Contents lists available at ScienceDirect

جــامـعــة الملكسعود King Saud University

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Integrative computational evaluation of genetic markers for Alzheimer's disease

Zhe Li^{a,b,1}, ZhenZhen Xiong^{c,1}, Lydia C. Manor^d, Hongbao Cao^{e,f}, Tao Li^{a,*}

^a The Mental Health Center and the Psychiatric Laboratory, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

^b The Mental Rehabilitation Center, Karamay Municipal People's Hospital, Karamay, Xinjiang 830054, China

^c School of Nursing, Chengdu Medical College, Chengdu, Sichuan 610083, China

^d Department of Bioinformatics, American Informatics Consultant LLC, Rockville, MD 20852, USA

^e Unit on Statistical Genomics, National Institute of Mental Health, NIH, Bethesda 20892, USA

^fDepartment of Genomics Research, R&D Solutions, Elsevier Inc., Rockville, MD 20852, USA

ARTICLE INFO

Article history: Received 1 May 2018 Revised 14 May 2018 Accepted 14 May 2018 Available online 18 May 2018

Keywords: Alzheimer's disease ResNet database Pathway enrichment analysis Sub-network enrichment analysis Gene-gene interaction analysis

ABSTRACT

Recent studies have reported hundreds of genes linked to Alzheimer's Disease (AD). However, many of these candidate genes may be not identified in different studies when analyses were replicated. Moreover, results could be controversial. Here, we proposed a computational workflow to curate and evaluate AD related genes. The method integrates large scale literature knowledge data and gene expression data that were acquired from postmortem human brain regions (AD case/control: 31/32 and 22/8). Pathway Enrichment, Sub-Network Enrichment, and Gene-Gene Interaction analysis were conducted to study the pathogenic profile of the candidate genes, with 4 metrics proposed and validated for each gene. By using our approach, a scalable AD genetic database was developed, including AD related genes, pathway, diseases and info of supporting references. The AD case/control classification supported the effectiveness of the 4 proposed metrics, which successfully identified 21 well-studied AD genes (i.g. TGFB1, CTNNB1, APP, IL1B, PSEN1, PTGS2, IL6, VEGFA, SOD1, AKT1, CDK5, TNF, GSK3B, TP53, CCL2, BDNF, NGF, IGF1, SIRT1, AGER and TLR) and highlighted one recently reported AD gene (i.g. ITGB1). The computational biology approach and the AD database developed in this study provide a valuable resource which may facilitate the understanding of the AD genetic profile.

© 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Alzheimer's disease (AD) is a chronic neurodegenerative disease that usually onsets slowly and progresses more rapidly over time (Burns and Iliffe, 2009). It is the leading cause of dementia, beginning with impaired memory, and most often onsets in people over 65 years of age (Mendez 2012). The global prevalence of AD as of 2015 was estimated to be as high as 48 million people worldwide

* Corresponding author at: The Mental Health Center and the Psychiatric Laboratory, West China Hospital, Sichuan University, No. 28 Dian Xin Nan Road, Chengdu, Sichuan 610041, China.

E-mail address: xuntao26@hotmail.com (T. Li).

¹ These authors contributed equally to this work.

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

(World Health Organization, 2015). Although the cause of most Alzheimer's cases largely remains unknown, about 70% of the risk is believed to come from a large network of genes (Ballard et al., 2011). As such, researches into the causes of AD are currently being explored.

In recent years, an increased number of genetic researches have been conducted revealing over a thousand altered genes linked to AD. For example, increased GSK3B activity and decreased phosphorylation of the gene have been repeatedly observed in AD cases (Cole et al., 2007; Koedam et al., 2013; Xu et al., 2016). Significantly increased expression levels of TP53, PTGS2 and TGFB1 were suggested by many independent studies to be associated with AD (Cenini et al., 2008; Lanni et al., 2012; Ramalho et al., 2008; Yoo et al., 2008; Wang et al., 2014; Luo et al., 2006). Observations from these previous studies are valuable in studying the genetic basis of the pathogenic development of the disease.

However, approximately one third of these AD-gene linkages were reported once with no further replication, and over 60% were supported by no more than three citations. Moreover, most of

https://doi.org/10.1016/j.sjbs.2018.05.019

1319-562X/© 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).





these studies had small sample sizes that were more susceptible to noise. Additionally, due to the variation in data collection and processing approaches, results from different studies were not always consistent. Meanwhile, there are dozens of new AD risk genes being reported every year, posing an increased need for further validation of these candidate genes to AD. While biological experiments were effective towards this validation task, they could be very costly. To address this issue, we propose a computational biology approach for a systematic evaluation of these AD candidate genes.

In recent years, Pathway Studio ResNet relation data have been widely used to study modeled relationships between proteins, genes, complexes, cells, tissues and diseases (http://pathwaystudio.gousinfo.com/Mendeley.html). In this study, we integrated large scale AD related ResNet literature knowledge data, independent gene expression data and related pathway/network information to study the functional profile of a large gene pool that has been reported to be linked to AD. The purpose of the study is to provide an easy-update computational evaluation workflow, through which an AD genetic database (AD_GD) could be generated to present a weighted landscape view of the genetic basis underlying the pathogenic development of AD. Our results support the hypothesis that AD candidate genes are functionally linked to each other, forming a large genetic network to regulate the pathogenic development of AD through multiple pathways.

2. Materials and methods

Fig. 1 presents the diagram of the proposed computational gene marker evaluation system, with detailed descriptions in the following sub-sections. Using our approach, a genetic database (AD_GD) was developed and deposited into an open source 'Bioinformatics Database' online available at http://database.gousinfo.com, including 1699 genes (with metric scores), 151 pathways and 114 diseases that are linked to AD. Also included in AD_GD are information of 27000+ supporting references for AD-gene relationships, including the titles and relevant sentences where the relations were identified. The AD_GD database is scalable and will be updated monthly or upon request using our approach.

2.1. ResNet literature knowledge data

ResNet relation data (AD-Gene) were acquired from the Pathway Studio ResNet[®] Mammalian database (http://pathwaystudio.gousinfo.com/ResNetDatabase.html) updated November 2016. The ResNet[®] Mammalian database are a group of real-time updated literature knowledge databases, including curated signaling, cellular processes and metabolic pathways, ontologies and annotations, as well as molecular interactions and functional relationships (http://pathwaystudio.gousinfo.com/ResNetDatabase. html). Modeled relation data are extracted from the 41M+ references covering entire PubMed abstracts and Elsevier and third party full text journals. The ResNet database employs an automated natural language processing-based information extraction system, MedScan, with precision of over 91% (Daraselia et al., 2004). Each relationship within the database is supported with one or more references. By far, Pathway Studio ResNet Databases is the largest database among known competitors in the field (Lorenzi et al., 2014).

