Identification of key genes and pathways for Alzheimer's disease via combined analysis of genome-wide expression profiling in the hippocampus

Mengsi Wu^{1,2}, Kechi Fang¹, Weixiao Wang^{1,2}, Wei Lin^{1,2}, Liyuan Guo^{1,2}✉, Jing Wang^{1,2}✉

¹ CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

² Department of Psychology, University of Chinese Academy of Sciences, Beijing 10049, China

Received: 8 August 2018 / Accepted: 17 January 2019 / Published online: 20 April 2019

Abstract In this study, combined analysis of expression profiling in the hippocampus of 76 patients with Alzheimer's disease (AD) and 40 healthy controls was performed. The effects of covariates (including age, gender, postmortem interval, and batch effect) were controlled, and differentially expressed genes (DEGs) were identified using a linear mixed-effects model. To explore the biological processes, functional pathway enrichment and protein–protein interaction (PPI) network analyses were performed on the DEGs. The extended genes with PPI to the DEGs were obtained. Finally, the DEGs and the extended genes were ranked using the convergent functional genomics method. Eighty DEGs with $q < 0.1$, including 67 downregulated and 13 upregulated genes, were identified. In the pathway enrichment analysis, the 80 DEGs were significantly enriched in one Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, GABAergic synapses, and 22 Gene Ontology terms. These genes were mainly involved in neuron, synaptic signaling and transmission, and vesicle metabolism. These processes are all linked to the pathological features of AD, demonstrating that the GABAergic system, neurons, and synaptic function might be affected in AD. In the PPI network, 180 extended genes were obtained, and the hub gene occupied in the most central position was *CDC42*. After prioritizing the candidate genes, 12 genes, including five DEGs (*ITGB5*, *RPH3A*, *GNAS*, *THY1*, and *SEPT6*) and seven extended genes (*JUN*, *GDI1*, *GNAI2*, *NEK6*, *UBE2D3*, *CDC42EP4*, and *ERCC3*), were found highly relevant to the progression of AD and recognized as promising biomarkers for its early diagnosis.

Keywords Alzheimer's disease, Combined analysis, Hippocampus, Gene expression, Differentially expressed genes, Microarray

Mengsi Wu and Kechi Fang have contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41048-019-0086-2>) contains supplementary material, which is available to authorized users.

✉ Correspondence: guoly@psych.ac.cn (L. Guo), wangjing@psych.ac.cn (J. Wang)

INTRODUCTION

Alzheimer's disease (AD) is an age-related neurodegenerative disease caused by central nervous system disorders. It accounts for 50%–75% of dementia patients. The common symptoms of AD are progressive deterioration of memory and cognitive decline, including degenerated learning, recall accuracy, and problem solving and changes in personality and behavior (Rosenberg *et al.* 2015). Many studies show that AD is a polygenic disease influenced by several susceptibility

genes with a small effect (van Cauwenberghe *et al.* 2016). However, the specific pathogenesis of AD remains unclear, and no effective treatment and prevention measures are still available.

To explore the molecular changes underlying AD, a number of genome-wide expression profiling experiments were performed on the postmortem brain tissues of AD patients (Blair *et al.* 2013; Blalock *et al.* 2004; Cooper-Knock *et al.* 2012; Liang *et al.* 2008a, b; Wang *et al.* 2016b). The hippocampus plays a critical role in memory and learning and is one of the earliest regions to be affected in AD patients (Mak *et al.* 2017; Weiner *et al.* 2017). Dysregulated genes and molecular pathways have been identified in a series of gene expression studies in the hippocampus of AD patients (Berchtold *et al.* 2013; Wang *et al.* 2016b). However, the findings in different studies have heterogeneity and low reproducibility, which are partly attributable to the different array types; small sample size; diverse analysis procedures in different cohorts; and other confounding factors, such as postmortem interval (PMI), age, and gender. To solve these issues, several studies sought to consolidate the knowledge of transcriptomic abnormalities via a combined analysis (Hu *et al.* 2015; Li *et al.* 2015). However, these studies (Hu *et al.* 2015; Li *et al.* 2015) had several limitations, including the following: (1) covariates, such as age, gender, PMI, and batch effect, were not considered when modeling; (2) compared with the combined-sample reanalysis of the individual-level data, the combined reanalysis of the summary statistics from multiple studies was relatively underpowered (Hess *et al.* 2016); and (3) new microarray-based gene expression studies of AD were conducted in the past two years.

Therefore, in this study, microarray-based transcriptomic studies in the hippocampus of AD patients were strictly screened, and only the datasets with detailed sample information and raw probe-level data generated from similar Affymetrix platforms were retained. A combined analysis of individual-level biological data from selected microarray studies was conducted for statistical modeling with proper correction for covariates and variances among studies. The differentially expressed genes (DEGs) in the hippocampus of AD patients and the age-matched healthy controls were best identified, thereby providing biological clues for the interpretation of the pathogenic mechanism of AD. Further validations of the DEGs were performed to test the robustness of these findings. Next, pathway enrichment and protein-protein interaction (PPI) network analyses for the DEGs were performed to explore the biological processes and interactions of the dysregulated genes, helping to elucidate the biological

underpinnings of AD. Furthermore, gene prioritization was conducted to discover more promising genes for subsequent experimental replication and identification of biomarkers from the large amount of candidate genes. The findings of the present study may contribute to characterizing intrinsic molecular processes underlying AD and implicating promising biomarkers for AD.

RESULTS

DEGs identified in the hippocampus of AD patients and age-matched controls

For our combined analysis, data from 116 samples, composed of 40 healthy controls and 76 AD cases, were obtained after quality control. Eight sample data were removed. After normalization, the expression matrices for each dataset were merged, and the combined gene expression matrix consisted of 116 samples and 22,277 probe sets. Detailed information of each dataset is shown in Table 1.

For the variables we considered, a significant difference was observed in gender between the AD cases and controls (p -value = 0.01647). Although age and PMI did not show statistical significance, these factors were still taken into account (Supplementary Table S1). After mixed-effect linear modeling, we identified 82 dysregulated probe sets with q -value < 0.1, in which 69 probe sets were downregulated, and 13 probe sets were upregulated. These probe sets mapped to 80 DEGs (67 downregulated genes and 13 upregulated genes) in the hippocampus of AD patients and healthy controls (Table 2). Two downregulated genes mapped by more than one probe set (*CDC42* and *IGF1*) implied higher confidence in the results of their expression changes.

Robustness and sensitivity of the DEGs

Jackknife cross-validation was used to validate the robustness of the findings. Each leave-out iteration resulted in a new list of DEGs (q -value < 0.10), which was subsequently compared with DEGs obtained from the combined analysis (Supplementary Table S2). Thirty-two DEGs (30 downregulated genes and two upregulated genes) were cross-validated by the jackknife method (Table 2).

