MEDICAL SCIENCE MONITOR

META-ANALYSIS

e-ISSN 1643-3750
© Med Sci Monit, 2017; 23: 2179-2188
DOI: 10.12659/MSM.900929

Received: 2016.08.05
Accepted: 2016.09.26
Published: 2017.05.08

A Novel Method to Identify Differential Pathways in Hippocampus Alzheimer's Disease

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

BE 1 Chun-Han Liu
AC 2 Lian Liu

1 Department of Neurology, Tianjin First Central Hospital, Tianjin, P.R. China
2 Department of Interventional Neuroradiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, P.R. China

Corresponding Author: Lian Liu, e-mail: lianliu2016@163.com
Source of support: Departmental sources

Background: Alzheimer's disease (AD) is the most common type of dementia. The objective of this paper is to propose a novel method to identify differential pathways in hippocampus AD.

Material/Methods: We proposed a combined method by merging existed methods. Firstly, pathways were identified by four known methods (DAVID, the neaGUI package, the pathway-based co-expressed method, and the pathway network approach), and differential pathways were evaluated through setting weight thresholds. Subsequently, we combined all pathways by a rank-based algorithm and called the method the combined method. Finally, common differential pathways across two or more of five methods were selected.

Results: Pathways obtained from different methods were also different. The combined method obtained 1639 pathways and 596 differential pathways, which included all pathways gained from the four existing methods; hence, the novel method solved the problem of inconsistent results. Besides, a total of 13 common pathways were identified, such as metabolism, immune system, and cell cycle.

Conclusions: We have proposed a novel method by combining four existing methods based on a rank product algorithm, and identified 13 significant differential pathways based on it. These differential pathways might provide insight into treatment and diagnosis of hippocampus AD.

MeSH Keywords: Alzheimer Disease • Neural Pathways • Protein Interaction Maps

Full-text PDF: http://www.medscimonit.com/abstract/index/idArt/900929

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

Background

Dementia is an umbrella term that describes a variety of diseases and conditions that develop when nerve cells in the brain die or no longer function normally, of which Alzheimer's disease (AD) is the most common type, accounting for an estimated 60% to 80% of cases [1]. The typical clinical presentation is progressive loss of memory and cognitive function, ultimately leading to a loss of independence and taking a heavy personal toll on the patient and the family [2]. The estimated annual incidence (rate of developing disease in 1 year) of AD appears to increase dramatically with age, from approximately 53 new cases/1000 people age 65 to 74 years, to 170 new cases/1000 people age 75 to 84 years, to 231 new cases/1000 people age 85 years and older (the "oldest-old") [3].

However, the cause of most AD cases is still mostly unknown, except for 1% to 5% of cases where genetic differences have been identified [4]. Mutations in one of three genes, amyloid precursor protein (*APP*), presenilins 1 (*PSEN1*), and presenilins 2 (*PSEN2*), have been identified to cause alterations in amyloid-beta (Aβ) processing and to lead to AD with complete penetrance [5,6]. The genes might be target markers for therapy of AD [7], but in complicated AD, a gene does not function alone; instead, multiple genes work together. It is important to correctly uncover and annotate all functional interactions among genes in the cell for any systems-level understanding of cellular functions [8–10]; thus, identification of pathways that are enriched in certain genes perhaps is a good way to reveal the pathological mechanism of hippocampus AD.

Currently, a variety of methods have been developed for the analysis of gene expression microarray data, but few methods exist for using these data to quantify the interrelated behavior of pathways [11]. The Database for Annotation, Visualization and Integrated Discovery (DAVID) and the neaGUI package are static, like a given Reactome pathway database, which may not reflect the specific conditioned pathway under study [12,13]. Hence, more and more researchers focus on pathways based on dynamic networks that consider network variations but produce many false-positive results resulting from various effects on expression of its interacting genes [14]. Meanwhile, identifying pathway changes or differential pathways will create an informative description of the biology that is occurring in a particular data set, making it possible to generate new hypotheses and to identify genetic signatures that provide insight into understanding, diagnosing, and treating disease [15,16].

Therefore, in the present work, we aimed to propose a novel method by combining static and dynamic methods that including DAVID, the neaGUI package, the pathway-based co-expressed method, and the pathway network approach to identify differential pathways between hippocampus AD and normal controls. Firstly, pathways were identified based on the four methods, but the results were inconsistent due to different methods. Subsequently, we combined all pathways based on a rank product (RP) algorithm and validated the feasibility of the combined method. Finally, common differential pathways among two or more of the five methods were identified, which might be potential biomarkers for detection and treatment in the progress of AD.

Material and Methods

In the present paper, a novel method to identify differential pathways in hippocampus AD is proposed, which combined together four known methods (DAVID, the neaGUI package, the pathway-based co-expressed method, and the pathway network approach) by a rank-based algorithm; the scheme flow is illustrated in Figure 1.

