Protein-protein interaction analysis of Alzheimer`s disease and NAFLD based on systems biology methods unhide common ancestor pathways

Reza Karbalaei1, Marzieh Allahyari2, Mostafa Rezei-Tavirani1, Hamid Asadzadeh-Aghdaei3, Mohammad Reza Zali4
1Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2Drug Design and Bioinformatics Unit, Medical Biotechnology Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran
3Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
4Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Aim: Analysis reconstruction networks from two diseases, NAFLD and Alzheimer`s diseases and their relationship based on systems biology methods.

Background: NAFLD and Alzheimer`s diseases are two complex diseases, with progressive prevalence and high cost for countries. There are some reports on relation and same spreading pathways of these two diseases. In addition, they have some similar risk factors, exclusively lifestyle such as feeding, exercises and so on. Therefore, systems biology approach can help to discover their relationship.

Methods: DisGeNET and STRING databases were sources of disease genes and constructing networks. Three plugins of Cytoscape software, including ClusterONE, ClueGO and CluePedia, were used to analyze and cluster networks and enrichment of pathways. An R package used to define best centrality method. Finally, based on degree and Betweenness, hubs and bottleneck nodes were defined.

Results: Common genes between NAFLD and Alzheimer`s disease were 190 genes that used construct a network with STRING database. The resulting network contained 182 nodes and 2591 edges and comprises from four clusters. Enrichment of these proteins and their first neighbors in network by OMIM database lead to carbohydrate metabolism, long chain fatty acid and regulation of JAK-STAT and IL-17 signaling pathways, respectively. Also seven genes selected as hub-bottleneck include: IL6, AKT1, TP53, TNF, JUN, VEGFA and PPARG. Enrichment of these proteins and their first neighbors in network by OMIM database lead to diabetes and obesity as ancestors of NAFLD and AD.

Conclusion: Systems biology methods, specifically PPI networks, can be useful for analyzing complicated related diseases. Finding Hub and bottlenecks proteins should be the goal of drug designing and introducing disease markers.

Keywords: Alzheimer`s disease (AD), Non-alcoholic fatty liver disease (NAFLD), Protein-protein interaction (PPI) network analysis, Hub-bottlenecks, Protein clusters.

Introduction

Alzheimer’s disease (AD) is a neurodegenerative disease that is one of the important disease in industrial countries. Based on Alzheimer’s Disease International Federation (ADI), at least 46.8 million people are affected by dementia worldwide, that anticipated to be 74.7 million by 2030 and 131.5 million by 2050 (1). This disease can be categorized in two forms: early-onset familial Alzheimer disease (EFAD) (2) and Late-onset Alzheimer’s disease (LOAD) or non-familial (3). EFAD form inheritance dominantly but LOAD form is a complex or multifactorial disease (4). Research on
AD showed that in addition to age and heredity, lifestyle is an important factor in the progression of this disease (5). On the other hand, lifestyle has an important role in producing some diseases such as obesity, diabetes and fatty liver (6). Fatty liver diseases divided into two forms: alcoholic fatty liver disease (AFLD) and Non-alcoholic fatty liver disease (NAFLD), that mainly occurs due to high using of Alcohol and fat (6). Non-alcoholic fatty liver disease (NAFLD) is one of the most important reasons for liver disease in the United States so that 30% of US population affected by NAFLD (7). Indeed, as well as AD, NAFLD depend on lifestyle and feeding. Our previous studies on AD that accomplished by meta-analysis in microarray data showed that NAFLD has an undoubted relation to AD (8). There are other studies about the relation of AD and NAFLD that focuses on some common genes (LRP1) (9), cross-sectional study (10) and AD-Transgenic model (11). Protein-protein interaction (PPI) network analysis is one of the major fields in systems biology in which analyzed complex interactome of proteins as the main source of data (12). Using systems biology method such as a comparison between gene sets of diseases, constructing PPI network and pathway enrichment can be helpful to decipher the shared mechanism of NAFLD and AD. In this study, we reported seven important shared proteins between these diseases that can be used not only as markers of disease but also as targets for drug designing. Also, pathways that shred between these diseases were introduces.