Furthermore, the DEGs were compared with the results of the AlzBase database. Seventy-five DEGs (63 downregulated genes and 12 upregulated genes) were in accordance with the finding of the AlzBase database. The detailed information of the total DEGs can be found in Table 2. Among these genes, 29 downregulated genes

Table 1 Combined gene expression datasets of AD in this study

Source	Series	References	Controls	AD	Array
GEO	GSE1297	Blalock <i>et al.</i> (2004)	9	20	Affymetrix Human Genome U133A
GEO	GSE48350	Berchtold <i>et al.</i> (2008)	21	16	Affymetrix Human Genome U133 Plus 2.0
GEO	GSE84422	Wang <i>et al.</i> (2016a, b, c)	10	40	Affymetrix Human Genome U133A
	Total		40	76	

(*GAD2*, *RPH3A*, *SST*, *GAD1*, *GABBR2*, *NUDT11*, *DLGAP2*, *PCLO*, *KALRN*, *WFDC1*, *AAK1*, *CDC42*, *PCSK1*, *RGS4*, *SYNGR3*, *IGF1*, *INA*, *GLS2*, *NCALD*, *CD200*, *F12*, *PRC1*, *LRRTM2*, *GAP43*, *TSPAN13*, *CKMT1A*, *CKMT1B*, *ADD2*, and *THY1*) and two upregulated genes (*TNFRSF11B* and *ITGB5*) were validated by the two methods (Table 2). Four other DEGs (*PTPN20*, *ADGB*, *KLHL18*, and *PHF24*) were first observed in our study (Table 2).

Biological classification and pathway enrichment analysis of the DEGs

By using DAVID, the 80 DEGs were significantly enriched in one KEGG pathway, GABAergic synapse, and 22 Gene Ontology (GO) terms had Benjamin-corrected p -value < 0.05 (Table 3). For the 80 DEGs identified between the AD patients and healthy controls, the significantly enriched KEGG pathway was “GABAergic synapses” pathway (Benjamin-corrected p -value = 0.039882). The DEGs involved in “GABAergic synapses” pathway were all downregulated, suggesting that the function of GABAergic synapses may be impaired in the pathogenesis of AD.

In the biological classification, the significant GO categories included those primarily involved in multiple aspects of synaptic function, notably synaptic signaling, transmission and processing, and synaptic vesicle metabolism (Table 3). For cellular component, the DEGs were significantly enriched in “synapse part”, “neuron part”, “synapse”, “presynapse”, “exocytic vesicle”, “transport vesicle”, “cytoplasmic, membrane-bounded vesicle”, “synaptic vesicle”, “secretory vesicle”, “neuron projection”, “exocytic vesicle membrane”, “synaptic vesicle membrane”, “postsynapse”, “transport vesicle membrane”, “cell junction”, “cytoplasmic vesicle part”, and “excitatory synapse.” For biological process, the DEGs were significantly enriched in “synaptic signaling”, “anterograde trans-synaptic signaling”, “trans-synaptic signaling”, “chemical synaptic transmission”, and “cell-cell signaling”. However, no molecular function was significantly enriched by DEGs. The results demonstrated that GABAergic system, neurons, and synaptic function might be involved in the occurrence and development of AD.

PPI network of the DEGs

As shown in Fig. 1, a PPI network composed of 250 nodes and 497 edges was obtained. Among the 250 nodes, 70 DEGs (ten upregulated genes and 60 down-regulated genes) and 180 extended genes interacting with the DEGs were observed. Notably, genes greater than ten degrees were 14 DEGs (*CDC42*, *RBL1*, *GNAS*, *CKMT1B*, *CKMT1A*, *AMPH*, *ACVR1B*, *CNR1*, *SEPT6*, *GAD1*, *PDIA2*, *MOB4*, *PRC1*, and *ACVR2A*) and one extended gene (*UBC*). The DEG *CDC42* occupies the most central position in the network because it has the highest degree.

Ranking of the DEGs and the extended genes

To identify more reliable candidate genes from a large number of AD-related genes for subsequent experimental validation and identification of biomarkers, a prioritized list of the DEGs and the extended genes was generated using a convergent functional genomics (CFG) method (Supplementary Table S3). For the 260 candidate genes (80 DEGs and 180 extended genes), 156 genes were listed as highly AD-related candidate genes when received at least two lines of AD-related evidence (CFG score > 1 , Fig. 2). Among the 156 highly AD-relevant genes, 82 genes (25 DEGs and 57 extended genes) showed early expression alteration in the hippocampus of 2-month-old AD mice compared with age-matched wild-type mice, implicating them as potential upstream regulators in AD, and 25 genes (eight DEGs and 17 extended genes) were supported by blood gene expression evidence, implicating them as potential blood biomarkers (Supplementary Table S3). In addition, 12 AD-relevant candidate genes, including five DEGs (*ITGB5*, *RPH3A*, *GNAS*, *THY1*, and *SEPT6*) and seven extended genes (*JUN*, *GDI1*, *GNAI2*, *NEK6*, *UBE2D3*, *CDC42EP4*, and *ERCC3*), exhibited early expression alteration in the hippocampus of 2-month-old AD mice and had blood gene expression evidence (Supplementary Table S3). The expression of these genes changed before the emergence of AD pathology and might be detected in the blood, suggesting that they