2.2. Enrichment and gene-gene interaction analysis

Pathway enrichment analysis (PEA) and sub-network enrichment analysis (SNEA) (http://pathwaystudio.gousinfo.com/SNEA. pdf) was conducted using Pathway Studio to identify genetic pathways and diseases potentially linked to AD (Sivachenko et al., 2007). Furthermore, a pathway based gene-gene interaction (GGI) analysis was conducted to generate weighted edges/linkage between genes. The weight of an edge is the number of pathways where both nodes were included.

2.3. Metrics analysis

For the gene network built through the aforementioned steps, 4 metrics were proposed for each node/gene, including 2 literature based metric scores (RScore and AScore), and 2 enrichment based metric scores (PScore and SScore). The logic is that, a gene is likely linked to AD if it satisfies one or more of the following conditions: the gene has been frequently observed in independent studies to be associated with AD (high RScore), plays roles within multiple pathways associated with AD (high PScore), and demonstrates strong functional linkage to many of other genes associated with AD (high SSCore). Additionally, an AScore was proposed to present the history of each AD-gene relation. The detailed definitions of the proposed metrics are described as follows.

2.3.1. Two literature metrics

The reference score (RScore) of a gene is defined as the reference number underlying a gene-disease relationship, as shown in Eq. (1).



Fig. 1. Diagram for the integrative computational marker evaluation approach for AD. First, literature based analysis were conducted to identify the AD related genes, then Gene-Gene Interaction Analysis, Enrichment analysis, and Metrics analysis were conducted on these gene and results were saved in the AD database. Finally, AD case/control classification were conducted to test the effectiveness of the identified genes, using gene expression datasets.

(1)

Rscore = The number of references underlying a realitionship

The age score (AScore) of a gene is defined as the earliest publication age of a gene-disease relationship, as shown in Eq. (2).

$$AScore = \max_{1 \le i \le n} ArtilcePubAge_i$$
(2)

where n is the total number of references supporting a gene-disease relation, and

$$ArticlePubAge = current \ date - publication \ date + 1$$
(3)

2.3.2. Two enrichment metrics

n

We define a network significance score (SScore) of a node as Freeman's formalized node degree centrality (Freeman, 2012), as defined in Eq. (4).

$$SScore = \sum_{j}^{n} x_{j} = number of nodes directly linked with current node$$
(4)

where n is the total number of nodes, and j represents all other nodes within the network; x is the adjacency matrix, in which the cell x is 1 if the jth node is connected with the current node, other with is 0.

Given a disease is associated with a set of genetic pathways \Re we define the *PScore* for the gene as Eq. (5).

 $PScore_k$ = The number of pathways from \Re including the *k*th gene (5)

2.4. Validation using independent gene expression data

We hypothesized that significant AD related genes should contribute to distinguishing AD patients from healthy controls. To validate the effectiveness of the selected genes and the proposed metrics, we performed a Euclidean distance-based multivariate classification (Wang et al., 2015) on two independent gene expression data sets (NCBI GEO: GSE29378 and GSE28146), followed by a leave-one-out (LOO) cross validation, using the overall gene set and the sub-sets selected by different scores as tentative markers. In each run of LOO, gene expression data of one subject is used for testing and the rest for training. A permutation of 5000 runs was then conducted to test the hypothesis that a randomly selected gene set in same size can reach equal or higher classification accuracy (CR).

For dataset GSE29378, RNA expression profile of 64 subjects (AD case/control: 31/32) were obtained from 60 µm sections of frozen human hippocampus using scalpel dissection. Of the 1699 genes evaluated, 1605 were included in the database. For GSE28146, brain sections from a total of 30 subjects were analyzed (case/control: 22/8), with 1621 out of 1699 genes included.

3. Results

3.1. AD genes for evaluation

AD-Gene literature knowledge data analysis identified 1699 AD genes, supported by 27128 scientific articles (**AD_GD** \rightarrow **Related Genes** and **AD_GD** \rightarrow **References for Disease-Gene Relation**). A scalable genetic database, AD_GD, was developed through our study, which is online available at 'Bioinformatics Database' (http://database.gousinfo.com).

Of the 1699 genes, 644 (37.90%) have been reported with one reference (RScore = 1), 262 (15.42%) with 2, 181 (10.65%) with 3,

and 433 (25.49%) with more than 5 references, as shown in Supplementary Fig. S1(a). Publication date statistics of the 27128 supporting references are presented in Supplementary Fig. S1(b), with novel genes reported in each year (Supplementary Fig. S1(c)). To note, these articles have an average publication age of only 7.4 years, indicating that most of the articles were published in recent years.

3.2. Enrichment analysis results

PEA showed that, 1453 out of 1669 genes got significantly enriched within 151 AD candidate pathways/gene sets (p-values < 1e-15, q = 0.001 for FDR; $AD_GD \rightarrow Related Pathways$). Not surprisingly, aging (GO: 0016280; overlap 140 genes; q-value = 2.28E-84) is the top one enriched gene group. In addition, 11 pathways/groups were related to the neuronal system (619 unique genes) (Gong and Lippa, 2010; Marcello et al., 2012), 10 to cell growth and proliferation (441 unique genes) (Hohman et al., 2015; Li and Yao, 2013), 8 to cell apoptosis (452 unique genes) (Behl, 2000), 5 to protein phosphorylation (246 unique genes) (Shapiro et al., 1991), 2 to protein kinase (209 unique genes) (Martin et al., 2013), 3 to brain function/development (99 unique genes) (Llorente-Vizcaíno and Cejudo-Bolívar, 2001), and 2 to immune system (268 unique genes) (Heneka et al., 2001). Due to lack of space, we only present the top 10 pathways enriched in Table 1 (p-value \leq 2.6e–55, including 755 out of 1699 genes).

A SNEA was also performed to identify the pathogenic significance of the reported genes to other disorders which are potentially related to AD. Interestingly, besides neuropathic related diseases (e.g., Parkinson's disease and Schizophrenia) and some types of cancers (e.g., breast cancer), AD seems to share major gene overlaps with many blood related diseases (e.g., diabetes mellitus, obesity, type 2 diabetes, atherosclerosis, ischemia, myocardial infarction, hyperglycemia and stroke). The full list of 113 disease related sub-networks enriched with p-value < 1e-150 (q = 0.001 for FDR; 1625 out of 1669 genes enriched; see **AD_GD** \rightarrow **Related Diseases**).

