

SCIENTIFIC REPORTS



OPEN

Analysis of microRNA and Gene Expression Profiles in Alzheimer's Disease: A Meta-Analysis Approach

Shirin Moradifard, Moslem Hoseinbeyki, Shahla Mohammad Ganji & Zarrin Minuchehr 

Understanding the molecular mechanisms underlying Alzheimer's disease (AD) is necessary for the diagnosis and treatment of this neurodegenerative disorder. It is therefore important to detect the most important genes and miRNAs, which are associated with molecular events, and studying their interactions for recognition of AD mechanisms. Here we focus on the genes and miRNAs expression profile, which we have detected the miRNA target genes involved in AD. These are the most quintessential to find the most important miRNA, to target genes and their important pathways. A total of 179 differentially expressed miRNAs (DEmiRs) and 1404 differentially expressed genes (DEGs) were obtained from a comprehensive meta-analysis. Also, regions specific genes with their molecular function in AD have been demonstrated. We then focused on miRNAs which regulated most genes in AD, alongside we analyzed their pathways. The miRNA-30a-5p and miRNA-335 elicited a major function in AD after analyzing the regulatory network, we showed they were the most regulatory miRNAs in the AD. In conclusion, we demonstrated the most important genes, miRNAs, miRNA-mRNA interactions and their related pathways in AD using Bioinformatics methods. Accordingly, our defined genes and miRNAs could be used for future molecular studies in the context of AD.

Alzheimer's disease (AD) is one of the ultimately fatal neurodegenerative diseases affecting more than 35 million people worldwide¹. In 2020, it has been predicted that the number of people affected with this disease will rise worldwide². In developed countries, Alzheimer's disease (AD) is the sixth leading cause of all deaths. While other major causes of death are declining, deaths caused by Alzheimer grows dramatically³. The neurodegenerative diseases such as AD is a type of disorder that neuronal function and structure is degenerated following by the death of neurons in the nervous system⁴. The greatest risk factor for neurodegeneration is age increase besides other risk factors. The pathophysiological process of AD patients before the clinical diagnosis is unknown⁴. The clinical pathogenesis of AD is said to be the accumulation of insoluble and extracellular amyloid-beta ($A\beta$) plaques and intracellular neurofibrillary tangles (NFT) in the brain⁵. This process has been shown intervened in long term potentiation (LTP), necessary in neuronal signaling, interfered in the signaling involved in pro-apoptotic signaling and results in neuronal loss⁶. The application of molecular mechanisms in the diagnosis of the AD is not clear, but it is thought that many factors are involved in the pathogenesis of AD. Therefore, such studies should have a high impact on AD diagnosis and treatment³.

miRNAs- single stranded non-coding RNAs- are small (18–25 nucleotides) and involved in the post-transcriptional regulation of gene expression⁷. Indeed, characterization of regulatory RNAs is one of the most important findings in the context of molecular biology in recent years⁸. In brief, miRNAs could bond partial complementarity to messenger RNA sequences, often in the 3' untranslated region (3'UTR). It has been shown that miRNAs participate and have implicated in neural development and differentiation as well as approximately 70% of all miRNAs expressions in the brain and they can function as biological regulators in neurons, for instance, neuronal differentiation, neurogenesis and synaptic plasticity^{5,9}. Therefore, it seems that miRNAs have a potential role in neurodegenerative diseases,⁵ and in particular AD. Evidence showed that miRNAs are involved in deregulation of neurodegenerative diseases⁴. Many studies demonstrated the expression of specific miRNA in the central nervous system (CNS) with different roles^{7,10}. Therefore a comprehensive study in miRNAs involved in neurodegenerative diseases could be conveniently used in innovative therapies⁴.

National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran. Shirin Moradifard and Moslem Hoseinbeyki contributed equally to this work. Correspondence and requests for materials should be addressed to S.M.G. (email: shahla@nigeb.ac.ir) or Z.M. (email: minuchehr@nigeb.ac.ir)