Table 2 Differentially expressed genes in AD with cross-validation

Downregulated DEGs						
Probe	Symbols	Gene name	<i>p</i> -value	<i>q</i> -value	Jackknife cross-validation	AlzBase
206780_at	<i>GAD2</i>	Glutamate decarboxylase 2	6.24×10^{-7}	1.17×10^{-2}	Y	Y
205230_at	<i>RPH3A</i>	Rabphilin 3A	8.59×10^{-6}	2.30×10^{-2}	Y	Y
213921_at	<i>SST</i>	Somatostatin	1.57×10^{-5}	2.43×10^{-2}	Y	Y
205278_at	<i>GAD1</i>	Glutamate decarboxylase 1	1.54×10^{-5}	2.43×10^{-2}	Y	Y
211679_x_at	<i>GABBR2</i>	Gamma-aminobutyric acid type B receptor subunit 2	3.58×10^{-5}	3.73×10^{-2}	Y	Y
219855_at	<i>NUDT11</i>	Nudix hydrolase 11	8.35×10^{-5}	5.04×10^{-2}	Y	Y
210227_at	<i>DLGAP2</i>	<i>DLG</i> associated protein 2	1.06×10^{-4}	5.69×10^{-2}	Y	Y
213558_at	<i>PCLO</i>	Piccolo presynaptic cytomatrix protein	4.16×10^{-4}	9.62×10^{-2}	Y	Y
205635_at	<i>KALRN</i>	Kalirin, <i>rhogef</i> kinase	3.36×10^{-6}	2.30×10^{-2}	Y	Y
219478_at	<i>WFDC1</i>	<i>WAP</i> four-disulfide core domain 1	5.64×10^{-6}	2.30×10^{-2}	Y	Y
214956_at	<i>AAK1</i>	<i>AP2</i> associated kinase 1	7.12×10^{-6}	2.30×10^{-2}	Y	Y
210232_at	<i>CDC42</i>	Cell division cycle 42	7.83×10^{-6}	2.30×10^{-2}	Y	Y
205825_at	<i>PCSK1</i>	Proprotein convertase subtilisin/kexin type 1	1.73×10^{-5}	2.43×10^{-2}	Y	Y
204337_at	<i>RGS4</i>	Regulator of G-protein signaling 4	1.82×10^{-5}	2.43×10^{-2}	Y	Y
205691_at	<i>SYNGR3</i>	Synaptogyrin 3	1.97×10^{-5}	2.43×10^{-2}	Y	Y
209540_at	<i>IGF1</i>	Insulin like growth factor 1	2.05×10^{-5}	2.43×10^{-2}	Y	Y
209541_at	<i>IGF1</i>	Insulin like growth factor 1	2.07×10^{-5}	2.43×10^{-2}	Y	Y
204465_s_at	<i>INA</i>	Internexin neuronal intermediate filament protein alpha	3.40×10^{-5}	3.73×10^{-2}	Y	Y
205531_s_at	<i>GLS2</i>	Glutaminase 2	4.40×10^{-5}	3.89×10^{-2}	Y	Y
211685_s_at	<i>NCALD</i>	Neurocalcin delta	4.42×10^{-5}	3.89×10^{-2}	Y	Y
214230_at	<i>CDC42</i>	Cell division cycle 42	4.77×10^{-5}	3.89×10^{-2}	Y	Y
209583_s_at	<i>CD200</i>	<i>CD200</i> molecule	5.17×10^{-5}	4.03×10^{-2}	Y	Y
205774_at	<i>F12</i>	Coagulation factor XII	5.38×10^{-5}	4.03×10^{-2}	Y	Y
218009_s_at	<i>PRC1</i>	Protein regulator of cytokinesis 1	7.63×10^{-5}	4.76×10^{-2}	Y	Y
206408_at	<i>LRRTM2</i>	Leucine-rich repeat transmembrane neuronal 2	9.52×10^{-5}	5.40×10^{-2}	Y	Y
216963_s_at	<i>GAP43</i>	Growth associated protein 43	1.11×10^{-4}	5.69×10^{-2}	Y	Y
217979_at	<i>TSPAN13</i>	Tetraspanin 13	1.24×10^{-4}	5.97×10^{-2}	Y	Y
202712_s_at	<i>CKMT1A</i>	Creatine kinase, mitochondrial 1A	1.62×10^{-4}	8.18×10^{-2}	Y	Y
202712_s_at	<i>CKMT1B</i>	Creatine kinase, mitochondrial 1B	1.62×10^{-4}	8.18×10^{-2}	Y	Y
205268_s_at	<i>ADD2</i>	Adducin 2	2.23×10^{-4}	7.90×10^{-2}	Y	Y
208851_s_at	<i>THY1</i>	<i>Thy-1</i> cell surface antigen	3.22×10^{-4}	9.06×10^{-2}	Y	Y
213666_at	<i>SEPT6</i>	Septin 6	6.35×10^{-5}	4.57×10^{-2}	Y	N
220359_s_at	<i>ARPP21</i>	<i>Camp</i> -regulated phosphoprotein 21	1.51×10^{-5}	2.43×10^{-2}	N	Y
220334_at	<i>RGS17</i>	Regulator of G-protein signaling 17	4.40×10^{-5}	3.89×10^{-2}	N	Y
213198_at	<i>ACVR1B</i>	Activin A receptor type 1B	4.63×10^{-5}	3.89×10^{-2}	N	Y
206941_x_at	<i>SEMA3E</i>	Semaphorin 3E	6.83×10^{-5}	4.57×10^{-2}	N	Y
205327_s_at	<i>ACVR2A</i>	Activin A receptor type 2A	7.08×10^{-5}	4.57×10^{-2}	N	Y
205625_s_at	<i>CALB1</i>	Calbindin 1	1.12×10^{-4}	5.69×10^{-2}	N	Y
206691_s_at	<i>PDIA2</i>	Protein disulfide isomerase family A member 2	1.24×10^{-4}	5.97×10^{-2}	N	Y
210247_at	<i>SYN2</i>	Synapsin II	1.37×10^{-4}	6.39×10^{-2}	N	Y
202919_at	<i>MOB4</i>	<i>MOB</i> family member 4, phocein	1.54×10^{-4}	6.84×10^{-2}	N	Y
215518_at	<i>STXBPL5L</i>	Syntaxin-binding protein 5 like	1.57×10^{-4}	6.84×10^{-2}	N	Y
220030_at	<i>STYK1</i>	Serine/threonine/tyrosine kinase 1	2.04×10^{-4}	7.90×10^{-2}	N	Y

