

RESEARCH ARTICLE

Open Access



Genome-wide network-based pathway analysis of CSF t-tau/A β ₁₋₄₂ ratio in the ADNI cohort

Wang Cong¹, Xianglian Meng^{1,2}, Jin Li¹, Qiushi Zhang^{1,3}, Feng Chen¹, Wenjie Liu¹, Ying Wang¹, Sipu Cheng¹, Xiaohui Yao^{4,5}, Jingwen Yan^{4,5,6}, Sungeun Kim^{4,6}, Andrew J. Saykin^{4,6}, Hong Liang^{1*}, Li Shen^{4,5*} and for the Alzheimer's Disease Neuroimaging Initiative

Abstract

Background: The cerebrospinal fluid (CSF) levels of total tau (t-tau) and A β ₁₋₄₂ are potential early diagnostic markers for probable Alzheimer's disease (AD). The influence of genetic variation on these CSF biomarkers has been investigated in candidate or genome-wide association studies (GWAS). However, the investigation of statistically modest associations in GWAS in the context of biological networks is still an under-explored topic in AD studies. The main objective of this study is to gain further biological insights via the integration of statistical gene associations in AD with physical protein interaction networks.

Results: The CSF and genotyping data of 843 study subjects (199 CN, 85 SMC, 239 EMCI, 207 LMCI, 113 AD) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) were analyzed. PLINK was used to perform GWAS on the t-tau/A β ₁₋₄₂ ratio using quality controlled genotype data, including 563,980 single nucleotide polymorphisms (SNPs), with age, sex and diagnosis as covariates. Gene-level *p*-values were obtained by VEGAS2. Genes with *p*-value ≤ 0.05 were mapped on to a protein-protein interaction (PPI) network (9,617 nodes, 39,240 edges, from the HPRD Database). We integrated a consensus model strategy into the iPINBPA network analysis framework, and named it as CM-iPINBPA. Four consensus modules (CMs) were discovered by CM-iPINBPA, and were functionally annotated using the pathway analysis tool Enrichr. The intersection of four CMs forms a common subnetwork of 29 genes, including those related to tau phosphorylation (*GSK3B*, *SUMO1*, *AKAP5*, *CALM1* and *DLG4*), amyloid beta production (*CASP8*, *PIK3R1*, *PPA1*, *PARP1*, *CSNK2A1*, *NGFR*, and *RHOA*), and AD (*BCL3*, *CFLAR*, *SMAD1*, and *HIF1A*).

Conclusions: This study coupled a consensus module (CM) strategy with the iPINBPA network analysis framework, and applied it to the GWAS of CSF t-tau/A β ₁₋₄₂ ratio in an AD study. The genome-wide network analysis yielded 4 enriched CMs that share not only genes related to tau phosphorylation or amyloid beta production but also multiple genes enriching several KEGG pathways such as Alzheimer's disease, colorectal cancer, gliomas, renal cell carcinoma, Huntington's disease, and others. This study demonstrated that integration of gene-level associations with CMs could yield statistically significant findings to offer valuable biological insights (e.g., functional interaction among the protein products of these genes) and suggest high confidence candidates for subsequent analyses.

Keywords: Alzheimer's disease, CSF biomarker, t-tau/A β ₁₋₄₂ ratio, Network analysis, Pathway analysis, Consensus module, CM-iPINBPA

* Correspondence: lh@hrbeu.edu.cn; shenli@iu.edu

¹College of Automation, Harbin Engineering University, 145 Nantong Street, BLDG 61-5029, Harbin 150001, China

⁴Department of Radiology and Imaging Sciences, Indiana University School of Medicine, 355 W 16th St, Suite 4100, Indianapolis, IN 46202, USA

Full list of author information is available at the end of the article



Background

Alzheimer's disease (AD) is a neurodegenerative disease and the most common form of dementia. Although its etiology is not completely understood, a genetic component of susceptibility to AD has been shown in the literature [1–6]. Cerebrospinal fluid (CSF) studies [7–10] have also been conducted in AD to identify differential biomarkers. Given that one hallmark of AD pathology is a cerebral accumulation of amyloid- β 1–42 peptide ($A\beta_{1-42}$) in amyloid plaques, the $A\beta_{1-42}$ level has been shown markedly reduced in CSF. In addition, the total tau (t-tau) protein level has been shown significantly elevated in the CSF of AD patients. As a result, the CSF t-tau/ $A\beta_{1-42}$ ratio has also been studied as a biomarker for differentiating AD from normal older adults [5, 11–13].

With the recent advances in high-throughput genotyping technologies, Genome-Wide Association Studies (GWAS) have been applied to the analysis of CSF biomarkers (e.g., [13, 14]) to identify relevant genetic markers, such as Single Nucleotide Polymorphisms (SNPs). While most studies examined genetic associations with CSF biomarkers at the individual SNP or gene level, mining higher level genetic associations using biological interaction networks is still an under-explored topic for the CSF biomarker studies in AD. Recently, many studies in other domains have demonstrated that integrative analyses of GWAS data and protein-protein interaction (PPI) networks can provide valuable biological insights. Some methods have been proposed to identify subnetworks enriched by GWAS results [15–19]. One tool is iPINBPA [16, 20–23], which is based on the fact that the genes identified in GWAS are more likely to physically interact as well as to belong to the same or related pathways.

With these observations, in this work, we performed a genome-wide network-based pathway analysis for CSF studies in AD. We analyzed an AD cohort from Alzheimer's Disease Neuroimaging Initiative (ADNI), used the CSF biomarker t-tau/ $A\beta_{1-42}$ ratio as the test phenotype or quantitative trait (QT), downloaded the PPI network from the Human Protein Reference Database (HPRD) (<http://www.hprd.org/>), and applied the iPINBPA analysis to the GWAS findings of the CSF t-tau/ $A\beta_{1-42}$ ratio. Our goal was to search for subnetworks or network modules enriched by the CSF GWAS findings, which may offer valuable biological insights and suggest high confidence candidates for subsequent analyses.

Methods

Figure 1 illustrates the work-flow of this study. The genotyping and CSF data were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. GWAS of the CSF QT was performed using the PLINK software [24]. This resulted in 563,980 SNPs

with associated p -values, which were then assigned to 22,179 genes. The gene assignment and gene-based p -values were calculated using the VEGAS2 software [25]. The nominally significant genes (i.e., gene-based p -values ≤ 0.05) were mapped onto the HPRD PPI network [26, 27] and analyzed using the iPINBPA method in order to identify the enriched subnetworks. The Enrichr pathway analysis tool [28] was applied to functionally annotate the subnetwork.

Subjects

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). One goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. For up-to-date information, see <http://adni.loni.usc.edu/>. Appropriate Institutional Review Boards approval occurred at each ADNI site and informed consent was obtained from each participant or authorized representative.

In this study, our analyses were concentrated on 843 ADNI subjects whose genotyping data (after quality control described below) and the baseline CSF biomarker data including t-tau and $A\beta_{1-42}$ were both available. This sample included 199 cognitively normal (CN) subjects, 85 subjects with significant memory concern (SMC), 239 subjects with early mild cognitive impairment (EMCI), 207 subjects with late mild cognitive impairment (LMCI), and 113 subjects with Alzheimer's disease (AD). Table 1 shows the demographic and clinical characteristics of these participants at the baseline, where the characteristics were analyzed with the statistical software IBM SPSS [29] Version 2 for differences across diagnostic groups using one-way analysis of variance (ANOVA) or Chi-square test.

CSF Biomarker Measurement as Quantitative Trait

The amyloid- β 1–42 peptide ($A\beta_{1-42}$) and total tau (t-tau) measured in the baseline CSF samples of the participants were downloaded from the ADNI database. The t-tau/ $A\beta_{1-42}$ ratio was computed and used as the quantitative trait in the GWAS.

Genotyping Data and Quality Control

The genotyping data of the participants were collected using either the Illumina 2.5 M array (a byproduct of the ADNI whole genome sequencing sample) or the Illumina OmniQuad array. For the present analyses, single nucleotide polymorphism (SNP) markers that were present on both arrays were included [6, 30, 31].