Gene data

Gene expression profiles

Three gene expression profiles of hippocampus AD patients and normal controls, with access numbers E-GEOD-1297 [17], E-GEOD-5281 [18,19], and E-GEOD-28146 [20], were downloaded from online ArrayExpress database. There were a total of 54 hippocampus AD samples and 30 samples from normal controls in the three data sets, and their characteristics are shown in Table 1.

Preprocessing for data

To control the qualities of data sets on the probe level, standard pre-treatments were performed for them. Firstly, background correction was carried out by using the Robust Multi-array Average (RMA) algorithm to eliminate the influence of nonspecific hybridization [21]. Subsequently, normalization was conducted according to a quantiles-based algorithm to make the distribution of probe intensities for each array in a set of arrays the same [22]. Next, we applied the Micro Array Suite (MAS) algorithm to revise perfect match and mismatch values [23]. Finally, the median polish method was selected to summarize expression values [21]. Expression structures were converted from the preprocessed data in AffyBatch formats and screened by the feature filter method to discard duplicated genes based on the genefilter package [24], and each probe ID was mapped to a gene symbol using the Annotate package [25]. A total of 12,434, 20,390, and 20,390 genes were obtained for E-GEOD-1297, E-GEOD-5281, and E-GEOD-28146, respectively.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

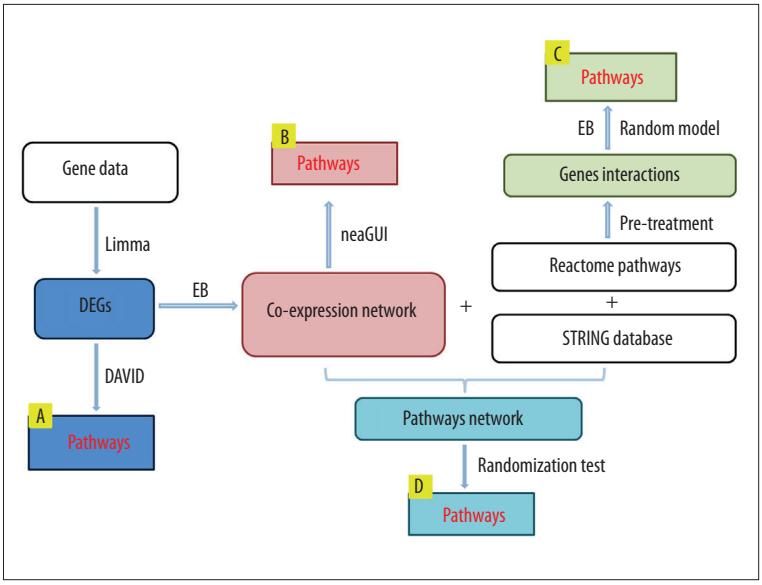

Figure 1. The scheme flow of four methods to identify differential pathways. (A) Database for Annotation, Visualization and Integrated Discovery (DAVID); (B) neaGUI; (C) the pathway-based co-expressed method; (D) the pathway network method. Figure abbreviations are as follows: Linear Models for Microarray data (Limma); differentially expressed genes (DEGs); empirical Bayes (EB); and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING).

Table 1. Characteristics of the three datasets.

Accession number	Sample size Total (disease/controls)	Platform
E-GEOD-1297	31 (22/9)	Affymetrix Human Genome U133A Array
E-GEOD-5281	23 (10/13)	Affymetrix Human Genome U133 Plus 2.0 Array
E-GEOD-28146	30 (22/8)	Affymetrix Human Genome U133 Plus 2.0 Array

Data merging

To remove the batch effects caused by the use of different experimentation plans and methodologies, we needed to merge the three preprocessed gene expression profiles into a single group. In this study, the COMBAT method in the inSilicoMerging package was utilized [26]. Measured gene expression values of gene *i* in sample *j* of the batch *k* could be expressed as:

$$x_{ij}^{k} = \alpha_i + C\beta_i + \gamma_i^k + \delta_i^k \varepsilon_{ij}^k$$

Where α_i was the overall gene expression, *C* was a design matrix for sample conditions, β_i was the vector of regression coefficients corresponding to *X*, γ_i^k and δ_i^k were the additive and multiplicative batch effects for gene *i* in batch *k*, respectively, and ε_{ij}^k were error terms. The merged data set contained 12,434 genes and was applied for further analysis of hippocampus AD samples.

Pathway data

Information from gene sets representing biological pathways was downloaded from the Reactome pathway database, and we acquired 1675 pathways. To make pathways more reliable and stable, we removed pathways with a gene number <2, and a total of 1639 pathways were obtained, which were denoted as background pathways. Reactome is an online curated resource for human pathway data and provides infrastructure for computation across the biologic reaction network [27].