3.3. GGI results

Fig. 2 presents the genetic network for AD, which was built through GGI analysis. The nodes of the network are 1453 out of 1669 genes that were enriched within the 151 AD target pathways. There were 319206 edges within the network, the weight of which are the numbers of pathways shared by the corresponding pair of nodes. The average node strength (sum of the number of genes directly connected) of the network was 219.69, and the node strength for the 246 unconnected genes was signed with 0.

Along with GGI, SScore and PScore were calculated for each gene (**AD_GD** \rightarrow **Related Genes**). The value of a PScore represents how many AD candidate pathways involve the gene, and an SScore represents how strong of a gene associated with other genes within the network.

3.4. Validation results

We hypothesized that, if our selected gene set (1669 genes) and the top genes selected by the proposed metric scores are significant to the pathogenesis of AD, they should lead to significant higher classification accuracy comparing to randomly selected genes. To test the hypothesis, classification and LOO cross validation were conducted on two independent public RNA expression dataset (NCBI GEO: GSE29378 and GSE28146), followed by a permutation test of 5000 runs.

For the LOO cross validation, the 1669 genes were first ranked by different metric scores, then the top n (n = 1, 2...) genes were

Table 1
Top 10 Molecular function pathways/groups enriched by 1669 genes reported

Pathway/gene set name	GO ID	# of Entities	Overlap	q-value	Jaccard similarity
Aging	0016280	254	140	2.28E-84	0.077
Neuronal cell body	0043025	466	172	4.7E-72	0.086
Neuron projection	0043005	378	153	2.8E-70	0.079
Response to lipopolysaccharide	0032496	252	126	4.9E-69	0.069
Response to hypoxia	0001666	259	127	8.11E-68	0.069
Response to organic cyclic compound	0014070	253	122	1.79E-64	0.067
Response to ethanol	0017036	161	94	4.21E-59	0.053
Negative regulation of apoptotic process	0006916	650	187	1.21E-56	0.086
Perinuclear region of cytoplasm	0048471	688	188	2.11E-55	0.085
Axon	0030424	318	125	2.56E-55	0.066

For each gene set, the p-value was calculated using Fisher's-Exact test against the hypothesis that a randomly selected gene group of same size (1669) can generate a same or higher overlap with the corresponding gene set (q = 0.001 for FDR correction). The Jaccard similarity (j_s) is a statistic used for comparing the similarity and diversity of sample sets, which is defined by $j_s(A, B) = \frac{A \cap B}{A \cup B}$ where A and B are two sample sets.