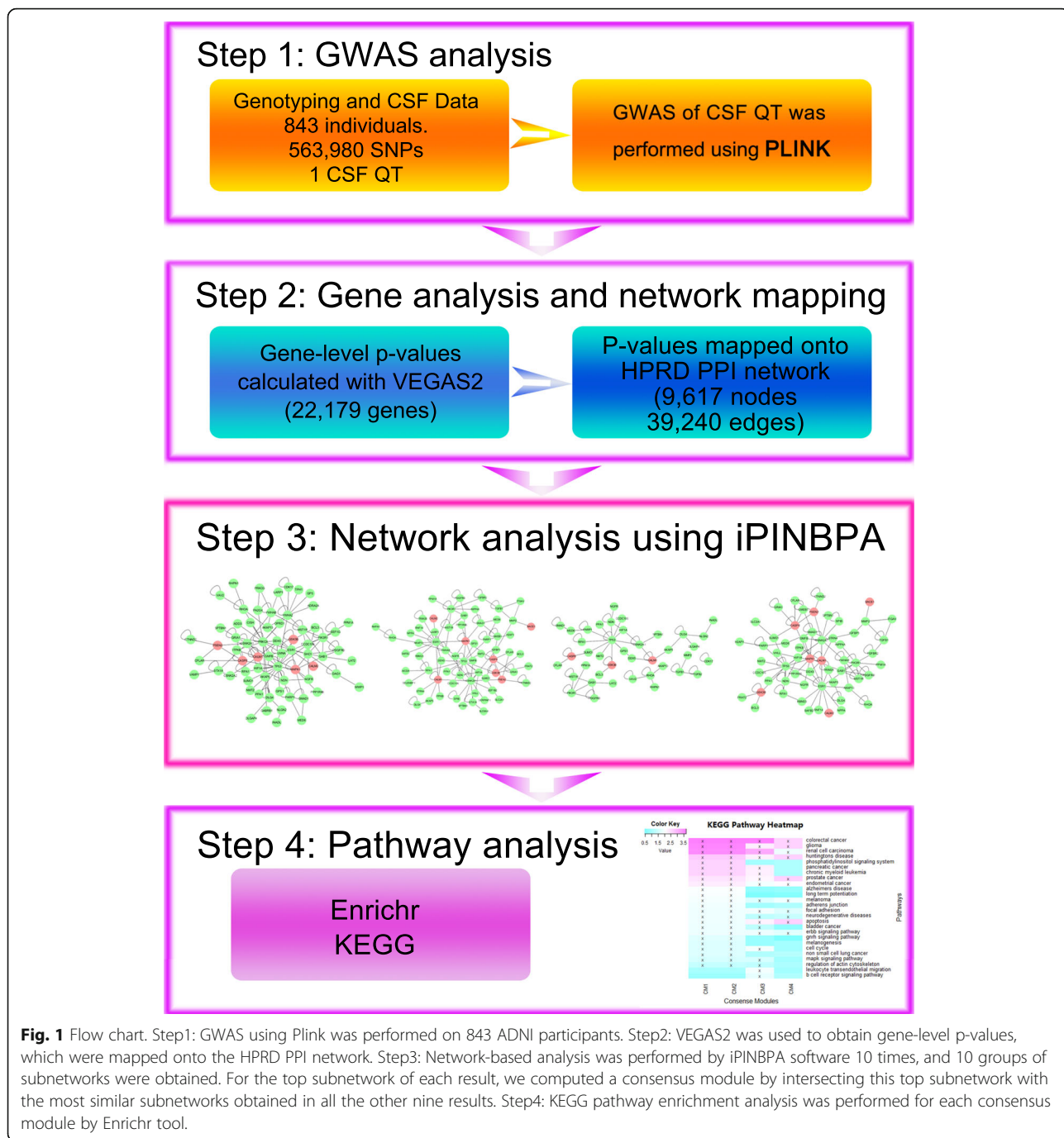


Fig. 1 Flow chart. Step1: GWAS using Plink was performed on 843 ADNI participants. Step2: VEGAS2 was used to obtain gene-level p-values, which were mapped onto the HPRD PPI network. Step3: Network-based analysis was performed by iPINBPA software 10 times, and 10 groups of subnetworks were obtained. For the top subnetwork of each result, we computed a consensus module by intersecting this top subnetwork with the most similar subnetworks obtained in all the other nine results. Step4: KEGG pathway enrichment analysis was performed for each consensus module by Enrichr tool.

Quality control (QC) was performed on the ADNI participants using the PLINK v1.07 software [24] (<http://zzz.bwh.harvard.edu/plink/>), following a similar procedure described in Li et al. [32]. Briefly, SNPs not meeting any of the following criteria were excluded: (1) SNPs available on both 2.5 M array and OmniQuad array, (2) call rate per SNP $\geq 95\%$; (3) minor allele frequency $\geq 5\%$ ($n = 1,845,510$ SNPs were excluded based on Criteria 1–3); and (4) Hardy-Weinberg equilibrium test of $p \geq 10^{-6}$ ($n = 198$ SNPs were excluded) using control subjects

only. Participants were excluded from the analysis if any of the following criteria were not satisfied: (1) call rate per subject $\geq 95\%$ (no participant was excluded), (2) sex check (1 participant was excluded), (3) identity check for related pairs (8 sibling pairs and 1 sibling triple were identified with PI_HAT >0.5 , 1 participant from each family was randomly selected and included in the study).

Population stratification analysis was performed using EIGENSTRAT [33], and confirmed using STRUCTURE

Table 1 Selected demographic and clinical characteristics of 843ADNI participants

	CN (N = 199)	SMC (N = 85)	EMCI (N = 239)	LMCI (N = 207)	AD (N = 113)	P-value
Age (years)	74.4 (5.79)	72.0(5.48)	71.4(7.30)	72.4(7.62)	75.2 (8.19)	$p < 0.001$
Women*	96(48%)	50(59%)	102(43%)	83(40%)	45(40%)	0.002*
Education (years)	16.2 (2.82)	16(2.79)	16.2(2.84)	16.4(2.53)	16.4(2.56)	0.764
APOE e4 allele present	47(24%)	31(36%)	99 (41%)	112(54%)	74 (65%)	$p < 0.001$
CDR-SOB	0.04(0.14)	0.08(0.18)	1.27(0.77)	1.65(0.94)	4.53 (1.70)	$p < 0.001$
Mini mental status examination	29.1(1.18)	29.0(1.2)	28.3(1.62)	27.5(1.75)	23.1 (2.05)	$p < 0.001$
Logical memory immediate recall	14.42(3.00)	14.44(3.34)	11.09(2.68)	7.18(3.06)	4.15 (2.70)	$p < 0.001$
Logical memory delayed recall	13.34(3.13)	13.29(3.31)	8.97(1.73)	3.94 (2.7)	1.52 (1.80)	$p < 0.001$
t-tau/Aβ ₁₋₄₂ Ratio (i.e., QT for GWAS)	0.40 (0.27)	0.37(0.24)	0.50(0.45)	0.70(0.47)	0.98 (0.49)	$p < 0.001$

AD Alzheimer’s disease, ADNI Alzheimer’s Disease Neuroimaging Initiative, CDR-SOB clinical dementia rating-sum of boxes, CN cognitively normal, SMC significant memory concern, EMCI early mild cognitive impairment, LMCI late mild cognitive impairment. Number (%) or mean (s.d.) was shown in each entry. P-values were assessed due to significant differences between diagnosis groups, which computed using one-way ANOVA (*except for gender using chi-square test)

[34]. It yielded 89 study participants who did not cluster with the remaining subjects and with the CEU HapMap samples who are primarily of European ancestry (non-Hispanic Caucasians). These 89 participants were excluded from the analysis. After QC, 563,980 SNPs and 843 individuals remained available for the subsequent GWAS, network and pathway analyses.

SNP-Level and Gene-Level GWAS Analyses

For GWAS, to examine the main effects, linear regression was implemented by PLINK to evaluate the association between each SNP and the t-tau/Aβ₁₋₄₂ ratio. An additive genetic model was tested with covariates, including age, gender, and diagnosis (through five binary dummy variables indicating CN, SMC, EMCI, LMCI, and AD). Then, the SNP-level p-values were obtained.

The VEGAS2 software [25] was used to assign 563,980 SNPs to 22,179 autosomal genes according to positions on the UCSC Genome Browser (out of 26,292 in hg19), and to compute gene-level p values. The software applies simulations from the multivariate normal distribution to employ information from a defined subset of markers within a gene as well as take into account linkage disequilibrium between the markers. To save running time, we use a multi-stage approach to adaptively determine the number of simulations per gene: (Stage 1) we run 10³ simulations for all the genes; (Stage 2) we run 10⁴ simulations only for genes with Stage 1 empirical p-values ≤ 0.1; (Stage 3) we run 10⁶ simulations only for genes with Stage 2 empirical p-values ≤ 0.001. We interpret an empirical p-value of 0 from 10⁶ simulations as $p < 10^{-6}$. Given 22,179 genes included in this analysis, a Bonferroni-corrected threshold is $p < 2.25 \times 10^{-6}$ (i.e., 0.05/22,179), which can be exceeded by the theoretically smallest empirical p-value shown above. A Manhattan

plot was generated using R (<http://www.r-project.org>) to visualize the gene-level GWAS results for our work.

Network-level Analysis

The Human PPI data ($n = 9,617$) were downloaded from the Human Protein Reference Database (HPRD, <http://www.hprd.org>); gene-level p-values obtained from the GWAS of the CSF t-tau/Aβ₁₋₄₂ ratio were mapped to the PPI network. The integrative protein interaction network-based pathway analysis (iPINBPA) software [22] was used to integrate GWAS findings with physical evidence of interaction at the protein level, and to detect new high-level associations (i.e., subnetworks of functionally interacted genes) with the CSF biomarker. Briefly, iPINBPA identifies enriched subnetworks using the following three steps.

In Step 1, using the GWAS findings, the nominally significant genes (i.e., $p \leq 0.05$) are treated as seed genes, and assigned with certain weights (e.g., in this work, 1 for seed genes, 0 for the rest). After that, a random walk with restart strategy is employed to smooth these weights over the entire network. Intuitively, the nodes in the network are weighted based on their connectivity to seed genes (i.e., guilt-by-association). Let n_k be a node on the PPI network mapped with gene-level p-value p_i . Let e_{ij} be the edge connecting n_i and n_j , and W_{ij} be the weight of e_{ij} . All the W_{ij} ’s form the adjacency matrix W . Extending Köhler’s approach [35], iPINBPA weights the edge e_{ij} as follows:

$$W_{ij} = \left((1-p_i) + (1-p_j) \right) / 2.$$

In addition, it normalizes the adjacency matrix W by its columns. After each step of random walk, a score vector is calculated as

$$P(t) = (1-r)W \cdot P(t-1) + rP(0),$$

where $P(t)$ is the score after walking t steps, and r is the restart ratio. In this work, we assign 1 to all the seed genes, and 0 to the rest. Upon the completion of the random walk after T steps, the vector $P(T)$ contains the node weights, which reflect the topological connections to the seed genes [36].

In Step 2, given a network A containing k nodes, iPINBPA defines its score by combining the gene-level p -values with node weights described above, using the Liptak-Stouffer method. Specifically, the network score of A is defined, via weighted Z transform test [37], as follows:

$$Z_A = \frac{\sum_{i \in A} P(T)_i z_i}{\sqrt{\sum_{i \in A} P(T)_i^2}}$$

A random sampling of gene sets of size $k \in [1, 500]$ for 1000 times was applied in iPINBPA [36] to determine the background distribution of the network score. Using this distribution, the adjusted network score of A is defined as:

$$S_A = \frac{Z_A - \mu_k}{\sigma_k},$$

where Z_A is the network score, and μ_k and σ_k are respectively the mean and the standard deviation of the background distribution of the network scores at size k .