Identification of pathways based on four methods

In this paper, we utilized four existing methods to explore differential pathways in hippocampus AD, which included DAVID, the neaGUI package, the pathway-based co-expressed method, and the pathway network approach.

DAVID

DAVID provides exploratory visualization tools that promote discovery through functional classification, biochemical pathway maps, and conserved protein domain architectures [12]. Before detecting pathways according to DAVID, differentially expressed genes (DEGs) between hippocampus AD and normal controls had to be evaluated. The Linear Models for Microarray data (Limma) package was selected to performed the evaluation [28], and genes that met the thresholds of $P<0.05$ and $|\log_2 \text{FoldChange}| > 2$ were identified as DEGs.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

Reactome pathway enrichment analysis for DEGs was performed by using the online tool DAVID, which implemented the Expression Analysis Systematic Explorer (EASE) test to calculate the P value for each pathway. EASE is an easy-to-use, customizable tool that allows investigators to systematically mine the mass of functional information associated with data generated by microarray [29]. The principle of EASE is shown as follows:

$$p = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{n}{a+c}}$$

Where $n=a'+b+c+d$ was the number of background genes; a' was the gene number of one gene set in the gene lists, and was replaced with $a=a'-1$; $a'+b$ was the number of genes in the gene list including at least one gene set; and $a'+c$ was the gene number of one gene list in the background genes. If there was one pathway of P<0.05, we considered it to be a differential pathway in hippocampus AD. Besides, the P value of a pathway was defined as its weight value in this paper.

NeaGUI package

NeaGUI is a user-friendly package developed to implement a method of network enrichment analysis (NEA) that extends the overlap statistics in gene enrichment analysis (GEA) to network links between genes in the experimental set and those in the functional categories [13]. The basis of this method was the network; hence a co-expression network for DEGs of hippocampus AD was constructed based on the empirical Bayes (EB) approach [30]. Expression values of DEGs were displayed as an *m*-by-*n* matrix, where *m* was the number of genes under consideration and *n* was the total number of microarrays over all conditions. Then these values were normalized and matrix *X* was obtained, of which the members took values in 1-*K*, where *K* was the total number of conditions; equal co-expression (EC)/differential co-expression (DC) classes also were defined. Based on *X*, we calculated intra-group correlations for all $l=m*(m-1)/2$ gene pairs; the resulting *Y* matrix of correlations was *l*-by-*K*. The Mclust algorithm [31] was employed to initialize the hyper-parameters to find the component normal mixture model that best fitted the correlations of *Y* after transformation. Finally, those who met a soft threshold of false discovery rate (FDR) ≤0.05 were selected to construct the co-expression network.

The neaGUI was composed of three parts: input specification, permutation, and output. The main input of neaGUI was a list of altered gene sets, which was usually a list of DEGs and gene network links. The gene network input could be a vector of gene pairs or a list representing the network link. In the case of vectors, each element had a combined name of two gene symbols with space separation. To specify known gene sets and provide detailed information about the pathways, the Reactome pathway database of group of genes describing their biological activities was utilized [27]. The default number of permutations was 100; note that this permutation procedure only needed to be performed once for each network list. The output contained the number of observed and expected network links, the number of genes, z-score, P-values based on network permutations, and the FDR, whereby pathways with P<0.05 were differential pathways. In addition, the P value of one pathway was regarded as its weight value.

Pathway-based co-expressed method

For this method, we performed co-expressed analysis for pathway genes and built a random model to screen pathways. In the first step, we calculated mean gene number (*G*) for 1639 background pathways. *G*=total gene number for pathways (73,099)/pathway number (1639)=44.6; here, we took *G*=44 for convenience. Secondly, for each pathway, its gene number was denoted by *A*, and the amount of its intersection with gene expression data was denoted by *B*; we selected those that satisfied $B>5$ and $B/A>0.5$ as the pathways for the present method, and 1271 pathways were gained. In the third step, the EB approach was implemented to conduct co-expressed analysis for genes of each pathway, and the number of possible constructed gene-gene interactions in one pathway was represented by *C*, which equaled $C=A*(A-1)/2$. Selecting FDR ≤0.05, we obtained the co-expressed gene interactions, and its amount was marked as *D*, and *D*/*C* was defined as the weight value for this pathway. Finally, to determine the differential pathways, we constructed a random model that comprised *G* genes extracted from gene expression data randomly. For the *G* genes, we carried out same operations in the third step, and obtained weight values for pathways. Capturing 10,000 times at random, 10,000 weight values were obtained, and we ranked them in descending order, set FDR for the fiftieth pathway to 0.05 (weight=0.105), and pathways with weight >0.105 were differential pathways.