Table 2 continued

Downregulated DEGs						
Probe	Symbols	Gene name	<i>p</i> -value	<i>q</i> -value	Jackknife cross-validation	AlzBase
204525_at	<i>PHF14</i>	<i>PHD</i> finger protein 14	2.25×10^{-4}	7.90×10^{-2}	N	Y
219660_s_at	<i>ATP8A2</i>	Atpase phospholipid transporting 8A2	2.26×10^{-4}	7.90×10^{-2}	N	Y
205924_at	<i>RAB3B</i>	<i>RAB3B</i> , member <i>RAS</i> oncogene family	2.28×10^{-4}	7.90×10^{-2}	N	Y
203159_at	<i>GLS</i>	Glutaminase	2.37×10^{-4}	8.06×10^{-2}	N	Y
214157_at	<i>GNAS</i>	<i>GNAS</i> complex locus	2.50×10^{-4}	8.37×10^{-2}	N	Y
213436_at	<i>CNR1</i>	Cannabinoid receptor 1	2.78×10^{-4}	8.78×10^{-2}	N	Y
214098_at	<i>KIAA1107</i>	<i>Kiaa1107</i>	2.83×10^{-4}	8.78×10^{-2}	N	Y
215081_at	<i>KIAA1024</i>	<i>Kiaa1024</i>	2.87×10^{-4}	8.78×10^{-2}	N	Y
207242_s_at	<i>GRIK1</i>	Glutamate ionotropic receptor kainate type subunit 1	2.93×10^{-4}	8.78×10^{-2}	N	Y
219825_at	<i>CYP26B1</i>	Cytochrome <i>P450</i> family 26 subfamily B member 1	2.99×10^{-4}	8.78×10^{-2}	N	Y
206051_at	<i>ELAVL4</i>	<i>ELAV</i> like RNA-binding protein 4	3.00×10^{-4}	8.78×10^{-2}	N	Y
219752_at	<i>RASAL1</i>	<i>RAS</i> protein activator like 1	3.04×10^{-4}	8.78×10^{-2}	N	Y
203769_s_at	<i>STS</i>	Steroid sulfatase (microsomal), isozyme S	3.05×10^{-4}	8.78×10^{-2}	N	Y
205257_s_at	<i>AMPH</i>	Amphiphysin	3.50×10^{-4}	9.34×10^{-2}	N	Y
218404_at	<i>SNX10</i>	Sorting nexin 10	3.56×10^{-4}	9.34×10^{-2}	N	Y
220182_at	<i>SLC25A23</i>	Solute carrier family 25 member 23	3.59×10^{-4}	9.34×10^{-2}	N	Y
220794_at	<i>GREM2</i>	Gremlin 2, DAN family <i>BMP</i> antagonist	3.85×10^{-4}	9.45×10^{-2}	N	Y
219896_at	<i>CALY</i>	Calcyon neuron-specific vesicular protein	3.85×10^{-4}	9.45×10^{-2}	N	Y
208017_s_at	<i>MCF2</i>	<i>MCF2</i> cell line derived transforming sequence	3.88×10^{-4}	9.45×10^{-2}	N	Y
213386_at	<i>TMEM246</i>	Transmembrane protein 246	3.90×10^{-4}	9.45×10^{-2}	N	Y
206089_at	<i>NELL1</i>	Neural <i>EGFL</i> like 1	3.93×10^{-4}	9.45×10^{-2}	N	Y
203001_s_at	<i>STMN2</i>	Stathmin 2	4.02×10^{-4}	9.46×10^{-2}	N	Y
205630_at	<i>CRH</i>	Corticotropin releasing hormone	4.30×10^{-4}	9.83×10^{-2}	N	Y
215172_at	<i>PTPN20</i>	Protein tyrosine phosphatase, non-receptor type 20	3.88×10^{-6}	2.30×10^{-2}	N	N
212882_at	<i>KLHL18</i>	Kelch like family member 18	3.24×10^{-4}	9.06×10^{-2}	N	N
213636_at	<i>PHF24</i>	<i>PHD</i> finger protein 24	3.83×10^{-4}	9.45×10^{-2}	N	N
Upregulated DEGs						
Probe	Symbols	Gene name	<i>p</i> -value	<i>q</i> -value	Jackknife cross-validation	AlzBase
204932_at	<i>TNFRSF11B</i>	<i>TNF</i> receptor superfamily member 11b	1.90×10^{-4}	7.76×10^{-2}	Y	Y
214020_x_at	<i>ITGB5</i>	Integrin subunit beta 5	2.13×10^{-4}	7.90×10^{-2}	Y	Y
220132_s_at	<i>CLEC2D</i>	C-type lectin domain family 2 member D	1.99×10^{-5}	2.43×10^{-2}	N	Y
220593_s_at	<i>CCDC40</i>	Coiled-coil domain containing 40	6.60×10^{-5}	4.57×10^{-2}	N	Y
202365_at	<i>UNC119B</i>	Unc-119 lipid-binding chaperone B	9.52×10^{-5}	5.40×10^{-2}	N	Y
204428_s_at	<i>LCAT</i>	Lecithin-cholesterol acyltransferase	1.40×10^{-4}	6.39×10^{-2}	N	Y
220317_at	<i>LRAT</i>	Lecithin retinol acyltransferase (phosphatidylcholine-retinol O-acyltransferase)	1.91×10^{-4}	7.76×10^{-2}	N	Y
58780_s_at	<i>ARHGEF40</i>	<i>Rho</i> guanine nucleotide exchange factor 40	2.06×10^{-4}	7.90×10^{-2}	N	Y
205296_at	<i>RBL1</i>	<i>RB</i> transcriptional corepressor like 1	2.22×10^{-4}	7.90×10^{-2}	N	Y
216897_s_at	<i>FAM76A</i>	Family with sequence similarity 76 member A	2.94×10^{-4}	8.78×10^{-2}	N	Y
202901_x_at	<i>CTSS</i>	Cathepsin S	3.39×10^{-4}	9.20×10^{-2}	N	Y
206693_at	<i>IL7</i>	Interleukin 7	4.04×10^{-4}	9.46×10^{-2}	N	Y

Table 2 continued

Upregulated DEGs						
Probe	Symbols	Gene name	<i>p</i> -value	<i>q</i> -value	Jackknife cross-validation	AlzBase
220614_s_at	<i>ADGB</i>	Androglobin	1.12×10^{-4}	5.69×10^{-2}	N	N

"Jackknife cross-validation" denotes DEGs validated by seven leave-one-out tests (Y) or not (N), "AlzBase" denotes DEGs identified in AlzBase database (Y) or not (N)

may serve as potential biomarkers for the early diagnosis of AD.

DISCUSSION

In this study, we investigated the mRNA expression changes in the hippocampus that were consistent across up to three independent cohorts of subjects to illustrate the pathogenesis of AD. Eighty DEGs were identified in the combined analysis, and 31 of them were validated by at least seven leave-one-out tests and confirmed by the results of the AlzBase database. The validations of the DEGs demonstrate the reliability of the results to a certain extent.

Pathway enrichment analysis was performed to interpret the function of these DEGs. KEGG pathway analysis for the 80 DEGs suggested that five downregulated genes were significantly enriched in one KEGG pathway "GABAergic synapses" (Benjamin-corrected *p*-value = 0.039882), indicating that the GABAergic synapse pathway might be impaired in AD patients. GABAergic synapses exert an inhibitory effect on the nervous system. Downregulated GABAergic synapses are intimately coupled with the loss of GABAergic inhibition (Kuzirian and Paradis 2011). A close linkage was observed between GABAergic neurotransmission and various aspects of AD pathology, including Ab toxicity (Bell *et al.* 2006), tau hyper-phosphorylation (Nilsen *et al.* 2013), and apoE4 effect (Li *et al.* 2009, 2016). Significantly lower levels of GABA inhibitory neurotransmitter (~33%) were observed in the AD cases, indicating deficient synaptic function and neuronal transmission in AD (Gueli and Taibi 2013). Furthermore, animal experiments of an AD model illustrated that impaired hippocampal neurogenesis in AD mice may be mediated by the dysfunction of GABAergic signaling or an imbalance between excitatory and inhibitory synapse (Li *et al.* 2009; Sun *et al.* 2009). Hence, the pathway of GABAergic synapses is important not only for the function of the hippocampus but also for the pathogenesis of AD. GO analysis indicated that the 80 DEGs

were mainly involved in the neuron, secretory vesicle, synaptic signaling, synaptic transmission, cell junction, and synaptic vesicle metabolism. The impairment of neuronal and synaptic functions has long been considered an important pathologic characteristic in neurodegenerative diseases, and decreased synaptic activity is also considered to be the most relevant pathological feature of cognitive impairment in AD (Marttinen *et al.* 2015). These results demonstrate that GABAergic system, neurons, and synaptic function might be affected in the pathogenesis of AD.

To explore the protein interactions of the 80 DEGs, a PPI network extending from the DEGs was constructed, in which 180 extended genes interacting with DEGs were obtained. In the PPI network with 250 nodes and 497 edges, the 15 genes with greater than ten degrees (including *CDC42*, *RBL1*, *GNAS*, *CKMT1B*, *CKMT1A*, *AMPH*, *ACVR1B*, *CNR1*, *SEPT6*, *GAD1*, *PDIA2*, *MOB4*, *PRC1*, *ACVR2A*, and *UBC*) were selected, which were mainly involved in TGF-beta and *Rap1* signaling pathway. Among these genes, the validated DEG *CDC42* was the hub gene with the highest degree. This gene codes the protein cell division cycle 42; it is a member of the Rho guanosine triphosphatase (GTPase) family and plays an important role in cell morphology, proliferation, cell migration, and cell progression. *CDC42* is reported to play a critical role in striatal neuron growth in the gene expression profiling analysis of Parkinson's disease (Gao *et al.* 2013). In addition, *CDC42* is dysregulated in a transcriptomic meta-analysis between AD and type 2 diabetes mellitus (Mirza *et al.* 2014). Therefore, although *CDC42* has not been reported in original articles, it may play a crucial role in the pathogenesis of AD.