In Step 3, a greedy algorithm was developed to search for modules with high network scores, i.e., those enriched in genes with low p -values and high weights. Details about the algorithm is available in Wang et al. [22].

In this work, the parameters were set in iPINBPA as follows: $r = 0.5$, $T = 5$, $\text{NetScore} > 3.0$, $\text{NetSize} \geq 5$, and $\text{MaxNetSize} \leq 300$. Given the stochastic nature of the iPINBPA algorithm, we ran iPINBPA ten times, respectively by setting the random seed value from 1 to 10. Consequently, we obtained ten groups of enriched subnetworks (GNs) identified by iPINBPA. Below, we couple a consensus module (CM) strategy with iPINBPA (named as CM-iPINBPA) to generate consensus findings from these analyses.

Given two subnetworks m and n , we use Dice's coefficient $DC(m, n)$ to measure their similarity:

$$DC(m, n) = \frac{2|m \cap n|}{|m| + |n|}.$$

In this work, we only focused on analyzing the top subnetwork (TN) in each iPINBPA run. Let TN_a be the top subnetwork identified in Run $a \in \{1, 2, \dots, 10\}$. We first find $SN_b(TN_a)$, which is the most similar subnetwork to TN_a in Run $b \in \{1, 2, \dots, 10\} \setminus \{a\}$. Clearly we have

$$SN_b(TN_a) = \text{argmax}_{sn} DC(TN_a, sn),$$

where sn is any subnetwork enriched in Run b . After that, we define the consensus module (CM) based on Run a as follows:

$$CM_a = TN_a \cap (\cap_{a \neq b, b \in \{1, 2, \dots, 10\}} SN_b(TN_a)).$$

Namely, CM_a is the intersection of TN_a and its most similar subnetworks identified in all the other runs.

Network Module Visualization and Functional Analysis

Cytoscape 3.2 [38] was used to visualize the example identified network modules. The Enrichr [28] pathway analysis tool (<http://amp.pharm.mssm.edu/Enrichr/>) and the Kyoto Encyclopedia of Genes and Genomes database (KEGG; <http://www.genome.jp/kegg/>) [39] was applied to functional analysis of the identified network modules. Heat map was plotted, using R 3.2.0 software, to demonstrate relations between consensus modules and relevant KEGG pathways.

Results

GWAS and gene-level analysis

The demographic and clinical characteristics for the 843 ADNI subjects in this study are presented in Table 1. The summary statistics for all diagnostic groups (CN, SMC, EMCI, LMCI, and AD) are given. Education level ($p = 0.764$) was not significantly different across the five diagnostic groups; however, gender demonstrated a significant difference ($p = 0.002$). Furthermore, as expected, age, *APOE* e4 status, clinical dementia rating–sum of boxes (CDR-SOB), mini mental status examination (MMSE), logical memory immediate recall, and logical memory delayed recall exhibited significant differences across the five groups ($p < 0.001$). Also as expected, the phenotype t-tau/ $A\beta_{1-42}$ ratio significantly differed across the diagnostic groups ($p < 0.001$).

The top SNP in the GWAS analysis was rs4420638 (chromosome 19, 14 kb away from the *APOC1* gene, $p = 2.576E-28$), which was previously reported by Lars Bertram et al. [40]. The SNP rs769449 within the *APOE* gene on chromosome 19 was also significant with $p = 4.98E-23$, and was previously reported by Soerensen et al. [41]. Similar to the results reported in our earlier paper [13], The *TOMM40* SNP rs2075650 (chromosome 19, $p = 4.23E-18$, was associated with t-tau/ $A\beta_{1-42}$ ratio.

Under the hypothesis that genes, rather than SNPs, are the functional units in biology [42], a gene-level association analysis of the t-tau/ $A\beta_{1-42}$ ratio was performed based on the SNP-level results by VEGAS2. Table 2 shows the top 10 genes identified by VEGAS2. Figure 2 shows the Manhattan plot of the gene-based GWAS results.

Network search for CMs

Consensus modules (CMs) were identified by CM-iPINBPA network analysis strategy. Subnetwork search was conducted on the GWAS findings using iPINBPA ten times by varying the seed value of random number generator, which ranged from 1 to 10. Table 3 summarizes the results of ten iPINBPA runs. *PRKCA* and *TP53* appeared to be the start nodes of the top subnetworks

Table 2 The top 10 genes identified by VEGAS2

Chr	Gene	nSNPs	Test	Pvalue	TopSNP	TopSNP-Pvalue
19	APOC1	6	274.96	1.00E-06	rs4420638	2.58E-28
19	APOE	7	188.92	1.00E-06	rs769449	4.98E-23
19	PVRL2	19	228.01	1.00E-06	rs2075650	4.23E-18
19	TOMM40	11	363.85	1.00E-06	rs769449	4.98E-23
17	PDK2	11	103.37	5.00E-06	rs3809762	1.86E-06
17	ITGA3	15	111.81	7.40E-05	rs3809762	1.86E-06
19	CBLC	2	24.35	1.12E-04	rs2965121	1.45E-04
17	CCL7	5	28.84	2.38E-04	rs991804	2.28E-04
19	KLK7	12	56.02	2.76E-04	rs11084043	7.74E-05
1	PTGER3	97	400.30	3.07E-04	rs7540868	1.89E-04

Chr: Chromosome; Gene: Gene name; nSNPs: Number of SNPs in the input file that map to the gene; Test: The sum of the individual chi-squared 1 degree of freedom SNP-association test statistic; Pvalue: The gene-based p-value considering the full set of SNPs; TopSNP: The name of the most significant SNP within the gene; TopSNP-Pvalue: The p-value for the most significant SNP with the gene

identified in multiple runs. The *PRKCA* gene was previously reported as being associated with an altered amyloid precursor protein (APP) secretion in fibroblasts from AD patients [43, 44]. Culmsee et al. demonstrated that *TP53* was a novel gene as a biomarker of AD and was related to neurodegenerative processes [45].

Table 4 shows the top subnetwork identified in each run, its most similar subnetworks in other runs coupled with the Dice's coefficient value, and the corresponding CM. For example, in Table 4, the Dice's coefficient between TN1 and SN1 (in GN₂) is 0.96335. Thanks to the overlapping subnetworks, only four unique CMs were identified (Table 4). These four CMs are shown in Fig. 3, where the reddish nodes represent the known AD genes from the KEGG AD pathway (hsa05010). CM1 ($S_A = 10.15$, $p < 0.001$) shown in Fig. 3(a) contains totally 67 genes, including KEGG AD genes *GSK3B*, *MAPK1*, *PSEN2*, *CALM1*, *CALM2* and *CASP8*. CM2 ($S_A = 10.39$, $p < 0.001$) shown in Fig. 3(b) contains totally 67 genes, including KEGG AD genes *BACE1*, *GSK3B*, *MAPK1*, *PSEN2*, *CALM1*, *CALM2* and *CASP8*. CM3 ($S_A = 7.99$, $p < 0.001$) shown in Fig. 3(c) contains 40 genes, including KEGG AD genes *GSK3B*, *CALM1* and *CASP8*. CM4 ($S_A = 10.46$, $p < 0.001$) shown in Fig. 3(d) contains 58 genes, including KEGG AD genes *BACE1*, *GSK3B*, *MAPK1*, *PSEN2*, *CALM1*, *CALM2* and *CASP8*.

In addition, the intersection of the four CMs was extracted and named as the common subnetwork. The common subnetwork is shown in Fig. 5, and contains total 29 genes, including 3 KEGG AD genes *CALM1*, *CASP8*, and *GSK3B*, and 26 other genes.

Pathway analysis of consensus modules and the common subnetwork

To test the hypothesis that CMs enriched by the GWAS findings might be significantly overrepresented in AD

Table 3 Results of 10 iPINBPA runs

SRN	# of Subnetworks	TN Size	TNStartNode	TN Score
1	1058	99	PRKCA	10.11
2	918	92	PRKCA	9.70
3	975	151	TP53	10.95
4	1100	133	-	10.52
5	977	108	PRKCA	9.85
6	936	135	-	10.64
7	1166	101	PRKCA	9.41
8	1050	147	-	11.27
9	972	67	TP53	8.04
10	955	146	-	11.25

SRN: The seed value used for an iPINBPA run

of Subnetworks: The number of subnetworks identified in an iPINBPA run

TN Size: The number of genes in the top subnetwork identified in an iPINBPA run

TN StartNode: The start node of the top subnetwork

TN Score: The score of the top subnetwork

and other relevant pathways, the Enrichr method was performed for pathway analysis. For the genes in each CM, a pathway enrichment analysis was conducted, and the nominally enriched pathways were identified with adjusted p -value ≤ 0.05 . Then, these identified pathways of the four CMs were plotted as a heat map shown in Fig. 4 to summarize the relationships between the pathways and CMs. Note that Fig. 4 lists only pathways enriched by at least one CM. Table 5 shows top 20 pathway enrichments analysis of the common subnetwork.

