MiR-15b and let-7a as Non-invasive Diagnostic Biomarkers of Alzheimer’s Disease Using an Artificial Neural Network

Masoumeh Darvishi Talemi, Leili Tapak, Alireza Rastgoo Haghi, Parisa Ahghari, Shadi Moradi, Saeid Afshar

1Department of Biology, Faculty of Basic Sciences, Bu-Ali Sina University, Hamedan, Iran
2Department of Biostatistics, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran
3Department of Pathology, Hamadan University of Medical Sciences, Hamadan, Iran
4Research Center for Molecular Medicine, Hamadan University of Medical Sciences, Hamadan, Iran
5Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Hamadan University of Medical Sciences, Hamadan, Iran

Abstract

**Background:** Gaining insight into the underlying molecular mechanisms of Alzheimer’s disease (AD) is crucial.

**Objectives:** This study aimed to employ a systems biology approach to identify new non-invasive diagnostic biomarkers for AD.

**Methods:** Gene expression data series GSE122063 and microRNA (miRNA) expression data series GSE90828 were obtained from the Gene Expression Omnibus database. The Limma package under R software was used to assess differentially expressed miRNAs and differentially expressed genes (DEGs). Afterward the protein-protein interaction (PPI) network was constructed by the STRING software and evaluated with Cytoscape software. The multilayer perceptron neural network (MLP-NN), a widely used artificial neural network (ANN), was employed to classify two groups.

**Results:** A total of 1388 DEGs were identified in AD patients compared to the control group, and 11 differentially expressed miRNAs were found in patients with mild cognitive impairment (MCI) in comparison to the control group. The results revealed that EGFR, identified as a hub gene, was targeted by miR-15b-3p and let-7a-5p, while TLR4, another hub gene, was targeted by miR-15b-3p. The MLP-NN constructed using both hsa-let-7a-5p and hsa-miR-15b-3p achieved a sensitivity of 0.857 and an area under the curve of 0.917 in detecting Alzheimer’s patients.

**Conclusion:** Our findings suggest that miR-15b-3p, by targeting EGFR and TLR4, and let-7a-5p, by targeting EGFR, may play a significant role in AD. Additionally, the constructed ANN utilizing the expression levels of plasma miR-15b-3p and let-7a-5p could serve as a potential non-invasive diagnostic tool with high sensitivity for AD detection.

**Keywords:** Alzheimer’s disease, miRNA, Biomarker, Artificial neural network

Background

Alzheimer’s disease (AD) accounts for about 70% of dementia cases and is characterized by severe cognitive decline, affecting memory, learning, occupational function, thinking ability, and language. It profoundly affects daily life, causing economic and social burdens on patients (1). AD, as a slowly progressive neurodegenerative disorder, is a disease that leads to the deterioration of brain cells. In addition, it stands as the primary underlying factor of dementia (2). Currently, there are approximately 50 million AD patients worldwide, and it is projected that this number will triple by 2050. AD places a significant burden on those who are affected by it, as well as on their families and the economy (3). This disease can be classified into two categories based on age of onset and pathological factors. Early-onset or familial AD (EOAD) presents before the age of 65 and may co-occur with conditions such as hypertension, metabolic syndrome, and diabetes mellitus. Late-onset AD (LOAD), on the other hand, manifests after 65 years and is associated with amyloid plaque formation (4). The progression of the disease is influenced by immutable factors such as age and genetic risk, as well as modifiable factors, including cardiovascular risk and lifestyle choices (5). The accumulation of tau protein and amyloid-β plays a pivotal role in AD development, occurring years before...
cognitive impairment emerges (6). The disease begins with abnormal processing of the amyloid-β precursor protein, resulting in the generation of β amyloids (7). Impaired degradation and decreased clearance at the blood-brain barrier result in an increased amount of toxic Aβ42. Subsequently, insoluble plaques precipitate in the brain, particularly in areas such as the medial temporal lobe, parietal lobe, and frontal lobe (8). Amyloid plaques lead to the blockade of neuronal membrane receptors such as AMPAR, NMDAR, mAChRs, and mAChR, resulting in the ultimate impairment of synaptic transmission. These plaques trigger the generation of reactive oxygen species (ROS), which in turn promotes mitochondrial oxidative stress and the activation of apoptotic-related pathways. Moreover, the production of ROS by amyloid plaques results in the hyperphosphorylation of tau and the disruption of microtubules due to the activation of ERK2, PFA, and PKC (9). Elevated Aβ amyloid concentration creates an excitotoxic environment, leading to neuronal injury, disruption of neuronal homeostasis, oxidative injury, and hyperphosphorylation of the microtubule-associated protein tau. These events contribute to cerebral atrophy (10). AD’s pathophysiology is multi-dimensional, reflecting intricate genetic heterogeneity (11). Preliminary biomarker studies in preclinical AD have shown familial or autosomal dominant manifestations confirmed by genetic tests, cerebrospinal fluid, blood, or brain abnormalities (12). Advances in genetic and genomic technologies have enhanced our understanding of AD’s genetic structure, with genes such as APP, PSEN1, PSEN2, and APOE having significant effects on disease susceptibility (13). Additionally, various studies have identified several genes as potential diagnostic biomarkers for AD (11,13-15). MicroRNAs (miRNAs) have also been suggested as potential candidates for early AD diagnosis (16). They are a group of non-coding RNAs, containing approximately 22 nucleotides, that regulate target genes (17-19). The evaluation of expression levels of circulating miRNAs in patients may aid in early AD diagnosis. Current studies suggest that miRNAs play an crucial role in the initiation and development of AD by influencing various processes, including Aβ metabolism, synaptic plasticity, immune-inflammatory responses, neuronal growth, differentiation, and apoptosis (20).