Data set selection flowchart

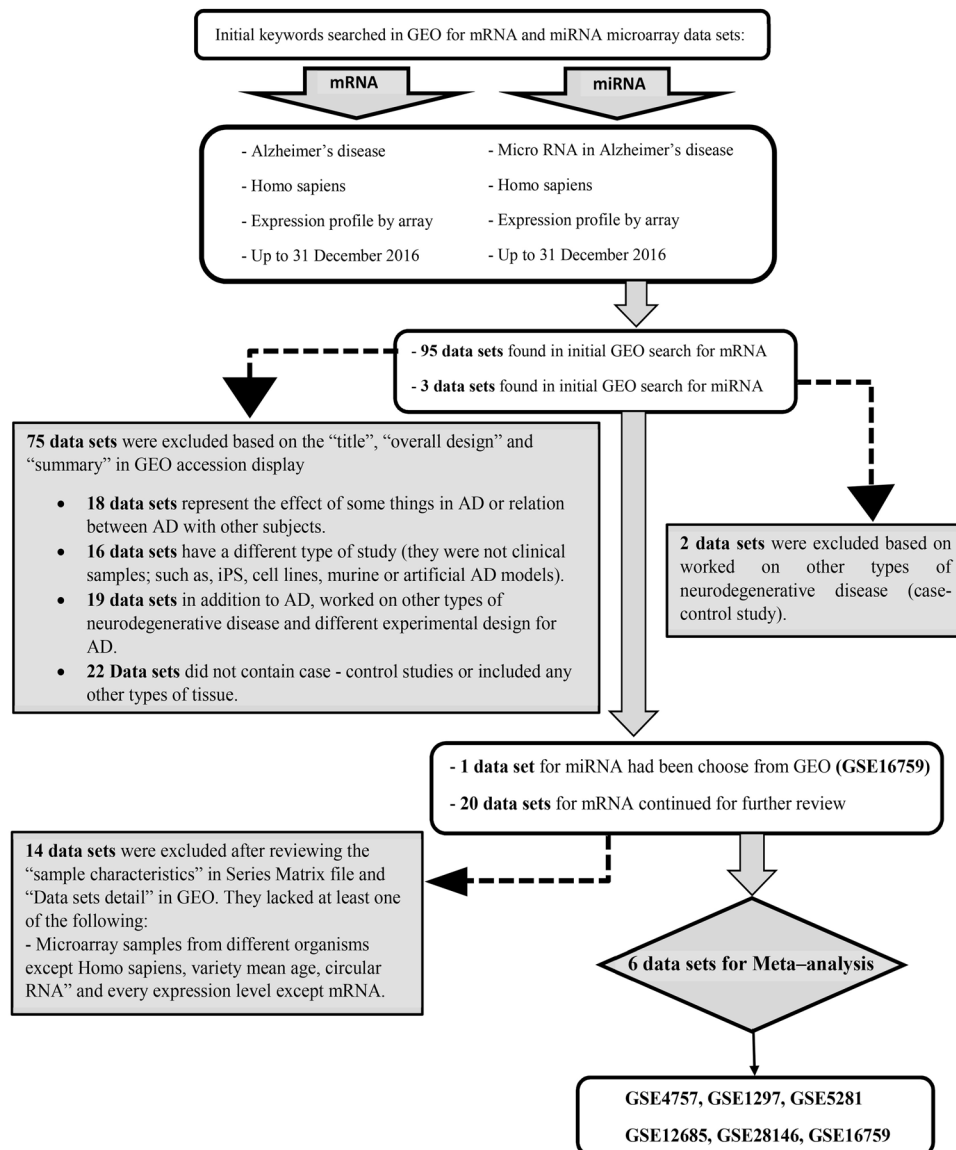


Figure 1. Data set selection flow chart. A total of 95 data sets from GEO was evaluated. Finally, 6 data sets for mRNA and 1 data set for miRNA were selected to be included in this meta-analysis.

The aim of this study is focused on miRNAs involved in AD and their target genes, the determination of the most important miRNAs, genes and their pathways in Alzheimer's disease. Here we investigated and identified AD related differentially expressed genes (DEGs) and miRNAs (DEmiRs), miRNA-mRNA interactions and signaling pathways. Our results showed the different expressions of some miRNA and their target genes in a disease group vs normal; that this miRNA and their target genes have important roles in AD and other neurodegenerative diseases. In addition, pathways are related to miRNAs-target genes showed that it is significantly linked to the AD.

Results

Gene and miRNA expression profile. *Differentially Expressed miRNAs and genes in AD.* Following the data sets selection according to our criteria (Fig. 1), analyzing the seven microarray data sets of genes and miRNAs expression profile according to the workflow (Table 1, Fig. 2) was done. After quality control and normalization, expression profile for each data set was created and a meta-analysis was performed. By using our criteria, differentially-expressed microRNA and genes were divided into Up- or Down-regulated miRNAs and genes. The P value < 0.05 and a fold-change ≥ 1.23 were set as the cut off values of DEgenes and DEmiRs. Our results in the meta-analysis showed a list of 1404 DEGs including 672 Up-regulated and 732 Down-regulated DEGs were selected and used in our subsequent analysis (supplementary Table S1). About DEmiRs expression analysis showed that a total of 179 differentially expressed miRNAs, 83 Down-regulated and 96 Up-regulated miRNAs in AD (supplementary Table S2). In Table 2 we summarized the most 30 Up- and Down-regulated DEmiRs which

	GEO accession number	Experiment	Sample size		Platform
			Controls	Patients	
mRNA					
1	GSE4757	Dunckley,... <i>et al.</i> ⁶¹	10	10	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
2	GSE12685	Williams,... <i>et al.</i> ⁶²	8	6	[HG-U133A] Affymetrix Human Genome U133A Array
3	GSE28146	Blalock,... <i>et al.</i> ⁵⁹	8	22	[HGU133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
4	GSE1297	Blalock,... <i>et al.</i> ⁶⁰	9	22	[HG-U133A] Affymetrix Human Genome U133A Array
5	GSE5281	Liang,... <i>et al.</i> ⁶³	74	87	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
		Liang,... <i>et al.</i> ⁶⁴			
6	GSE16759	Nunez-Iglesias,... <i>et al.</i> ⁴⁷	4	4	[HGU133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
miRNA					
7	GSE16759	Nunez-Iglesias,... <i>et al.</i> ⁴⁷	4	4	USC/XJZ Human 0.9 K miRNA940v1.0

Table 1. Microarray data sets used in this study and their experimental design. Gene Expression Omnibus dataset (GEO) is the one of the international public repositories in the context of high-throughput data; Controls: control samples; Patients: Alzheimer's disease patients.