BLMH	MGLL	162	PIN1	BCL2L1	ANK3	SIRT4	DNMT38	6H3KBP1	CNTN4	GGA1	BECN1	TRPV1A	LOX5AP	FIS1	RG64	SNX3 r	m_Nirp1a	CDOD	VCAM1	SSTR4	PYCARD	CBL	PI4K2A	ANGPT 1	SLC30A1	ERN1	APBB1	LIMK2	HRH1	GAS7	S100B	AIRM1	GRIA1 C	SSTM3	KCNC4	FDPS	CD36
LPL	NRGAN	UQOR	C1SPPL2	B CNP	MMP1	\$100A12	BHLHB9	LAMP1	HMOX1	TPT1	RAB11A	NPPB	PICALM	PPP3R1	SREBF2	PCSK9	APBA1	DMD	SST	ATP7A	MARK1	FPR1	CLIN3	PPARA	ERG	C4A	SLC1A3H	A-ORB5	HTRA2	NPW1R	POK1 S	LOIBAS	СКВ	TPH2	ENPPAD	AMDECI	IRS2
PSEN1	IL21	ADAN	19 ARPP1	9SEMA3/	WNT5A	UBOLNI	VIM	MSRA	TMED 10	GAP43	FZD1	PRBX6	MARCKS	RHEB	CASP9	RPSA	GABRA2	MMP9	RAB6A	PUK2	CaSRIN	FRSF11	BLTA	CHN1	DUST	APOE	CALR	марк9	CYBB	ADEY1	GC I	AYO18A	MME N	APK1	TIMP1	IL33 HL	AORA
TPH1	UTP11		5D LPA	NOTCH	1 GIP	GAB2	PIP	LOAT	AGPS	HTR4	CRTR	GRIN2C	CCR5	LON2 N	DUFA13	BORBS2	SLOBA3	ABO	MEPIB	MEM106	BETS2	HTR2C	IL12B	PTK3	MGMT	HSPB6	DEFB1 (CHRM 158	ERPINC	MAPK8	TFAP4	AKAP9	ER300	AZU1	SNCG C	YRIP2	GJA1
A2M	ACO1	RIPK	1 RTN4R	L2PLA2G3	IGR2R	DAB1	LEP	MEF2C	CSNK1D	CRP	APP	PENK	AB6A2	CRWAB	HSPG2	TGFBR2	DUG3	PRKCA	P2RX4	PSPH	CYRIAI E	ELAVL4	PNP	KUC1	ENPEP	OPRL1	CD025C	ESR2	CYP2J2	PSMB8	VDAC1	ROEK2	GNAI1 (SSK3B	CNR2	REN I	NEU1
GS	GRASE	CASE	2SERPIN	FPLA2GE	B PEBP1	LWZ	DRD1	GES I	HOMERS	EREN2	CB388	PPBP	ABCG5	CXGL12	ECE2	MMP13	TREML2	MFI2	PTPRC	ARTN	MARK14	9/1/16	SRPK2	PRICE	IRAK1	HRAS	KAERN H	MGCR	SELP	AIF1	CFLAR	TUR2 P	RKAR2A	PEMT	ТРК1	TFAM	SOX2
APBB2	F12	TXN	C9orf7	2 GBPD	LRP2	F3	HPRT1	EIR4E	VEP	PDE78	IG#1R	EPHA1	CASP12	RARB	SEPP1	GLP1R	PTGER3	RAB3A	PNMT	HSPA2S	ERPINE2	NME1	PSEN2	FADS2	CTIGF	ACAT1	SYNJ1 T	GEBR1	MADD	MAPT	SIRT2	PPIB	WF1 S	SPHK2	VIMP H	юокз	ARC
PRNP	KCNA	AKTI	S1 APBA	FORAD	DRD2	REG1A	HCRTR2	NOS3	NAIP (CYRII9A1	RAIC1	OLIG2	PTK2B	MARK10	SNAP25	BAIK1	RUNX1	NLRP1	STXBP5	COX4I1	COX1	87	APH1B	AGER	ALB	P28X7	EPM2A	GPR84	THBS4	NTRK3	CD33	HIE1A	BAD F	TPN1 E	IF2AK2	EDD9	ILIA
HSPD1	S1PR5	P2R	-	SERPINE	EI GOA3	ABCG1	TIMP2	OPRD1	ACAT2	NLRP3	PTGDS	CDKIN1B	AHR	THRA	CR1	ST13	KUF2	PARK2	BIN1	PFKFB3	С₹Н	РМСН	ITGB1 H	HSD11B1	TP53 1	MNAT3	GFBP1	TOELIP	CRL1 C	DKN1A	AGP4	PSMB9N	FRSF10A	MMP2	COR1 C	TNND2	AR
PMP22	NELL2	SMPE	02 P2RY1	2 BAGE1	CRMP1	PTGS2	EEF1A1	NRC2	ADRBK1	ANXA5	PZP	SLC1A2	TRADD	TJP1	NOS2	IRS1	PABPC4	MMEL1	CRAT	GPX1	MPO /	ABCG2	CXCR24	DAMTS	CA8S4	CDK8	CXCR4 F	LA2G7	TLR10	HMGB1P	GRMC1	TNC	CASP1	TBK1 P	PP2R2B	GRAP (CLIC1
GDF11	LRAT	ODC	1 YWHA	G ALPL	FZD8	NOTCH4	NOTCH3	CDH23	LOKL1	CXGL10	CHIT1	DAKX	KCNIP4	SLC8A1	PGR	NCL	UBB	S061	SMPD1	MKL1	VIP H	HMOX2	OPRM1	CHRINA4	IE4 I	NFATC4	CX3CR1	CXER3 E	ONMT3A	CRH H	ISP90B1	IGF1 N	EUROD14	POC4	JUN	C D 14 G	RIN3A
MGEAS	NPC1	PINK	1 GATA	1 PTGS1	TBXA2R	CCL15	CNTN2	CD74	HTR3A	CTNNB1	STIM1	DRD5	GSK3A	GRC1	TNK1	PIK3R1	DVL2	SLC25A4	PRDX3	TGFB2	S100A8	EFMD2	LIPG	S100A7	SRF F	HOMER1	SLC5A7P	LA2G4A	HLA-A	ROEK1	IL34	RXRA	TH T	XNRD2	PWY	PON3 C	OPS5
NGF	DI02	STAT	TI FTMT	MFGE8	ICAM1	PRA	MAPK3	SREBF1	LINGO1	SPAST	CCND1	AKT1	CHGA	SCN8AF	APGEF	AHCY	EGR1	PVALB	DRD4	CD200R1	NEFH N	AP2K6	ADRB2	CASP5	SSTR5	CAPN2 E	CLIZL11	MMP7	BBC3	CASP8	KRAS	NGB	CCNG1 N	RXN3	PCNA F	RKCZ	HBG2
FZD5	PPP1C	A GRB	2HSD17B	100001	PEGAMI	1POU2F2	ITSN15	ERPING	INFKB1	PROS1	HGF	ACHE	MT2A	IL23R	GRM5	NPY	NFB2L2	FAAH	B2M	PRKRA	KIR6B	KL	ATG5	HCRTR1	BDKRB2	AP@C3	LDOC1	IL46	NR4H2	DNMT1	KCNQ1	PKN1	ITPKB S	LOB1A24	MBRA1	SRC (CONF
LOH1A	2ADRA2	A GAL	LRP6	PLA2GS	SNDUFV2	NAPG	AOIN1	SMAD3	CAPN1	ADAM10	HSPA8	SYK	MFN2	PRTN3	BCL2	CLPTM1	EEF2	MYC	PTCH1	SHANK1	MEF2A	SOID2 F	PPP2CB	ALOX15	GEI1	IL6R	CCL3 A	RNTL2	/P926AF	PPICBP	ARGC1	FABP3	LOX	GRPR (CRERI	LTF	NTF3
CEBPD	m_Sod	2 SIRT	1 PARK	7 RB1		STC1	PRND	INPPL1	РКМ	NTF4	EGFR	SYNPO	LRPAP1	NTSR1	NTRK2	KCNC1	DYRK1A	NDRG2	SLO1A1	SLC7A11	IP8K3 A	DAM17	SYT1	PLG	SRR	ABCG4	IGF2	ADAP1	AGRN	TFEP2	MT3	GSN			KCNIP3 R	B1CC1	ECE1
VAMP1	OTC	NLGN	I CXEL	I IL22	CA2	ABCC1	FGF205	SERTAD1	FMR1	APCS	TSPOS		3 LIF	CP	CES	TIMP3	HDC	SNCB	HM13	CDE42	CCL4	ICAM3	LAMA1	NAT8		ADORA1	CNTF	DNM2 S	SRD5A1	PRKOC	PEA15	GAS6	MARK2	CD40 0	LDN11	TLN1 H	INMT
CDC88	A TRH	ARPO	3 APC	ALDH2	PLVAP	LHCGR	RELN	SLC22A6	RTN3	PRKCI	DCN	DUG1 F	LARG2A	NR3C1	SFTPC	CARTPT	FOS	TYROBP	YWHAE	MAOB	PANK1 C	YP2D6	ASAH1F	PPP2R5C	FYN	CDH2	EDNRB	TSC2	HBB I	NR4H3	GHRH	LCK	BAX	ARIG2			HEXB
CDK1	SIGMAR	RI COK	5 GNB2L	1 MSR1	AOC3	CXEL8	CHI3L1	IL IRN	DDIT3	MZF1	IL1B	GUUL	PDE5A	TRPM2	PPRSCA	WASF1	CD40LG	LGI3	DOCK8	PTGER4	ADRA28	GUSBT	VFRSF1	AAMBP	NDUFS8	FOXO3	TLR5	NTN1 C	DPYSL2	PPIL2	GAPDH	CST3 S	H3GLB1	RBB4	PDED1	NGFR A	DIPOQ
