Investigating the Relationship between Diabetes and Alzheimer’s Disease: A Network Systems Biology Approach

Tammane R. Sahrawat 1*, Jyoti Dwivedi 1

Abstract

Ageing is associated with a number of diseases. Alzheimer’s disease (AD) and diabetes are among such most common diseases. These two diseases are considered to be fundamentally similar disorders because they share some common elements, though they differ in the time of onset, tissues affected as well as the magnitudes of their specific traits. The present study was undertaken to prospect the association between the genes involved in Diabetes and AD; and their common pathophysiology. Using a network system biology approach, the genes common between Diabetes and AD were retrieved from DisGeNET database. The common genes were analysed using in silico tool, Cytoscape’s various plug-ins, ClusterONE, CytoHubba, ClueGO and CluePedia. Eleven genes which can act as potential marker for both Diabetes and AD namely IL4, ICAM1, ALB, INS, CSF2, IL6, TNF, IL10, GAPDH, TLR4, and AKT have been identified in the present study. This is the first study of its kind in which relationship between Diabetes and AD has been investigated to identify their common genes, which can help in better understanding of pathophysiology of these age-related diseases.

Keywords Systems biology; pathophysiology; neurodegenerative disease; biomarkers

Enquête sur la relation entre le diabète et la maladie d'Alzheimer : Une approche de biologie systémique en réseau

Résumé

Le vieillissement est associé à un certain nombre de maladies. La maladie d’Alzheimer (MA) et le diabète font partie des maladies les plus courantes. Ces deux maladies sont considérées comme des troubles fondamentalement similaires car elles partagent certains éléments communs, bien qu'elles diffèrent quant au moment de leur apparition, aux tissus affectés ainsi qu'à l'ampleur de leurs...
caractéristiques spécifiques. La présente étude a été entreprise pour examiner l'association entre les gènes impliqués dans le diabète et la MA; et leur physiopathologie commune. En utilisant une approche de biologie systémique en réseau, les gènes communs au diabète et à la MA ont été extraits de la base de données DisGeNET. Les gènes communs ont été analysés en utilisant l'outil in silico, les différents plug-ins de Cytoscape, ClusterONE, CytoHubba, ClueGO et CluePedia. Onze gènes qui peuvent agir comme marqueurs potentiels du diabète et de la MA, à savoir IL4, ICAM1, ALB, INS, CSF2, IL6, TNF, IL10, GAPDH, TLR4 et AKT, ont été identifiés dans la présente étude. C'est la première étude de ce type dans laquelle la relation entre le diabète et la MA a été étudiée afin d'identifier leurs gènes communs, ce qui peut aider à mieux comprendre la physiopathologie de ces maladies liées à l'âge.

Mots-clés Biologie des systèmes ; pathophysiologie ; maladie neurodégénérative ; biomarqueurs

Introduction

An interactome is a dynamic network of active proteins involved in various metabolic, signaling and biochemical reactions. Network biology has emerged as a powerful instrument for studying complicated diseases. Both network biology and systems biology approaches are commonly constructed on the basis of physical or functional interactions between molecules, genes or proteins that are represented in the form of an interaction network. These interaction networks aid in the investigation of cellular communication between nodes and various binary relationships (Schork 1997, p. S103).

The composition and topology of networks which includes nodes and edges in complex networks are closely linked to important cellular functions. Systemic networks analysis can be used for the combined, integrated and functional analysis of high-throughput experimental data which can be useful in drug discovery (Hu et al. 2016, p. 615). Study of disease development patterns of patients with more than one disease and particularly disease coincidence can assist us to understand the connection between the internal factors and external factors like diet, lifestyle (Zou et al. 2013, p. 2).

Unraveling the molecular pathways through which inherited traits affect a phenotype is not an easy task and moreover, in complex diseases, these inherited traits may be different (Cho et al. 2012, p. e1002820). In recent years, studies of complex diseases are done using systems biology approaches and particularly network based approaches (Karbalaei et al. 2018, p. 27). Many studies have been performed using network biology approaches to understand the relationship between different diseases and especially different cancers. Such studies help to identify common genes to different cancers and their respective pathways which in turn provide relevant information to the development of secondary cancers and their metastasis (Sahrawat et al. 2017, p. 783).

