The Third International Symposium on Optimization and Systems Biology (OSB'09)
Zhangjiajie, China, September 20–22, 2009

Crosstalk of Dysfunctional Pathways in Alzheimer's Disease Brains*

Zhi-Ping Liu Luonan Chen

Department of Electrical Engineering and Electronics, Osaka Sangyo University, Osaka 574-8530, Japan

Abstract Alzheimer's disease (AD) is a major neurodegenerative disorder leading to dementia in the elderly. We proposed a network-based systems biology approach to detect the dysfunctional crosstalk of pathways in six brain regions of AD disease patients. We identified the relationship between AD pathways and its neighbor pathways to investigate their intersections and relationships. The significance of pathways developed in the fatal disorders and the overlapping strength indicate the impacts of neighbor pathways to AD development. The crosstalk with AD pathways provides evidences for the particular tasks of important neighbor pathways as well as the associations between them.

Keywords Gene expression; dysfunctional pathway; clustering; crosstalk; Alzheimer's disease

1 Introduction

Alzheimer's disease (AD) is a debilitating neurodegenerative progressive and fatal disease. The genetic mechanisms of AD are far from being clearly clarified although there are several popular hypotheses about its pathogenesis [1]. In the post-genomic era, the large-scale of genome-wide data have provided unprecedented opportunity to analyze complex diseases [2]. The recent advances in gene expression microarray technologies present genome wide scale data for AD [3]. To date, a key challenge has been to identify the biological processes or signaling pathways that play significant roles in the disorder. Improving the understanding of the complexity of molecular pathways underlying AD phenotypes is essential to uncover the dysfunctional processes of disease progression in different brain regions. Moreover, systems biology approaches such as network-based methods have been applied to elucidate the mechanism of diseases [2, 4]. The information of human transcriptome and interactome can potentially bridge the gap between genotype and phenotype [5]. The availability of protein-protein interaction data provides the screenings of protein cooperations in the cell survival and growth. The interrelated transcriptomic and interactomic information could be integrated together to investigate pathways [6] and the crosstalk of pathway interactions [7, 8].

*This work was supported by the National Natural Science Foundation of China (NSFC) under Grant No. 10631070 and No. 60873205. Part of the authors are also supported by the Grant No. 5039052006CB from the Ministry of Science and Technology, China. The work is also supported by NSFC-JSPS collaborative project No. 10711140116.

In this paper, we present a network-based analysis of the crosstalk between AD pathways and its related pathways in different AD brain regions by integrating ensemble protein interactions and region-specific gene expression profiles. First, we collect gene set of KEGG AD pathway and that of its related pathways, which contain at least one overlapping gene with KEGG AD gene set. Then we propose a scoring scheme to identify the dysfunctions of AD related protein interactions combining gene differential expression and coexpression data. The crosstalk between these pathways are measured by their overlapping significance in the specific regions. We cluster the pathways by their identified relationships each other in different brain regions individually. These related pathways are simultaneously sorted by the defined score of dysregulation. We demonstrate that the brains process complex crosstalk of pathways when affected with AD on the protein level.

2 Results

2.1 Clusters of dysregulated pathway

We build the network of pathways by integrating KEGG AD pathway and all its neighbor pathways of protein interactions. Figure 1(a) shows the network. The red part is the KEGG AD pathway. Figure 1(b) shows the global topology of AD pathway and its neighbor pathways. The red node represents the red part of the AD pathway in Figure 1(a).

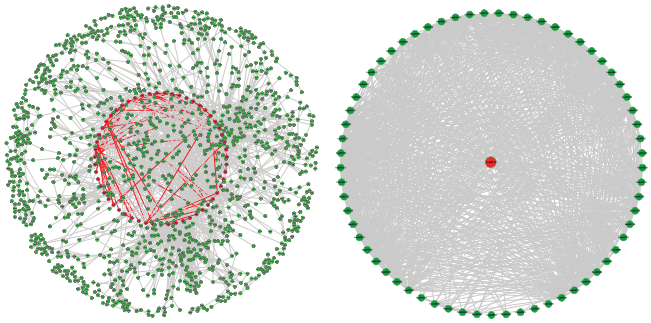

Figure 1: Network of pathways. (a) is the protein interaction network constructing KEGG AD pathway and its neighbor pathways. (b) is topological linkages between these pathways. A pathway is represented by a node.