RBFOX	TRAF2	DAP	A VLOLA	NYD88	ATP50	CYP7B1	NRG1	1138	DDIT4	RIPK2	XIAP	USF1	LRBK2	NDRG4	PLSCR1	SLC80A4	NB1	CSF2	S100A6	MSRB1	NRN1 N	ITNR1B	SAA1	HTR1A	HSPH1	GNB3	ITGB5 S	CARB1	ATF3	L10RA	FSHR S	LC2A12	OGG1	KBKB	HAMP	CTSD H	IDAC2
ID01	PTEN	MEO	K2 MT1F	PRICAC	A IL17A	SOD1	GLO1	TNPAIP1	HPX	ALAS1	NRGN	COL5	CTEG	CEU	IL10	FXN	TPM1	PAWR	HPSE	APH1A	BMP6	PDE4A	SNIX1	PL6G1	HDAC3	ARLEIPEN	EUROG1	PPY	SCN1A	RELA	NPTX1	FGA	CASP4	JAR2 F	KBP1A	DRP4 CN	TNAP1
CALB1	ATF4	ADN	P OPA1	ADCYAR	ALOOCO	DADRB1	AG02	ERCC2	NRG3	NINJ2	LRP1B	THY1	FUT1 S	SLC29A1	CREB1	MKI67	SYT9	PIKOCG	TREM2	SIM2	APOA1	PAK1 I	PRKCG	APEX1	YV1	MBP	PROCRIN	IMNAT2	CIQA	LIPA C	YP11At	HANK3	PCSK1 S	LC8A2	CXCL9 C	RIN2A C	CNE1
THED	LMAP18	CASE		ARSA	PRKACE	B NEFM	CREBBP	ND2	CHRM2	SHANK2	CAST	BMP4	PER2	RTN4R	CTSS	HMBS	IL45	DDOST	VDR	AHSG	AGTR2 S	LOB0A3	CAV2 F	PRKAR28	FKBP4	COMT	MX1	TACR3	POLG I	NCSTN	AS¥N2	GAD2	SFTPD	MIOR	APBB3	TNF C	EBPB
LEPR	TRAFE	TNFRS	F21 RYR3	RARRES	BOL25A	1KEAP1	RHO	NBIL1 C	CYPE7A	PPP2R1A	TSHZ3	MRC1	CCKAR	GPR3	FN1	IOMER2	NCF1	SLC11A1	MTHER	VASP	PLD3	ABCB1	MOK	CHEK2	TGM2	LHB I	MAP2K1	EGF I	NCAM1	SMAD2	ABEC9	MTR	WWC1 N	DUFS3	ALBOA P	PP2CA U	INC5C
DBI	DVL1	FBXC	2 SIRTS	CEBPA	CUK1	HGN1	IL1RL1	PC	CIB1 S	SPARCL1	INSR	MAIOA	FEN1	CHMP2B	BPTF	PTGES2	GRM1	TAP2	DIPORI	NRXN2	F2 /	APOC2	NAE1	ANXA1	CYP3A4	PPID	HMGA1	CSTB	SV2A	EPHB2	GRIA2	OPTN C	ACYBP H	KAT2B	OPRK1	PRL A	LOX5
TRPI	CDK2	GPH	N KLK8	TGFB1	NMNATO	DKSRAP	2IF2AK4	PTGES	HT82A	GRIN1	IL13	CHMP2A	DKK1	IGFBP7	CSK1N	PNRLA6	KCNJ10	CLOCK	KIAA0196	G C4B	SPHK1	GRM2	STMN1	PDIA6	CSF1R	CAV3	SDC1	PLAUR	PARD3 S	TABIAN	TP6V1B2	ARPC2	IGFBP3 F	ROP1	GSTP1	CCL2 MA	PK8IP1
SUMO	P4HB	BTNL	2 CHEK	1 BCL2L2	2 S100A9	APBA3	APOD	FFAR1	DAPK3	PUK1	DNM1L	THRB		DORA2A	CASP3	GRIN2B	CH25H	MED12	HSF1	SYNJ2	STK11	TFRC	PROX1	RYR2	PAK3	TLR9	REST	CETP	SGMS2 S	SH3GL2	TIMP4	F2RL1	GPR6	CTEL	CB4 Z	NR703 H	SPA1L
ILEST	SORLI	CYP2	C9 AVP	PDPK1	SLO17A	6LRRTM3	PUD2	BSG	KITLG	PON1	CL9TN1	MMP12	INS	RGMA	MDM4	DRD3	KLKB1	SET	ACTA2	WNT3A	USF2	MEF	RTN4	SMPD3	NRD1	FGF1	RPS6	STAT3	EPHX1	SOAT1	FAS	VEGFAH	ISD17B1	WASL	DCX	TM2B	TON2
RAB7A	DUL1	NRIP	APOC	1 ENO1	ASS1	PANK2	ITPR1	DBN1	C5AR1	CAMK1	MMP3	IREB2	TUR3	NLRC4	SELL	ABCA5	NR3C2	DDAH1	SIRPB1	NAT8B	PPP3CBS	TARD3	KCND2	TFR2	XBP1	TREB /	ABCA12T	NFSF10	NRF1 S	SLC6A2	FLOT1 F	PPICC	SNK2A2	STUB1	CYCS C	DC25BK	AA1033
LRP5	BGN	ENTP	D2SLC30/	7 CDK4	EDN1	SP1	MYOCD	GALR2	MRE11A	CTNND1	KCNB1	GRB14	CDK7	RGN	CD5	BCHE	CD2AP	SPAG16	CTF1	CA3	IGFBP2	NUOT1	AGTR1	CAMK2A	SRP1	SH2B1	APOB N	FATC1	MSN SE	ERPINA1	OGT I	DHCR24	SGPL1 7	RPC4	PLTP S	PTBN1	IFNG
APOAS	ANP32	A CCR		2 STAR	KCNN2	MOBP	GDNF	11/24	RORA	COX5A	CONC	ABAT	ESR1 (CYPEC19	CRTC1	SPPL2A	CBS	HORT	ABCA1	DNAJB6	PPR2R4	C4BPA	MPL	RGS2	GSTM1	TLR8	BACE2	LOLR	CD59	AGP1	PLP1	SYN1		IDAC6	SIRT6 E	IRCA1	ADK
CUTA	SELM	BAJA	P2 OGLN	CHAT	PPARG	РТК2	DDB1	PROC	FGF2	SORBS3	SNK12	SHC3 (CYP46A1	MSMO1	ENO2	CGR2	MARK4	TBP	PDE88	TF S	SLO47A7	GSR	SRI	GRIK2	KISS1	ANK1 F	SENEN	ARNTL	TACR2	HDAC1	NEFL	AGT	GMFB G	SNRH1	D8H /	PEP2 X	REC1
TP73	IAPP	DYR	2 DTNBP	1 CORT	PIK8C3	ARRB2	STM1A	CL9TN3	GRIA3	LRP1	EDNRA	CSF3	PDE9A	BAG5	PDIA3	HADH	HRH3	CDK5R1	LRP8	GRK5	HTR6	GU62	SPI1	UCHL1	GRN	RHOA	PLD1	INA	HSPB1	HSPA5	TOF3	PID1	NQO2	TPP1	CSPG4 F	ARP1	OXT
COR2	B ERBB2	wwm	R1 MT1G	OLR1	KIF11	9/2/16	HSPA9	YWHAZE	RPS6KB1	GBA	RAF1	HDAC5	CASK	CX3CL1	ANG	BGLAP	CDC25A	DUG4	MAP2	P2RY2	M6PR	GAD1	TRPA1	SORCS3	PRDX2	PDGFB (DNAUC5	KLK6 S	SPTLC1 C	SRIN2D	E2F1	HFE	SORT1	HRH2 H	CNU11	RAK4 S	LC2A2
SYPL1	SELE	PLEC	01 GFRA	1 IL12A	TIAF1	CHRNA3	S100A1	CAT	NRXN1	ATP7B	CHRINA7	ERO	TRIML2	P2RX6	TRIP4	CARD8	H2AFX	FPR2	R6	AK1	SLG4A1	ABL1	CSF1	9/3/16 A	LDH5A1	CCL8	HP	GDF15H	LA-ORB1	WIHAQ	GSTO1	ONMBP	POLB C	OL6A1	SNCA S	PTLC2 E	PHA4
CALCA	SLC40A	1 118	HTT	KLK1	RAB29	ATF2	PRECO	NFRSF1	BUSP22	PTGER1	NTSR2	NFKBIA	LMNA	CNIR1	UBE2I	NPPA	CTSB	CAEM3	NGF2	IL1R2	MMP14	PDE4D F	RANBP9	NES	SHC1	GNAS A	DAM12	FASLGN	AP3K1N	EUROGE	APGEFS	RUNE2	HDAC8SE		LRP10 E	F2AK3 C	CD244
GH1	ABCA7	OCIAL	D2 SLC2A	3 SP3	UGCG	SIAH1	HSPA1A	SPOCK1	LGMN	EEF2K	GSAP	FGF9	FERMT2	HK1	KAT5	FURIN	LIPC	LIMK1	GPX4	FTH1	CFH C	DKN2A	EPG5	F2RL3 C	YSLTR	MARSKS	SLOGA4	VDAC2C	NTNAP2	COR3 C	HRNB2	LAMP2	VP\$35 H	NRNPKS	SLO2A1	NOS1 H	ISPB8
SLC6A3	SYP	IDE	SNX2	PRKARI	BPVRL2	CIZ1	VGF	ARIG1	ATP6A1	CASP6	HTR IBT	MEM163	6H3RF3	RCAN1	DKRB1	IL11	TXN2	PER3	NPTX2	AGE	HDAC100	COLIA1	CONB1	ZFMX3	PON2	THBS1	APLP1 A	LOX12	CAV1	PDYN	DNM1	PPARD	PLAU F	FNA1	9/4/16	CIR :	SETX
TLR4	TNMD	PLXN	A4 TKT	wwox	STMN2	RBTN	SMOC2	ENSA																													