Recently, the application of machine learning methods in medicine has garnered significant attention (21). Among these techniques, artificial neural networks (ANNs) have stood out for their capacity to model nonlinear correlations between variables. ANNs have been employed in various medical applications, including diagnosis, screening, and image processing (22).

The aim of molecular biology studies in AD is to acquire a deeper understanding of the fundamental mechanisms of disease risk and develop targeted treatments to intercept or delay its onset. In this research, we adopted a systems biology approach to explore the molecular pathways involved in AD and discover new non-invasive diagnostic biomarkers for the disease. Additionally, we utilized ANNs for the classification of Alzheimer’s patients and healthy controls.

Materials and Methods
Gene and miRNA Expression Data
In this system biology study, gene expression data were retrieved from the Gene Expression Omnibus (GEO) database (under accession number GSE122063) using the platform GPL16699. This data series consists of 136 frontal and temporal cortex samples, comprising 36 vascular dementia (VaD) cases, 56 AD cases, and 44 non-demented controls. For miRNA expression data, information was retrieved from GEO database (under accession number GSE90828) utilizing platform GPL22741. This data series includes 53 plasma samples, comprising 30 control samples and 23 mild cognitive impairment (MCI) samples.

The Limma package in R software was employed to assess the differentially expressed genes (DEGs) between 56 AD samples and 44 non-demented control samples. Similarly, the Limma package was used to evaluate the differentially expressed miRNAs between 23 MCI samples and 30 control samples, with the criterion of Abs(logFC) ≥ 1 and P ≤ 0.05. MicroRNAs target prediction.

The miRWalk software (http://zmf.umm.uniheidelberg.de/apps/zmf/mirwalk2/) was applied to determine the valid target genes of differentially expressed miRNAs.

Functional/Enrichment Analysis
The DAVID software (https://david.ncifcrf.gov/) was employed for conducting gene ontology (GO) and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway enrichment analysis. The GO analysis covered biological processes (BP), molecular functions, and cellular components.

Network Construction
The protein-protein interaction (PPI) network of the selected DEGs was created through String software (https://string-db.org/), with a minimum required interaction score of 0.7. Subsequently, the constructed network was assessed using Cytoscape software (version 3.6). To identify hub genes, the degree of the network was analyzed using the cytoHubba plugin within Cytoscape software. A bipartite miRNA-mRNA graph was also created by merging the PPI network with the miRNA-target gene network.

Artificial Neural Network
In this research, a multilayer perceptron neural network (MLP-NN), widely used in ANNs, was utilized for the classification of two groups. The MLP-NN comprises the input, output, and hidden layers. Activation functions within the MLP-NN facilitate data transformation from one layer to the next. For our investigation, the input layer incorporated the expression levels of selected miRNAs, and a single hidden layer was utilized as well. The
output layer represented the binary response variable for Alzheimer’s presence, indicating the probable outcomes for Alzheimer’s and healthy controls. To achieve a composite nonlinear mapping between the input and output layers, several nodes were empirically identified in the hidden layer to optimize network performance. The hyperbolic tangent as the activation function was used for the hidden layer, while the “softmax” function was employed in the output layer.