Pathway network approach

A total of 787,896 human protein-protein interactions (PPIs) were downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database. Extracting interactions that contained genes of 1639 background pathways, PPIs for each pathway were obtained, and we called them the pathway network. Meanwhile, by taking intersections between the EB co-expression network and the pathway network, we gained the amount of intersections, count(*i*), where *i* stood for the *i*th pathway. To evaluate the weight value for each pathway, a random network with 11,187 edges (the same as the co-expression network) was constructed; these edges were captured from the possible PPIs built by 370 DEGs (370*369/2=68,265)

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

Table 2. Differential pathways with P<0.05 based on the Database for Annotation, Visualization and Integrated Discovery (DAVID).

Pathway	P value
HIV Infection	2.30E-08
Signaling by Wnt	2.21E-07
Integration of energy metabolism	5.47E-07
Regulation of activated PAK-2p34 by proteasome mediated degradation	5.83E-07
Metabolism of amino acids and derivatives	1.75E-06
Cdc20: Phospho-APC/C mediated degradation of Cyclin A	2.79E-05
Metabolism of carbohydrates	1.28E-04
Apoptosis	5.60E-04
DNA replication	7.40E-04
Cell cycle checkpoints	2.25E-03
Cell cycle checkpoints mitotic	5.30E-03
Pyruvate metabolism and citric acid (TCA) cycle	9.03E-03

and took intersections with each pathway, respectively. The random network was constructed 10,000 times in the above way, and we could obtain the intersection number between each random network and pathway, count(*ij*). For each pathway, the weight (*W*) was calculated as following:

$$W = \frac{\sum count(ij) \geq count(i)}{10000}$$

Where *i* stood for the *i*th pathway, and *j* represented the *j*th random network. The pathway with a smaller weight value (W<0.05) might be a differential pathway in hippocampus AD.

Combination of the four methods

The weight of each pathway was obtained by the above four methods. Considering that there were differences in the results obtained by different approaches, all weights should be analyzed further to keep the results uniform at the same standard. Therefore, the rank products (RP) algorithm was utilized to convert the weights of pathways in our study [32]. Consider a microarray experiment with two replicates (*A* and *B*), each examining *n* genes. In that case, the RP for a certain pathway *g* would be

$$RP_g = (rank_g^{replicate\ A} / n) \times (rank_g^{replicate\ B} / n)$$

This could be interpreted as a P value (=RP value), as it described the probability of observing pathway *g* at a certain rank

Table 3. Top 20 differential pathways with P=1.98E-02 based on neaGUI package.

Pathway	Intersected amount with DEGs
Immune system	45
Disease	41
Gene expression	37
Metabolism	37
Adaptive immune system	34
Infectious disease	34
Innate immune system	32
Cell cycle	31
Cell cycle mitotic	27
HIV infection	24
Metabolism of proteins	24
Signaling by Wnt	23
Host Interactions of HIV factors	22
Class I MHC mediated antigen processing & presentation	19
Metabolism of amino acids and derivatives	19
Antigen processing: Ubiquitination & proteasome degradation	18
M Phase	18
Apoptosis	17
Axon guidance	17
Downstream signaling events of B Cell receptor (BCR)	17

($rank_g^{replicate\ A}$) or better in the first replicate and at another rank ($rank_g^{replicate\ B}$) or better in the second replicate.

Subsequently, for each pathway *g*, one could also calculate a conservative estimate of the percentage of false-positives (PFPs) if this pathway (and all pathways with RP values smaller than this cutoff) would be considered as significantly differentially expressed:

$$q_g = \frac{E(RP_g)}{rank(g)}$$

Here, rank(*g*) denoted the position of pathway *g* in a list of all pathways sorted by increasing RP value; i.e., it was the number of pathways accepted as significantly regulated. The q_g was considered as the weight value for pathway in the combined method, and pathways that met a q_g value <0.05 were defined as differential pathways.

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

Table 4. Top 20 differential pathways according to pathway co-expressed method.

Pathway	Weight (W)
RSK activation	0.467
Synthesis of 12-eicosatetraenoic acid derivatives	0.460
Hormone ligand-binding receptors	0.444
Synthesis of PIPs at the late endosome membrane	0.356
Uptake and function of anthrax toxins	0.335
CYP2E1 reactions	0.334
Highly calcium permeable nicotinic acetylcholine receptors	0.333
Regulation of signaling by NODAL	0.332
mTORC1-mediated signalling	0.289
S6K1-mediated signalling	0.288
Anchoring fibril formation	0.287
Crosslinking of collagen fibrils	0.286
Ligand-independent caspase activation via DCC	0.285
Formation of ATP by chemiosmotic coupling	0.267
Release of eIF4E	0.267
Ligand-gated ion channel transport	0.262
Xenobiotics	0.255
Viral mRNA translation	0.252
Synthesis of PIPs at the Golgi membrane	0.248
Glucocorticoid biosynthesis	0.238

Results

DAVID

A total of 370 DEGs were identified between hippocampus AD and normal controls based on the Limma package with the thresholds of P<0.05 and $|\log_2 FoldChange| > 2$. Based on DAVID, 35 pathways were identified, among which 12 were differential pathways with P<0.05 (Table 2). The most significant five pathways were HIV infection (P=2.30E-08), signaling by Wnt (P=2.21E-07), integration of energy metabolism (P=5.47E-07), regulation of activated PAK-2p34 by proteasome mediated degradation (P=5.83E-07), and metabolism of amino acids and derivatives (P=1.75E-06).