Discussion

Gene-level analysis

In the GWAS analysis, the gene-level p -values were determined and shown in Fig. 2. The use of the CSF t-tau/ $A\beta_{1-42}$ ratio as a quantitative trait (QT) in this study enabled us to examine the effect of genes previously known to be associated with the QT as well as to identify novel genes. Table 2 lists the top ten genes obtained by converting SNPs into gene-wise p -values. Given the Bonferroni-corrected threshold of $p < 2.25 \times 10^{-6}$ (i.e., 0.05/22,179), we found five significant genes. As expected, significant associations were identified between loci on chromosome 19 and the CSF t-tau/ $A\beta_{1-42}$ ratio (e.g., *APOC1*, *APOE*, *PVRL2*, *TOMM40*, $p = 1 \times 10^{-6}$, see Fig. 2). Among these genes, apolipoprotein C1 (ApoC1) encoded by the *APOC1* gene is associated with amyloid β plaques; the *APOE* and *TOMM40* (rs769449) genes code for proteins related to the clearance of $A\beta$ and mitochondrial functions [5, 13]; and the *PVRL2* gene was previously reported as related to risk factors that contribute to AD pathogenesis [46]. *PDK2* ($p = 5 \times 10^{-6}$) shows a trend towards the significance, and the overexpression of this gene may be related to cancer and diabetes [47]. Additionally, *CCL7* mRNA is highly increased

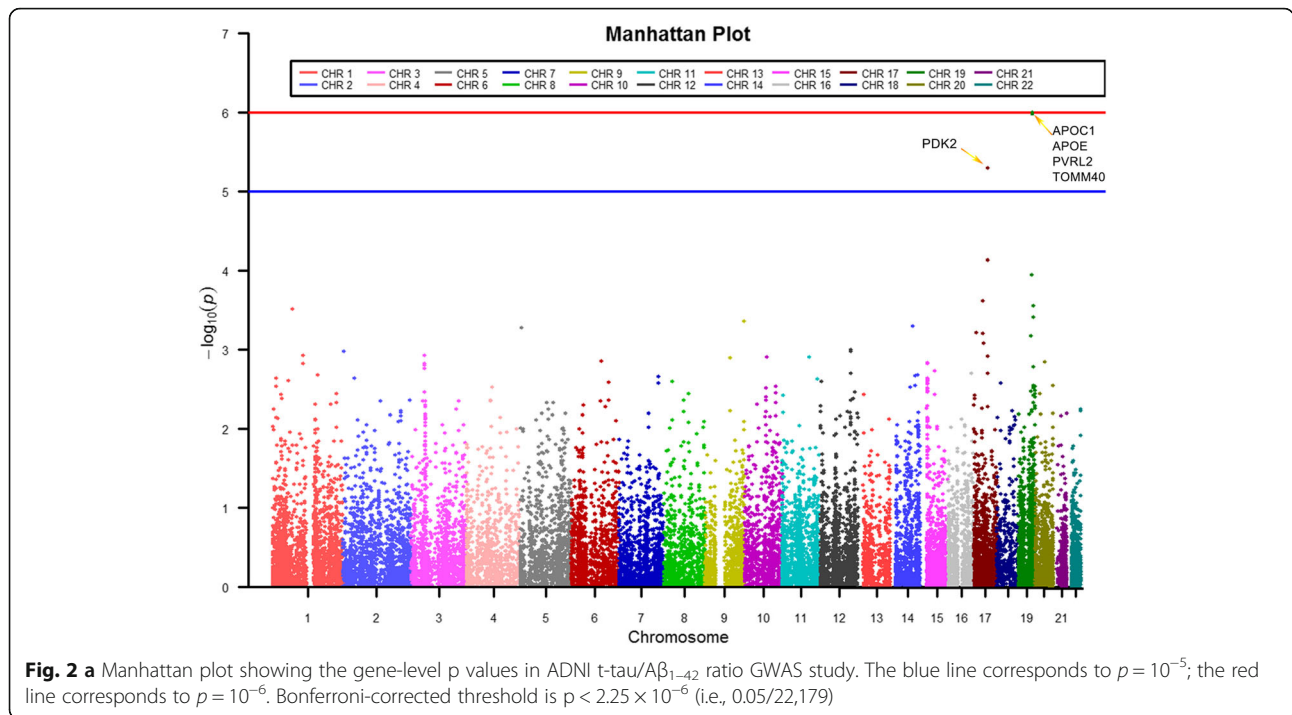
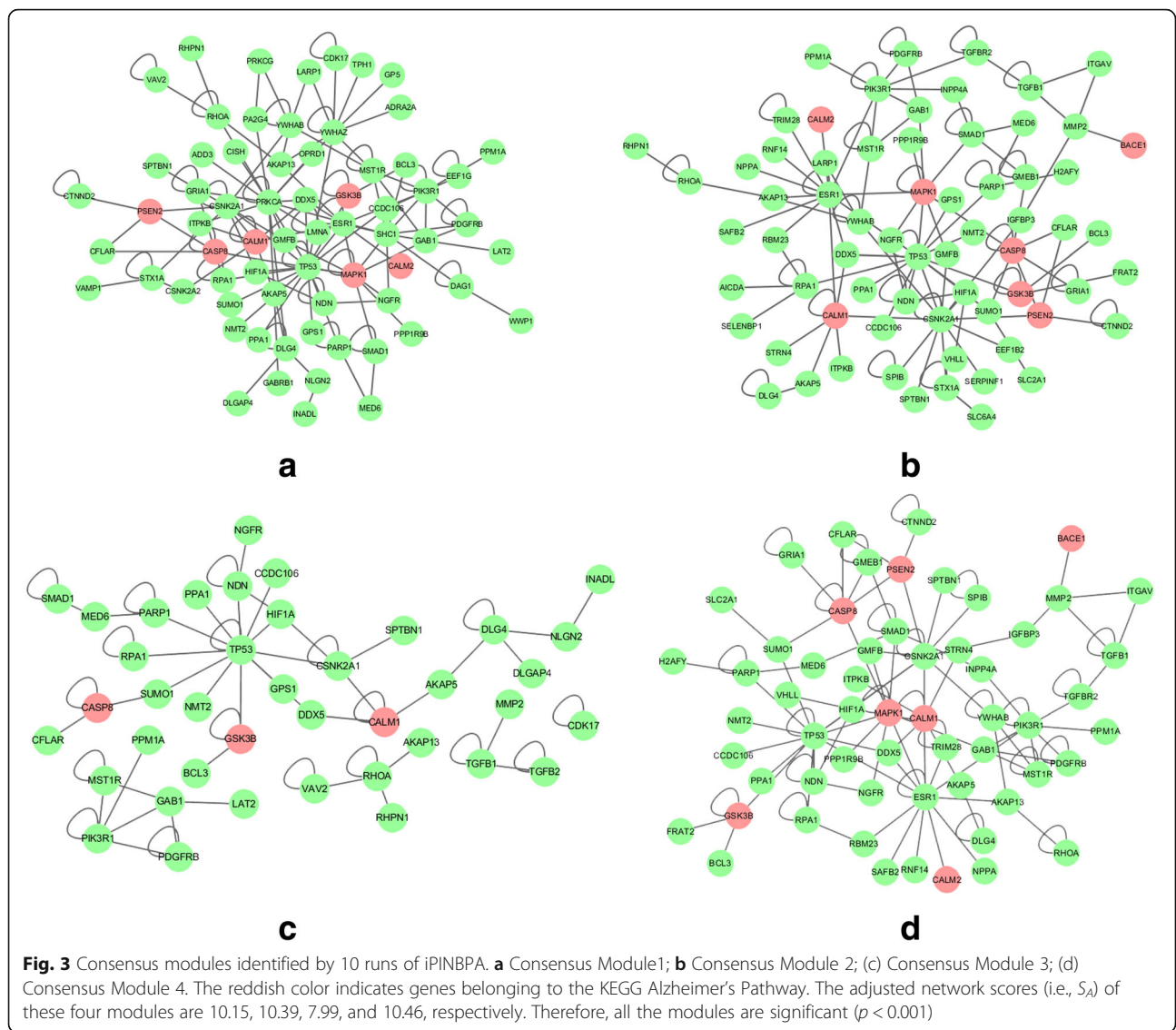


Table 4 The characteristics of the identified consensus modules in 10 iPINBPA runs

CM	RunA	Ta: the top subnetwork in RunA. Sb: the most similar subnetwork to Ta in RunB										
		RunB	1	2	3	4	5	6	7	8	9	10
1	1	Rank of Sb in RunB	1	1	4	3	1	6	1	3	8	4
		DC(Ta, Sb)	1.00	0.96	0.88	0.88	0.96	0.96	0.99	0.82	0.81	0.96
1	2	Rank of Sb in RunB	1	1	4	3	1	6	1	3	8	4
		DC(Ta, Sb)	0.96	1.00	0.85	0.85	0.92	1.00	0.95	0.78	0.84	0.92
2	3	Rank of Sb in RunB	3	2	1	2	3	2	3	2	1	5
		DC(Ta, Sb)	0.78	0.77	1.00	0.93	0.83	0.91	0.81	0.97	0.61	0.83
3	4	Rank of Sb in RunB	2	3	2	1	2	1	2	1	13	1
		DC(Ta, Sb)	0.85	0.82	0.97	1.00	0.89	0.99	0.86	0.95	0.67	0.95
1	5	Rank of Sb in RunB	1	1	4	3	1	6	1	3	8	4
		DC(Ta, Sb)	0.96	0.92	0.93	0.93	1.00	0.92	0.97	0.86	0.77	1.00
3	6	Rank of Sb in RunB	2	3	2	1	2	1	2	1	13	1
		DC(Ta, Sb)	0.84	0.82	0.98	0.99	0.88	1.00	0.86	0.96	0.66	0.96
1	7	Rank of Sb in RunB	1	1	4	3	1	6	1	3	8	4
		DC(Ta, Sb)	0.99	0.95	0.89	0.89	0.97	0.95	1.00	0.83	0.80	0.97
3	8	Rank of Sb in RunB	2	3	2	1	2	1	2	1	13	1
		DC(Ta, Sb)	0.80	0.78	0.98	0.95	0.84	0.96	0.81	1.00	0.63	1.00
4	9	Rank of Sb in RunB	3	2	1	2	3	8	3	2	1	5
		DC(Ta, Sb)	0.82	0.83	0.61	0.68	0.77	0.73	0.79	0.64	1.00	0.77
3	10	Rank of Sb in RunB	2	3	2	1	2	1	2	1	13	1
		DC(Ta, Sb)	0.80	0.78	0.98	0.95	0.85	0.96	0.82	1.00	0.63	1.00

CM consensus module id, DC Dice's coefficient



by β_{1-42} stimulation [48]. The *CCL7* gene was previously reported as related to the chemotaxis of pro-inflammatory cells to the inflamed location [49]. The *PTGER3* gene was previously reported as being related to the inflammatory response [50].