Fig. 2. Gene-gene interaction network for AD. The network contains 1453 out of 1669 genes AD target genes that enriched within the 151 AD target pathways. The weight of an edge between two nodes is the number of pathways shared by both nodes. The larger the size of a node, the larger the number of AD candidate pathways including the gene (high PScore); the brighter the color, the larger number of AD candidate genes associated with gene (high SScore). 216 out of 1669 genes were not included in the network as they were not enriched within the top 151 AD candidate pathways.

Table 2

Permutation test on top genes corresponding to highest CRs.

Data sets	Items	RScore	AScore	PScore	SScore	All genes
GSE29378 (31/32)	Max CR (%) #Gene	80.95 66	79.37 109	82.54 20	74.60 57	60.32 1605
	p-value	0.0022	0.004	0.0004	0.03	0.96
GSE28146 (22/8)	Max CR (%) #Gene	80.00 18	90.00 102	83.33 25	90.00 22	73.33 1621
	p-value	0.017	0.001	0.0074	0.0002	0.90

used as input variables for classification and LOO cross validation. Table 2 and Fig. 3 presents the results with the maximum classification ratios (CRs) marked at the position of corresponding number of genes.

From Fig. 3 we see that the top genes selected by different scores (in descending order) can lead to the highest classification accuracies, which are significantly higher than the average CRs of randomly selected gene set in same size, while adding more genes with lower score may not necessarily lead to improved CRs.

3.5. Cross metrics analysis

Results from AD case/control classification (Table 2 and Fig. 3) showed that, the top genes selected using each of the four proposed metrics led to significantly higher CR compared to randomly selected gene sets, demonstrating the effectiveness of the proposed metrics. Therefore, it is worthy to study the overlaps among these top genes. Cross metrics analysis of the top 5% (86 genes, corresponding to the number of genes reported this years, 2016) of 1699 genes selected using different scores showed that (see Veen



Fig. 3. Validation of different metrics through a LOO cross validation. (a) Results from GSE29378; (b) Results from GSE28146. Mean of CRs by randomly selected genes are displayed in a dash-grey line (Legend: Random). The maximum CR by different metrics are presented at the corresponding positions.

diagram at Supplementary Fig. S2), there was strong overlap between PScore group and SScore group (63/86). Among these 63 genes, 21 were also identified to be within RScore group, including TGFB1, CTNNB1, APP, IL1B, PSEN1, PTGS2, IL6, VEGFA, SOD1, AKT1, CDK5, TNF, GSK3B, TP53, CCL2, BDNF, NGF, IGF1, SIRT1, AGER and TLR4, with RScore = 542 ± 16 refrences, PScore = 29 ± 6 pathways, SScore = 969 ± 85 connected genes. Network analysis using Pathway Studio showed that, these 21 genes also demonstrated strong correlation with the top disorders that are linked to AD (Fig. 4, highlighted in red). The genes related to these diseases present significant overlap with these genes linked to AD (see **AD_GD** \rightarrow **Related Diseases**). On the other hand, only one gene, ITGB1, was identified to be the overlap of AScore, PScore and SScore groups (Fig. 4, highlighted in yellow), which also linked to several other diseases (e.g., diabetes, stroke and breast cancer) that are genetically linked to AD.



Fig. 4. AD genes selected by cross metrics analysis and their relation with other diseases. The 21 genes that were overlap in RScore, PScore and SScore groups are highlighted in green; Gene ITGB1 that was the overlap in AScore, PScore and SScore groups and is highlighted in yellow. The network was built using the 'network building' module of Pathway Studio.

4. Discussion

Recent studies proposed over a thousand of AD risk genes with dozens of novel targets identified each year. However, over half of these AD target genes were lack of replication and results were not always consistent, posing an increasing need of a systematic evaluate approach to test the significance of these genes as a network to AD. In this study, we integrated large scale literature knowledge data, gene expression data and related pathways and disease-sub networks to evaluate 1669 AD candidate genes. Four metric scores have been proposed and validated. A scalable genetic database, AD_GD, was developed through our study, which is online available at 'Bioinformatics Database' (http://database.gousinfo.com).