Methods

DisGeNET is a discovery database that gathered genes and variants associated with human diseases and publicly available (13). The related genes of NAFLD and AD were exported from DisGeNET database and common genes between two diseases used to construct PPI network by Search Tool for the Retrieval of Interacting Genes/Proteins (STRING). STRING is a database for predicted protein-protein interactions at EMBL clusters the extracted results from many protein-protein interactions databases, like Mint, BioGrid, etc. It also uses the information from KEGG pathways and Reactome to provide the best annotations for the interactions of one protein (14). The common network was constructed by importing shared genes in STRING database and clustered by ClusterONE plugin of Cytoscape software (15) that finds overlapping protein complexes in a protein interaction network loaded into Cytoscape. (overlap threshold = 1, node penalty = 0, haircut threshold = 0) (16). Pathway enrichment and the relation between pathways were accomplished using ClueGO and CluePedia plugins of Cytoscape software (17, 18). To find the best centrality method for selection of the most important nodes, we use an R package named CINNA (19, 20). A network is composed of nodes (e.g., genes or proteins) and edges/links (e.g., co-expression relationships or physical interactions). In network biology terms, degree, and Betweenness are important centrality parameters that are useful for analysis network topology. Edges/links of a node are called the degree of that node. Nodes with high degree are called hubs and nodes that achieve top-ten, or top-five percent of betweenness centrality are called bottlenecks (both based on researcher’s definition) (21). So, nodes that are simultaneously hubs and bottlenecks are named hub-bottlenecks (22). Average degree (A.D) and standard deviation (SD) of degrees were calculated, and nodes with a degree above 2SD + A.D were selected as hub proteins in each network. Also, the top five percent of betweenness centrality measures were selected as bottleneck proteins. Shared genes, hubs and bottleneck proteins of these two networks were extracted and used for further analysis. We used Cytoscape to analyze networks and extract hubs, hub-bottlenecks, and their first neighbors (23).

Results

From DisGeNET, 332 and 1200 genes were extracted for NAFLD and Alzheimer’s diseases, respectively. Totally, 189 genes were shared between the two lists were shared and were named common genes. The common genes network that was constructed using STRING database has 182 nodes and 2591 edges and four clusters (Figure 1). Cluster analyzing by ClueGO and CluePedia plugins showed that there are 29 meaningful pathways based on statistical analysis and there is no duplication between them. Cluster one mainly includes carbohydrate metabolism pathways and their related signaling, and
the main category of this cluster was AMPK signaling pathway. In cluster two long chain fatty acid and their extract metabolic process include arachidonic acid, xenobiotics and calciferol were enriched. Finally, enrichment of cluster three lead to signaling pathways such as regulation of JAK-STAT cascade, IL-17 signaling pathway and AGE-RAGE signaling pathway in diabetic complications. Due to low number of nodes in cluster four, pathway enrichment was meaningless (table 1).

Based on CINNA package results, degree and Betweenness centrality methods were the best qualified methods for this network. In next step, the network was analyzed by Cytoscape to define hubs, hub-bottleneck. Results showed that IL6, AKT1, TP53, TNF, JUN, VEGFA, PPARG, MAPK3, IGF1, and LEP are hubs that first seven proteins were also bottlenecks, so...
selected as hub-bottlenecks (table 2). By extracting these hub-bottlenecks and their first neighbors from the network, we reach to a new interesting network that contains 82% of nodes (150 nodes) and 92% of edges (2367 edges) from the main network. So by analyzing them in OMIM as the main database for disease, we reached to diabetes and obesity (table 3).

**Discussion**

Systems biology methods such as PPI network analysis and pathway enrichment have been used broadly to discover main proteins and pathways underlying complex diseases (24). Different types of cancers, various kinds of neurodegenerative diseases and disorders and also many cellular conditions are analyzed via protein-protein interaction method (25-30). The relation between NAFLD and AD is becoming increasingly recognized (9-11). In this study, we used the complete genes list of the two diseases (NAFLD and AD) that may have shared mechanism based on risk factors and previous studies (8). According to network clustering and further pathways enrichment, 42 pathways were enriched. Altogether, three main group of pathways are candidate as key pathways in both AD and NAFLD: carbohydrate metabolism, long fatty acid metabolism, and IL signaling pathways. Previous studies indicate evidence about all mentioned relation...
except the role of long fatty acid metabolism in AD (8, 31-37).