Diabetes is the most common endocrine disorder caused by unusual insulin secretion. It is estimated that currently 425 million adults are affected with diabetes and by 2045 it will rise to 625 million (International Diabetes Federation 2017, para 1). Lack of insulin results in increased blood glucose levels, a condition referred as hyperglycemia and also results in impaired metabolism of carbohydrates, fat and proteins. Therefore, diabetes represents a cluster of metabolic disorders and can be divided into two major clinical types according to the etiopathology of the disorder namely- Type 1 diabetes and Type 2 diabetes (Kavakiotis et al. 2017, p. 104). The most common form of Diabetes is Type 2 diabetes, characterized by insulin resistance and is mainly affected by lifestyle, physical activity, eating habits and genetic factors. Type 1 diabetes is an autoimmune disease characterized by destruction of beta cells of pancreas that produce insulin (Ndisang et al. 2017, p. 1).

Alzheimer's Disease (AD) is one of the most common forms of dementia and various factors which triggers this complex disease are age, genetic, and environmental factors. Etiopathology of AD includes formation of thick intra-neuronal neurofibrillary tangles composed of hyperphosphorylated Tau protein and extracellular amyloid plaques (Singh et al. 2016, p. 2). It is estimated that
approximately 50 million people are affected by this disease worldwide (World Health Organisation 2019, para 1).

Recent studies indicate possibility of a direct relation between Diabetes and AD with reports of insulin being related to memory function and regulation of tau phosphorylation in patients with AD. Cell-culture and in vivo experiments also indicate that insulin-degrading enzyme degrades amyloid β (that form extracellular plaque deposits which are a hallmark of AD). Genetic linkage studies have also reported that a locus on chromosome 10 is situated near the insulin-degrading enzyme gene which has been linked to late-onset of AD (Arvanitakis et al. 2004, p. 665).

There are no reports on the association of Diabetes and AD from the systems network biology perspective and consequently a lacuna exists in knowledge of common genes and associated pathways between these two diseases. Therefore, the present study was undertaken to investigate the correlation between Diabetes and AD by comparing their networks using in silico network biology approach to identify genes common to both these diseases which would give meaningful insights to understand the pathophysiology and association of these diseases.

Materials and Methods

Retrieval of genes
The common genes which have been reported to be involved in Diabetes and AD were retrieved from DisGeNET database version v6.0 (Piñero et al. 2016, p. D833).

Construction of Gene Interaction Network
To study physical and functional interactions of common genes, a gene network was constructed using STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database (Szklarczyk et al. 2010, p. D561).

Analysis of Network
Physical and functional interactions of genes common between Diabetes and AD were analysed using various plug-ins of Cytoscape 3.7.1 software (Lopes et al. 2010, p. 2347).

ClusterONE
ClusterONE (Clustering with Overlapping Neighbourhood Expansion) finds overlapping regions in a protein-protein interaction network loaded in cytoscape. It was used to find highly connected regions in the form of clusters (Wan et al. 2013, p. 618).

CytoHubba
CytoHubba plugin is used to find hub genes present in the network. It provides a simple interface to analyze a network with eleven scoring methods. The scoring method used is Maximal Clique Centrality (MCC) as its precision quality is greater as compared to other methods. Hub genes were shown according to their rank (Chin et al. 2014, p. S11).

ClueGO
ClueGO plugin was used to find functionally annotated terms related to corresponding genes. It integrates Gene Ontology (GO) terms from various databases such as KEGG, Wiki pathways, Biocarta etc. and it also provides significance of the terms with their respective P-values (Bindea et al. 2009, p. 1091).

CluePedia
CluePedia is a Cytoscape plugin used to find potential biomarkers associated to the pathways. It also provides cellular location for the genes which can act as potential biomarkers. It calculates statistical dependencies from experimental data (Bindea et al. 2013, p. 661).
Results and Discussion

Genes involved in Diabetes and AD were retrieved from DisGeNET database in which 1982 and 1268 gene entries were present for Diabetes and AD respectively. Out of all these genes, 614 genes were common in both the diseases.

Construction of gene interaction network

Network construction was done using STRING database. Out of the 614 genes common between Diabetes and AD, only 591 were specific to *Homo Sapiens*. The protein-protein interaction network from STRING consists of 591 nodes (representing the genes) and 16675 edges (representing the interconnection between genes). Figure 1 shows the network obtained from STRING.