PEA results showed that most genes within the network (1453 out of 1699) were significantly enriched (FDR corrected p-value < 1e–15) in the pathways previously implicated with AD, including pathways/groups related to aging, neuronal system pathways, cell growth and proliferation, cell apoptosis, protein phosphorylation brain function/development and immune system (Gong and Lippa, 2010; Marcello et al., 2012; Hohman et al., 2015; Li and Yao, 2013; Behl, 2000; Shapiro et al., 1991; Martin et al., 2013; Llorente-Vizcaíno and Cejudo-Bolívar, 2001; Heneka et al., 2001). These observations support the hypothesis that, most of the AD target genes are functionally linked to each other and play roles within multiple pathways associated with AD.

In addition to PEA, we performed a SNEA, which can provide high levels of confidence when interpreting experimentallyderived genetic data against the background of previously published results (http://pathwaystudio.gousinfo.com/SNEA.pdf). SNEA results demonstrated that over 95% (1625) of the 1669 ADgenes were as well identified as causal genes for other disorders that were linked to AD (**AD_GD** \rightarrow **Related Diseases**).

For a quantitative measure of the significance of these 1669 AD candidate genes, we proposed 4 metrics: (1) publication frequency (RScore), (2) novelties (AScore), (3) Number of associated AD candidate pathways (PScore), and (4) Network centrality (SScore). We hypothesized that if a gene satisfies one or more of the following conditions, it has high possibility to be linked to AD: The gene is frequently identified by independent studies to be linked to AD (high RScore), plays roles within multiple AD pathways (high PScore), and is functionally linked to multiple AD genes (high SScore).

The effectiveness of our 4 proposed metrics were supported by the AD case/control classification study using two independent gene expression data sets (GSE29378 and GSE28146). Results of the LOO cross validation and permutation process showed that, the top genes by the 4 proposed metrics can lead to significantly higher classification ratio than using randomly selected gene sets (Table 2 and Fig. 3). While using the identified gene set as a whole (1605 and 1621 out of 1669 for GSE29378 and GSE28146, respectively) showed no significant efficiency in terms of AD prediction (permutation p-value > 0.9; see Table 2), suggesting the necessity of using our network metrics for further analysis of the candidate AD genes when dealing with specific experiment data. Notably, for each score, the number of top genes corresponded to the maximum CRs for the two data sets were different. This may reflect the group-wise variation in terms of sample size (63 vs. 30) and clinical parameter dissimilarities (e.g., age, gender). The difference may be also caused by the unique variation of different subjects' genome in case of AD.

Cross metrics analysis showed that 21 genes were overlapped within RScore, SScore and PScore groups (Fig. 4, highlighted in green). These genes were frequently identified by different studies to be linked to AD (RScore = 542 ± 16 references), play roles with in multiple AD candidate pathways (PScore = 29 ± 6 pathways), and

demonstrate strong network centrality (SScore = 969 ± 85 direct gene connections). Therefore, our results suggest that they are among the top AD risk genes that likely pose biological significance with the disease. As a matter of fact, these genes were also identified to play role within many other disorders that were linked to AD, such as diabetes mellitus, obesity, Parkinson's disease, Schizophrenia, and breast cancer (Fig. 4). These results support the effectiveness of the proposed metric scores in the identification of top genes for AD.

Additionally, there was one newly reported gene, ITGB1 (AScore = 1), also demonstrated high SScore and PScore (Fig. 4, highlighted in yellow). Although ITGB1 were not frequently replicated in their association with AD (RScore = 1 reference), and presented less relationships with other AD related mental disorders, it demonstrated high interaction with other genes within the genetic network (SScore = 1050 directly connected genes) and play role within multiple pathways implicated with AD (PScore = 28 pathways). Therefore, our study suggests that it may be worthy of further study. In fact, activation of integrin *β*1 (ITGB1) has been reported to regulate the synthesis of enterovirus 71-induced and NADPH oxidase-driven reactive oxygen species (ROS) (Tung et al., 2011), which are closely involved in pathogeneses of AD (Aliev et al., 2003). It also revealed that adhesion of HeLa cells to $\beta 1$ integrin clustering can increase the release of arachidonic acid (Xu and Clark, 1997) that has therapeutic functions against AD (Huang and Cheung, 2011). These findings support our observation that ITGB1 may play roles for pathogenic development of AD, demonstrating the effectiveness of our proposed PScore and SScore in identifying novel genes for the disease. To note that, although in this study we evaluated 1669 known AD candidate genes acquired from ResNet database, which already received literature support for their association with AD, our proposed PScore and SScore can be applied to any given genes and therefore could be used for evaluation and discovery of novel target genes for AD.

The genetic database built through our approach, namely AD_GD, is scalable and can be automatically updated using the computational workflow proposed in this study. Any novel AD-gene relationships can be added to update the database. Moreover, further network analysis with more experiment data may extract additional meaningful features that can be added into our proposed system to gain improved evaluation of existing and/or novel AD genes.

To our knowledge, this is the first study integrating large scale literature knowledge data, experiment data and related pathway/ network data for a systematical evaluation of AD candidate genes. The computational biology approach of this study provides a comprehensive weighted genetic network and genetic database for AD, which may help in the evaluation and prioritization of AD genes for further study in the field.

Acknowledgement

We would like to thank Mckenzie Ritter for her contribution to the English writing of this manuscript. This work was partly funded by National Nature Science Foundation of China (81130024 and 81630030, Tao Li), National Key Technology R & D Program of the Ministry of Science and Technology of China during the 12th Five-Year Plan (2012BAI01B06, Tao Li), National Key Research and Development Program of the Ministry of Science and Technology of China (2016YFC0904300, Tao Li), Science and Technology Department of Sichuan Province of China (2017SZ0049, Zhe Li; 2018ZR0216, Zhenzhen Xiong), and Health and Family Planning Commission of Sichuan Province of China (17PJ080, Zhe Li). The Fundamental Research Funds for The Central Universities of China (2017SCU11072, Zhe Li).

Conflict of interest

The author HC is with Elsevier Inc., the company that owns the software Pathway Studio used in this study.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.sjbs.2018.05.019.