After we weight the network by integrating gene differential expression and coexpression data in six different regions, we can evaluate the interaction significance between these pathways. The overlapping score and its significance can be used as the similarity measurements of the pathways. Figure 2 gives the clustering result of these pathways based on their interaction significance in EC region. In Figure 2, the red cluster is the module of KEGG AD pathway (ID: hsa05010) and its closest interaction pathways. Different colored clusters describe different modules of clustering. The grey color is the part of pathways which can not be grouped to any individual cluster distinctly. We can find that several pathways of neurodegenerative diseases are grouped together, such as Parkinson's disease (hsa05012), Huntington's disease (hsa05040) and Dentatorubropallidoluysian atrophy (hsa05050). This result indicates the close relationship between these

neuron diseases. The significant interaction between these pathways provides evidences for their interrelations. We also do the similar clustering in other five regions (Data is not shown). Pathways are grouped into similar clusters as that in EC region. This provides more evidences not only for the closely related neuron diseases [9], but also for the effectiveness of our method. Moreover, we identified some closed related pathways with AD in the same cluster, e.g. Oxidative phosphorylation (hsa:00190), p53 signaling pathway (hsa04115) and Apoptosis (hsa04210). Genetic evidences have supported the results that oxidative phosphorylation is closely related to AD [10]. Functionally important reductions in oxidative phosphorylation enzyme activities appear to occur in AD and may be related to β-amyloid accumulation, which is the main phenotype in AD patient brains. Thus this kind of dysfunctional defects could play an important role in the pathophysiology of AD. The critical role of p53 is evidenced by the fact that it is mutated in a very large fraction of tumors. It is an important transcriptional activator whose activity is regulated by phosphorylation. p53 is also known as a potential biomarker for AD for its special alternation presented in Aβ accumulation [11]. The signaling pathway can cause the cell to enter apoptosis process (pathway ID: hsa04210). The brains of AD patients contain dying neurons displaying apoptosis [12]. The close relationship between pathways of apoptosis and AD indicates that apoptosis plays an important role in the progression of the disease. From this viewpoint, the therapies of AD must consider the apoptosis features of the neuron cell. These clusters of pathways provide us a global relationship between these pathways.

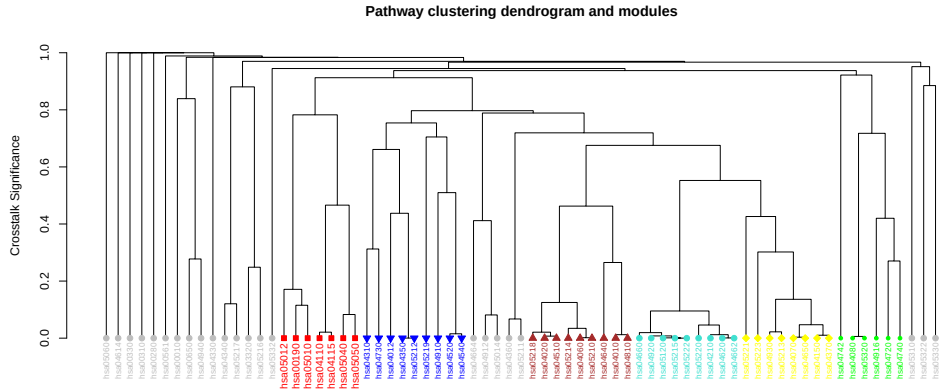

Figure 2: Clustering of the pathways in EC region. Different color represents different modules. The red one represents the cluster in which KEGG AD pathway has been grouped.

2.2 Dysfunctional crosstalk between pathways

From the clusters, we identified the modules of these pathways. The crosstalk of AD pathways not only underlies the same cluster of pathway grouped in the former section, but also underlies dysfunctional relationship of all its neighbor pathways. As to these related pathways, we also defined the dysregulated score and its significance criterion (see Methods). We identified the significance of their dysfunctions individually in different