![Figure 1: Protein-Protein Interaction (PPI) network of common genes between Diabetes and Alzheimer's Disease obtained from STRING](image)

Clustering of Genes

Network obtained from STRING was imported in Cytoscape. This network was then analyzed using Cytoscape. Clustering was done using Cytoscape plug-in ClusterONE and a main network (Figure 2) was obtained in which red nodes [squared shaped] represent highly significant genes, while yellow [diamond shaped] and grey nodes [circles] represents least significant nodes and outliers respectively. Sixty five clusters were obtained, out of which only three clusters had significant P-value (<0.05) (Table 1).
Table 1: Clusters with significant P-value

<table>
<thead>
<tr>
<th>CLUSTER DETAILS</th>
<th>Nodes - 170</th>
<th>Density - 0.146</th>
<th>Quality - 0.502</th>
<th>P-value - 0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLUSTER DETAILS</td>
<td>Nodes - 148</td>
<td>Density - 0.500</td>
<td>Quality - 0.466</td>
<td>P-value - 0.000</td>
</tr>
<tr>
<td>CLUSTER DETAILS</td>
<td>Nodes - 115</td>
<td>Density - 0.500</td>
<td>Quality - 0.320</td>
<td>P-value - 0.042</td>
</tr>
</tbody>
</table>

Figure 2: ClusterONE results obtained by analyzing the Network

Hub genes

Hub genes were identified using CytoHubba plug-in of Cytoscape that uses Maximum Clique Centrality (MCC) method. The network obtained from ClusterONE was analysed with CytoHubba. Fifteen genes were found to be significant amongst the common genes of Diabetes and AD namely Tumor necrosis factor (TNF), Interleukin 6 (IL6), Vascular endothelial growth factor (VEGFA), all three of them were ranked as 1, Interleukin 10 (IL10), Albumin (ALB), Interleukin 8 (CXCL8), Insulin (INS), Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), chemokine ligand 2 (CCL2), Interleukin 1 beta (IL1B), Protein kinase B (AKT1), Interleukin 4 (IL4), Toll-like receptor 4 (TLR4), Colony Stimulating Factor (CSF2), and Intercellular Adhesion Molecule (ICAM1) (Table 2).
Table 2: Ranks of Hub genes obtained from CytoHubba plug-in of Cytoscape.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Gene</th>
<th>Rank</th>
<th>Gene</th>
<th>Rank</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TNF, IL6 and VEGFA</td>
<td>4</td>
<td>IL10</td>
<td>8</td>
<td>GAPDH</td>
</tr>
<tr>
<td>5</td>
<td>ALB</td>
<td>9</td>
<td>CCL2</td>
<td>13</td>
<td>TLR4</td>
</tr>
<tr>
<td>6</td>
<td>CXCL8</td>
<td>10</td>
<td>IL1B</td>
<td>14</td>
<td>CSF2</td>
</tr>
<tr>
<td>7</td>
<td>INS</td>
<td>11</td>
<td>AKT1</td>
<td>15</td>
<td>ICAM1</td>
</tr>
</tbody>
</table>

The hub genes obtained from CytoHubba are also present in the three clusters shortlisted using Cytoscape plug-in ClusterONE. Therefore, the identified hub genes, may act as promising candidates for further functional validation studies and potentially represent important points in the network for therapeutic interventions.

**Enrichment of genes**

All the 15 hub genes obtained from CytoHubba were analyzed using Cytoscape plug-in ClueGo that allows analysis of interrelations of terms and functional groups in the biological networks. In Figure 3, each node represents a significant pathway and edge represents a pathway crosstalk (representing significant overlap of the component genes between two linked pathways).

A total of 198 pathways were obtained which are divided into eight GO groups that represents the pathways with GO term, GOID, P-value and their associated genes. The colours in the nodes are based on their respective GO groups and shared genes between them respectively. The label in the nodes reflects the group leading terms such as Amoebiasis, TNF signaling Pathway, Malaria, Rheumatoid Arthritis *etc.*

ClueGo also provides pie-chart which shows the summarised output of all GO groups represented by different colours along with their occurrence in the group with their respective percentages (Figure 4). The biological role of the genes visualised with ClueGO are associated with GO and KEGG terms—neurotransmitter biosynthetic process, HIF-1 signaling pathway, IL-17 signaling pathway, TNF signaling pathway, Non-alcoholic fatty liver disease (NAFLD), AGE-RAGE signaling pathway in diabetic complications, Toll-like receptor signaling pathway *etc.*
Identification of potential biomarkers

Cytoscape plug-in CluePedia automatically cite the cellular location of markers with highest interaction score for all the selected nodes with the help of GO terms. It also maps the markers on pre-defined cellular compartments. From the 15 hub genes, cellular location was retrieved from CluePedia only for 11 genes which can act as potential marker for both the Diabetes and AD namely IL4, ICAM1, ALB, INS, CSF2, IL6, TNF, IL10, GAPDH, TLR4, and AKT1 (Figure 5).