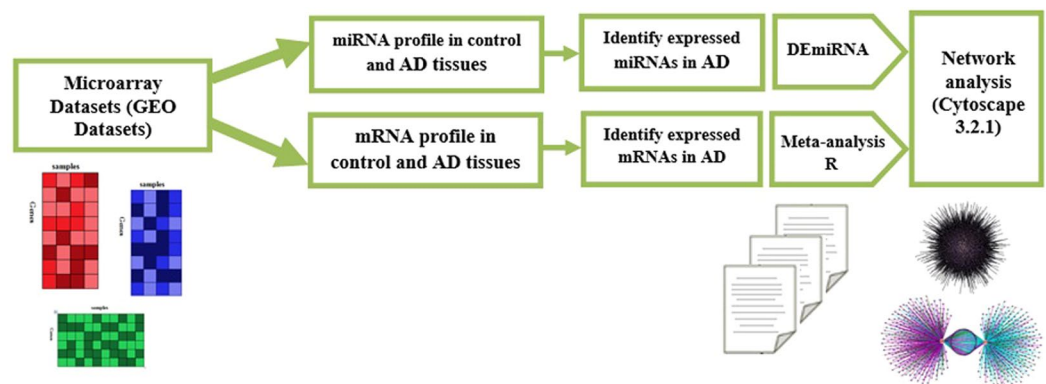


Figure 2. Workflow and analysis process.

by their roles interfere in AD. For visualizing the Differentially Expressed miRNAs and genes, we sorted them. We also categorized the top Up- and Down-regulated DE genes in AD in Table 3.

For further analyzing on the meta-analysis results (Up- and Down-regulated genes), we used ClueGO a Bioinformatics tool for clustering pathways and gene ontology terms. This tool visualized the interactions between genes and clusters; therefore, we can separate them easily based on their importance. The most important among them had been chosen by “Term P value Corrected with Bonferroni step down.” The KEGG pathway analysis of Up- and Down-regulated genes that were highly enriched have been shown here: the significant pathways in Up-regulated genes were ECM-receptor interaction and Cell adhesion molecules (CAMs) but for Down-regulated genes, which were involved in 10 pathways were oxidative phosphorylation, Parkinson's disease, Synaptic vesicle cycle, Alzheimer's disease, Epithelial cell signaling in *Helicobacter pylori* (HPI) infection, Huntington's disease and citrate cycle (TCA cycle) were highly significant. These pathways and the number of genes involved, have been shown in Fig. 3 by individual curves (supplementary Tables S3, S4). The significant pathways of Up- and Down-regulated genes with their interactions have been visualized by ClueGo software in Figs 4, 5.

Study of brain regions and sex-related differences in gene expression in AD. Analyzing the expression profile of brain regions with six data sets was done in which each data set represented one region in the brain and one of them (GSE5281) consists of more than one region; likewise, the sample size was insufficient for further studying of brain regions independently. Analyzing the gene expression profile on the six data sets according to brain regions separately and based on control vs AD was done. The sub meta-analysis of hippocampus and entorhinal cortex have been analyzed; but, the expression profile of other regions was utilized to continue the reviewing. The top DEGs, the top region specific genes and the genes which were in common with meta-analysis results in each brain region (based on the resold P value < 0.001) was categorized in supplementary Table S5. Also, the significant molecular functions of region specific genes (P value < 0.001) in eight brain regions were demonstrated in

Down-regulated			Up-regulated		
ID	P. Value	FC	ID	P. Value	FC
hsa-mir-12497 RNAz	0.003394	0.0263887	hsa-mir-19790 RNAz	0.0026067	3.4836785
hsa-miR-424	0.0011322	0.079282	hsa-mir-27120 RNAz	0.0011101	3.9197762
hsa-mir-44608 RNAz	0.0026307	0.0920438	hsa-mir-35456 RNAz	0.0032355	3.9345788
hsa-miR-181c	0.0015209	0.1018621	hsa-miR-122a	0.0029606	4.0567087
hsa-miR-301	0.0012602	0.1034073	hsa-miR-134	0.003466	4.499232
hsa-miR-656	0.0013878	0.1617726	hsa-mir-10939 RNAz	0.001997	5.3955761
hsa-miR-186	0.0025921	0.1711834	hsa-mir-10912 RNAz	0.0012295	5.7425097
hsa-mir-20546 RNAz	0.0013675	0.1773218	hsa-miR-617	0.0003481	6.1780417
hsa-mir-02532 RNAz	0.0031668	0.2263602	hsa-mir-30184 RNAz	0.0000352	6.3673885
hsa-mir-42448 RNAz	0.002298	0.2375547	hsa-mir-03996 RNAz	0.0031868	7.653148
hsa-miR-101	0.000485	0.2680839	hsa-miR-188	0.0004821	12.61189
hsa-miR-551b	0.0013837	0.2809868	hsa-mir-06383 RNAz	0.0010481	16.090057
hsa-mir-08570 RNAz	0.0019679	0.2865049	hsa-mir-23974 RNAz	0.0018586	23.14629
hsa-miR-29b	0.0014182	0.3212273	hsa-mir-40796 RNAz	0.0001212	25.918338
hsa-miR-189	0.0016735	0.4098583	hsa-miR-601	0.0013881	39.057145

Table 2. Top 30 DE miRNAs obtained from microarray expression in AD patients. All the top miRNAs were ranked based on the P value. This table showed the miRNAs expression based on (≥ 1.23 fold-change) criteria and divided into Up- or Down-regulated miRNAs. FC: Fold Change.