**Results**

**Identification of Differentially Expressed Genes and Differentially Expressed MicroRNAs**

The results obtained from the Limma package revealed a total of 1388 DEGs that were selected in AD patients compared to the control group. Additionally, 11 differentially expressed miRNAs were identified in MCI patients when compared to the control group. These miRNAs included hsa-miR-339-3p, hsa-miR-374a-5p, miR-15b-3p, hsa-miR-151-3p, miR-93-3p, hsa-miR-652-5p, hsa-miR-27b-5p, hsa-let-7a-5p, hsa-miR-374b-5p, hsa-miR-1974, and hsa-miR-146a-5p.

**Protein-Protein Interaction Network and MicroRNA-mRNA Network Analysis**

The depicted PPI network for 1388 DEGs was visualized and analyzed using Cytoscape. According to the outcomes of the network analysis conducted with cytoHubba, the top 10 hub genes with the highest degree were selected, including SNAP25, SYP, EGFR, BDNF, PTPRC, SYT1, SLC32A1, GAD2, GFAP, and TLR4.

Furthermore, upon merging and analyzing the constructed network for DEGs and the miRNA-target gene network, it was found that among the 11 differentially expressed miRNAs, only 6 miRNAs (miR-652-5p, let7a-5p, miR-15b-3p, miR-27b-5p, miR-374b-5p, and miR-93-3p) had valid targets that were also chosen as DEGs (Figure 1). Specifically, EGFR and SETD7 were the common targets of miR-15b-3p and let7a-5p, while SV2B was the common target of miR-93-3p and miR-27b-5p. In addition, EGFR, as one of the top 10 hub genes, was targeted by both let7a-5p and miR-15b-3p (Figure 2). Similarly, SYP and TLR4, also among the top 10 hub genes, were targets of miR-374b-5p and miR-15b-3p, respectively (Figure 2).

**Gene Ontology and the Kyoto Encyclopedia of Genes and Genome Pathway Analysis**

The results of the GO enrichment analysis revealed significant enrichments in several BP terms, including chemical synaptic transmission, neurotransmitter secretion, glutamate secretion, inflammatory response, and neuropeptide signaling pathway. For cellular component terms, enrichments were observed in cell junction, synaptic vesicle membrane, axon, plasma membrane, synapse, integral component of the plasma membrane, neuronal cell body, neuron projection, synaptic vesicle, and postsynaptic membrane. In terms of molecular functions, enrichments were found in calcium ion binding, calmodulin binding, and neuropeptide hormone activity.

Additionally, the KEGG pathway analysis demonstrated enrichments in several pathways, including retrograde endocannabinoid signaling, gamma-aminobutyric acid (GABA)ergic synapse, nicotine addiction, neuroactive ligand-receptor interaction, staphylococcus aureus infection, cholinergic synapse, glutamatergic synapse, calcium signaling pathway, synaptic vesicle cycle, and morphine addiction (Figure 3).

![Figure 1. Constructed Network for DEGs and miRNA-target Genes. Note. DEG: Differentially expressed gene; miRNA: MicroRNA. Among 11 differentially expressed miRNAs, only 6 miRNAs had valid targets selected as DEGs. EGFR and SETD7 were the common targets of miR-15b-3p and let7a-5p. In addition, SV2B was the common target of miR-93-3p and miR-27b-5p](image-url)
Figure 2. The Constructed PPI Network Merged With the MiRNA-mRNA Interaction Network. Note: PPI: Protein-protein interaction; miRNA: MicroRNA. The top 10 hub genes with the highest degree, including SNAP25, SYT, EGFR, BDNF, PTPRC, SYT1, SLC32A1, GAD2, GFAP, and TLR4, are shown in green color. EGFR, as a top 10 hub gene, was a target of let-7a-5p and miR-15b-3p. Further, SYT and TLR4 as top 10 hub genes were the targets of miR-374b-5p and miR-15b-3p, respectively.