Six Hub- Bottleneck nodes are important targets for both NAFLD and AD. High level secretion of peripheral IL-6 may be responsible for acute-phase proteins that observed in AD patients (38) and high levels of IL-6 were detected in NAFLD patients (39). AKT activity in temporal cortex of Alzheimer patients were significantly increased (40) and activated the PI3-K/Akt kinase pathway triggers NAFLD (41). TP53 that known as P53, up-regulated in Alzheimer's disease (42) and inhibition of attenuates signs of NAFLD (43). Inhibition of TNF alpha decrease amyloid plaques and tau phosphorylation in the mouse brain, and so risk of AD (44), and this protein involved in the pathophysiology of NAFLD (45). Inhibition of JUN is a therapeutic strategy to stop progression of AD (46) and expression of this protein Increased in NAFLD (47). Abnormal regulation of VEGFA expression implicated in AD (48) and involved in pathophysiology of NAFLD (49). Finally, PPARG is a potential therapeutic targets for both AD and NAFLD (50, 51).

Analyzing main nodes and their first neighbors by OMIM database showed that diabetes and obesity were the results of this enrichment. We can conclude that diabetes and obesity are common ancestors of AD and NAFLD. These results showed that application of systems biology methods unhide unravels the secret behind impact of common proteins on treatment of NAFLD and AD also needs to be further assessed.

Conflict of interests

The authors declare that they have no conflict of interest.

References

1. Realdon O, Rossetto F, Nalin M, Baroni I, Cabinio M, Fioravanti R, et al. Technology-enhanced multi-domain at home continuum of care program with respect to usual care for people with cognitive impairment: the Ability TelearHABILITYtation study protocol for a randomized controlled trial. BMC Psychiatry 2016;16:425.

2. Bird TD. Early-onset familial Alzheimer disease. 2012.

3. Hardy J. A hundred years of Alzheimer's disease research. Neuron 2006;52:3-13.

4. Rhinn H, Fujita R, Qiang L, Cheng R, Lee JH, Abeliovich A. Integrative genomics identifies APOE e4 effectors in Alzheimer's disease. Nature 2013;500:45-50.

5. Ghareeb DA, Hafez HS, Hussien HM, Kabapy NF. Non-alcoholic fatty liver induces insulin resistance and metabolic disorders with development of brain damage and dysfunction. Metab Brain Dis 2011;26:253-67.

Table 2. Hub and Bottleneck genes of Common network genes between NAFLD and AD are shown. Hub-bottleneck genes are bolded.

| R | Hubs | Degree | Bottlenecks | Betweenness score |
|---|------|--------|-------------|------------------|
| 1 | IL6  | 109    | TP53        | 0.087078         |
| 2 | AKT1 | 99     | IL6         | 0.06467          |
| 3 | TP53 | 98     | AKT1        | 0.054262         |
| 4 | TNF  | 97     | TNF         | 0.047973         |
| 5 | JUN  | 89     | JUN         | 0.034409         |
| 6 | VEGFA| 83     | PPARG       | 0.030972         |
| 7 | PPARG| 79     | INSR        | 0.026792         |
| 8 | MAPK3| 78     | VEGFA       | 0.024631         |
| 9 | IGFI | 77     | LPL         | 0.021859         |
| 10| LEP  | 75     | CAT         | 0.020219         |

Table 3. OMIM result from enrichment of Hub-bottleneck genes and their first neighbors are represented.

| disease name                     | Adjusted P-value | Genes                                                                 |
|----------------------------------|------------------|----------------------------------------------------------------------|
| diabetes                         | 3.41E-10         | TCF7L2;IL6;ACE;IRS1;HNF4A;INSR;UCP3;IRS2;PPARG;GCK                   |
| diabetes_mellitus_type_2         | 1.35E-09         | TCF7L2;IRS1;HNF4A;INSR;UCP3;IRS2;PPARG;GCK                           |
| obesity                          | 3.87E-05         | LEP;UCP3;LEPR;PPARG;PPARGC1B                                        |
| diabetes_mellitus                | 4.25E-05         | TCF7L2;IRS1;HNF4A;INSR;IRS2;GCK                                      |
6. Reddy JK, Rao MS. Lipid metabolism and liver inflammation. II. Fatty liver disease and fatty acid oxidation. Am J Physiol Gastrointest Liver Physiol 2006;290:G852-8.

7. Than NN, Newsome PN. A concise review of non-alcoholic fatty liver disease. Atherosclerosis 2015;239:192-202.

8. Karbalaei R, Torkzaban B, Rezaei-Tavirani M, Azimzadeh S. Multi-regional Neurodegeneration in Alzheimer's disease: Meta-analysis and data integration of transcriptomics data. bioRxiv 2018:245571.