Gene symbols	Ave FC	Meta-analysis score	Gene symbols	Ave FC	Meta-analysis score
Down regulated			Down regulated		
<i>CARTPT</i>	0.615281653	4.06737E-06	<i>DYNLL1</i>	0.79243	0.000235
<i>ATP5F1</i>	0.77653477	4.03204E-05	<i>TMEM189</i>	0.760451	0.000258
<i>RNFT2</i>	0.65113094	6.13025E-05			
<i>USO1</i>	0.698005449	9.30377E-05	Up-regulated		
<i>COL5A2</i>	0.53826377	9.61705E-05	<i>GFAP</i>	1.856253	2.34E-06
<i>CALM1</i>	0.710120806	0.000113127	<i>TGFB111</i>	2.041183	2.05E-05
<i>DZIP3</i>	0.650845423	0.000127364	<i>APLNR</i>	1.912765	2.18E-05
<i>PLK2</i>	0.647124318	0.000137089	<i>ATP10A</i>	1.346906	3.42E-05
<i>BEX4</i>	0.69732923	0.000140728	<i>AQP1</i>	1.648286	0.000159
<i>MDH1</i>	0.694880723	0.000144727	<i>WAS</i>	1.323854	0.000162
<i>ATP2B3</i>	0.686082434	0.000156006	<i>DOCK2</i>	1.491173	0.000187
<i>ATP6V1E1</i>	0.721259284	0.000186009	<i>LRRFIP1</i>	1.420411	0.000192
<i>PFKM</i>	0.77010002	0.000211936	<i>IGFBP5</i>	1.399666	0.000198
<i>RAN</i>	0.640432884	0.000216575	<i>NAV2</i>	1.347951	0.000225
<i>CRH</i>	0.686953124	0.000220206	<i>FDFT1</i>	1.633492	0.000232
<i>MEST</i>	0.791186	0.000223	<i>MAFF</i>	1.4246	0.00027

Table 3. Top 30 differentially expressed genes (DEGs) identified in the meta-analysis of AD studies. All the top genes were ranked based on meta-analysis scores (this score calculated according to the P value of each gene in their datasets with R). Ave FC: average Fold change.

Fig. 6. The hierarchical clustering analysis was used for the top 30 DE genes in the meta-analysis result compared to the expression profile of all brain regions in Fig. 7a.

The sub meta-analysis of male and female with four data sets (GSE16759 due to the low sample size and GSE4757 due to its unknown gender was excluded) have been done. Our results in the context of sub meta-analysis on male and female showed the lists of 1903 DEGs in male (830 Up-regulated and 1073 Downregulated DEgenes) which 1210 genes were male specific and 2333 DEGs in female (1125 Up-regulated genes and 1208 Down-regulated DEgenes) that 1640 genes were female specific respectively (all details about gender specific genes and significant DEGs have been categorized in Supplementary Table S6). Thus, male and female specific genes were specified in AD; the female specific genes are involved in pathways associated with neurodegenerative diseases such as Oxidative phosphorylation, Alzheimer's disease, Huntington's disease, Parkinson's disease pathways which demonstrated the importance of these genes in neurodegenerative including AD. In about male specific genes, no significant pathway was determined. Since, the aim of our study in this meta-analysis project was to exclude the interventive factors and categorized the top DEGs among AD across brain regions; for comparing the effects of gender on our meta-analysis result, we used T-test and one-way analysis of variance (ANOVA), these statistical analyses just among common DEGs of meta-analysis results, female and male was

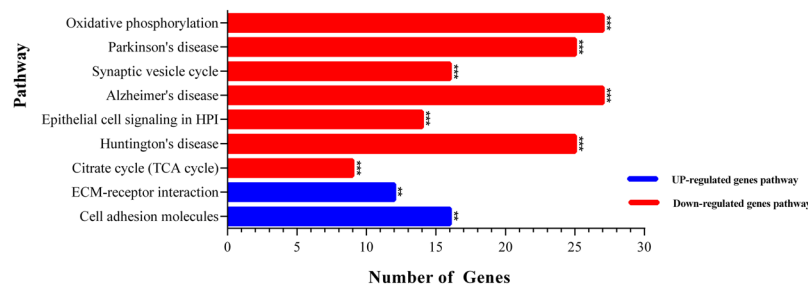


Figure 3. Top biological pathways and terms by ClueGO software in meta-analysis of Up- and Down-regulated genes in AD patients. The significant Up- and Down-regulated genes in AD patients were 672 and 732 different expression genes respectively that generated by meta-analysis and transferred into the ClueGO software (kappa score = 0.4). Pathways or terms of the associated genes were ranked based on the P value corrected with Bonferroni step down (*show the P value < 0.05).