Network search for CMs and functional validation

Although iPINBPA has been successfully applied in several previous studies [16, 20–23], we observed that different subnetworks could be obtained by using different random seed values. To overcome this limitation, we proposed to examine the consensus modules discovered by multiple iPINBPA analyses. In other words, we focused on examining the shared subnetworks among multiple iPINBPA runs, which turned out to be more stable patterns. A two-stage strategy was employed. First, ten groups of subnetworks were generated by running

iPINBPA ten times with varying random seed values ranging from 1 to 10. After comparing these ten sets of results, we identified ten CMs, one from each run (defined as the intersection of the top subnetwork from the current run and the most similar subnetworks from all the other runs). As a result, there are four unique CMs based on ten identified ones.

The genes in the CMs might not show a direct statistical significance but could interact with some genes identified in our GWAS. These genes can demonstrate indirect association with the studied QT, and may potentially provide valuable biological interpretation. For example, Consensus Module 1 contains the protein gamma-aminobutyric acid (GABA) A receptor (*GABRB1*) gene. *GABRB1* codes for the β_1 subunit of gamma-aminobutyric acid A (GABA_A) receptors [1]. The *GABRB1* gene has been demonstrated to be

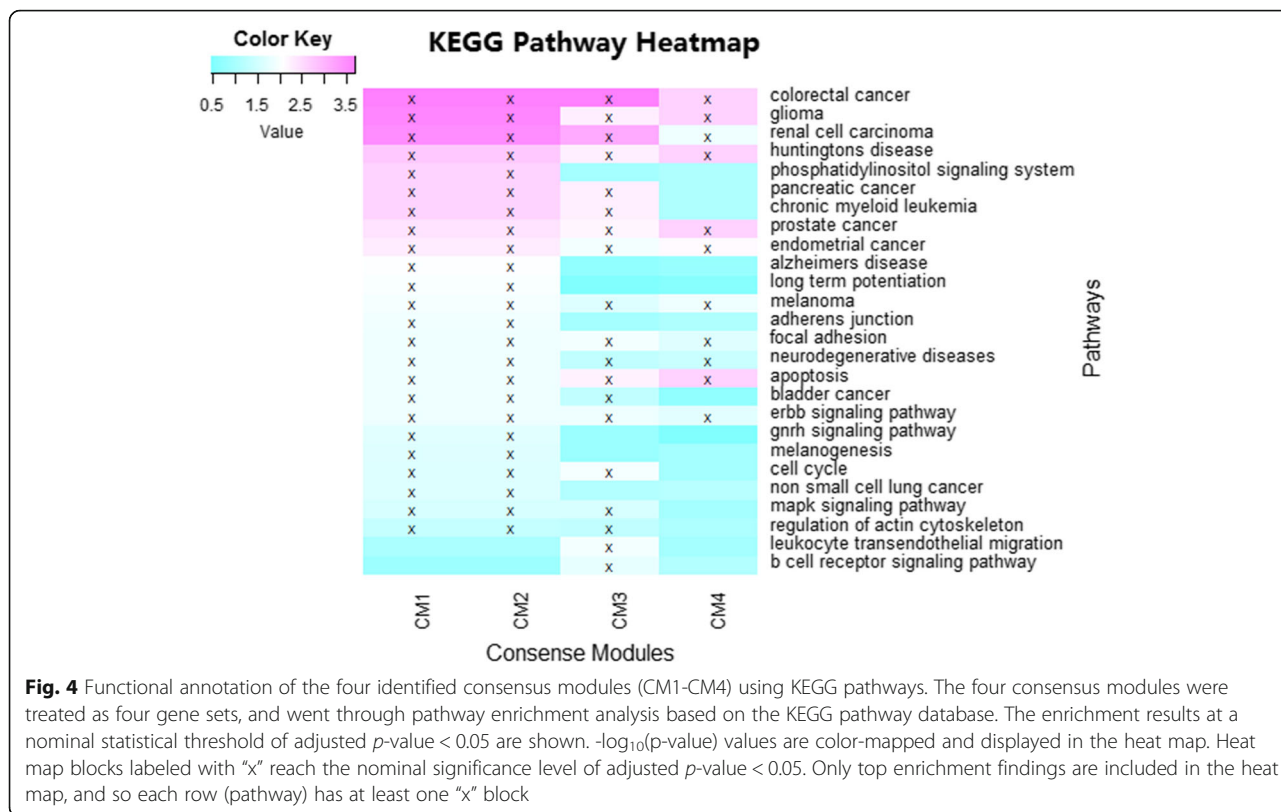


Fig. 4 Functional annotation of the four identified consensus modules (CM1-CM4) using KEGG pathways. The four consensus modules were treated as four gene sets, and went through pathway enrichment analysis based on the KEGG pathway database. The enrichment results at a nominal statistical threshold of adjusted p -value < 0.05 are shown. $-\log_{10}(p\text{-value})$ values are color-mapped and displayed in the heat map. Heat map blocks labeled with "x" reach the nominal significance level of adjusted p -value < 0.05. Only top enrichment findings are included in the heat map, and so each row (pathway) has at least one "x" block

involved in the thalamus structure and its interactive effects on intelligence [51]. The *GABRB1* gene had also been associated with many neuropsychological diseases, such as schizophrenia, major depression, bipolar disorder, and Alzheimer’s disease [52].

In this study, we hypothesize that trait prioritized CM with high replication might have strong functional associations with $t\text{-tau}/A\beta_{1-42}$ ratio. We gathered these identified pathways for 4 CMs to plot heatmap to explore and refine the relationships between pathways and CMs (Fig. 4). In Fig. 4, it was observed that four pathways, including colorectal cancer, gliomas, renal cell carcinoma, and Huntington’s disease, were commonly enriched in all the consensus modules. The neurodegenerative symptoms of neuron death affect many diseases, including Alzheimer’s, Parkinson’s, and Huntington’s diseases. Below we briefly discuss a few additional pathways identified in Fig. 4. In AD, focal adhesion complexes regulating neuronal viability can be used in treatment to adjust neuronal survival [53]. The adherens junction has been demonstrated as maintaining blood–brain barrier integrity, and the adherens junction pathways is highly associated with neurodegenerative diseases [54]. Apoptosis is an important pathway in Alzheimer’s disease that is associated with neuronal loss [55]. The change in the MAPK signaling pathway contributes to significant change in neurotrophin [56]. Some cancer-related

pathways were found, such as colorectal cancer, pancreatic cancer, prostate cancer, endometrial cancer, bladder cancer, and so on. Some prior studies have been performed to examine the relationship between cancer and neurologic disease [57, 58].

With these observations, the genes in the CMs may provide valuable information to suggest novel molecular mechanisms for subsequent analyses. Compared with the standard iPINBPA method, CM-iPINBPA network analysis strategy for mining consensus models could offer more stable results.

Common subnetwork and functional validation

The common subnetwork is the intersection of the four identified common modules, and consists of 29 genes (Fig. 5). Among these genes, the *GSK3B*, *SUMO1*, *AKAP5*, *CALM1* and *DLG4* genes have been identified to be involved in tau phosphorylation, and over-phosphorylation of the tau protein can form the tangles in the brain of AD patients [59–63]. Additionally, the *CASP8*, *PIK3R1*, *PPA1*, *PARP1*, *CNSK2A1*, *NGFR*, and *RHOA* genes have been demonstrated to be involved in amyloid beta peptide production [64–70]. The *BCL3*, *CFLAR*, *SMAD1*, and *HIF1A* genes have been identified to be associated with late-onset Alzheimer’s disease. The common subnetwork also contains the following genes *TP53*, *DDX5*, *NDN*, *MST1R*, *CCDC106*, *NMT2*, *RPA1*,

Table 5 Top 20 pathway enrichments analysis of the common consensus subnetwork

Pathway	Overlap	P-value	Adjusted P-value	Z-score	Combined Score	Genes
glioma	4/62	7.19E-05	0.002277	-2.14	13.00	PDGFRB; PIK3R1;CALM1;TP53
apoptosis	4/81	1.94E-04	0.002277	-1.85	11.28	CASP8; CFLAR;PIK3R1;TP53
huntingtons disease	3/31	2.20E-04	0.002277	-1.77	10.77	CASP8;CALM1;TP53
colorectal cancer	4/84	2.22E-04	0.002277	-1.84	11.18	PDGFRB; GSK3B;PIK3R1;TP53
prostate cancer	4/86	2.42E-04	0.002277	-1.80	10.97	PDGFRB; GSK3B;PIK3R1;TP53
endometrial cancer	3/52	9.23E-04	0.007232	-1.65	8.13	GSK3B;PIK3R1;TP53
renal cell carcinoma	3/69	2.02E-03	0.01286	-1.61	6.99	GAB1;PIK3R1;HIF1A
melanoma	3/71	2.19E-03	0.01286	-1.49	6.50	PDGFRB;PIK3R1;TP53
erbB signaling pathway	3/85	3.59E-03	0.018769	-1.53	6.08	GSK3B;GAB1;PIK3R1
focal adhesion	4/192	4.51E-03	0.02118	-1.60	6.16	PDGFRB; GSK3B;PIK3R1;RHOA
neurodegenerative diseases	2/38	9.01E-03	0.038512	-0.87	2.83	NGFR;CASP8
non small cell lung cancer	2/53	1.66E-02	0.064347	-1.21	3.32	PIK3R1;TP53
basal cell carcinoma	2/55	1.78E-02	0.064347	-0.92	2.53	GSK3B;TP53
b cell receptor signaling pathway	2/62	2.22E-02	0.074377	-0.91	2.37	GSK3B;PIK3R1
phosphatidylinositol signaling system	2/72	2.91E-02	0.081711	-1.00	2.50	PIK3R1;CALM1
pancreatic cancer	2/73	2.98E-02	0.081711	-0.99	2.49	PIK3R1;TP53
chronic myeloid leukemia	2/75	3.13E-02	0.081711	-0.93	2.33	PIK3R1;TP53
adherens junction	2/75	3.13E-02	0.081711	-0.78	1.96	CSNK2A1;RHOA
regulation of actin cytoskeleton	3/201	3.54E-02	0.087678	-0.96	2.34	PDGFRB; PIK3R1;RHOA
small cell lung cancer	2/85	3.92E-02	0.092081	-0.73	1.74	PIK3R1;TP53