9. Kim DG, Krenz A, Toussaint LE, Maurer KJ, Robinson SA, Yan A, et al. Non-alcoholic fatty liver disease induces signs of Alzheimer's disease (AD) in wild-type mice and accelerates pathological signs of AD in an AD model. J Neuroinflammation 2016;13:1.

10. Seo SW, Gottesman RF, Clark JM, Hernaez R, Chang Y, Kim C, et al. Nonalcoholic fatty liver disease is associated with cognitive function in adults. Neurology 2016;86:1136-42.

11. de la Monte SM, Tong M. Brain metabolic dysfunction at the core of Alzheimer's disease. Biochem Pharmacol 2014;88:548-59.

12. Safaei A, Arefi Oskouei A, Mohebbi SR, Rezaei-Tavirani M, Mahboubi M, Peyvandi M, et al. Metabolic analysis of human cirrhosis, hepatocellular carcinoma, non-alcoholic fatty liver disease and non-alcoholic steatohepatitis diseases. Gastroenterol Hepatol Bed Bench 2016;9:158-73.

13. Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Res 2017;45:D833-9.

14. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Mingué P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 2011;39:D651-8.

15. Nepusz T, Yu H, Paccanaro A. Detecting overlapping protein complexes in protein-protein interaction networks. Nat Methods 2012;9:471-2.

16. Bhosle A, Chandra N. Structural analysis of dihydrofolate reductases enables rationalization of antifolate binding affinities and suggests repurposing possibilities. FEBS J 2016;283:1139-67.

17. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091-3.

18. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics 2013;29:661-3.

19. Ashtiani M, Mirzaie M, Jafari M. CINNA: An R package for deciphering Central Informative Nodes in Network Analysis. bioRxiv 2017:168757.

20. Ashtiani M, Salehzadeh A, Razaghi-Moghadam Z, Hennig H, Wolkenhauer O, Mirzaie M, et al. Selection of most relevant centrality measures: A systematic survey on protein-protein interaction networks. bioRxiv 2017:149492.

21. Safaei A, Rezaei Tavirani M, Arefi Oskouei A, Zamanian Azodi M, Mohebbi SR, Nikzamir AR. Protein-protein interaction network analysis of cirrhosis liver disease. Gastroenterol Hepatol Bed Bench 2016;9:114-23.

22. Zamanian-Azodi M, Mortazavi-Tabatabaeei SA, Mansouri V, Vafaei R. Metabolite-protein interaction (MPI) network analysis of obsessive-compulsive disorder (OCD) from reported metabolites. Arvand J Health Med Sci 2016;1:112-20.

23. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.

24. Hunter DJ. Gene-environment interactions in human diseases. Nat Rev Genet 2005;6:287.

25. Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. Protein-protein interaction networks (PPI) and complex diseases. Gastroenterol Hepatol Bed Bench 2014;7:17-31.

26. Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, Hasanzadeh H, Rezaei Tavirani M, Seyyedi SS. Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer. Gastroenterol Hepatol Bed Bench 2015;8:215-24.

27. Rezaei-Tavirani M, Zamanian-Azodi M, Rajabi S, Masoudi-Nejad A, Rostami-Nejad M, RahmatiRad S. Protein Clustering and Interactome Analysis in Parkinson and Alzheimer's Diseases. Arch Iran Med 2016;19:101-9.