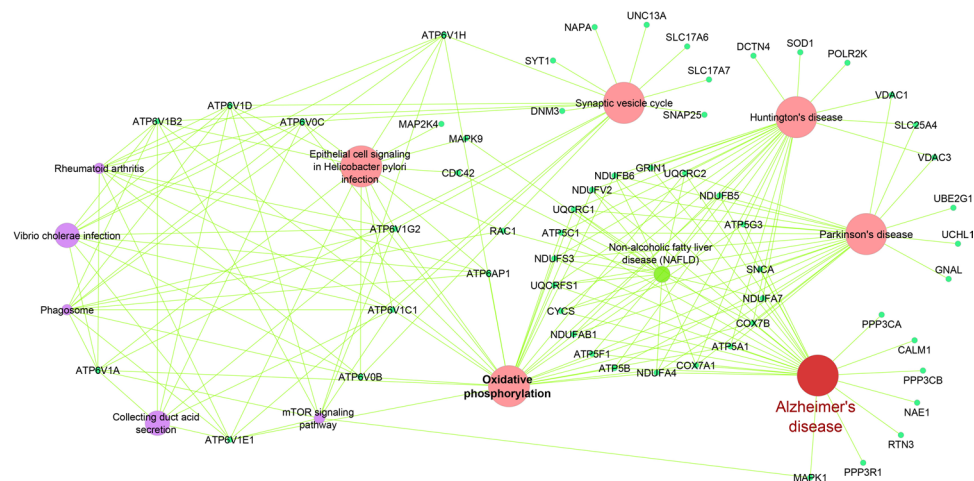


Figure 4. Alzheimer's disease pathway and enriched gene ontology with Down-regulated genes. Among 10 significant pathways of Down-regulated meta-analysis genes, Alzheimer's disease pathway and enriched gene ontology with other top significant pathways in this study have been visualized in ClueGO software (kappa score = 0.4). It demonstrated the genes that are common among Parkinson's disease, Alzheimer's disease, and Huntington's disease; also, the genes that are only involved in Alzheimer's disease and other significant pathways.

done; the outputs of statistical tests demonstrated that gender's sub meta-analysis results do not have a significant difference in gene expression of the meta-analysis result (P value > 0.05). The top 30 DEGs in meta-analysis result which were common with female, male or either of them, based on LogFC have been visualized in Fig. 7b.

Investigating the gene expression profiles in AD with RNA-seq. In this study, we tried to select data sets, which as much as possible were matched and comparable with microarray sampling brain regions. Hence, studying on differential gene expressions between AD vs Control in three RNA-seq data sets (GSE53697 sampling from part of the dorsolateral prefrontal cortex (BA9), GSE67333 sampling from hippocampi brain regions and GSE57152 sampling from superior temporalis gyrus) have been analyzed^{11,12}. The results demonstrated that the 1102 DEGs in GSE67333 (614 Up-regulated and 488 Down-regulated genes), 1032 DEGs in GSE53697 (324 Up-regulated genes and 708 Down-regulated genes) and 898 DEGs in GSE57152 (591 Up-regulated and 307 Down-regulated genes) with a threshold P value < 0.05 and FC ≥ 1.23 that were categorized in supplementary Table S7. DEGs between RNA-seq and microarray data were analyzed. 95 common genes between the dorsolateral prefrontal cortex (BA9) and frontal cortex brain regions, 61 common genes on the hippocampus in RNA-seq and microarray data sets; also, 240 genes between superior temporalis gyrus and medial temporalis gyrus in RNA-seq and microarray data sets respectively have been obtained (P value < 0.05). The highly significant common Up- and Down-regulated genes between RNA-seq and microarray have been categorized in Table 4.

Gene-miRNA regulatory network in AD. Our results demonstrate the different expression value for each type of genes, so we could divide the Up- and Down-regulated genes. All of these meta-analysis results and miRNAs were transferred into the Cytoscape 3.2.1, and the network was constructed. The regulatory network constructed by cyTargetLinker¹³ using three databases (TargetScan, MicroCosm V5, mirTarbase). The regulatory