Table 5. Top 20 differential pathways based on pathway network analysis.

Pathway	W value	Count
Metabolism	0	16
Disease	0	12
Immune system	0	12
Adaptive immune system	0	9
Gene expression	0	9
Innate immune system	0	9
Infectious disease	0	8
Cell cycle	0	7
HIV infection	0	6
Host interactions of HIV factors	0	6
Metabolism of proteins	0	6
Signaling by Rho GTPases	0	6
Apoptosis	0	5
Cell cycle checkpoints	0	5
Programmed cell death	0	5
beta-catenin independent WNT signaling	0	4
PCP/CE pathway	0	4
Regulation of mRNA stability by proteins that bind AU-rich elements	0	4
Diseases of signal transduction	0.001	4
RHO GTPase effectors	0.001	4

NeaGUI package

There were 358 nodes and 11,187 edges in the co-expression network that was constructed by the EB approach. When inputting the network into the neaGUI package, 1606 pathways were obtained in total. Among them, 923 were differential pathways with P<0.05; in a detailed analysis, P values of 842 pathways were the same (P=1.98E-02), and the others were 3.96E-02. To further investigate properties of pathways with the same P values, we examined their interactions with DEGs: the larger the number of intersected DEGs in one pathway, the more significant the pathway was. Table 3 lists the top 20 pathways with P=1.98E-02. Immune system, disease, gene expression, metabolism, and adaptive immune system were the top 5 pathways, including more DEGs (45, 41, 37, 37,and 34, respectively).

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

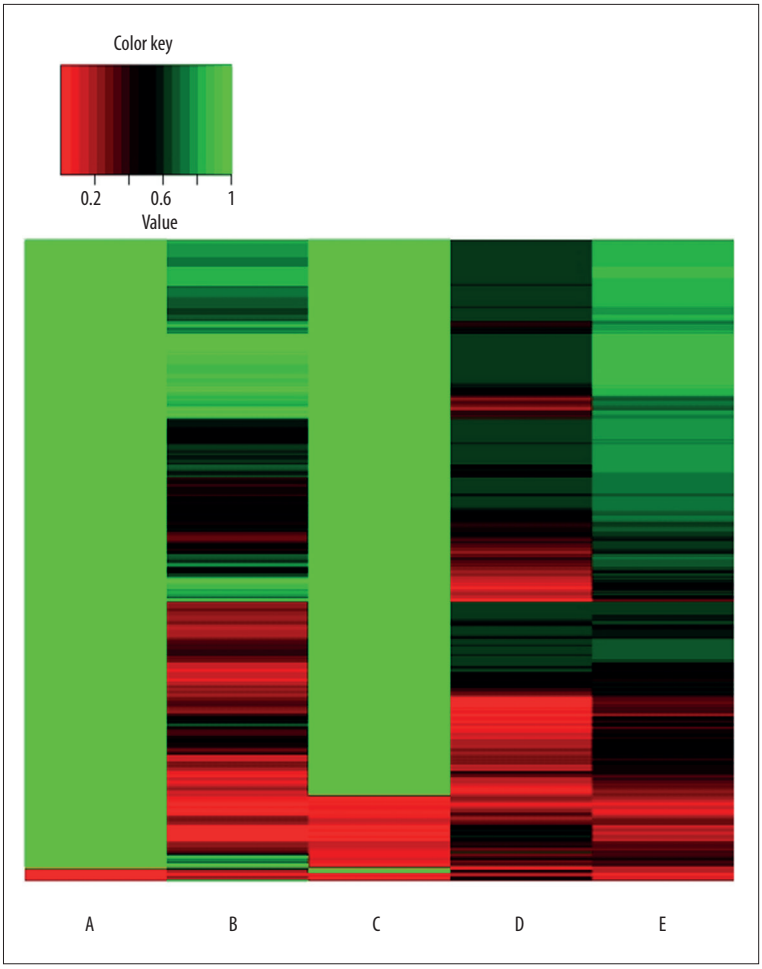

Figure 2. The heatmap for rank product (RP) values of 1639 pathways obtained from the five methods: Database for Annotation, Visualization and Integrated Discovery (DAVID) (A), the neaGUI package (B), the pathway-based co-expressed method (C), the pathway network approach (D), and the combined method (E).