Pathway: The name of KEGG pathway

Overlap: The number of overlapping genes compared with the number of input genes

P-value: P-value was computed using the Fisher exact test

Adjusted P-value: Adjusted P-value was a corrected p-value to the Fisher exact test

Z-score: Computed by assessing the deviation from the expected rank

Combined score: Computed by taking the log of the p-value from the Fisher exact test and multiplying that by the z-score of the deviation from the expected rank

Genes: The overlapping genes between the input and the pathway

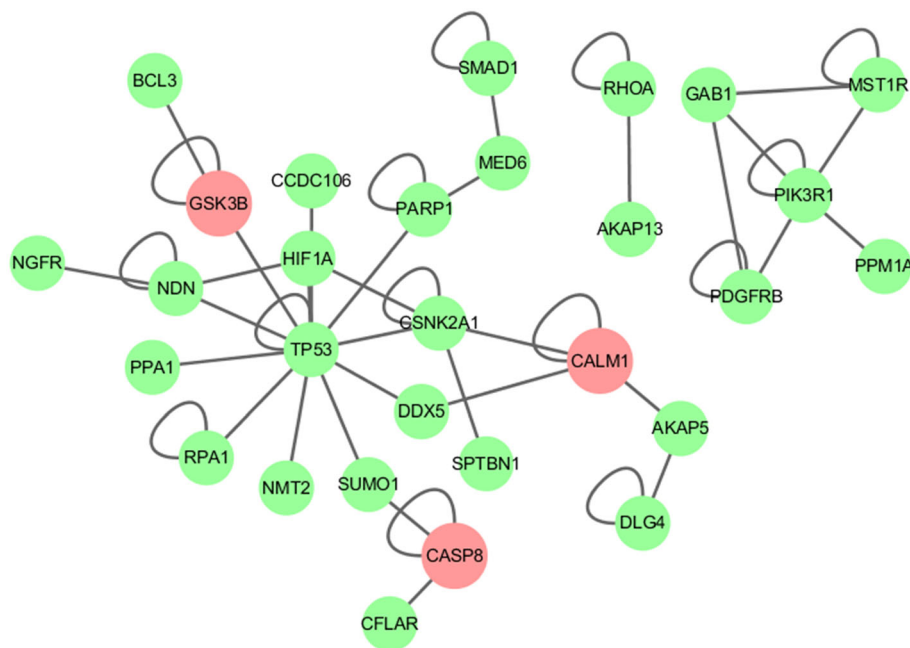


Fig. 5 The common subnetwork. This subnetwork consists of only overlapping genes of all consensus modules. The reddish color indicates genes belonging to the KEGG Alzheimer's Pathway

AKAP13, *GAB1*, *PPM1A*, *SPTBN1* and *MED6*, which interact with themselves and other genes. These findings offer valuable biological insights and suggest promising candidates for subsequent analyses.

Table 5 shows the top twenty pathways enriched by the common subnetwork. Some significant pathways were observed, such as glioma, apoptosis, Huntington's disease, renal cell carcinoma, melanoma, the *erbB* signaling pathway, focal adhesion, neurodegenerative diseases, and so on. Twelve genes (*PDGFRB*, *PIK3R1*, *CALM1*, *CFLAR*, *TP53*, *RHOA*, *CASP8*, *HIF1A*, *GSK3B*, *GAB1*, *NGFR* and *CSNK2A1*) were involved in the top twenty pathways. The gene *PDGFRB* has been confirmed as causative of primary familial brain calcifications (PFBC) [71]. The gene *PIK3R1* has been shown to be involved in Alzheimer's disease [72]. *CALM1* encodes for tau protein and regulates the subcellular localization and function of calmodulin in neurons [73]. *CFLAR* suppresses death receptor-induced apoptosis and TCR activation, which induces cell death by inhibiting caspase-8 activation [74]. *RHOA* was implicated in A β neurotoxicity, and the activation generates cytoskeletal changes [75]. *NGFR* ligands play an important role in preventing fundamental tau-related pathologic mechanisms in Alzheimer's disease [76]. These pathways appear relevant given the results in prior studies. For other genes identified in this pathway analysis, it warrants further investigation to demonstrate the role they play.

Limitations

Due to the limited number of subjects available to us, we were only able to perform a discovery study in this work. When more data become available in the future, replication studies in independent cohorts are required to evaluate and validate the identified network modules in our study. In addition, in this work, we reported the results using the default parameter setting provided by the software tool and according to Lili et al. [22], except the random seed value. We ran iPINBPA multiple times by using different random seed values and then extracting the consensus patterns to stabilize the results. As to other parameters, we also briefly tested each of those by varying its value. For most of these parameters, we obtained very similar results. The most sensitive parameter is the restart ratio used in the step called "random walk with restart" for prioritizing phenotype-associated genes. It is expected that different restart ratios will assign different weights to network nodes and subsequently produce different scores for network components. The determination of the optimal restart ratio warrants a separate and focused investigation and is an interesting future direction.

Conclusions

Network-based methods form a new generation of enrichment analysis strategy, and they can overcome the

limitation of traditional enrichment analysis where only a fixed set of pre-defined pathways are examined. In this study, a genome-wide network-based pathway analysis of the CSF biomarker of the t-tau/A β ₁₋₄₂ ratio was performed, using a sample of 843 subjects from the ADNI database. To our knowledge, this is the first genome-wide network-based pathway study on the CSF biomarker t-tau/A β ₁₋₄₂ ratio in Alzheimer's disease. Due to the stochastic nature of the iPINBPA method, we employed a consensus module (CM) strategy to run iPINBPA multiple times and aimed to identify CMs from these different runs. We identified 4 CMs. These CMs contain not only genes from KEGG AD pathways, including *BACE1*, *GSK3B*, *MAPK1*, *PSEN2*, *CALM1*, *CALM2*, *CASP8*, and *SK3B*; but also interesting genes with relevant biological functions such as *GABRB1*, *MMP2*, *CDK17*, and *IGFBP3*. In sum, besides confirming previous findings (e.g., *APOE*, *TOMM40*, *APOC1*), this study has also suggested new susceptible genes, CMs and pathways underlying Alzheimer's disease.

Abbreviations

AD: Alzheimer's disease; ADNI: Alzheimer's Disease Neuroimaging Initiative; ANOVA: Analysis of variance; A β ₁₋₄₂: Amyloid- β 1-42 peptide; CDR-SOB: Clinical dementia rating-sum of boxes; CM: Consensus module; CM-iPINBPA: Consensus module (CM) strategy with iPINBPA; CMs: Consensus modules; CN: Cognitively normal; CSF: Cerebrospinal fluid; DC: Dice's coefficient; EMCI: Early mild cognitive impairment; GNs: Groups of enriched subnetworks; GWAS: Genome-Wide Association Studies; HPRD: Human Protein Reference Database; iPINBPA: Integrative protein interaction network-based pathway analysis; LMCI: Late mild cognitive impairment; MCI: Mild cognitive impairment; MMSE: Mini mental status examination; MRI: Magnetic resonance imaging; PET: Positron emission tomography; PPI: Protein-protein interaction; QT: Quantitative trait; SMC: Significant memory concern; SNPs: Single Nucleotide Polymorphisms; TN: Top subnetwork; t-tau: Total tau

Acknowledgements

Alzheimer's Disease Neuroimaging Initiative
Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Funding

Data analysis, result interpretation and manuscript writing were supported in part by National Key Scientific Instrument and Equipment Development Projects of China (2012YQ04014010), Fundamental Research Funds for the Central Universities (HEUCF160412) and the Natural Science Foundation of Heilongjiang Province (F201241). Data analysis, method development, and manuscript editing were supported in part by NIH R01 LM011360, R01 EB022574, RC2 AG036535, R01 AG19771, P30 AG10133, and NSF IIS-1117335 at IU. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson

& Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Availability of data and materials

The genotyping and CSF data were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>). Application for access to the ADNI data can be submitted by anyone at <http://adni.loni.usc.edu/data-samples/access-data/>. The process includes completion of an online application form and acceptance of Data Use Agreement. We have received administrative approval for access to the ADNI database. The Human PPI data were downloaded from the public Human Protein Reference Database (<http://www.hprd.org/>).

Authors' contributions

W.C., J.L., H.L. and L.S. led and supervised the research. X.M., H.L., L.S., W.C., and J.L. designed the research and wrote the article. X.M. performed the gene analysis, network mapping, iPINBPA analysis, and pathway analysis. Q.Z. and F.C. prepared the genotyping and CSF biomarker data, and performed quality control and GWAS. W.L., Y.W., and S.C. helped with research design and data analysis. X.Y., J.Y., S.K., A.J.S. and L.S. provided guidance and consultation on the genotyping and biomarker details about ADNI data, data preprocessing, quality control, population stratification, GWAS protocol, gene analysis, network analysis and pathway analysis. All the authors reviewed, commented and approved the manuscript.