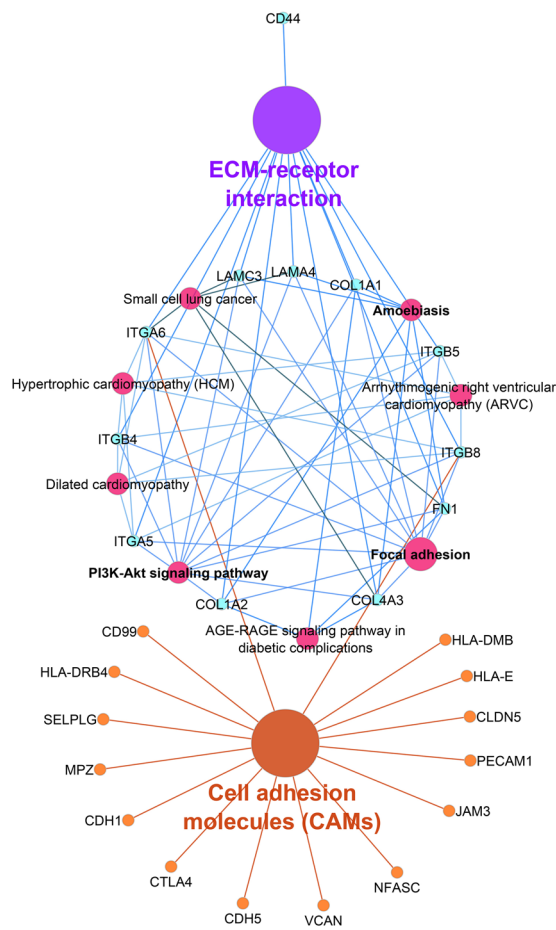


Figure 5. Enriched gene ontology and pathways that generated by Up-regulated genes in the meta-analysis. Enriched gene ontology and pathways have been visualized in CluGO software (κ score = 0.4). Each node showed the biological process in their pathways. The color of each node represents the functional group that they belong to it. Edges demonstrate the gene-term and term-term interactions. Both significant pathways of Up-regulated meta-analysis genes, the ECM-receptor interaction and Cell adhesion molecules (CAMs) with involved genes have been shown in this pathway. The size of each node demonstrated its value in these pathways.

miRNA-target genes network included 18189 genes, 181 miRNA, and 113796 edges. We mapped the expression data that included meta-analysis result and DEMiRs to the regulatory network; we filtered the target genes that were not included in our meta-analysis and did not have an expression data then based on expression value, the node size and color change. This regulatory network included 2438 genes, 181 miRNA and 16886 edges that were differentially expressed in AD (Fig. 8).

Construction of Gene Regulatory subnetwork for miRNA-target genes. Surveys conducted by centiscape¹⁴ showed our network hubs were miR-30a-5p and miR-335 which regulated most genes involved in AD in our meta-analysis results. Subsequently, hierarchical clustering analysis was done in order to find significant differences in both differentially expressed miRNAs and genes expression between the AD and Control groups for the hubs of our network - miR-30a-5p and miR-335 (Fig. 9a,b). Each hierarchical cluster for miR-30a-5p and miR-335 demonstrated the expression profile of near neighbor target genes. Furthermore, Venn diagram by VENNY 2.1 tool¹⁵ has been drawn for common genes between both miRNAs (Fig. 9c). The results showed that 71 genes are common target genes for both miRNAs.

Expression of miR-30a-5p ($\log_{2}FC -0.61128$) and miR-335 ($\log_{2}FC -0.91861$) decreased, but both have Up- and Down-regulated target genes. miR-30a-5p have 355 near neighbor target genes that 135 genes with positive $\log_{2}FC$ and 220 genes with negative $\log_{2}FC$, for miR-335 among 396 near neighbor target genes have 213 positive $\log_{2}FC$ and 183 negative $\log_{2}FC$ genes in AD. The circulate subnetwork of miR-30a-5p and miR-335 with immediate neighbors had been constructed by Cytoscape 3.2.1, here we demonstrated the most important target genes. The node size and color (red-blue) represented the expression value of Down- and Up-regulated respectively. Also, the edges' colors showed the source of prediction target genes (Fig. 10). The validated target genes of miR-30a-5p and miR-335 vs the predicted ones have been categorized in supplementary Table S8.

