PROTEIN-PROTEIN INTERACTION NETWORK OF ALZHEIMER’S DISEASE FROM HUMAN HIPPOCAMPAL PROTEOME

Hakimeh Zali1, Mostafa Rezaei Tavirani2*, Majid Rezaei Tavirani3

1School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
2Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
3Ilam University of Medical Sciences, Ilam, Iran
*Correspondence author: E-mail: rezaei.tavirani@ibb.ut.ac.ir.

ABSTRACT: Alzheimer’s disease (AD) is the most common dementia characterized by tangles and plaques that hippocampus as one of the primary regions of the brain affected by AD. In this study investigated network-based Protein-Protein Interactions (PPI) for hippocampal proteins that altered in AD patient's comparison with healthy control. Altered protein data were extracted from beforehand investigation which contains 245 proteins. Out of which, 105 were found to be exclusively expressed in AD; whereas 140 proteins were detected down-regulated or silenced. The differentially expressed proteins and related networks were explored using cytoscape and the PPI analysis methods such as MCODE and CluGO. AD network contains 3851 nods and 3480 edges. Important hubs are ASC, Smad8, Smad2, SnoN, KIAA1196, Smad1, RNF11, Smad3, ATF7ip and SHBG. Network analysis illustrated 22 clusters (protein complex) with distinctive seed genes. Gene ontology categories based on CluGO analysis demonstrated increasing in apoptosis, oxidoreductase activity, glutathione transferase activity and immune system process, otherwise was seen decreasing in GTPase activity and glucose metabolism. In sum up, network analysis could help to comprehend AD mechanism and discover potential biomarkers which may be helpful for diagnosis, prognosis and treatment prediction also it is necessary to find that part of these biomarkers would be able to detect in peripheral blood and cerebrospinal fluid as more accessible biomarker.

Abbreviation: Alzheimer disease(AD), Protein-Protein Interactions (PPI), Amyloid-β( Aβ), Long-term potentiation (LTP), Gene Ontology (GO), Huntington's Disease (HD), Parkinson's Disease(PD), Amyotrophic Lateral Sclerosis(ALS), N-methyl-D-aspartate (NMDA), 3-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA), Molecular Complex Detection (MCODE), Kyoto Encyclopedia of Genes and Genomes (KEGG), Toll-like receptors (TLRs), Central Nervous System(CNS)

Keywords: Alzheimer’s Disease, Protein-Protein Interactions Network, Seed, Protein Complex, Cluster, Biomarker.

INTRODUCTION: Alzheimer’s disease is the most common form of senile dementia that characterized by a progressive cognitive impairment and a brain specific neuropathology such as the abundant occurrence of amyloid β (Aβ) plaques, neurofibrillary tangles, and neuronal and synaptic loss (1). Since its internal ethology have not fulfill elucidated, it is the big challenge for researchers to clarify its mechanism by different methods on those parts of brain suffer from AD. Expression proteomics is one of these methods, even if it is in its infancy, the technology has been demanded to depict an efficient tool in discovering new biomarkers or a panel of markers for more precise diagnosis of complex human diseases. Expression proteomics has also been foresaid to be the solution in early stages diagnosis and follow-up of the progression of the AD. Since the prevalent findings for uncovering the secrets of AD pathogenesis stand in the tip of the iceberg, search for diagnostic markers and drug targets is one
of the major focuses of research in this field (2). Applying bioinformatics methods based on network analysis is the way to help finding biomarkers in neurological disease, so in this study the hippocampal proteome of AD patients will analysis by bioinformatics methods.

**HIPPOCAMPUS**

The hippocampus, well defined and confined areas of axonal sprouting and synaptic remodeling, has been determined to be an appropriate model system for the study of learning, memory and neuronal plasticity following lesions such as epilepsy, AD, and ischemic brain damage (3). Before the appearance of significant clinical symptoms, the neuropathological changes in AD are thought to begin primarily in the entorhinal and transentorhinal cortex, hippocampus, and then to progress to the association cortices of the temporal, parietal and frontal lobes (4). In particularly, the hippocampal formation is thought to be a major location of the memory impairment seen in AD (5).