Furthermore, to identify key genes for early diagnosis and treatment of AD, we prioritized the DEGs and the extended genes by using the CFG method. The results suggested that 12 highly AD-relevant genes, including five DEGs (*ITGB5*, *RPH3A*, *GNAS*, *THY1*, and *SEPT6*) and seven extended genes (*JUN*, *GDI1*, *GNAI2*, *NEK6*, *UBE2D3*, *CDC42EP4*, and *ERCC3*), might be promising for evaluating early diagnostic biomarkers in AD. Among them, three genes, including two validated DEGs (*ITGB5*,

Table 3 Significantly enriched KEGG pathway and GO terms of the 80 DEGs

Category	Term	Genes	Benjamin correction
KEGG			
KEGG_PATHWAY	hsa04727 ~ GABAergic synapse	<i>GLS2, GAD2, GLS, GABBR2, GAD1</i>	3.99×10^{-2}
GO terms			
GOTERM_CC_FAT	GO:0044456 ~ synapse part	<i>MOB4, RAB3B, GRIK1, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, SYNGR3, AMPH, GAD2, AAK1, LRRTM2, SYN2, SEPT6, GAD1, ADD2, GAP43, KALRN</i>	6.79×10^{-8}
GOTERM_CC_FAT	GO:0097458 ~ neuron part	<i>RAB3B, MOB4, GABBR2, CALB1, AMPH, CDC42, GAD2, AAK1, CNR1, SYN2, GAD1, DLGAP2, STMN2, RGS17, RPH3A, PCLO, SYNGR3, THY1, ATP8A2, CRH, GNAS, SEPT6, SST, GAP43, ADD2, KALRN</i>	9.01×10^{-8}
GOTERM_CC_FAT	GO:0045202 ~ synapse	<i>MOB4, RAB3B, GRIK1, DLGAP2, GABBR2, RGS17, RPH3A, PCLO, CALB1, SYNGR3, AMPH, GAD2, AAK1, LRRTM2, SYN2, SEPT6, GAD1, ADD2, GAP43, KALRN</i>	1.03×10^{-7}
GOTERM_BP_FAT	GO:0099536 ~ synaptic signaling	<i>RAB3B, GRIK1, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, AMPH, GAD2, GLS, LRRTM2, CNR1, SYN2, CRH, GAD1, SST, KALRN</i>	1.62×10^{-5}
GOTERM_BP_FAT	GO:0098916 ~ anterograde trans-synaptic signaling	<i>RAB3B, GRIK1, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, AMPH, GAD2, GLS, LRRTM2, CNR1, SYN2, CRH, GAD1, SST, KALRN</i>	1.62×10^{-5}
GOTERM_BP_FAT	GO:0099537 ~ trans-synaptic signaling	<i>RAB3B, GRIK1, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, AMPH, GAD2, GLS, LRRTM2, CNR1, SYN2, CRH, GAD1, SST, KALRN</i>	1.62×10^{-5}
GOTERM_BP_FAT	GO:0007268 ~ chemical synaptic transmission	<i>RAB3B, GRIK1, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, AMPH, GAD2, GLS, LRRTM2, CNR1, SYN2, CRH, GAD1, SST, KALRN</i>	1.62×10^{-5}
GOTERM_CC_FAT	GO:0098793 ~ presynapse	<i>RAB3B, GAD2, AAK1, SYN2, RPH3A, SEPT6, GAD1, CALB1, PCLO, SYNGR3, AMPH</i>	4.94×10^{-5}
GOTERM_BP_FAT	GO:0007267 ~ cell-cell signaling	<i>RAB3B, GRIK1, IL7, DLGAP2, GABBR2, RPH3A, PCLO, CALB1, AMPH, STXBP5L, CDC42, PCSK1, GAD2, CNR1, GLS, LRRTM2, SYN2, CRH, GNAS, GAD1, SST, KALRN</i>	1.91×10^{-3}
GOTERM_CC_FAT	GO:0070382 ~ exocytic vesicle	<i>RAB3B, GAD2, SYN2, IGF1, RPH3A, SEPT6, SYNGR3, AMPH</i>	2.25×10^{-4}
GOTERM_CC_FAT	GO:0030133 ~ transport vesicle	<i>PCSK1, RAB3B, GAD2, NCALD, SYN2, IGF1, GNAS, RPH3A, SEPT6, SYNGR3, AMPH</i>	2.06×10^{-4}
GOTERM_CC_FAT	GO:0016023 ~ cytoplasmic, membrane-bounded vesicle	<i>RAB3B, CALY, NCALD, ITGB5, IGF1, RPH3A, SYNGR3, AMPH, STXBP5L, CDC42, PCSK1, GAD2, AAK1, SYN2, GNAS, SEPT6, GAD1, ADD2</i>	9.71×10^{-4}
GOTERM_CC_FAT	GO:0008021 ~ synaptic vesicle	<i>RAB3B, GAD2, SYN2, RPH3A, SEPT6, SYNGR3, AMPH</i>	9.31×10^{-4}
GOTERM_CC_FAT	GO:0099503 ~ secretory vesicle	<i>CDC42, PCSK1, RAB3B, GAD2, SYN2, IGF1, RPH3A, SEPT6, SYNGR3, AMPH, STXBP5L</i>	1.82×10^{-3}
GOTERM_CC_FAT	GO:0043005 ~ neuron projection	<i>MOB4, STMN2, GABBR2, RGS17, RPH3A, CALB1, THY1, CDC42, GAD2, AAK1, CNR1, CRH, GNAS, SEPT6, GAP43</i>	3.65×10^{-3}
GOTERM_CC_FAT	GO:0099501 ~ exocytic vesicle membrane	<i>GAD2, SYN2, RPH3A, SYNGR3, AMPH</i>	4.71×10^{-3}
GOTERM_CC_FAT	GO:0030672 ~ synaptic vesicle membrane	<i>GAD2, SYN2, RPH3A, SYNGR3, AMPH</i>	4.71×10^{-3}
GOTERM_CC_FAT	GO:0098794 ~ postsynapse	<i>MOB4, GRIK1, DLGAP2, LRRTM2, GABBR2, PCLO, GAP43, ADD2, KALRN</i>	8.05×10^{-3}
GOTERM_CC_FAT	GO:0030658 ~ transport vesicle membrane	<i>GAD2, NCALD, SYN2, RPH3A, SYNGR3, AMPH</i>	1.72×10^{-2}
GOTERM_CC_FAT	GO:0030054 ~ cell junction	<i>GRIK1, DLGAP2, ITGB5, GABBR2, RGS17, RPH3A, PCLO, SYNGR3, AMPH, THY1, CDC42, GAD2, LRRTM2, SYN2, CD200, GAP43</i>	3.04×10^{-2}
GOTERM_CC_FAT	GO:0044433 ~ cytoplasmic vesicle part	<i>PCSK1, GAD2, CALY, NCALD, SYN2, IGF1, RPH3A, GAD1, SYNGR3, AMPH</i>	3.89×10^{-2}
GOTERM_CC_FAT	GO:0060076 ~ excitatory synapse	<i>DLGAP2, LRRTM2, PCLO, GAP43, ADD2, KALRN</i>	3.67×10^{-2}