regions. Table 1 lists the top five rank pathways in every region. The dysfunction score of every pathway neighbored to AD pathway represents the activation status during the AD progression in the form of pathway of protein cooperations. We found that Apoptosis (hsa04210), Notch signaling pathway (hsa04330), Wnt signaling pathway (hsa04310), and Cytokine-cytokine receptor interactionare (hsa04060) are listed on the top of six regions respectively. From the rank, we give a quantitative measure of the simultaneous activation when AD pathway performs its dysfunctions of toxic processes. From the pathways in the same cluster with AD pathway, we can find that some of them are also the most significant pathways. Apoptosis is the pathway not only with high interaction significance, but also with high dysfunction significance. The result gives more evidence for strong relationship between the apoptosis of the cell and AD process [12]. Especially, Notching signaling pathway is listed on the first one twice in the six regions. The pathway is crucial in communication between the cells, which involves important processes during embryonic and adult life [13]. When developing medicines for AD, the side effect for Notch signaling pathway is an important factor because it functions to inhibit the drug targets [13]. Wnt signaling pathway often involves Ca2+ signaling, which leads to transient increases in cytoplasmic free calcium that subsequently activates the calcium kinase and the phosphatase calcineurin [14]. It is now known that the Ca2+ is often crucial to AD, which is known as the calcium hypothesis [15]. If the calcium level is not properly controlled, it would lead to nerve cell dysfunction and death. Cytokines are also crucial to cell death. Proinflammatory cytokines may lead to neuronal death and dysfunction by variant mechanisms. Inhibition of cytokines has been tested as a therapy method for the treatment of AD [16]. The significant neighbored pathways indicate their crosstalk between those pathways and AD pathway. The critical influences from the neighbor pathways are identified from the dysfunctional crosstalk.

Table 1: Rank of related pathways. The top 5 pathways which significant in every region are listed individually.

Region	ID	Size	Edge	Node	P-value	Description
EC	hsa04210	89	156	72	0.000001	Apoptosis
	hsa04060	279	259	220	0.000002	Cytokine-cytokine receptor interaction
	hsa04620	102	110	69	0.000695	Toll-like receptor signaling pathway
	hsa05220	75	181	69	0.003802	Chronic myeloid leukemia
	hsa05120	68	54	40	0.003899	Epithelial cell signaling in Helicobacter pylori infection
HIP	hsa04310	152	152	83	0.000000	Wnt signaling pathway
	hsa05214	65	113	52	0.000000	Glioma
	hsa04060	279	259	220	0.000001	Cytokine-cytokine receptor interaction
	hsa04012	87	200	71	0.000002	ErbB signaling pathway
	hsa05215	90	181	72	0.000002	Prostate cancer
MTG	hsa04330	46	59	34	0.000000	Notch signaling pathway
	hsa05120	68	54	40	0.000000	Epithelial cell signaling in Helicobacter pylori infection
	hsa04310	152	152	83	0.000003	Wnt signaling pathway
	hsa04520	78	105	59	0.000024	Adherens junction
	hsa04210	89	156	72	0.000388	Apoptosis
PC	hsa04060	279	259	220	0.000000	Cytokine-cytokine receptor interaction
	hsa04640	87	49	62	0.000000	Hematopoietic cell lineage
	hsa04115	69	72	46	0.000049	p53 signaling pathway
	hsa04210	89	156	72	0.000689	Apoptosis
	hsa05217	55	33	26	0.001264	Basal cell carcinoma
SFG	hsa04330	46	59	34	0.000000	Notch signaling pathway
	hsa04510	203	340	150	0.000000	Focal adhesion
	hsa04920	67	79	39	0.000002	Adipocytokine signaling pathway
	hsa04210	89	156	72	0.000009	Apoptosis
	hsa05120	68	54	40	0.000190	Epithelial cell signaling in Helicobacter pylori infection
VCX	hsa04060	279	259	220	0.000000	Cytokine-cytokine receptor interaction
	hsa04640	87	49	62	0.000000	Hematopoietic cell lineage
	hsa04340	57	14	11	0.000121	Hedgehog signaling pathway
	hsa04510	203	340	150	0.000279	Focal adhesion
	hsa04662	65	93	51	0.002533	B cell receptor signaling pathway

As to the crosstalk of GO functional relationship between these pathways, we identified the hypergeometric significant GO functions in every pathways. The results of linkage for the top three GO terms in part of the pathways are shown in Figure 3. The most significant biological processes in AD pathway are proteolysis (GO:0006508), membrane protein ectodomain proteolysis (GO:0006509) and electron transport chain (GO:0022900). The functions correspond to the main process of Aβ accumulation the AD progression in the EC region. The significant functions of neighbor pathways provide a flow of transporting of molecules (macromolecules, small molecules, protons and other ions) of substances into, out of, within or between cells (GO:0006810, GO:0015992). Other pathways introduce one or more phosphate groups into a phosphoinositide (GO:0046854). Especially, the Apoptosis pathway (hsa04210) enrichs the conversion of proteins, and induces or sustains apoptosis to an active form (GO:0008633). From the significant GO enrichments, we know the crosstalk of GO biological processes during the disease development in EC region. The GO significance among these pathways provides high correlated functions in these pathways which provide implications that the close related pathways are crucial to the dysfunction of AD pathway during the disease progression accumulating abnormally folded A-beta and tau proteins in the brain regions.