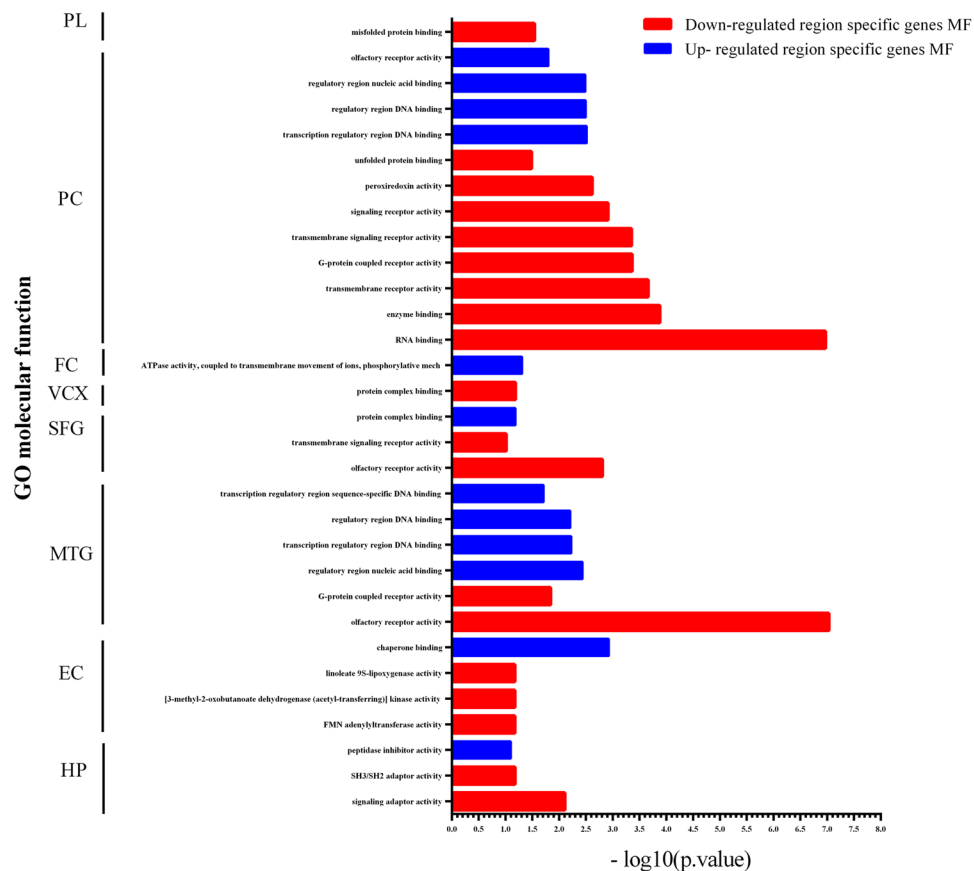


Figure 6. The molecular function of region specific genes in AD. Molecular function (MF) of all region specific genes in each brain region based on cut off P value < 0.001 have been selected. The significant MF for Up and Down-regulated region specific genes were shown. HP: hippocampus, EC: Entorhinal cortex, MTG: Medial temporal gyrus, PC: Posterior singlet cortex, SFG: superior frontal gyrus, VCX: primary visual cortex, FC: frontal cortex. PL: parietal lobe.

KEGG pathway Analysis of miRNA-335 and miRNA-30a-5p target genes. According to a KEGG pathway^{16–18} analysis of ClueGo plugging software, we demonstrated the significant functions and pathways of near neighbor target genes of miR-30a-5p and miR-335. All of them sorted based on the Bonferroni step-down (Table 5). Also, visualized by CluGo in Fig. 11a and b. Our results represented the significant pathways of miR-30a-5p near neighbor is Long term potentiation (LTP), for miR-335 the near neighbor target genes interfered in the Steroid biosynthesis pathway but the role of this pathway was not significant.

Discussion

Alzheimer's disease is a general neurodegenerative disorder and its pathophysiological process mostly begins before clinical diagnosis. During the pre-clinical process most AD patients show no symptoms believed is¹⁹. Therefore, molecular study of AD could have an important role in detecting genes involved in the disease. On the other hand, a variety of researchers showed that miRNAs and expression of target genes were tightly associated with molecular events in neurodegenerative diseases such as AD⁴. The miRNAs in the central nervous system have been shown that contributes to the regulation of development, survival, function, and plasticity²⁰. Moreover, any disruption and alterations in microRNAs and their expression in neurons, leading to neurodegenerative diseases such as Alzheimer's disease (AD)²⁰. It is noteworthy that miRNAs have a high abundance in the central nervous system, and mostly their expression patterns are brain-specific²¹. According to our results, the DEGs and DE miRs both can be a potential candidate for biomarkers in AD. Other studies such as qRT-PCR can also be suggested for confirming them.

Differentially expressed genes and their relation to sex have been acknowledged in some publications^{22,23}. Albeit, it is very important to know that there is a difference between male and female in the brain structure and function; but, sex-related differences in gene expression are depended on the brain tissues and age²². A large number of genes are varied in the developmental stages of the brain²⁴; this is due to the majority of brain development and organization and the influence of gender occurring in adulthood²⁵. The risk of AD is depended on age and gender that in women is higher than the men. Among the proposed reasons for this difference, the mitochondria and hormonal dissection^{26,27} could be mentioned, where, hormone therapy and estrogenic compounds worked as a protection in front of the amyloid-beta toxicity²⁶. Based on our results there are more than one hundred genes which specifically expressed in male and female and the other possible mechanism for gender response against