**NEUROPATHOLOGY OF ALZHEIMER DISEASE**

Neuroimaging investigations and subsequent post-mortem evidences have shown that AD characterizes a loss of neurons and synapses in the cerebral cortex and certain subcortical regions. This loss leads to gross atrophy (6) that visually revealed by microscopic appearance of diffuse neuritic plaques and neurofibrillary tangles (4, 7). A major reduction in the levels of acetylcholine, a central neurotransmitter, is followed by degeneration (8). Loss of white matter, cerebral amyloid angiopathy, inflammation and oxidative damage also exist in AD brain (9).

Long-term potentiation (LTP) that shows a strong correlation with learning and memory as well as synaptic plasticity (10) has revealed significant deficits in basal synaptic transmission and/or LTP in APP-transgenic mice carrying human AD mutations. These deficits were shown to occur in the hippocampus well before the appearance of any detectable Aβ deposits (11). There is now persuasive evidence to recommend that the maintenance of hippocampal LTP is repressed by synthetic human Aβ-derived diffusible ligands (12) and soluble, low-number oligomers of naturally secreted human Aβ (13, 14). Aβ oligomers impair synaptic plasticity by altering the balance between LTP and long-term depression and by reducing the number of dendritic spines.

Furthermore, excess build-up of Aβ and Aβ oligomers can induce neurotransmitter receptor internalization [such as N-methyl-D-aspartate (NMDA) and 𝛼-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors] and inhibition of voltage-gated calcium channels, and nicotinic acetylcholine receptors (15). Over the past few decades, many studies have identified and characterized pathways within AD related to altered processing and accumulation of Aβ and tau which support the roles of mitochondrial dysfunction (altered energy and oxidative metabolism, altered antiapoptotic response), endoplasmic reticulum stress (unfolded protein response), inflammation (imbalance in oxidative metabolism and glutamate excitotoxicity), and systemic alterations in Ca2+ homeostasis leading to cell death in AD (16,17).

**GENETICS AND PROTEOMICS IN ALZHEIMER’S DISEASE**

Familial early onset AD genes are included APP, PSEN1 and PSEN2 (18-20) and the most well-known genetic risk factor for late onset AD is apolipoprotein E4 (APOE). Indeed the top 10 risk genes in AD are APOE, BIN1, CLU, ABCA7, CR1, PICALM, MS4A6A, CD33, MS4A4E, CD2AP (21-23).

Apart from genetic aberrations, protein expression changes have been reported. During the last decade human brain tissue proteomics investigations gradually increased (24-26). Sultana R et al. introduced 18 proteins in hippocampus region which differentially expressed in AD patient and related to the different cellular functions in AD pathology(27) whereas in recent proteomic study has demonstrated a total of 197 proteins differentially abundant in AD versus controls, after examining the temporal lobe region (28).

Oxidatively modified proteins in AD brain as well can be associated with tau and Aβ pathology, i.e. peptidyl prolyl cis-trans isomerase, link to the cell dysfunction influencing energy metabolism, i.e. ATP synthase subunits, altered redox regulation, i.e. peroxiredoxins, mitochondrial function, i.e. voltage-dependent anion-selective channel protein 1, proteasomal activity, i.e. ubiquitin carboxy-terminal hydrolase L-1, excitotoxicity, i.e. glutamine synthase, synaptic alterations, i.e. gamma synaptosomal-associated protein and regulation of cell death, i.e. heat shock proteins(24,29).

**PROTEIN-PROTEIN INTERACTION ANALYSIS**
Since genes do not act as individual units, they cooperate in overlapping pathways, the deregulation of which is a hallmark of diseases like AD. In addition, gene clustering based on topology and functions illustrate correlated expression patterns (30, 31). Because of importance of networks in system biology in recently years, developed quantitative tools for analyzing them. Analyzing the network properties of gene-expression data might reveal the organizational pattern of gene expression in disease, which might, in turn, help us to identify new potential drug targets (32).

During the last decade, there has been an exponential increase in the number of studies analyzing AD in different part of brain tissue; in this study, data were extracted from Begcevic et al. investigation (33). They identified different expressed proteins in hippocampus tissues of AD in comparing to normal. Among 245 regulated genes, 105 were up-regulated or new expression and 140 were down-regulated or suppressed.

In order to carry out a retrospective meta-analysis of the functional annotations using UniProt accession numbers (http://www.uniprot.org), a publicly available web-based tool, to search for annotations that are significantly associated to the list of AD related proteins.