28. Safari-Alighiarloo N, Taghizadeh M, Tabatabaeei SM, Shahsavari S, Namaki S, Khodakarim S, et al. Identification of new key genes for type 1 diabetes through construction and analysis of protein-protein interaction networks based on blood and pancreatic islet transcriptomes. J Diabetes 2017;9:764-77.
29. Zali H, Rezaei Tavirani M. Meningioma protein-protein interaction network. Arch Iran Med;17:262-72.
30. Abbaszadeh HA, Peyvandi AA, Sadeghi Y, Safaei A, Zamanian-Azodi M, Khoramgah MS, et al. Er:YAG Laser and Cyclosporin A Effect on Cell Cycle Regulation of Human Gingival Fibroblast Cells. J Lasers Med Sci 2017;8:143-9.
31. Harley IT, Stankiewicz TE, Giles DA, Softic S, Flick LM, Cappelletti M, et al. IL-17 signaling accelerates the progression of nonalcoholic fatty liver disease in mice. Hepatology 2014;59:1830-9.
32. Jin JJ, Kim HD, Maxwell JA, Li L, Fukushima K. Toll-like receptor 4-dependent upregulation of cytokines in a transgenic mouse model of Alzheimer's disease. J Neuroinflammation 2008;5:23.
33. Yang W, He Y, Liu S, Gan L, Zhang Z, Wang J, et al. Integrative transcriptomic analysis of NAFLD animal model reveals dysregulated genes and pathways in metabolism. Gene 2016;595:99-108.
34. Kang H, Greenson JK, Omo JT, Chao C, Peterman D, Anderson L, et al. Metabolic syndrome is associated with greater histologic severity, higher carbohydrate, and lower fat diet in patients with NAFLD. Am J Gastroenterol 2006;101:2247-53.
35. Karbalaei R, Piran M, Rezaei-Tavirani M, Asadzadeh-Aghdaei H, Heidari MH. A systems biology analysis protein-protein interaction of NASP and IBD based on comprehensive gene information. Gastroenterol Hepatol Bed Bench 2017;10:194-201.
36. Salminen A, Kaarniranta K, Haapasalo A, Soininen H, Hiltunen M. AMP-activated protein kinase: a potential player in Alzheimer's disease. J Neurochem 2011;118:460-74.
37. Kohjima M, Higuchi N, Kato M, Kotoh K, Yoshimoto T, Fujino T, et al. SREBP-1c, regulated by the insulin and AMPK signaling pathways, plays a role in nonalcoholic fatty liver disease. Int J Mol Med 2008;21:507-11.
38. Cojocaru IM, Cojocaru M, Miu G, Sapira V. Study of interleukin-6 production in Alzheimer's disease. Rom J Intern Med 2011;49:55-8.
39. Jorge AS, Andrade JM, Paraíso AF, Jorge GC, Silveira CM, de Souza LR, et al. Body mass index and the visceral adipose tissue expression of IL-6 and TNF-alpha are associated with the morphological severity of non-alcoholic fatty liver disease in individuals with class III obesity. Obes Res Clin Pract 2016. pii: S1871-403X(16)30006-0.
40. Rickle A, Bogdanovic N, Volkman I, Winblad B, Ravid C, Cowburn RF. Akt activity in Alzheimer's disease and other neurodegenerative disorders. Neuroreport 2004;15:955-9.
41. Matsuda S, Kobayashi M, Kitagishi Y. Roles for PI3K/AKT/PTEN Pathway in Cell Signaling of Nonalcoholic Fatty Liver Disease. ISRN Endocrinol 2013;2013:472432.
42. Hooper C, Meimariou E, Tavassoli M, Melino G, Lovestone S, Killick R. p53 is upregulated in Alzheimer's disease and induces tau phosphorylation in HEK293a cells. Neurosci Lett 2007;418:34-7.
43. Derdak Z, Villegas KA, Harb R, Wu AM, Sousa A, Wands JR. Inhibition of p53 attenuates steatosis and liver injury in a mouse model of non-alcoholic fatty liver disease. J Hepatol 2013;58:785-91.
44. McCaulley ME, Grush KA. Alzheimer's Disease: Exploring the Role of Inflammation and Implications for Treatment. Int J Alzheimers Dis 2015;2015:515248.
45. Tilg H. The role of cytokines in non-alcoholic fatty liver disease. Dig Dis 2010;28:179-85.
46. Scip A, Tozzi A, Abaza A, Cardinetti D, Colombo I, Calabresi P, et al. c-Jun N-terminal kinase has a key role in Alzheimer disease synaptic dysfunction in vivo. Cell Death Dis 2014;5:e1019.
47. Dorn C, Engelmann JC, Saugspier M, Koch A, Hartmann A, Müller M, et al. Increased expression of c-Jun in nonalcoholic fatty liver disease. Lab Invest 2014;94:394-408.
48. Storkebaum E, Carmeliet P. VEGF: a critical player in neurodegeneration. J Clin Invest 2004;113:14-8.
49. Coulson S, Francque S, Colle I, Verrijken A, Blomme B, Heindryckx F, et al. Evaluation of inflammatory and angiogenic factors in patients with non-alcoholic fatty liver disease. Cytokine 2012;59:442-9.
50. Tailleux A, Wouters K, Staels B. Roles of PPARs in NAFLD: potential therapeutic targets. Biochim Biophys Acta 2012;1821:809-18.
51. Kummer MP, Heneka MT. PPARs in Alzheimer's disease. PPAR research 2008;2008:403896.