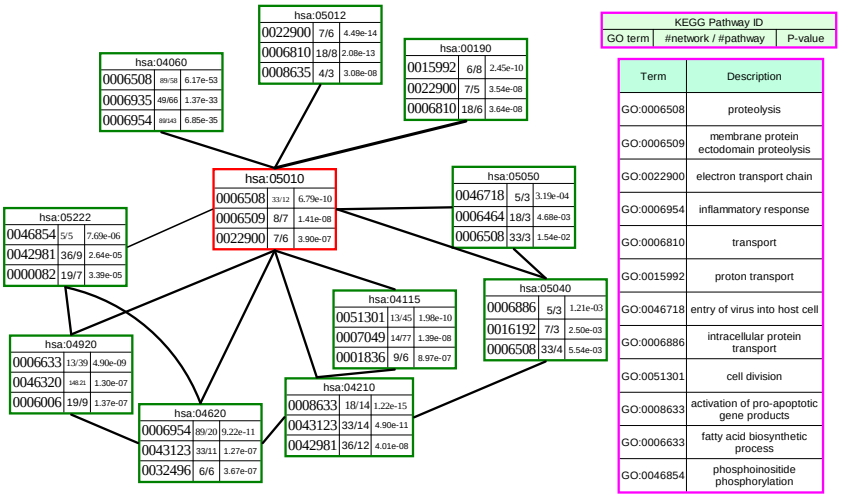

Figure 3: Top 3 GO enrichments in the crosstalk of some dysfunctional pathways.

3 Conclusion

In this work, we identified the clusters of pathways related to AD pathway in different brain regions based on a novel network-based method. In the ensemble protein network, we first build the network of pathways by integrating KEGG AD related gene sets. The crosstalk of interaction significance between these pathways and their dysfunctional score were detected by corresponding transciptome information in six disease brain regions. The analysis of dysfunctional crosstalk with AD pathway is based on the pathway clusters and all its neighbor pathways. Some of the crosstalk that we identified

between the pathway are consistent with our knowledge for AD. Some of them provides valuable alternatives for AD mechanism, especially from the pathway relationship perspective. The interaction and crosstalk of these dysfunctional pathways provides more insights for the AD progression in variant brain regions. This work provides evidence that the complication of AD needs more comprehensive and detailed analysis complementing for the traditional approaches from the systematic perspective, especially for the complex disease of neurodegenerative disorder.

4 Methods

4.1 Data sources

We downloaded the KEGG AD gene set and detected all its neighbor gene sets of pathway which have at least one overlapping gene. There are total 77 neighbored pathway gene sets in KEGG intersecting with 176 KEGG AD genes. All these gene-related proteins were obtained from NCBI (`www.ncbi.nlm.nih.gov`) and are represented by their NCBI Entrez Gene IDs. Then we construct ensemble PPI by integrating the interactions existing in three of five human PPI databases, i.e. HPRD (`www.hprd.org`), BIND (`www.bind.ca`), BioGrid (`www.thebiogrid.org`), IntAct (`www.ebi.ac.uk/intact`) and MINT (`mint.bio.uniroma2.it/mint`).

We used the data of gene expression profiling research on AD patients with normal controls in [3], which were downloaded from NCBI GEO (`www.ncbi.nlm.nih.gov/geo`) database (ID:GSE5281). These data were used to study AD patients [3] and contained six distinct regions in human brains. Temporarily, these genes with no corresponding proteins in the integrated PPI network were not be included in the analysis as well as the isolated proteins in every gene sets and those in the PPI network without a corresponding gene in the expression experiment. This ended up with an KEGG AD pathway with 54 nodes and 79 edges and the neighbor 70 pathways. All these pathways contain 1401 proteins and 2888 interactions among them.

4.2 Mapping gene expression to PPI network

We used Welch's t-test to determine the differential gene between control and disease brains. The correlated expressions were measured by Pearson correlated coefficient test of pair genes in the disease cases individually. Then we mapped the p-values to the protein network. We used Fisher's method [17] to define a function as the combination of statistical significance of an interaction by a scoring scheme in the formula: $\text{Score}(e(x,y)) = f(\text{diff}(x), \text{corr}(x,y), \text{diff}(y)) = -2\Sigma_{i=1}^{k} log_e(p_i)$, where $\text{diff}(x)$ and $\text{diff}(y)$ are differential expression p-values of node x and node y, respectively. $\text{corr}(x,y)$ represents their correlation p-value. f is a general data integration method to combine the p-values. The expressions for different specific regions of AD were mapped respectively and we obtained different condition-based weighted pathways of protein networks.