These genes were found to be enriched in pathways of various GO groups which includes pathways like neurotransmitter biosynthetic process, HIF-1 signaling pathway, IL-17 signaling pathway, TNF signaling pathway, Non-alcoholic fatty liver disease (NAFLD), Toll receptor signaling pathway etc.
The results of the present study were also validated by literature review of previous reports. In a study to identify relationship between Diabetes and AD using *in vitro* experiments performed by Nan-Qu Huang et al. (2016, p. 234), it has been reported that TLR4 is involved in the physiological and pathological progress of Diabetes and AD. While identifying the connection between Diabetes and AD using molecular biology approach performed by Felice et al. (2014, p. 2262) and Liu et al. (2015, p. 619), it was reported that TNF and AKT1 are involved in pathophysiology of both Diabetes and AD respectively.

In a study performed by Kikodze et al. (2013, p. 29) using molecular biology approach, it was reported that interleukins IL4, IL6, IL10 play an important role in Diabetes whereas in an independent study on genetic polymorphisms performed by Su et al. (2016, p. 569) it was found that IL4, IL6, IL10 also play an important role in pathophysiology of AD. In a study performed by Bhat et al. (2017, p. 677) using mass spectrometry based approaches, it has been reported that albumin (ALB) has evolved to be one of the best candidates in the pursuit of diagnostic markers for prediction of pre-diabetes and diabetic complications. Another study by Costa et al. (2018, p. 1395) reports that post-translational modifications of albumin cause oxidative stress in the brain and peripheral systems which is considered a major player in pathophysiology of AD.

![Figure 5: Potential Markers and their cellular location obtained from CluePedia Plug-in of Cytoscape](image-url)

Two separate studies by Seydel et al. (2008, p. 377) and Li et al. (2015, p. 6548) using molecular biology approach and microarray respectively were conducted and identified the role of CSF2 in Diabetes and Alzheimer’s Disease. *In vitro* analysis performed by Matthew et al. (2008, p. 272), they found GAPDH to be involved in pathogenesis of Diabetic complications while Butterfield et al. (2010, p. 369) reported that GAPDH interacts with neurodegenerative disease-associated proteins.

In a molecular biology study performed by Alberto et al. (2002, p. 14), it was reported that insulin (INS) has such a central role in pathogenesis of both forms of diabetes whereas a study performed by Mullins et al. (2017, p. 1933) using MPRAGE images that were segmented by brain tissue type; and voxel-based morphometry (VBM) analysis for gray matter, found an important role of INS in pathophysiology of AD. In a study performed by Gu et al. (2013) using genomics approach, it was reported that ICAM1 act as a biomarker and target for prediction and treatment of diabetes and diabetic nephropathy while Frohman et al. (1991, p. 105) reported that ICAM1 is involved in pathophysiology of AD.
Some evidence from epidemiological and basic sciences has been reported by Akter et al. (2011, p. 365) that Type 2 Diabetes Mellitus (T2DM) and AD may share common pathophysiology and has even hypothesized that AD might be ‘type 3 diabetes’. But the relationship between Diabetes and AD has not been investigated till date. This is a first study of its kind to identify potential biomarkers, common to both diabetes and AD.

**Conclusion**
In complex networks, the identification of regulatory hubs involves multi-targeted Systems biology approaches. From the present study we have been able to identify IL4, ICAM1, ALB, INS, CSF2, IL6, TNF, IL10, GAPDH, TLR4 and AKT1 which can act as potential biomarker for both Diabetes and AD.

**Recommendation**
Identification of common biomarkers amongst various diseases, using a network-systems biology approach, can play an important role to understand the pathophysiology of diseases and in design of their diagnosis and treatment strategies.

**Acknowledgement**
The authors would like to thank their parent institute Panjab University Chandigarh, India for providing the infrastructure for carrying out the research.

**Conflict of Interest**
There is no conflict of interest related to this study.

**References**