Figure 3. The Results of the KEGG Pathway and GO Enrichment Analysis. Note: KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene ontology. Five selected biological process terms with a $P$-value less than 0.05, the top ten cellular component terms, tree molecular functions with a $P$-value less than 0.05, and the top ten KEGG pathways have been shown.
Classification of Two Groups With Artificial Neural Networks

Based on the data in Table 1, the area under the receiver operating characteristic (ROC) curve (AUC), specificity, and sensitivity were calculated for hsa-miR-374b-5p, hsa-miR-15b-3p, and hsa-let-7a-5p, as well as the combination of hsa-let-7a-5p and hsa-miR-15b-3p, using the MLP-NN method. All three AUC values were found to be significantly greater than 0.5, which is the reference value indicating random prediction.

Among the scenarios, the MLP-NN constructed using both hsa-let-7a-5p and hsa-miR-15b-3p exhibited a sensitivity of 0.857 and an AUC of 0.917 in detecting Alzheimer’s patients, which were superior to the other scenarios (Figure 4).

Discussion

Diagnosing AD in its early stages can be pivotal in preventing disease progression, expediting the treatment process, and potentially improving patients’ cognition and overall quality of life. The conventional methods of diagnosing AD have limitations, such as late detection, invasiveness, high costs, complex procedures, time-consuming processes, low sensitivity, and the possibility of false-positive results. Therefore, the identification of novel non-invasive diagnostic biomarkers with high accuracy is of paramount importance.

The results of the current study revealed that six miRNAs, namely, miR-652-5p, let7a-5p, miR-15b-3p, miR-27b-5p, miR-374b-5p, and miR-93-3p, had valid targets that were also selected as DEGs. Among these, miR-652-5p, a conserved member of the miR-652 family, plays a role in controlling several cellular processes, including apoptosis, proliferation, angiogenesis, and glycolysis (23,24). Wang et al indicated that miR-652-5p is associated with AD progression (25). Another miRNA, let-7a-5p, Table 1. Sensitivity, Specificity, and Area Under the ROC for the Three Genes Using a Multilayer Perceptron Neural Network

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-374b-5p</td>
<td>0.728</td>
<td>0.714</td>
<td>0.667</td>
</tr>
<tr>
<td>hsa-miR-15b-3p</td>
<td>0.802</td>
<td>0.857</td>
<td>0.696</td>
</tr>
<tr>
<td>hsa-let-7a-5p</td>
<td>0.778</td>
<td>0.750</td>
<td>0.833</td>
</tr>
<tr>
<td>All miRNAs</td>
<td>0.751</td>
<td>0.600</td>
<td>0.913</td>
</tr>
<tr>
<td>hsa-let-7a-5p and hsa-miR-15b-3p</td>
<td>0.848</td>
<td>0.917</td>
<td>0.692</td>
</tr>
</tbody>
</table>

Note: ROC curve: Receiver operating characteristic curve. AUR: Area under the curve.

Figure 4. The Area Under the ROC Curve of (a) hsa-let-7a-5p and hsa-miR-15b-3p, (b) hsa-let-7a-5p, (c) hsa-miR-374b-5p, and (d) hsa-miR-15b-3p. Note: ROC curve: Receiver operating characteristic curve.
a member of the let-7 family, is involved in regulating critical BPs such as apoptosis and inflammation. It plays a role in controlling inflammatory injury in microglia and neuronal autophagy (26). Furthermore, the findings of Zhao et al demonstrated that miR-15b-3p targets WNT5a, thereby regulating neuronal and astrocyte differentiation (27). mir-27b-5p is known to be involved in the control of several BPs, such as proliferation, apoptosis, and migration (28,29). MiR-374b-5p plays a crucial role in neuronal stem cell differentiation and proliferation and has been associated with neurodegenerative changes and neuroinflammation in AD (30). Additionally, mir-93-3p, which is located on chromosome 13q31.3, was found to be dysregulated in cerebrospinal fluid after certain conditions. Moreover, the expression level of this miRNA was found to be reduced in the serum and cerebrospinal fluid of patients with neurosyphilis (31,32).