4.3 Interaction significance between pathways

We define an interaction function to evaluate the significance of all non-empty overlaps between two pathways, i.e. $O_{ij} = P_i \cap P_j$, $i, j \in S, O_{ij} \neq \varnothing$, where P_i and P_j are two pathways, and O_{ij} is their overlapping. Specifically, the interaction score between

one pathway and one of its neighbor pathways is estimated by their overlapping status of weighted pathways in the following formula:

$$C(P_i, P_j) = \sum_{e \in O_{ij}} S(e).$$

The overlapping score is the summation of the overlapping edge score between pathways. To estimate the significance of the overlapping during the different disease brain regions individually, we random sample 10^6 times of the same size two pathways in the edge sets of two pathways and calculated their overlapping scores. The frequency larger than C is used as the interaction significance p-value. From the scores, a distance matrix of these pathways are built for clustering them into pathway groups. Hierarchical clustering is implemented for grouping the pathways into clusters with dynamic hierarchy tree cut (`www.r-project.org`).

4.4 Dysregulation of related pathways

To define the dysfunction of a pathway P, we summarize all the scores of edges $S(e)$ of every pathway, i.e.

$$S_P = \sum_{e \in P} S(e).$$

To estimate a p-value for significance of this pathway, we iteratively compute similar scores 10^6 times on randomly generated pathways of the same size as that of pathway P. The frequency of scores that are larger than S_P is used as the significant p-value of pathway P to describe its dysregulation in specific region. In every region, we get a ranked list of dysregulated pathways.

References

[1] Goedert M, Spillantini MG (2006) A century of Alzheimer's disease. Science 314: 777–781

[2] Ideker T, Sharan R (2008) Protein networks in disease. Genome Res 18: 644–552

[3] Liang WS, Reiman EM, Valla J, Dunckley T, Beach TG, Grover A, Niedzielko TL, Schneider LE, Mastroeni D, Caselli R, et al (2008) Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci USA 105: 4441–4446

[4] Zhao XM, Wang RS, Chen L, Aihara K (2008) Uncovering signal transduction networks from high-throughput data by integer linear programming. Nucleic Acids Res 36:e48

[5] Wachi S, Yoneda K, Wu R (2005) Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. Bioinformatics 21: 4205–4208

[6] Ideker T, Ozier O, Schwikowski B, Siegel AF (2002) Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 18(Suppl): S233–S240

[7] Li Y, Agarwal P, Rajagopalan D (2008) A global pathway crosstalk network. Bioinformatics 24: 1442–1447

[8] Francesconi M, Remondini D, Neretti N, Sedivy JM, Cooper LN, Verondini E, Milanesi L, Castellani G (2008) Reconstructing networks of pathways via significance analysis of their intersections. BMC Bioinformatics 9(Suppl 4): S9

[9] Limviphuvadh V, Tanaka S, Goto S, Ueda K, Kanehisa M (2007) The commonality of protein interaction networks determined in neurodegenerative disorders (NDDs). Bioinformatics. 23: 2129–2138

[10] Selkoe DJ (2001) Alzheimer's disease: genes, proteins, and therapy. Physiol Rev 81: 741–766

[11] Lanni C, Uberti D, Racchi M, Govoni S, Memo M (2007) Unfolded p53: a potential biomarker for Alzheimer's disease. J Alzheimers Dis 12: 93–99

[12] Ballatore C, Lee VM, Trojanowski JQ (2007) Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. Nat Rev Neurosci 8: 663–672

[13] Bolos V, Grego-Bessa J, de la Pompa JL (2007) Notch signaling in development and cancer. Endocr Rev 28: 339–363

[14] Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36: D480–D484

[15] Cheung KH, Shineman D, Muller M, Cardenas C, Mei L, Yang J, Tomita T, Iwatsubo T, Lee VM, Foskett JK (2008) Mechanism of Ca2+ disruption in Alzheimer's disease by presenilin regulation of InsP3 receptor channel gating. Neuron 58: 871–883

[16] Tobinick E, Gross H, Weinberger A, Cohen H (2006) TNF-alpha modulation for treatment of Alzheimer's disease: a 6-month pilot study. MedGenMed 8:2

[17] Fisher RA (1948) Combining independent tests of significance. American Statistician 2(5): 30