Pathway-based co-expressed method

This method mainly focused on pathways genes and employed EB to perform co-expressed interactions among these genes, and differential pathways were identified using a random model for pathways. The results showed that there were 1271 pathways for hippocampus AD in total, of which 174 were differential pathways with weight >0.105. As shown in Table 4, RSK activation (W=0.467), synthesis of 12-eicosatetraenoic acid derivatives (W=0.460), hormone ligand-binding receptors (W=0.444), synthesis of PIPs at the late endosome membrane (W=0.356), and uptake and function of anthrax toxins (W=0.335) were the most significant differential pathways.

Pathway network approach

We constructed a pathway network based on STRING PPIs, background pathways, and the co-expression network. There were 206 pathways, of which 156 pathways met W<0.05 and were differential pathways. The top 20 differential pathways are displayed in Table 5. The results showed that 18 of 20 had W=0, perhaps because the weight values were too small to calculate and were approximatively taken as 0. Fortunately, the count value was another measure to detect differential pathways, and top 5 differential pathways were metabolism (count=16), disease (count=12), immune system (count=12), adaptive immune system (count=9), and gene expression (count=9).

Combined method

By combining the above four methods utilizing the RP algorithm, we obtained a total of 1639 pathways. The heat map between DAVID, the neaGUI package, the pathway-based co-expressed method, the pathway network approach, and the combined method is illustrated in Figure 2. We found that the distribution of RP values for the 1639 pathways obtained from the combined method was more even than the others, which suggested that the combined method was applicable to more pathways. Among the 1639 pathways, we identified 596 differential pathways, which contained all pathways identified by the four methods. It indicated that the combined method could avoid inconsistent results caused by different methods and be used for identifying differential pathways in hippocampus AD. Interestingly, we found that there were common

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

Table 6. Common differential pathways based on the five methods.

Pathway	Methods				
	DAVID	neaGUI package	Pathway based co-expressed	Pathway network	Combined
Metabolism	√	√	√	√	√
Immune system	√	√	√	√	√
Cell cycle	√	√	√	√	√
Metabolism of proteins		√		√	√
Signal transduction			√		√
Adaptive immune system		√		√	√
Infectious disease		√		√	√
Innate immune system		√		√	√
Gene expression			√	√	√
Cell cycle checkpoints mitotic			√		√
HIV infection	√	√		√	√
Disease				√	√
Signaling by Wnt	√	√			√

"√" indicated that one pathway was identified by the method.

pathways among two, three, or four of the five methods, and these common pathways might play more significant roles than pathways only obtained from one single approach in the progression of hippocampus AD. A total of 13 common differential pathways were discovered (Table 6), especially metabolism, immune system, and cell cycle, which were the common pathways among the five methods.

Discussion

In the current study, to better understand the molecular mechanisms of hippocampus AD, a new algorithm that combined multiple existing approaches was developed. Firstly, we identified pathways in hippocampus AD according to four known methods, DAVID, the neaGUI package, the pathway based co-expressed method, and the pathway network approach, but the results were inconsistent. To solve this problem, we developed a novel method to convert and combine the weight values of pathways obtained from the four methods. Then a new matrix with a combined weight value of each pathway was produced, and we proceeded to sort again using a rank-based method. Pathways gained from the combined method included all pathways identified from the four methods, which suggested that the combined approach could settle the inconsistent results caused by different methods to a great extent.

Differential pathways obtained from different methods were also different, but there were common differential pathways across two, three, or four of the five methods that were possibly more important than other differential pathways. A total of 13 common differential pathways were found, such as metabolism, immune system, and cell cycle. In addition, metabolism of proteins was also a significant differential pathway, so we may infer that metabolism-related pathways play a key role in hippocampus AD. Besides, this metabolism pathway network also was a main component of the EB co-expression network and pathway network, which validated the feasibility of the novel method. Epidemiological evidence has implicated global disorders of metabolism (such as obesity and type II diabetes mellitus) in cognitive aging and AD [33,34]. It had been demonstrated that manipulations that improve global energy metabolism (such as caloric restriction and exercise) may be effective in attenuating the atrophy associated with brain aging and AD in humans and animals [35–37]. Suzanne and Tong [38] illustrated that AD was fundamentally a metabolic disease, especially brain metabolic dysfunction with molecular and biochemical features. In addition, glucose metabolism abnormality played a critical role in AD pathophysiological alterations through the induction of multiple pathogenic factors such as oxidative stress, mitochondrial dysfunction, and so forth [39]. Furthermore, changes in metabolites and metabolic pathways associated with AD and using multiple analytical approaches offered support for metabolomics analysis of plasma for AD diagnosis [40]. Therefore,

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

we deduced that AD was closely correlated to metabolism, and confirmed the validity of our combined method.