PPIs represent as the basic skeleton for the self-organization and homeostasis of living organisms (34). In this study, information on human PPI networks from significant genes was obtained from BIND databases. The PPI network was visualized using the Cytoscape 3 software (35). The PPI networks of the significantly expressed genes between the AD pattern and the control are shown in Figure 1. AD network contain 3851 nods and 3480 edges. Previously Korolainen et al. (24) visualized interaction Network of 109 proteins of AD data obtained from literature survey by using Cytoscape software platform (35). A total of 823 direct interactions partners and 11 697 interconnections among them were extracted using MiMI Plugin (36).

Degree distribution was presented in figure 2. It represent distribution model such other biological network showing scale-free topology. The nodes with the most degree (hub) are ASC, Smad8, Smad2, SnoN, KIAA1196, Smad1, RNF11, Smad3, ATF7ip and SHBG. Most of these hubs act as transcription factor and participates in a wide range of critical processes including morphogenesis, cell-fate determination, proliferation, differentiation and apoptosis (37).

We integrated the databases and networks and used a Molecular COmplex DEtection (MCODE) to analyze the characteristics of the networks. MCODE also makes the visualization of large networks manageable by extracting the dense regions around a protein of interest based on network topology (38). Interactomes with a score greater than 2.0 and at least two nodes were taken as significant predictions.

MCODE efficiently finds densely connected regions in PPI network that many of which contribute to known molecular complexes and implies that large amounts of available knowledge is buried in large protein interaction networks. Structured molecular interaction data resources such as BIND will be crucial in making these resources (38). Analyzing network with MCODE to further study of complex revealed 22 sub-networks described in table 1. The PPI sub-networks based on the differently expressed genes made up of highly connected regions in AD pattern versus control comparison represented in figure 3. The second stage in MCOD algorithm recognize seeds a complex with the highest weighted vertex (forward and outward) and whose weight is above a given threshold (38). The seed nodes of these complexes included NF-L, RXX beta, Perilamin A, EAP20, Xm2, PPARBP, Upf2, SMURF2, SEPT8, NP-005070.1, PMP70, 14-3-3E, PEX14, LRRC7, Pax3, TRIP1, Smad4, Borg1, CENP6, CENP4, MCM4 and CHMP4C. None of these are located in the hippocampus but most of them are brain tissue specific proteins. Pax3 present in neural tube defects, PEX14 in zellweger syndrome and Smad4 is involved in a wide range of diseases (39).

Gene ontology categories were further analyzed to identify the function of up and down-regulated of AD proteome. ClueGO v2.0.5, Cytoscape plug-in tool that visualizes the non-redundant biological terms for large clusters of genes in a functionally grouped network, was used to statistically evaluate groups of proteins with respect to the existing annotations of the Gene Ontology. (40). ClueGO use kappa Score that shows the relationships between the terms based on their overlapping genes. It is used for creating the network and can be used for creating the groups. Initially, a term-gen matrix containing the selected terms and their associated genes is created. The gene ontology analysis of network performed by ClueGO is depicted in figure 4-9. The Gene Ontology (GO) projects (41) aims to capture the increasing knowledge on gene function in a controlled vocabulary applicable to all the organisms. GO describes gene products in terms of their associated biological processes, cellular components and molecular functions. There exists a hierarchical relationship between the terms. Because of complexity of hierarchy structure, the terms can be in several different levels. The specificity of the terms fluctuates in the tree: from very general terms (in first levels of GO) to very specific ones.
According to figure 4, the most biological process were found in up regulated contain positive regulation of apoptosis, central nervous system (CNS) neuron development and platelet aggregation and down-regulated genes contribute to pentose phosphate shunt, oxidative branch, glutathione metabolic process, negative regulation of ERK1 and ERK2 cascade, microtubule polymerization and de-polymerization and negative regulation of protein complex assembly. It was previously illustrated biological process category, proteins related to oxidation/reduction, glycolysis, anti-apoptosis, transport, nervous system development, and protein folding represented significantly altered pathways in AD (24). In current study one of important biological pathway related to pathogenesis of AD is apoptosis. Apoptosis plays crucial role during normal development and tissue homeostasis but its inconsistent regulation is linked to neurodegenerative diseases (i.e. AD, PD, HD and ALS), ischaemic stroke, AIDS, cancer and autoimmune disorders (42). Intracellular or extracellular protein aggregation in such diseases is connected to the cell death and neurodegeneration (43). It has been thought such aggregates are associated to caspase activation, at least in the cases of AD, HD and ALS, and lead to apoptosis (44). Cell death in AD might be contribute to oxidized proteins that cause decline in the antioxidant capacity or increased inflammatory processes (45). Oxidative stress and modified antioxidant defense systems are involved in the pathogenesis of AD (46). Activated microglia, particularly during gliosis and inflammation, are responsible for substantial production of reactive oxygen species. The main cell organ responsible for oxidative stress is mitochondrion because of the electron transport chain (24). Glutamate is the most abundant excitatory neurotransmitter in the brain that acts through activation of glutamate receptors. These include ionotropic glutamate receptors (NMDA, AMPA, Kainate) and metabotropic glutamate receptors (Quisqualate-B). Excessive release of glutamate from presynaptic and glial cells into the extracellular space triggers excitotoxicity. This neurotransmitter then over-activates glutamate receptors, especially N23 methyl-D-aspartate (NMDA) receptors, leading to excessive Ca2+ (and Na+) influx into the cell. Glutamate induced excitotoxicity has been suggested to cause either necrosis or apoptosis (47). Aponecrosis represent the molecular, morphological and dynamic features of both apoptosis and necrosis (48). In addition, there are studies indicating that in AD brains the typical neurofibrillary tangles and neuritic plaques may be outcome of aberrant cell cycle events (49, 50) so, neuronal cells that suffer from cell cycle distractions are compelled to one of two fates; they either die via apoptosis or they produce Alzheimer type pathology (50).