Furthermore, our findings confirmed that among the identified hub genes, EGFR and SETD7 were the most common targets of miR-15b-3p and let-7a-5p, respectively. Additionally, SV2B was identified as a common target of miR-93-3p and miR-27b-5p. Furthermore, SYP and TLR4 were found to be the targets of miR-374b-5p and miR-15b-3p, respectively. Notably, SETD7, functioning as a lysine methyltransferase, has implications in the post-translational modification of tau protein, influencing its subcellular localization (33). SV2B plays a crucial role in regulating neurotransmitter release, which is widely dispersed everywhere in brain. Moreover, this gene is crucial in mitigating the toxicity associated with amyloid (34). The protein encoded by EGFR is a transmembrane tyrosine kinase receptor that regulates various cellular processes, including apoptosis, proliferation, differentiation, and adhesion. This gene is crucial for the survival of neuronal cells, and its polymorphisms have been associated with AD (35). SYP encodes an integral synaptic vesicle membrane protein that is involved in synaptic plasticity and cognitive impairment (36). TLR4, as a member of the toll-like receptor family, is involved in controlling neuronal plasticity, proliferation of neuronal precursor cells, and neuroinflammation (37).

The GO enrichment analysis indicated that chemical synaptic transmission, neurotransmitter secretion, glutamate secretion, inflammatory response, neuropeptide signaling pathway, calcium ion binding, calmodulin binding, and neuropeptide hormone activity terms were significantly enriched. The results of similar studies revealed that chemical synaptic transmission, neurotransmitter secretion (38), inflammatory response (39), neuropeptide signaling pathway (40), calcium ion binding (41), calmodulin binding (42), and neuropeptide hormone activity (43) were associated with AD.

Furthermore, the KEGG pathway analysis indicated that retrograde endocannabinoid signaling, GABAergic synapse, nicotine addiction, neuroactive ligand-receptor interaction, staphylococcus aureus infection, cholinergic synapse, glutamatergic synapse, calcium signaling pathway, synaptic vesicle cycle, and morphine addiction terms were significantly enriched. The results of related studies in the molecular mechanism of AD showed that GABAergic synapse (44), retrograde endocannabinoid signaling (45), nicotine addiction, morphine addiction (46), Staphylococcus aureus infection (47), cholinergic synapse (48), glutamatergic synapse pathway (49), calcium signaling pathway (50), and synaptic vesicle cycle (51) were significantly enriched.

Finally, the findings of the current study revealed that the AUC and sensitivity of the MLP-NN constructed by both hsa-let-7a-5p and hsa-miR-15b-3p in detecting Alzheimer’s patients were remarkably high, indicating their potential as non-invasive diagnostic biomarkers for AD through the ANN approach.

**Conclusion**

In summary, the constructed ANN, utilizing the expression levels of plasma miR-15b-3p and let-7a-5p, could serve as a non-invasive diagnostic tool with high sensitivity. Further, our results suggest that miR-15b-3p, by targeting EGFR and TLR4, and let-7a-5p, by targeting EGFR, may play essential roles in AD.

**Authors’ Contribution**

**Conceptualization:** Masoumeh Darvishi Talemi, Leili Tapak, Shadi Moradi, Saeid Afshar.

**Data curation:** Masoumeh Darvishi Talemi, Leili Tapak, Saeid Afshar.

**Investigation:** Masoumeh Darvishi Talemi, Leili Tapak, Alireza Rastgoo Haghi, Parisa Aghhari, Shadi Moradi, Saeid Afshar.

**Methodology:** Masoumeh Darvishi Talemi, Leili Tapak, Saeid Afshar.

**Project administration:** Leili Tapak, Saeid Afshar.

**Supervision:** Saeid Afshar.

**Validation:** Masoumeh Darvishi Talemi, Leili Tapak, Saeid Afshar.

**Visualization:** Masoumeh Darvishi Talemi, Leili Tapak, Shadi Moradi, Saeid Afshar.

**Writing—original draft:** Masoumeh Darvishi Talemi, Leili Tapak, Alireza Rastgoo Haghi, Parisa Aghhari, Shadi Moradi, Saeid Afshar.

**Writing—review & editing:** Masoumeh Darvishi Talemi, Leili Tapak, Alireza Rastgoo Haghi, Parisa Aghhari, Shadi Moradi, Saeid Afshar.

**Competing Interests**

The authors declare no conflict of interests.

**Ethical Approval**

This study was approved by the Ethics Committee of the Hamadan University of Medical Sciences (Ethics No.: IR.UMSHA.REC.1401.401).

**Funding**

The study was funded by the Vice-chancellor for Research and Technology, Hamadan University of Medical Sciences, Hamadan, Iran (grant number: 140105183660).

**References**


3. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard...
Darvishi Talemi et al.