Among 13 common differential pathways, 3 (immune system, adaptive immune system, and innate immune system) were related to the immune system. An increasing number of studies on AD have reported alterations in systemic immune responses, including changes in lymphocyte and macrophage distribution and activation, the presence of autoantibodies, or abnormal inflammatory factors and cytokine production [41–45]. For instance, Parker et al. demonstrated that peripheral immune response may be stronger in later stages of AD pathophysiology, when dementia had developed [42]. Zhang et al. suggested that AD pathogenesis might be influenced by systemic immunologic dysfunction and provided potential immunologic targets for therapeutic intervention [46]. Inflammatory components related to AD neuroinflammation include brain cells such as microglia and astrocytes, the complement system, and cytokines and chemokines [47]. Triggering receptor expressed on myeloid cells 2 protein (*TREM2*) was expressed on myeloid cells including microglia, which are a major part of the innate immune system in the central nervous system and are also involved in stimulating adaptive immunity to AD [48]. Increased expression of Toll-like receptor 2 (*TLR2*) and *TLR4* on peripheral blood mononuclear cells was detected in AD patients, which indicated that TLRs play a key role in inflammatory neurodegeneration, binding the highly hydrophobic amyloid peptides in AD [49]. Together these data have led to the speculation that AD might be a systemic inflammatory disorder resulting in changes associated with cognitive dysfunction.

Conclusions

In conclusion, we have proposed a novel method by combining DAVID, the neaGUI package, the pathway-based co-expressed method, and the pathway network approach utilizing the RP-based algorithm, validated its feasibility, and identified 13 significant differential pathways based on it. These differential pathways might provide insight into treatment and diagnosis of hippocampus AD.

Statement

This research received no specific grants from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest

We declare that we have no competing interests.

References:

1. Alzheimer's Association: 2013 Alzheimer's disease facts and figures. Alzheimers Dement, 2013; 9: 208–45
2. Bateman RJ, Xiong C, Benzinger TL et al: Clinical and biomarker changes in dominantly inherited Alzheimer's disease. New Engl J Med, 2012; 367: 795–804
3. Hebert LE, Beckett LA, Scherr PA, Evans DA: Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050. Alzheimer Dis Assoc Disord, 2001; 15: 169–73
4. McKhann GM, Knopman DS, Chertkow H et al: The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement, 2011;7: 263-9.
5. Waring SC, Rosenberg RN: Genome-wide association studies in Alzheimer disease. Arch Neurol, 2008; 65: 329–34
6. Wang Z, Jiang Y, Wang X et al: Butyrylcholinesterase K variant and Alzheimer's disease risk: a meta-analysis. Med Sci Monit, 2015; 21: 1408–13
7. Gao L, Yan Z, Deng J et al: Polymorphisms ofCHATbut notTFAMorVR22are associated with Alzheimer disease risk. Med Sci Monit, 2016; 22: 1924–35
8. Szklarczyk D, Franceschini A, Kuhn M et al: The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res, 2011; 39: D561–68
9. Bao J, Wang XJ, Mao ZF. Associations between genetic variants in 19p13 and 19q13 regions and susceptibility to Alzheimer disease: A meta-analysis. Med Sci Monit, 2015; 22: 234–43
10. Zhang T, Jia Y: Meta-analysis of ubiquilin1 gene polymorphism and Alzheimer's disease risk. Med Sci Monit, 2014; 20: 2250–55
11. Skafidas E, Testa R, Zantomio D et al: Predicting the diagnosis of autism spectrum disorder using gene pathway analysis. Mol Psychiatry, 2014; 19: 504–10
12. da Huang W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc, 2008; 4: 44–57
13. Pramana S: neaGUI: An R package to perform the network enrichment analysis (NEA). R package version 2013; 1
14. Wu C, Zhu J, Zhang X: Integrating gene expression and protein-protein interaction network to prioritize cancer-associated genes. BMC Bioinformatics, 2012; 13: 182
15. Doniger S, Salomonis N, Dahlquist K et al: MAPPFinder: Using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. Genome Biol, 2003; 4: R7
16. Dawson JA, Kendziorski C: An empirical bayesian approach for identifying differential coexpression in high-throughput experiments. Biometrics, 2012; 68: 455–65
17. Blalock EM, Geddes JW, Chen KC et al: Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proc Natl Acad Sci USA, 2004; 101: 2173–78
18. Liang WS, Dunckley T, Beach TG et al: Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol Genomics, 2007; 28: 311–22
19. Liang WS, Reiman EM, Valla J et al: Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci USA, 2008; 105: 4441–46
20. Blalock EM, Buechel HM, Popovic J et al: Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. J Chem Neuroanat, 2011; 42: 118–26
21. Irizarry RA, Bolstad BM, Collin F et al: Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res, 2003; 31: e15
22. Bolstad BM, Irizarry RA, Astrand M, Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, 2003; 19: 185–93
23. Bolstad B: affy: Built-in processing methods. 2013.
24. Lee J, Kim D-W: Efficient multivariate feature filter using conditional mutual information. Electronics Letters, 2012; 48: 161–62