According figure 5, the most molecular functions were related to the up regulated genes that display as alpha catenin binding, coenzyme binding, dimethylargeninase activity, oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, glutathione transferase activity and S100 protein binding and down-regulated genes are linked to the GTPase activity, endopeptidase regulator activity, steroid hormone receptor binding and NADH dehydrogenase activity.

As represent in figure 6, the most cellular component were related to the up regulated genes that display in neurofilament, microvillus membrane and phosphopyruvate hydratase complex and down-regulated genes are linked to the membrane coat, kinesin complex, mitoconderial respiratory chain, chaperonin containing t complex, 6-phosphofructokinase complex, pyruvate hydrogenase complex heterogenous nuclar ribonucleoprotein complex and axonal growth cone. Recently Kroinen et al. determined the most altered proteins in 2-DE studies are located in cytoplasm and mitochondrion. These kind of proteins related to energy metabolism are located in cytoplasm and mitochondria, and apoptosis is believed to be under mitochondrial regulation (22).

As depicted in figure 7, there not exist any relationship between up regulated genes and immune system process but down-regulated genes are linked to the regulated innate immune system. Innate immunity is the first line of defense against invading organisms (51). Microglia's, CNS resident macrophages, is the main cell involved in the innate immune system and stimulating adaptive immunity that express several Toll-like receptors (TLRs). TLR4 was identified in relating to the AD pathology. It was also determined that stimulation of the innate immune system through TLR9 in AD model mice is an effective and safe method to reduce the amyloid burden and tau-related pathology (52).

Kyoto Encyclopedia of Genes and Genomes (KEGG) (53) is a database of biological systems that integrates genomic, chemical and systemic functional information. The terms are analyzed trough the perspective of their associated genes so that, the genes from both clusters could be associated with a term, but in different proportions. It is considered that a term as specific for one of the clusters if the percentage of associated genes from this cluster is higher than the selected threshold (i.e. %66) (35). Therefore, charts with specific terms for each cluster are provided. The common terms are included in a separate chart.

According to figure 8, on the network, the different proportion of the genes from the analyzed clusters is represented with a color gradient from green, for the first cluster genes, to red for the second cluster. The
visualization of the groups on the network can be switched with the one of the uploaded clusters distribution on the selected terms. Clusters distribution network based on KEGG database with terms with up/down regulated genes are shown in red/green, respectively in figure 8. Common pathway in both group contain glycolysis/gluconeogenesis, pentose phosphate pathway, pyruvate metabolism, tryptophan metabolism, glyoxylate and dicarboxylate metabolism, pathogenic E.coli infection and amyotrophic lateral sclerosis. Prior studies from three decades of research have noted the importance of glucose hypometabolism in the pathology of Alzheimer’s brain that occurs early in specific region, and correlates with other clinical features (54, 55). In the other hand Bigl M et al. illustrated that increased activity of some glycolytic enzymes might be the result of the reactive astrocytosis developing in the course of AD. So that a significant increase in specific activity of pyruvate kinase and lactate dehydrogenase was found in frontal and temporal cortex of AD brains, whereas the activities of aldolase and hexokinase are not changed. Glucose 6-phosphate dehydrogenase activity was significantly reduced in hippocampus. (56).