CC: cellular component; BP: biological process

Fig. 1 PPI network formed by the DEGs and their interacting genes. The *red nodes* all represent the DEGs in our finding, in which the circular nodes represent the downregulated DEGs and the diamond nodes indicate the upregulated DEGs. The *gray triangle nodes* are the extended genes interacting with the DEGs. The *red edges* are interactions among the DEGs, whereas the *gray edges* are interactions between the DEGs and the extended genes. The node size in each panel is proportional to the degree of the node

RPH3A) and one extended gene (*JUN*), had the highest CFG scores. *ITGB5* encodes a beta subunit of integrin, which participates in cell adhesion and cell-surface-mediated signaling. It not only supports tumorigenesis but also enhances tumor growth (Reynolds *et al.* 2002). Moreover, studies have shown that dysregulated *ITGB5* gene is correlated with diabetic nephropathy (Wang *et al.* 2016c) and might play an important role in the progression of AD. *RPH3A* is a small G-protein that acts in the exocytosis of neurotransmitters and hormones and is involved in neurotransmitter release and synaptic vesicle traffic. Previous studies demonstrate that the expression level of *RPH3A* is downregulated in different brain regions in a transgenic mouse model of Huntington disease and may be correlated with the symptoms of the neurodegenerative disorder (Smith *et al.* 2007). Although *JUN* was not dysregulated in the combined analysis of genome-wide expression profiling in the hippocampus, it was found to closely interact with eight genes in the PPI network, including the hub gene *CDC42*. *JUN* encodes a virus-like protein, which regulates gene expression in response to cell stimulation by interacting directly with target DNA sequences. Several studies have shown that the transcription factor *JUN* is

essential for neuronal microtubule assembly and apoptosis (Nateri *et al.* 2004) and plays a very key regulatory role in the unfolded protein response in acute myeloid leukemia, which can serve as a promising therapeutic target in this disease (Zhou *et al.* 2017). Recently, PPI analysis of AD and non-alcoholic fatty liver disease (NAFLD) has revealed that *JUN* is one of the hub-bottleneck proteins in the PPI network and is an important target for both AD and NAFLD (Karbalaei *et al.* 2018; Paquet *et al.* 2017).

Our study makes the statistical improvement to directly combine raw probe-level data from different studies and control several confounding factors to identify a number of best-estimated DEGs between AD patients and age-matched controls. Based on these DEGs, we further elaborated the associated biological pathways and potential biomarkers, shedding new light on the interpretation of the pathogenic mechanisms and early diagnosis underlying AD. However, given that all functional evidences were obtained via bioinformatics analysis, future independent validation studies and essential functional assays are necessary for consolidating the current conclusions and characterizing the putative impact of the candidate genes in AD.

Fig. 2 Probability pyramid representing the results of gene prioritization for the 80 DEGs and the 180 extended genes. The highest CFG score is 6. Colors represent different genes in our study. Blue represents the DEGs, whereas red means the extended genes

MATERIALS AND METHODS

Dataset selection

Microarray-based gene expression data of subjects with AD were obtained from Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) and ArrayExpress (<https://www.ebi.ac.uk/arrayexpress>). The eligible studies were searched with keywords “Alzheimer’s disease” and we set organism as *Homo sapiens*, and array type as “Expression profiling by array” in GEO or “Transcription profiling by array” in ArrayExpress. Raw probe-level data (CEL files) that focused on gene expression profiling in the hippocampus of a cohort of neuropathological healthy subjects and a cohort of AD subjects were collected. Information on covariates, including age, gender, PMI, and batch effect, was required for this study. To avoid deviation among the different microarray platforms, only data generated from two similar Affymetrix platforms, HGU 133a (Human Genome U133A) and HGU 133p 2.0 (Human Genome U133 Plus 2.0), were finally used. Ultimately, after removing the duplicate individuals, raw data from three independent datasets were retained.

Data quality control and pre-processing

To reduce the bias due to different analytical methods, each individual dataset was reprocessed and normalized independently using the R Bioconductor affy package with the default settings for robust multi-array average (RMA) normalization (Irizarry *et al.* 2003). All data were background-adjusted, normalized, and log-transformed. The microarrays were assessed for data quality using the SimpleAffy package (Wilson and Miller 2005). The scale factor, average background, percent present, the 3'/5' intensity ratio of GAPDH, and the 3'/5' intensity ratio of beta-actin provided by SimpleAffy were all evaluated to determine the quality of the RNA samples and their subsequent labeling and hybridization. The default values were selected according to the recommendations of Affymetrix, SimpleAffy, and Larsson and Sandberg (Larsson and Sandberg 2006). Samples beyond the default values were suspected as unqualified and were subsequently identified by RLE and NUSE boxplot using the affyPLM package (Bolstad *et al.* 2005). If samples in the RLE and NUSE plots were far from 1, they were removed. With regard to the probe sets, only the intersection of probe sets from the two Affymetrix platforms was utilized. Ultimately, data from

eight samples were removed, including one control and seven AD patients. The merged gene expression value matrix contained 116 samples and 22,777 probe sets.

Statistical modeling

The expression and sample characteristics from each study were merged, and the gene expression value for each probe was calculated using a standard linear mixed-effects model. Statistical modeling was conducted by the lme4 package in R (Bates *et al.* 2015). In the combined analysis, disease, age, and PMI were used as fixed effects, whereas gender and batch effect were used as random effects (Wang *et al.* 2016a). The likelihood ratio test was used to calculate the statistical significance (*p*-values) by comparing this model with the null model containing all factors in the original model, except disease. For each probe set, the *t*-statistic for the disease effect calculated in the linear mixed-effects model manifested the direction of gene expression, *i.e.*, upregulated or downregulated. For the multiple test correction, the *p*-values for the dysregulated signatures were further adjusted using the *q*-value method to control the false discovery rate (Storey and Tibshirani 2003), and a more permissive *q*-value < 0.10 was used to gain more transcripts for comparison and further analysis. Then, the AnnotationDbi package (<http://www.bioconductor.org/packages/release/bioc/html/AnnotationDbi.html>) was utilized to annotate the probe sets with gene symbols, EntrezIDs, and gene names, thereby obtaining DEGs between the AD patients and age-matched healthy controls.

Validation of the identified DEGs

To test the robustness of the findings, a jackknife cross-validation was used to conduct the "leave-one-out" test. The procedure was that the data from the total samples, the cases or the control samples in each dataset, were sequentially removed, and the same analysis procedure was applied for the remaining data. Next, the results were compared with the findings of the combined analysis to explore the overlapping genes. This step helped to identify the most important genes that were not dependent on a single study, as well as each study's contribution to the final results.