References

- Aliev, G., Obrenovich, M.E., Smith, M.A., Perry, G., 2003. Hypoperfusion, mitochondria failure, oxidative stress, and alzheimer disease. J. Biomed. Biotechnol. 2003 (3), 162–163.
- Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., Jones, E., 2011. Alzheimer's disease. Lancet 377 (9770), 1019–1031.
- Behl, C., 2000. Apoptosis and Alzheimer's disease. J. Neural Transm. (Vienna) 107 (11), 1325-1344.
- Burns, A., Iliffe, S., 2009. Alzheimer's disease. BMJ 338, b158.
- Cenini, G., Sultana, R., Memo, M., Butterfield, D.A., 2008. Effects of oxidative and nitrosative stress in brain on p53 proapoptotic protein in amnestic mild cognitive impairment and Alzheimer disease. Free Radic. Biol. Med. 45 (1), 81– 85.
- Cole, A.R., Astell, A., Green, C., Sutherland, C., 2007. Molecular connexions between dementia and diabetes. Neurosci. Biobehav. Rev. 31 (7), 1046–1063.
- Daraselia, N., Yuryev, A., Egorov, S., Novichkova, S., Nikitin, A., Mazo, I., 2004. Extracting human protein interactions from MEDLINE using a full-sentence parser. Bioinformatics 20, 604–611.
- Freeman, L.C., 2012. Centrality in social networks conceptual clarification. Soc. Netw. 1, 215–239.
- Gong, Y., Lippa, C.F., 2010. Review: disruption of the postsynaptic density in Alzheimer's disease and other neurodegenerative dementias. Am J Alzheimers Dis Other Demen. 25 (7), 547–555.
- Heneka M.T., Golenbock D.T., Latz E., 2001. Innate immunity in Alzheimer's disease. Nat Immunol. 2015;16(3):229–236.
- Hohman, T.J., Bell, S.P., Jefferson, A.L., 2015. The role of vascular endothelial growth factor in neurodegeneration and cognitive decline: exploring interactions with biomarkers of Alzheimer disease. JAMA Neurol. 72 (5), 520–529.
- Huang, J.J., Cheung, P.C., 2011. +UVA treatment increases the degree of unsaturation in microalgal fatty acids and total carotenoid content in *Nitzschia closterium* (Bacillariophyceae) and *Isochrysis zhangjiangensis* (Chrysophyceae). Food Chem. 129 (3), 783–791.
- Koedam, E.L., van der Vlies, A.E., van der Flier, W.M., Verwey, N.A., Koene, T., Scheltens, P., et al., 2013. Cognitive correlates of cerebrospinal fluid

biomarkers in frontotemporal dementia. Alzheimers Dement. 9 (3), 269-275.

- Lanni, C., Racchi, M., Memo, M., Govoni, S., Uberti, D., 2012. p53 at the crossroads between cancer and neurodegeneration. Free Radic. Biol. Med. 52 (9), 1727– 1733.
- Li, J.S., Yao, Z.X., 2013. Modulation of FGF receptor signaling as an intervention and potential therapy for myelin breakdown in Alzheimer's disease. Med. Hypotheses 80 (4), 341–344.
- Llorente-Vizcaíno, A., Cejudo-Bolívar, J.C., 2001. Memories and Alzheimer's disease. Rev. Neurol. 32 (12), 1163–1172.
- Lorenzi, P.L., Claerhout, S., Mills, G.B., Weinstein, J.N., 2014. A curated census of autophagy-modulating proteins and small molecules: candidate targets for cancer therapy. Autophagy 10, 1316–1326.
- Luo, J., Lin, A.H., Masliah, E., Wyss-Coray, T., 2006. Bioluminescence imaging of Smad signaling in living mice shows correlation with excitotoxic neurodegeneration. Proc. Natl. Acad. Sci. USA 103 (48), 18326–18331.
- Marcello, E., Epis, R., Saraceno, C., Di Luca, M., 2012. Synaptic dysfunction in Alzheimer's disease. Adv. Exp. Med. Biol. 970, 573–601.
- Martin, L., Latypova, X., Wilson, C.M., Magnaudeix, A., Perrin, M.L., Yardin, C., et al., 2013. Tau protein kinases: involvement in Alzheimer's disease. Ageing Res. Rev. 12 (1), 289–309.
- Mendez, M.F., 2012. Early-onset Alzheimer's disease: nonamnestic subtypes and type 2 AD. Arch. Med. Res. 43 (8), 677–685.
- Ramalho, R.M., Viana, R.J., Low, W.C., Steer, C.J., Rodrigues, C.M., 2008. Bile acids and apoptosis modulation: an emerging role in experimental Alzheimer's disease. Trends Mol. Med. 14 (2), 54–62.
- Shapiro, I.P., Masliah, E., Saitoh, T., 1991. Altered protein tyrosine phosphorylation in Alzheimer's disease. J. Neurochem. 56 (4), 1154–1162.
- Sivachenko, A.Y., Yuryev, A., Daraselia, N., Mazo, I., 2007. Molecular networks in microarray analysis. J. Bioinform. Comput. Biol. 5, 429–456.
- Tung, W.H., Hsieh, H.L., Lee, I.T., Yang, C.M., 2011. Enterovirus 71 induces integrin β1/EGFR-Rac1-dependent oxidative stress in SK-N-SH cells: role of HO-1/CO in viral replication. J. Cell. Physiol. 226 (12), 3316–3329.
- Wang, J., Cao, H., Liao, Y., Liu, W., Tan, L., Tang, Y., et al., 2015. Three dysconnectivity patterns in treatment-resistant schizophrenia patients and their unaffected siblings. Neuroimage Clin. 8, 95–103.
- Wang, J., Pang, T., Hafko, R., Benicky, J., Sanchez-Lemus, E., Saavedra, J.M., 2014. Telmisartan ameliorates glutamate-induced neurotoxicity: roles of AT(1) receptor blockade and PPARγ activation. Neuropharmacology 79, 249–261.
- World Health Organization. Dementia Fact sheet No. 362. March 2015. Archived from the original on 18 March 2015. Retrieved 13 January 2016.
- Xu, J., Clark, R.A., 1997. A three-dimensional collagen lattice induces protein kinase C-zeta activity: role in alpha2 integrin and collagenase mRNA expression. J. Cell Biol. 136 (2), 473–483.
- Xu, Z.P., Yang, S.L., Zhao, S., Zheng, C.H., Li, H.H., Zhang, Y., et al., 2016. Biomarkers for early diagnostic of mild cognitive impairment in type-2 diabetes patients: a multicentre, retrospective, nested case-control study. EBioMedicine 5, 105–113.
- Yoo, H.J., Cho, I.H., Park, M., Cho, E., Cho, S.C., Kim, B.N., et al., 2008. Association between PTGS2 polymorphism and autism spectrum disorders in Korean trios. Neurosci. Res. 62 (1), 66–69.