Competing interests

The authors have no actual or potential conflicts of interest including any financial, personal, or other relationships with other people or organizations that could inappropriately influence (bias) our work.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study procedures were approved by the institutional review boards of all participating centers (https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf), and written informed consent was obtained from all participants or their authorized representatives. Ethics approval was obtained from the institutional review boards of each institution involved: Oregon Health and Science University; University of Southern California; University of California—San Diego; University of Michigan; Mayo Clinic, Rochester; Baylor College of Medicine; Columbia University Medical Center; Washington University, St. Louis; University of Alabama at Birmingham; Mount Sinai School of Medicine; Rush University Medical Center; Wien Center; Johns Hopkins University; New York University; Duke University Medical Center; University of Pennsylvania; University of Kentucky; University of Pittsburgh; University of Rochester Medical Center; University of California, Irvine; University of Texas Southwestern Medical School; Emory University; University of Kansas, Medical Center; University of California, Los Angeles; Mayo Clinic, Jacksonville; Indiana University; Yale University School of Medicine; McGill University, Montreal-Jewish General Hospital; Sunnybrook Health Sciences, Ontario; U.B.C.Clinic for AD & Related Disorders; Cognitive Neurology—St. Joseph's, Ontario; Cleveland Clinic Lou Ruvo Center for Brain Health; Northwestern University; Premiere Research Inst (Palm Beach Neurology); Georgetown University Medical Center; Brigham and Women's Hospital; Stanford University; Banner Sun Health Research Institute; Boston University; Howard University; Case Western Reserve University; University of California, Davis—Sacramento; Neurological Care of CNY; Parkwood Hospital; University of Wisconsin; University of California, Irvine—BIC; Banner Alzheimer's Institute; Dent Neurologic Institute; Ohio State University; Albany Medical College; Hartford Hospital, Olin

Neuropsychiatry Research Center; Dartmouth-Hitchcock Medical Center; Wake Forest University Health Sciences; Rhode Island Hospital; Butler Hospital; UC San Francisco; Medical University South Carolina; St. Joseph's Health Care Nathan Kline Institute; University of Iowa College of Medicine; Cornell University; and University of South Florida: USF Health Byrd Alzheimer's Institute.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹College of Automation, Harbin Engineering University, 145 Nantong Street, BLDG 61-5029, Harbin 150001, China. ²Harbin Huade University, No.288 Xue Yuan Rd. Limin Development Zone, Harbin 150025, China. ³College of Information Engineering, Northeast Dianli University, 169 Changchun Street, Jilin City, Jilin 132012, China. ⁴Department of Radiology and Imaging Sciences, Indiana University School of Medicine, 355 W 16th St, Suite 4100, Indianapolis, IN 46202, USA. ⁵School of Informatics and Computing, Indiana University, 719 Indiana Avenue, Indianapolis, IN 46202, USA. ⁶Indiana University Network Science Institute, Bloomington, IN 47405, USA.

Received: 15 October 2016 Accepted: 16 May 2017

Published online: 30 May 2017

References

1. Kertes DA, Kalsi G, Prescott CA, Kuo PH, Patterson DG, Walsh D, Kendler KS, Riley BP. Neurotransmitter and neuromodulator genes associated with a history of depressive symptoms in individuals with alcohol dependence. *Alcohol Clin Exp Res.* 2011;35(3):496–505.
2. Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack Jr CR, Feldman HH, Bokde AL, Alexander GE, Scheltens P, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging.* 2012;33(7):1203–14.
3. Verbeek MM, Olde Rikkert MG. Cerebrospinal fluid biomarkers in the evaluation of Alzheimer disease. *Clin Chem.* 2008;54(10):1589–91.
4. Fernandez-Martinez M, Elcoroaristizabal Martin X, Blanco Martin E, Galdos Alcelay L, Ugarriza Serrano I, Gomez Busto F, Alvarez-Alvarez M, Molano Salazar A, Bereincua Gandarias R, Ingles Borda S, et al. Oestrogen receptor polymorphisms are an associated risk factor for mild cognitive impairment and Alzheimer disease in women APOE {varepsilon}4 carriers: a case-control study. *BMJ Open.* 2013;3(9):e003200.
5. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, Roses AD. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A.* 1993;90(5):1977–81.
6. Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort. *Neuroimage.* 2010;53(3):1051–63.
7. Blennow K, Hampel H. CSF markers for incipient Alzheimer's disease. *Lancet Neurol.* 2003;2(10):605–13.
8. Galasko D. Cerebrospinal fluid biomarkers in Alzheimer disease: a fractional improvement? *Arch Neurol.* 2003;60(9):1195–6.
9. Brys M, Pirraglia E, Rich K, Rolstad S, Mosconi L, Switalski R, Glodzik-Sobanska L, De Santi S, Zinkowski R, Mehta P, et al. Prediction and longitudinal study of CSF biomarkers in mild cognitive impairment. *Neurobiol Aging.* 2009;30(5):682–90.
10. Kauwe JS, Bailey MH, Ridge PG, Perry R, Wadsworth ME, Hoyt KL, Staley LA, Karch CM, Harari O, Cruchaga C, et al. Genome-wide association study of CSF levels of 59 Alzheimer's disease candidate proteins: significant associations with proteins involved in amyloid processing and inflammation. *PLoS Genet.* 2014;10(10):e1004758.
11. Goedert M. Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Ann N Y Acad Sci.* 1996;777:121–31.
12. Laske C, Stransky E, Fritsche A, Eschweiler GW, Leyhe T. Inverse association of cortisol serum levels with T-tau, P-tau 181 and P-tau 231 peptide levels and T-tau/Abeta 1–42 ratios in CSF in patients with mild Alzheimer's disease dementia. *Eur Arch Psychiatry Clin Neurosci.* 2009;259(2):80–5.

13. Kim S, Swaminathan S, Shen L, Risacher SL, Nho K, Foroud T, Shaw LM, Trojanowski JQ, Potkin SG, Huentelman MJ, et al. Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. *Neurology*. 2011;76(1):69–79.
14. Cruchaga C, Ebbert MT, Kauwe JS. Genetic discoveries in AD using CSF amyloid and tau. *Curr Genet Med Rep*. 2014;2(1):23–9.
15. Jia P, Zheng S, Long J, Zheng W, Zhao Z. dmGWAS: dense module searching for genome-wide association studies in protein-protein interaction networks. *Bioinformatics*. 2011;27(1):95–102.
16. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, Wu W, Uitdehaag BM, Kappos L, Gene MSAC, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet*. 2009;18(11):2078–90.
17. Peng G, Luo L, Siu H, Zhu Y, Hu P, Hong S, Zhao J, Zhou X, Reveille JD, Jin L, et al. Gene and pathway-based second-wave analysis of genome-wide association studies. *Eur J Hum Genet*. 2010;18(1):111–7.
18. Liu G, Jiang Y, Chen X, Zhang R, Ma G, Feng R, Zhang L, Liao M, Miao Y, Chen Z, et al. Measles contributes to rheumatoid arthritis: evidence from pathway and network analyses of genome-wide association studies. *PLoS One*. 2013;8(10):e75951.
19. Perez-Palma E, Bustos BI, Villaman CF, Alarcon MA, Avila ME, Ugarte GD, Reyes AE, Opazo C, De Ferrari GV, Alzheimer's Disease Neuroimaging I, et al. Overrepresentation of glutamate signaling in Alzheimer's disease: network-based pathway enrichment using meta-analysis of genome-wide association studies. *PLoS One*. 2014;9(4):e95413.
20. International Multiple Sclerosis Genetics C. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet*. 2013;92(6):854–65.
21. Wang L, Matsushita T, Madireddy L, Mousavi P, Baranzini SE. PINBPA: cytoscape app for network analysis of GWAS data. *Bioinformatics*. 2015;31(2):262–4.
22. Wang L, Mousavi P, Baranzini SE: iPINBPA: an integrative network-based functional module discovery tool for genome-wide association studies. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing* 2015:255–266.
23. Monge-Argiles JA, Munoz-Ruiz C, Pampliega-Perez A, Gomez-Lopez MJ, Sanchez-Paya J, Rodriguez Borja E, Ruiz-Vegara M, Montoya-Gutierrez FJ, Leiva-Santana C. Biomarkers of Alzheimer's disease in the cerebrospinal fluid of Spanish patients with mild cognitive impairment. *Neurochem Res*. 2011;36(6):986–93.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
25. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Investigators A, Hayward NK, Montgomery GW, Visscher PM, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet*. 2010;87(1):139–45.
26. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, et al. Human Protein Reference Database—2009 update. *Nucleic Acids Res*. 2009;37(Database issue):D767–772.
27. Goel R, Harsha HC, Pandey A, Prasad TS. Human Protein Reference Database and Human Proteinpedia as resources for phosphoproteome analysis. *Mol Biosyst*. 2012;8(2):453–63.
28. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*. 2013;14:128.
29. Tran H, Turingan I, Tran K, Zajkowska M, Lam V, Hawthorne W. Potential benefits of single-port compared to multiport laparoscopic inguinal herniorrhaphy: a prospective randomized controlled study. *Hernia*. 2014;18(5):731–44.
30. Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, Risacher SL, Nho K, Huentelman MJ, Craig DW, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. *Alzheimer's Dement*. 2010;6(3):265–73.
31. Shen L, Thompson PM, Potkin SG, Bertram L, Farrer LA, Foroud TM, Green RC, Hu X, Huentelman MJ, Kim S, et al. Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. *Brain Imaging Behav*. 2014;8(2):183–207.
32. Li J, Zhang Q, Chen F, Yan J, Kim S, Wang L, Feng W, Saykin AJ, Liang H, Shen L. Genetic Interactions Explain Variance in Cingulate Amyloid Burden: An AV-45 PET Genome-Wide Association and Interaction Study in the ADNI Cohort. *Biomed Res Int*. 2015;2015:647389.
33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904–9.
34. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–59.
35. Kohler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet*. 2008;82(4):949–58.
36. Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*. 2002;18 Suppl 1:S233–240.
37. Whitlock MC. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J Evol Biol*. 2005;18(5):1368–73.
38. Demchak B, Hull T, Reich M, Liefeld T, Smoot M, Ideker T, Mesirov JP. Cytoscape: the network visualization tool for GenomeSpace workflows. *F1000Res*. 2014;3:151.
39. Arakawa K, Kono N, Yamada Y, Mori H, Tomita M. KEGG-based pathway visualization tool for complex omics data. In *Silico Biol*. 2005;5(4):419–23.
40. Bertram L, Lange C, Mullin K, Parkinson M, Hsiao M, Hogan MF, Schjeide BM, Hooli B, Divito J, Ionita I, et al. Genome-wide association analysis reveals putative Alzheimer's disease susceptibility loci in addition to APOE. *Am J Hum Genet*. 2008;83(5):623–32.
41. Soerensen M, Dato S, Tan Q, Thinggaard M, Kleindorp R, Beekman M, Suchiman HE, Jacobsen R, McGue M, Stevnsner T, et al. Evidence from case-control and longitudinal studies supports associations of genetic variation in APOE, CETP, and IL6 with human longevity. *Age*. 2013;35(2):487–500.
42. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet*. 2011;12(1):56–68.
43. Wang X, Chen Y, Wang X, Lu L. Genetic regulatory network analysis for app based on genetical genomics approach. *Exp Aging Res*. 2010;36(1):79–93.
44. Benussi L, Govoni S, Gasparini L, Binetti G, Trabucchi M, Bianchetti A, Racchi M. Specific role for protein kinase C alpha in the constitutive and regulated secretion of amyloid precursor protein in human skin fibroblasts. *Neurosci Lett*. 1998;240(2):97–101.
45. Culmsee C, Mattson MP. p53 in neuronal apoptosis. *Biochem Biophys Res Commun*. 2005;331(3):761–77.
46. Xiao Q, Liu ZJ, Tao S, Sun YM, Jiang D, Li HL, Chen H, Liu X, Lapin B, Wang CH, et al. Risk prediction for sporadic Alzheimer's disease using genetic risk score in the Han Chinese population. *Oncotarget*. 2015;6(35):36955–64.
47. Tso SC, Qi X, Gui WJ, Wu CY, Chuang JL, Wernstedt-Asterholm I, Morlock LK, Owens KR, Scherer PE, Williams NS, et al. Structure-guided development of specific pyruvate dehydrogenase kinase inhibitors targeting the ATP-binding pocket. *J Biol Chem*. 2014;289(7):4432–43.
48. Ito S, Sawada M, Haneda M, Ishida Y, Isobe K. Amyloid-beta peptides induce several chemokine mRNA expressions in the primary microglia and Ra2 cell line via the PI3K/Akt and/or ERK pathway. *Neurosci Res*. 2006;56(3):294–9.
49. Stuart MJ, Singhal G, Baune BT. Systematic Review of the Neurobiological Relevance of Chemokines to Psychiatric Disorders. *Front Cell Neurosci*. 2015;9:357.
50. Surmiak M, Kaczor M, Sanak M. Proinflammatory genes expression in granulocytes activated by native proteinase-binding fragments of anti-proteinase 3 IgG. *J Physiol Pharmacol*. 2015;66(4):609–15.
51. Zhu B, Chen C, Xue G, Lei X, Li J, Moyzis RK, Dong Q, Lin C. The GABRB1 gene is associated with thalamus volume and modulates the association between thalamus volume and intelligence. *Neuroimage*. 2014;102(Pt 2):756–63.
52. Mueller TM, Remedies CE, Haroutunian V, Meador-Woodruff JH. Abnormal subcellular localization of GABAA receptor subunits in schizophrenia brain. *Transl Psychiatry*. 2015;5:e612.
53. Caltagarone J, Jing Z, Bowser R. Focal adhesions regulate Abeta signaling and cell death in Alzheimer's disease. *Biochim Biophys Acta*. 2007;1772(4):438–45.
54. Abbott NJ, Patabendige AA, Dolman DE, Yusof SR, Begley DJ. Structure and function of the blood-brain barrier. *Neurobiol Dis*. 2010;37(1):13–25.
55. Sastry PS, Rao KS. Apoptosis and the nervous system. *J Neurochem*. 2000;74(1):1–20.
56. Armbrrecht HJ, Siddiqui AM, Green M, Farr SA, Kumar VB, Banks WA, Patrick P, Shah GN, Morley JE. Antisense against Amyloid-beta Protein Precursor Reverses Memory Deficits and Alters Gene Expression in Neurotropic and Insulin-Signaling Pathways in SAMP8 Mice. *J Alzheimers Dis*. 2015;46(2):535–48.
57. Moran LB, Graeber MB. Towards a pathway definition of Parkinson's disease: a complex disorder with links to cancer, diabetes and inflammation. *Neurogenetics*. 2008;9(1):1–13.
58. Paulsen RD, Soni DV, Wollman R, Hahn AT, Yee MC, Guan A, Hesley JA, Miller SC, Cromwell EF, Solow-Cordero DE, et al. A genome-wide siRNA

- screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell*. 2009;35(2):228–39.
59. Schaffer BA, Bertram L, Miller BL, Mullin K, Weintraub S, Johnson N, Bigio EH, Mesulam M, Wiedau-Pazos M, Jackson GR, et al. Association of GSK3B with Alzheimer disease and frontotemporal dementia. *Arch Neurol*. 2008;65(10):1368–74.
 60. Lee L, Dale E, Staniszewski A, Zhang H, Saeed F, Sakurai M, Fa M, Orozco I, Michelassi F, Akpan N, et al. Regulation of synaptic plasticity and cognition by SUMO in normal physiology and Alzheimer's disease. *Sci Rep*. 2014;4:7190.
 61. Jicha GA, Weaver C, Lane E, Vianna C, Kress Y, Rockwood J, Davies P. cAMP-dependent protein kinase phosphorylations on tau in Alzheimer's disease. *J Neurosci*. 1999;19(17):7486–94.
 62. Esteras N, Munoz U, Alquezar C, Bartolome F, Bermejo-Pareja F, Martin-Quero A. Altered calmodulin degradation and signaling in non-neuronal cells from Alzheimer's disease patients. *Curr Alzheimer Res*. 2012;9(3):267–77.
 63. Mondragon-Rodriguez S, Trillaud-Doppia E, Dudilot A, Bourgeois C, Lauzon M, Leclerc N, Boehm J. Interaction of endogenous tau protein with synaptic proteins is regulated by N-methyl-D-aspartate receptor-dependent tau phosphorylation. *J Biol Chem*. 2012;287(38):32040–53.
 64. Matsui T, Ramasamy K, Ingelsson M, Fukumoto H, Conrad C, Frosch MP, Irizarry MC, Yuan J, Hyman BT. Coordinated expression of caspase 8, 3 and 7 mRNA in temporal cortex of Alzheimer disease: relationship to formic acid extractable abeta42 levels. *J Neuropathol Exp Neurol*. 2006;65(5):508–15.
 65. Zhang D, Hu X, Qian L, Chen SH, Zhou H, Wilson B, Miller DS, Hong JS. Microglial MAC1 receptor and PI3K are essential in mediating beta-amyloid peptide-induced microglial activation and subsequent neurotoxicity. *J Neuroinflammation*. 2011;8(1):3.
 66. Tezuka Y, Okada M, Tada Y, Yamauchi J, Nishigori H, Sanbe A. Regulation of neurite growth by inorganic pyrophosphatase 1 via JNK dephosphorylation. *PLoS One*. 2013;8(4):e61649.
 67. Turunc Bayrakdar E, Uyanikgil Y, Kanit L, Koylu E, Yalcin A. Nicotinamide treatment reduces the levels of oxidative stress, apoptosis, and PARP-1 activity in Abeta(1–42)-induced rat model of Alzheimer's disease. *Free Radic Res*. 2014;48(2):146–58.
 68. Adachi A, Kano F, Saido TC, Murata M. Visual screening and analysis for kinase-regulated membrane trafficking pathways that are involved in extensive beta-amyloid secretion. *Genes Cells*. 2009;14(3):355–69.
 69. Bengoechea TG, Chen Z, O'Leary DA, Masliah E, Lee KF. p75 reduces beta-amyloid-induced sympathetic innervation deficits in an Alzheimer's disease mouse model. *Proc Natl Acad Sci U S A*. 2009;106(19):7870–5.
 70. Kukar T, Murphy MP, Eriksen JL, Sagi SA, Weggen S, Smith TE, Ladd T, Khan MA, Kache R, Beard J, et al. Diverse compounds mimic Alzheimer disease-causing mutations by augmenting Abeta42 production. *Nat Med*. 2005;11(5):545–50.
 71. Legati A, Giovannini D, Nicolas G, Lopez-Sanchez U, Quintans B, Oliveira JR, Sears RL, Ramos EM, Spiteri E, Sobrido MJ, et al. Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. *Nat Genet*. 2015;47(6):579–81.
 72. Ramanan VK, Risacher SL, Nho K, Kim S, Swaminathan S, Shen L, Foroud TM, Hakonarson H, Huentelman MJ, Aisen PS, et al. APOE and BChE as modulators of cerebral amyloid deposition: a florbetapir PET genome-wide association study. *Mol Psychiatry*. 2014;19(3):351–7.
 73. Barreda EG, Avila J. Tau regulates the subcellular localization of calmodulin. *Biochem Biophys Res Commun*. 2011;408(3):500–4.
 74. He MX, He YW. c-FLIP protects T lymphocytes from apoptosis in the intrinsic pathway. *J Immunol*. 2015;194(7):3444–51.
 75. Pianu B, Lefort R, Thuilliere L, Tabourier E, Bartolini F. The Abeta(1)–(4)(2) peptide regulates microtubule stability independently of tau. *J Cell Sci*. 2014;127(Pt 5):1117–27.
 76. Nguyen TV, Shen L, Vander Griend L, Quach LN, Belichenko NP, Saw N, Yang T, Shamloo M, Wyss-Coray T, Massa SM, et al. Small molecule p75NTR ligands reduce pathological phosphorylation and misfolding of tau, inflammatory changes, cholinergic degeneration, and cognitive deficits in AbetaPP(L/S) transgenic mice. *J Alzheimers Dis*. 2014;42(2):459–83.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

