mazel B, Abhyankar A, Nitschke P, Quintana-Murci L, et al. HGCS: an online tool for prioritizing disease-causing gene variants by biological distance. BMC Genomics 2014;15:256.
[49] Itan Y, Zhang SY, Vogt G, Abhyankar A, Herman M, Nitschke P, et al. The human gene connectome as a map of short cuts for morbid allele discovery. Proc Natl Acad Sci U S A 2013;110:5558–63.
[50] Matarin M, Salih DA, Yasvoina M, Cummings DM, Guelfi S, Liu W, et al. A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. Cell Rep 2015;10:633–44.
[51] Zhang DF, Li J, Wu H, Cui Y, Bi R, Zhou HJ, et al. CFH variants affect structural and functional brain changes and genetic risk of Alzheimer's disease. Neuropsychopharmacology 2016;41:1034–45.
[52] Xiang Q, Bi R, Xu M, Zhang DF, Tan L, Zhang C, et al. Rare genetic variants of the transthyretin gene are associated with Alzheimer's disease in Han Chinese. Mol Neurobiol 2017;54:5192–200.
[53] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[54] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
[55] Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics 2010;26:2927–8.
[56] Landreth GE, Reed-Geaghan EG. Toll-like receptors in Alzheimer's disease. Curr Top Microbiol Immunol 2009;336:137–53.
[57] Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Ramsey K, et al. Neuronal gene expression in non-demented individuals with intermediate Alzheimer's Disease neuropathology. Neurobiol Aging 2010;31:549–66.
[58] Antonell A, Llado A, Altirriba J, Botta-Orfila T, Balasa M, Fernandez M, et al. A preliminary study of the whole-genome expression profile of sporadic and monogenic early-onset Alzheimer's disease. Neurobiol Aging 2013;34:1772–8.
[59] Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A 2015;112:7285–90.
[60] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13.
[61] Karch CM, Goate AM. Alzheimer's disease risk genes and mechanisms of disease pathogenesis. Biol Psychiatry 2015;77:43–51.
[62] Kurronen A, Pihlaja R, Pollari E, Kanninen K, Storvik M, Wong G, et al. Adult and neonatal astrocytes exhibit diverse gene expression profiles in response to beta amyloid *ex vivo*. World J Neurosci 2012;2:57–67.
[63] Wakutani Y, Ghani M, Tokuhiro S, Bohm C, Chen F, Sato C, et al. Misprocessing of APP and accumulation of β-Amyloid causes early alteration of pathways implicated in late-onset Alzheimer disease. Available at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31372. Accessed September 27, 2017.
[64] Piccolo S, Dupont S, Cordenonsi M. The biology of YAP/TAZ: hippo signaling and beyond. Physiol Rev 2014;94:1287–312.
[65] Puthiyedth N, Riveros C, Berretta R, Moscato P. Identification of differentially expressed genes through integrated study of Alzheimer's disease affected brain regions. PLoS One 2016;11:e0152342.
[66] Taminau J, Lazar C, Meganck S, Nowe A. Comparison of merging and meta-analysis as alternative approaches for integrative gene expression analysis. ISRN Bioinform 2014;2014:345106.
[67] De Strooper B, Karran E. The cellular phase of Alzheimer's disease. Cell 2016;164:603–15.
[68] Miller JA, Oldham MC, Geschwind DH. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J Neurosci 2008;28:1410–20.
[69] Liang D, Han G, Feng X, Sun J, Duan Y, Lei H. Concerted perturbation observed in a hub network in Alzheimer's disease. PLoS One 2012;7:e40498.
[70] Kochubey O, Lou X, Schneggenburger R. Regulation of transmitter release by Ca(2+) and synaptotagmin: insights from a large CNS synapse. Trends Neurosci 2011;34:237–46.
[71] Mukherjee S, Russell JC, Carr DT, Burgess JD, Allen M, Serie DJ, et al. 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. Alzheimers Dement 2017. http://dx.doi.org/10.1016/j.jalz.2017.01.016.
[72] Fernandez-L A, Northcott PA, Dalton J, Fraga C, Ellison D, Angers S, et al. YAP1 is amplified and up-regulated in hedgehog-associated medulloblastomas and mediates Sonic hedgehog-driven neural precursor proliferation. Genes Dev 2009;23:2729–41.
[73] Musah S, Wrighton PJ, Zaltsman Y, Zhong X, Zorn S, Parlato MB, et al. Substratum-induced differentiation of human pluripotent stem cells reveals the coactivator YAP is a potent regulator of neuronal specification. Proc Natl Acad Sci U S A 2014;111:13805–10.
[74] Huang Z, Hu J, Pan J, Wang Y, Hu G, Zhou J, et al. YAP stabilizes SMAD1 and promotes BMP2-induced neocortical astrocytic differentiation. Development 2016;143:2398–409.
[75] Huang Z, Wang Y, Hu G, Zhou J, Mei L, Xiong WC. YAP is a critical inducer of SOCS3, preventing reactive astrogliosis. Cereb Cortex 2016;26:2299–310.
[76] Swistowski A, Zhang Q, Orcholski ME, Crippen D, Vitelli C, Kurakin A, et al. Novel mediators of amyloid precursor protein signaling. J Neurosci 2009;29:15703–12.
[77] Orcholski ME, Zhang Q, Bredesen DE. Signaling via amyloid precursor-like proteins APLP1 and APLP2. J Alzheimers Dis 2011;23:689–99.
[78] Plouffe SW, Hong AW, Guan KL. Disease implications of the Hippo/YAP pathway. Trends Mol Med 2015;21:212–22.
[79] Yu HB, Johnson R, Kunarso G, Stanton LW. Coassembly of REST and its cofactors at sites of gene repression in embryonic stem cells. Genome Res 2011;21:1284–93.
[80] Arnold P, Scholer A, Pachkov M, Balwierz PJ, Jorgensen H, Stadler MB, et al. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. Genome Res 2013;23:60–73.
[81] Moroishi T, Hansen CG, Guan KL. The emerging roles of YAP and TAZ in cancer. Nat Rev Cancer 2015;15:73–9.
[82] Yu C, Ji SY, Dang YJ, Sha QQ, Yuan YF, Zhou JJ, et al. Oocyte-expressed yes-associated protein is a key activator of the early zygotic genome in mouse. Cell Res 2016;26:275–87.
[83] Mo JS, Meng Z, Kim YC, Park HW, Hansen CG, Kim S, et al. Cellular energy stress induces AMPK-mediated regulation of YAP and the Hippo pathway. Nat Cell Biol 2015;17:500–10.
[84] Miesfeld JB, Gestri G, Clark BS, Flinn MA, Poole RJ, Bader JR, et al. Yap and Taz regulate retinal pigment epithelial cell fate. Development 2015;142:3021–32.

[85] Kim JY, Park R, Lee JH, Shin J, Nickas J, Kim S, et al. Yap is essential for retinal progenitor cell cycle progression and RPE cell fate acquisition in the developing mouse eye. Dev Biol 2016;419:336–47.
[86] Yimlamai D, Fowl BH, Camargo FD. Emerging evidence on the role of the Hippo/YAP pathway in liver physiology and cancer. J Hepatol 2015;63:1491–501.
[87] Cebola I, Rodriguez-Segui SA, Cho CH, Bessa J, Rovira M, Luengo M, et al. TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. Nat Cell Biol 2015;17:615–26.
[88] Stein C, Bardet AF, Roma G, Bergling S, Clay I, Ruchti A, et al. YAP1 exerts its transcriptional control via TEAD-mediated activation of enhancers. PLoS Genet 2015;11:e1005465.