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]

25. Allen JD, Wang S, Chen M et al: Probe mapping across multiple microarray platforms. Brief Bioinform, 2012; 13(5): 547–54
26. Taminau J, Meganck S, Lazar C: Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. BMC Bioinformatics, 2012; 13: 335
27. Croft D, O'Kelly G, Wu G et al: Reactome: A database of reactions, pathways and biological processes. Nucleic Acids Res, 2011; 39(Database issue): D691–97
28. Diboun I, Wernisch L, Orengo CA , Koltzenburg M: Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC Genomics, 2006; 7: 252
29. Hosack DA, Dennis G Jr., Sherman BT et al: Identifying biological themes within lists of genes with EASE. Genome Biol, 2003; 4: R70
30. Dawson JA, Ye S, Kendziorski C: R/EBcoexpress: An empirical Bayesian framework for discovering differential co-expression. Bioinformatics, 2012; 28: 1939–40
31. Fraley C, Raftery AE: Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association, 2002; 97: 611–31
32. Breitling R, Armengaud P, Amtmann A, Herzyk P: Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett, 2004; 573: 83–92
33. Luchsinger JA, Reitz C, Patel B et al: Relation of diabetes to mild cognitive impairment. Arch Neurol, 2007; 64: 570–75
34. Toro P, Schönknecht P, Schröder J: Type II diabetes in mild cognitive impairment and Alzheimer's disease: Results from a prospective population-based study in Germany. J Alzheimers Dis, 2009; 16(4): 687–91
35. Erickson KI, Prakash RS, Voss MW et al: Aerobic fitness is associated with hippocampal volume in elderly humans. Hippocampus, 2009; 19: 1030–39
36. Kannangara TS, Lucero MJ, Gil-Mohapel J et al: Running reduces stress and enhances cell genesis in aged mice. Neurobiol Aging, 2011; 32: 2279–86
37. Kapogiannis D, Mattson MP: Disrupted energy metabolism and neuronal circuit dysfunction in cognitive impairment and Alzheimer's disease. Lancet Neurol, 2011; 10: 187–98
38. Suzanne M, Tong M: Brain metabolic dysfunction at the core of Alzheimer's disease. Biochem Pharmacol, 2014; 88: 548–59
39. Chen Z, Zhong C: Decoding Alzheimer's disease from perturbed cerebral glucose metabolism: implications for diagnostic and therapeutic strategies. Prog Neurobiol, 2013; 108: 21–43
40. Trushina E, Dutta T, Persson X-MT et al: Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. PLoS One, 2013; 8(5): e63644
41. Kim S-M, Song J, Kim S et al: Identification of peripheral inflammatory markers between normal control and Alzheimer's disease. BMC Neurology, 2011; 11: 51
42. Parker DC, Mielke MM, Yu Q et al: Plasma neopterin level as a marker of peripheral immune activation in amnestic mild cognitive impairment and Alzheimer's disease. Int J Geriatr Psychiatry, 2013; 28: 149–54
43. Pellicanò M, Larbi A, Goldeck D et al: Immune profiling of Alzheimer patients. J Neuroimmunol, 2012; 242: 52–59
44. Hochstrasser T, Marksteiner J, Defrancesco M et al: Two blood monocytic biomarkers (CCL15 and p21) combined with the mini-mental state examination discriminate Alzheimer's disease patients from healthy subjects. Dement Geriatr Cogn Dis Extra, 2011; 1(1): 297–309
45. Pellicano M, Bulati M, Buffa S et al: Systemic immune responses in Alzheimer's disease: *In vitro* mononuclear cell activation and cytokine production. J Alzheimers Dis, 2010; 21: 181
46. Zhang R, Miller RG, Madison C et al: Systemic immune system alterations in early stages of Alzheimer's disease. J Neuroimmunol, 2013; 256: 38–42
47. Rubio-Perez JM, Morillas-Ruiz JM: A review: Inflammatory process in Alzheimer's disease, role of cytokines. ScientificWorldJournal, 2012; 2012: 756357
48. Boutajangout A, Wisniewski T: The innate immune system in Alzheimer's disease. Int J Cell Biol, 2013; 2013: 576383
49. Zhang W, Wang L-Z, Yu J-T et al: Increased expressions of TLR2 and TLR4 on peripheral blood mononuclear cells from patients with Alzheimer's disease. J Neurol Sci, 2012; 315: 67–71

Indexed in: [Current Contents/Clinical Medicine] [SCI Expanded] [ISI Alerting System] [ISI Journals Master List] [Index Medicus/MEDLINE] [EMBASE/Excerpta Medica] [Chemical Abstracts/CAS] [Index Copernicus]