To assess the sensitivity of DEGs, the findings were contrasted with the results of the AlzBase database (<http://alz.big.ac.cn/>). AlzBase is an integrative database for dysregulated genes in AD pathogenesis that are discovered from studies with animal models or neuronal cell lines (Bai *et al.* 2016). This database collects the frequency of the dysregulated genes compiled from

published studies of the AD brain transcriptome, which may have a high priority to be pursued further. Our signatures were compared with the genes from the AlzBase database for a better understanding of our results.

Functional and network analysis

To identify the functional categories and biological processes in the hippocampus, we performed pathway analyses using DAVID (version 6.8) (da Huang *et al.* 2009). The KEGG pathways and GO terms (including cellular component, biological process, and molecular function) were selected in the enrichment analysis, and after a Benjamin multiple test correction, the cutoff of the significance was set to $q < 0.05$. The gene interaction network among these DEGs was constructed from the largest PPI database, InWeb_InBioMap (version 2016_09_12, <https://www.intomics.com/inbio/map/#downloads>) (Li *et al.* 2017). With DEGs as the seed genes, the extended genes were introduced on the basis of the experimentally validated interaction in the InWeb_InBioMap database. Each extended node gene was required to have at least two direct interactions with the DEGs. The DEGs and the extended genes were mapped into the PPI network to explore the molecular mechanism of AD. The biological graph visualization tool Cytoscape (version 3.5.1, <http://www.cytoscape.org/>) software was used to visualize the PPI networks. In the PPI network, the number of genes directly linked to a node was defined as the degrees of the node, and the node with the higher degree (degree > 10) was defined as the hub gene.

Gene prioritization of the candidate genes

AlzData (<http://www.alzdata.org/>) integrates five lines of evidence associated with AD, including GWAS, PPI, brain expressional quantitative trait loci, expression correlation with AD pathology in AD mice, and early alteration in 2-month-old AD mouse brain, to prioritize candidate genes for further characterization using a CFG method (Xu *et al.* 2018). In addition to the evidence collected in the AlzData database, we explored the expression pattern of the blood of AD patients. Dysregulated genes between the blood of AD patients and aged-matched healthy elderly controls were also collected from the results of publications (Chen *et al.* 2011; Fehlbaum-Beurdeley *et al.* 2010; Maes, *et al.* 2007; Sood *et al.* 2015). The DEGs and the extended genes obtained from the PPI network were scored based on the evidence collected from the AlzData database (CFG score). One point would be assigned if the gene was supported

by the above-mentioned evidence. Otherwise, the gene would be assigned zero points. The CFG score of each gene ranged from 0 to 6 points; the higher the score, the more promising the gene is.

Acknowledgements This work was supported by the National Basic Research Program (973 Program) (2015CB351702) and the CAS Key Laboratory of Mental Health, Institute of Psychology. We thank all the groups who released the original datasets for sharing.

Compliance with Ethical Standards

Conflict of interest Mengsi Wu, Kechi Fang, Weixiao Wang, Wei Lin, Liyuan Guo, and Jing Wang declare that they have no conflicts of interests.

Human and animal rights and informed consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bai Z, Han G, Xie B, Wang J, Song F, Peng X, Lei H (2016) AlzBase: an integrative database for gene dysregulation in Alzheimer's disease. *Mol Neurobiol* 53:310–319
- Bates D, Machler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- Bell KF, Ducatenzeiler A, Ribeiro-da-Silva A, Duff K, Bennett DA, Cuello AC (2006) The amyloid pathology progresses in a neurotransmitter-specific manner. *Neurobiol Aging* 27:1644–1657
- Berchtold NC, Cribbs DH, Coleman PD, Rogers J, Head E, Kim R, Beach T, Miller C, Troncoso J, Trojanowski JQ, Zielke HR, Cotman CW (2008) Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proc Natl Acad Sci USA* 105:15605–15610
- Berchtold NC, Coleman PD, Cribbs DH, Rogers J, Gillen DL, Cotman CW (2013) Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. *Neurobiol Aging* 34:1653–1661
- Blair LJ, Nordhues BA, Hill SE, Scaglione KM, O'Leary JC 3rd, Fontaine SN, Breydo L, Zhang B, Li P, Wang L, Cotman C, Paulson HL, Muschol M, Uversky VN, Klengel T, Binder EB, Kaye R, Golde TE, Berchtold N, Dickey CA (2013) Accelerated neurodegeneration through chaperone-mediated oligomerization of tau. *J Clin Invest* 123:4158–4169
- Blalock EM, Geddes JW, Chen KC, Porter NM, Markesbery WR, Landfield PW (2004) Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc Natl Acad Sci USA* 101:2173–2178
- Bolstad BM, Collin F, Brettschneider J, Simpson K, Cope L, Irizarry RA, Speed TP (2005) Quality assessment of Affymetrix GeneChip data. In: Gentleman R et al (eds) *Bioinformatics and computational biology solution using R and bioconductor*. Springer, New York, pp 33–47
- Chen K-D, Chang P-T, Ping Y-H, Lee H-C, Yeh C-W, Wang P-N (2011) Gene expression profiling of peripheral blood leukocytes identifies and validates ABCB1 as a novel biomarker for Alzheimer's disease. *Neurobiol Dis* 43:698–705
- Cooper-Knock J, Kirby J, Ferraiuolo L, Heath PR, Rattray M, Shaw PJ (2012) Gene expression profiling in human neurodegenerative disease. *Nat Rev Neurol* 8:518–530
- da Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57
- Fehlbaum-Beurdeley P, Prado ACJ-L, Pallares D, Carriere J, Soucaille C, Rouet F, Drouin D, Sol O, Jordan H, Wu D, Lei L, Einstein R, Schweighoffer F, Bracco L (2010) Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimers Dement* 6:25–38
- Gao L, Gao H, Zhou H, Xu Y (2013) Gene expression profiling analysis of the putamen for the investigation of compensatory mechanisms in Parkinson's disease. *BMC Neurol* 13:181
- Gueli MC, Taibi G (2013) Alzheimer's disease: amino acid levels and brain metabolic status. *Neurol Sci* 34:1575–1579
- Hess JL, Tylee DS, Barve R, de Jong S, Ophoff RA, Kumarasinghe N, Tooney P, Schall U, Gardiner E, Beveridge NJ, Scott RJ, Yasawardene S, Perera A, Mendis J, Carr V, Kelly B, Cairns M, Neurobehavioural Genetics Unit, Tsuang MT, Glatt SJ (2016) Transcriptome-wide mega-analyses reveal joint dysregulation of immunologic genes and transcription regulators in brain and blood in schizophrenia. *Schizophr Res* 176:114–124
- Hu W, Lin X, Chen K (2015) Integrated analysis of differential gene expression profiles in hippocampi to identify candidate genes involved in Alzheimer's disease. *Mol Med Rep* 12:6679–6687
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)* 4:249–264
- Karbalaei R, Allahyari M, Rezaei-Tavirani M, Asadzadeh-Aghdaei H, Zali MR (2018) Protein-protein interaction analysis of Alzheimer's disease and NAFLD based on systems biology methods unhide common ancestor pathways. *Gastroenterol Hepatol Bed Bench* 11:27–33
- Kuzirian MS, Paradis S (2011) Emerging themes in GABAergic synapse development. *Prog Neurobiol* 95:68–87
- Larsson O, Sandberg R (2006) Lack of correct data format and comparability limits future integrative microarray research. *Nat Biotechnol* 24:1322–1323
- Li G, Bien-Ly N, Andrews-Zwilling Y, Xu Q, Bernardo A, Ring K, Halabisky B, Deng C, Mahley RW, Huang Y (2009) GABAergic interneuron dysfunction impairs hippocampal neurogenesis in adult apolipoprotein E4 knockin mice. *Cell Stem Cell* 5:634–645
- 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* 5:12393
- Li Y, Sun H, Chen Z, Xu H, Bu G, Zheng H (2016) Implications of GABAergic neurotransmission in Alzheimer's disease. *Front Aging Neurosci* 8:331
- Li T, Wernersson R, Hansen RB, Horn H, Mercer J, Slodkiewicz G, Workman CT, Rigina O, Rapacki K, Staerfeldt HH, Brunak S, Jensen TS, Lage K (2017) A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat Methods* 14:61–64

- Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Ramsey K, Caselli RJ, Kukull WA, McKeel D, Morris JC, Hulette CM, Schmechel D, Reiman EM, Rogers J, Stephan DA (2008a) Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. *Physiol Genomics* 33:240–256
- Liang WS, Reiman EM, Valla J, Dunckley T, Beach TG, Grover A, Niedzielko TL, Schneider LE, Mastroeni D, Caselli R, Kukull W, Morris JC, Hulette CM, Schmechel D, Rogers J, Stephan DA (2008b) Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proc Natl Acad Sci USA* 105:4441–4446
- Maes OC, Xu S, Yu B, Chertkow HM, Wang E, Schipper HM (2007) Transcriptional profiling of Alzheimer blood mononuclear cells by microarray. *Neurobiol Aging* 28:1795–1809
- Mak E, Gabel S, Su L, Williams GB, Arnold R, Passamonti L, Vazquez Rodriguez P, Surendranathan A, Bevan-Jones WR, Rowe JB, O'Brien JT (2017) Multi-modal MRI investigation of volumetric and microstructural changes in the hippocampus and its subfields in mild cognitive impairment, Alzheimer's disease, and dementia with Lewy bodies. *Int Psychogeriatr* 29:545–555
- Marttinen M, Kurkinen KM, Soininen H, Haapasalo A, Hiltunen M (2015) Synaptic dysfunction and septin protein family members in neurodegenerative diseases. *Mol Neurodegener* 10:16
- Mirza Z, Kamal MA, Buzenadah AM, Al-Qahtani MH, Karim S (2014) Establishing genomic/transcriptomic links between Alzheimer's disease and type 2 diabetes mellitus by meta-analysis approach. *CNS Neurol Disord-Drug Targets* 13:501–516
- Nateri AS, Riera-Sans L, Da Costa C, Behrens A (2004) The ubiquitin ligase SCFFbw7 antagonizes apoptotic JNK signaling. *Science (New York, NY)* 303:1374–1378
- Nilsen LH, Rae C, Ittner LM, Gotz J, Sonnewald U (2013) Glutamate metabolism is impaired in transgenic mice with tau hyperphosphorylation. *J Cereb Blood Flow Metab* 33:684–691
- Paquet C, Nicoll JA, Love S, Mouton-Liger F, Holmes C, Hugon J, Boche D (2017) Downregulated apoptosis and autophagy after anti-Aβ immunotherapy in Alzheimer's disease. *Brain Pathol* 28(5):603–610
- Reynolds LE, Wyder L, Lively JC, Taverna D, Robinson SD, Huang X, Sheppard D, Hynes RO, Hodivala-Dilke KM (2002) Enhanced pathological angiogenesis in mice lacking β3 integrin or β3 and β5 integrins. *Nat Med* 8:27–34
- Rosenberg PB, Nowrangi MA, Lyketsos CG (2015) Neuropsychiatric symptoms in Alzheimer's disease: What might be associated brain circuits? *Mol Aspects Med* 43–44:25–37
- Smith R, Klein P, Koc-Schmitz Y, Waldvogel HJ, Faull RL, Brundin P, Plomann M, Li JY (2007) Loss of SNAP-25 and rabphilin 3a in sensory-motor cortex in Huntington's disease. *J Neurochem* 103:115–123
- Sood S, Gallagher IJ, Lunnon K, Rullman E, Keohane A, Crossland H, Phillips BE, Cederholm T, Jensen T, van Loon LJ, Lannfelt L, Kraus WE, Atherton PJ, Howard R, Gustafsson T, Hodges A, Timmons JA (2015) A novel multi-tissue RNA diagnostic of healthy ageing relates to cognitive health status. *Genome Biol* 16:185
- Storey JD, Tibshirani R (2003) Statistical significance for genome-wide studies. *Proc Natl Acad Sci USA* 100:9440–9445
- Sun B, Halabisky B, Zhou Y, Palop JJ, Yu G, Mucke L, Gan L (2009) Imbalance between GABAergic and glutamatergic transmission impairs adult neurogenesis in an animal model of Alzheimer's disease. *Cell Stem Cell* 5:624–633
- van Cauwenberghe C, van Broeckhoven C, Sleegers K (2016) The genetic landscape of Alzheimer disease: clinical implications and perspectives. *Genet Med* 18:421–430
- Wang J, Qu S, Wang W, Guo L, Zhang K, Chang S, Wang J (2016a) A combined analysis of genome-wide expression profiling of bipolar disorder in human prefrontal cortex. *J Psychiatr Res* 82:23–29
- Wang M, Roussos P, McKenzie A, Zhou X, Kajiwarra Y, Brennand KJ, De Luca GC, Crary JF, Casaccia P, Buxbaum JD, Ehrlich M, Gandy S, Goate A, Katsel P, Schadt E, Haroutunian V, Zhang B (2016b) Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med* 8:104
- Wang Z, Wang Z, Zhou Z, Ren Y (2016c) Crucial genes associated with diabetic nephropathy explored by microarray analysis. *BMC Nephrol* 17:128
- Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Clifford RM, Jagust W, Morris JC, Petersen RC, Saykin AJ, Shaw LM, Toga AW, Trojanowski JQ, Alzheimer's Dis N (2017) Recent publications from the Alzheimer's disease neuroimaging initiative: reviewing progress toward improved AD clinical trials. *Alzheimers Dement* 13:E1–E85
- Wilson CL, Miller CJ (2005) Simpleaffy: a BioConductor package for affymetrix quality control and data analysis. *Bioinformatics* 21:3683–3685
- Xu M, Zhang DF, Luo R, Wu Y, Zhou H, Kong LL, Bi R, Yao YG (2018) A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement* 14:215–229
- Zhou C, Martinez E, Di Marcantonio D, Solanki-Patel N, Aghayev T, Peri S, Ferraro F, Skorski T, Scholl C, Frohling S, Balachandran S, Wiest DL, Sykes SM (2017) JUN is a key transcriptional regulator of the unfolded protein response in acute myeloid leukemia. *Leukemia* 31:1196–1205