In AD patients, KEGG pathways related to up-regulated and down-regulated genes represent in figure 9a and b respectively. Up-regulated pie chart contained amino acid metabolism such as phenylalanine, glutathione, glycine, serine and threonine and down-regulated gene pie chart contain endocrine and other factor-related calcium reabsorption and AD.

In sum up, network analysis could help to comprehend AD mechanism and discover potential biomarkers which may be helpful for diagnosis, prognosis and treatment prediction. Study of AD-hippocampal specific genes, ether up-regulated or down-regulated, lead to determine a number of proteins that are AD index, so for utilizing these proteins as monitoring or prognosis markers need to find them in cerebrospinal fluid (CSF) or peripheral blood.

ACKNOWLEDGMENTS
We gratefully acknowledge Proteomics Research Center of Shahid Beheshti University of Medical Sciences for financial support. This paper was derived from Ph.D. thesis of Hakimeh Zali.

Figure 1: PPI Network of AD based on cytoscape 3 software.
Figure 2: degree distribution of network of AD based on network analysis in cytoscape 3 software

Table 1: The PPI subnetworks were clustered as highly connected regions in AD network based on MCOD

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Score (Density*#Nodes)</th>
<th>Nodes</th>
<th>Edges</th>
<th>Seed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>15</td>
<td>38</td>
<td>NF-L</td>
</tr>
<tr>
<td>2</td>
<td>4.4</td>
<td>6</td>
<td>11</td>
<td>RXR beta</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>PerlaminA</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>EAP20</td>
</tr>
<tr>
<td>5</td>
<td>3.38</td>
<td>14</td>
<td>24</td>
<td>Xm2</td>
</tr>
<tr>
<td>6</td>
<td>3.06</td>
<td>16</td>
<td>23</td>
<td>PPARBP</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>Upf2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>7</td>
<td>11</td>
<td>SMURF2</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>SEPT8</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>NP-005070.1</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>PMP70</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>14-3-3E</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>PEX14</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>LRRRC7</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>Pax3</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>TRIP1</td>
</tr>
<tr>
<td>17</td>
<td>2.8</td>
<td>11</td>
<td>16</td>
<td>Smad4</td>
</tr>
<tr>
<td>18</td>
<td>2.66</td>
<td>4</td>
<td>4</td>
<td>Borg1</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>CSN6</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>CSN4</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>MCM4</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>CHMP4C</td>
</tr>
</tbody>
</table>
Figure 3: The PPI subnetworks based on the differently expressed genes made up of highly connected regions in Alzheimer proteome pattern versus control comparison. Cluster 1, 2, 4, 5, 8, 12 and 14 represented as a, b, c, d, e, f and g respectively. Yellow ellipses represent seed nodes. Pink ellipses represent neighbor nodes. All edges represent interactions.

(a)

Figure 4: The biological process of gene ontology analysis from up-regulated (a) and down-regulated (b) genes in hippocampus of AD patient compare to normal.

(b)
**Figure 5:** The molecular function of gene ontology analysis from up-regulated (a) and down-regulated (b) genes in hippocampus of AD patient compare to normal.

**Figure 6:** The cellular component of gene ontology analysis from up-regulated (a) and down-regulated (b) genes in hippocampus of AD patient compare to normal.
Figure 7: Immune system process of gene ontology analysis from down-regulated genes in hippocampus of AD patient compared to normal.

Figure 8: Clusters distribution network. Terms with up/down regulated genes are shown in red/green, respectively. The color gradient shows the gene proportion of each cluster associated with the term. Equal proportions of the two clusters are represented in white.
Figure 9: KEGG pathway of gene ontology analysis from up-regulated (a) and down-regulated (b) genes in hippocampus of AD patient compared to normal

REFERENCES