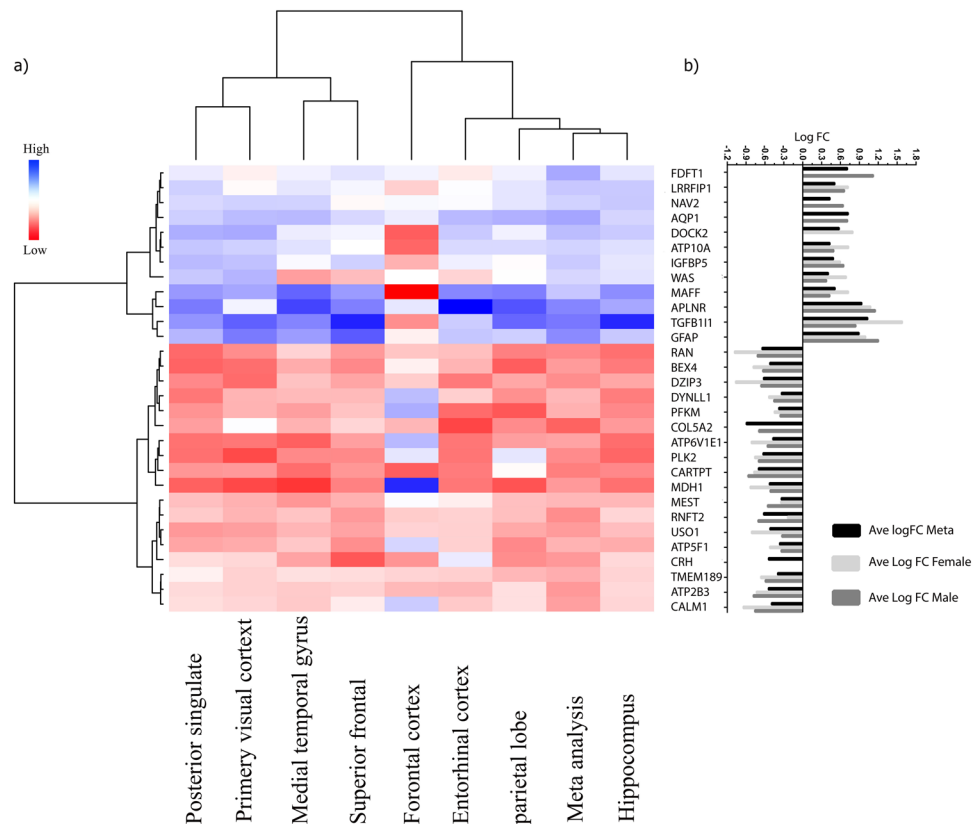


Figure 7. Comparing the gene expression pattern of top 30 genes from meta-analysis (a) in different brain regions (b) with Female and Male. (a) The heat map panel showed the top 30 differentially expressed genes in meta-analysis results versus different brain regions (posterior singlet cortex, primary visual cortex, medial temporal gyrus, superior frontal gyrus, frontal cortex, entorhinal cortex, hippocampus, parietal lobe vs DEGs in meta-analysis results). This panel compared the gene expression pattern in each type of brain regions between AD and control patients in six data sets (GSE12685, GSE4757, GSE5281, GSE1297, GSE28146, and GSE16759). The red color showed the low expression value, and blue color showed higher expression value. (b) This figure compares the LogFC of each gene among female, male and meta-analysis result. The Y-axis shows a selected top 30 gene names and the X-axis features black, light gray and gray bars which represent the LogFC by meta-analysis, Female and Male.

AD could be resulted from their associated pathways. Meanwhile, in this study, we found the significant pathways with female specific genes which are related to AD and indicated their role and function in this neurodegenerative disease. The statistical analysis on the common genes between male, female and meta-analysis demonstrated that there are no significant differences between the gender and meta-analysis results. Therefore, the common genes between meta-analysis, and sex could represent the new window of study in AD.

In our study on DE genes of brain regions, we analyzed eight brain regions (posterior singlet cortex, primary visual cortex, medial temporal gyrus, superior frontal gyrus, frontal cortex, entorhinal cortex, hippocampus, parietal lobe) based on comparing the DEGs in each brain region and meta-analysis. The molecular function (MF) of top Up- and Down-regulated region specific genes showed that their function in some cases was in common among SFG, VCX, PC, and MTG. The most significant MF, among region specific genes were olfactory receptor activity and RNA binding (Fig. 6). Studies revealed that olfactory receptors (ORs) have been expressed in the human brain regions and the gene expression patterns have altered in several neurodegenerative diseases such as PD and AD. The differentially expressed OR genes (Up- or Down-regulated genes) have been reported in AD²⁸. In our results, we demonstrated the expression of top genes which are related to these MF and played their function in each brain region. Also, aggregation of RNA binding proteins in the cytoplasm under stress conditions produced the stress granules (SGs). Accumulation of these SGs in the brain have a pathological impact in AD and other neurodegenerative diseases; due to altered the normal physiological activities of RNA binding proteins^{29,30}.

As far as we are concerned, our research categorized the top and significant DEGs in the different brain regions and compared and analyzed the results using a meta-analysis output. Our findings can be helpful for understanding the most important DEGs in AD and making a connection between the gene expression level and higher level information about their functions, interactions and pathways. In the further study, we would like to analyze the region specific genes by using more data sets with a high sample size that are specialized on each specific brain region. These will increase the accuracy and will avoid false positive data in the study.

	Up-regulated					Down-regulated				
	Gene Symbol	RNA-seq		microarray		Gene Symbol	RNA-seq		microarray	
		logFC	P value	logFC	P value		logFC	P value	logFC	P value
Common genes in FC and BA9 of DL-PFC	<i>NRN1</i>	2.2334	0.0198	0.8070	0.0213	<i>PRELP</i>	-3.9582	0.00005	-0.3136175	0.0099
	<i>GPRASP1</i>	1.5928	0.0209	0.9407	0.0354	<i>EPAS1</i>	-2.1622	0.000205	-0.39160792	0.0159
	<i>ESRRG</i>	1.3135	0.0250	0.2804	0.0228	<i>KIF1C</i>	-3.0637	0.001339	-0.44908542	0.0246
	<i>TPBG</i>	1.6981	0.0252	0.5743	0.0011	<i>FKBP10</i>	-1.9021	0.001747	-0.29188125	0.0483
	<i>KCNK1</i>	1.4959	0.0252	0.6657	0.0004	<i>KLF4</i>	-3.3412	0.002665	-0.3612275	0.0067
	<i>HTR2A</i>	1.9082	0.0253	1.3014	0.0007	<i>NQO1</i>	-3.7408	0.003207	-0.31109333	0.0476
	<i>GSTO1</i>	1.7010	0.0261	0.6422	0.0087	<i>ELN</i>	-2.1514	0.003593	-0.3799525	0.0417
	<i>TOMM20</i>	1.7740	0.0301	0.6976	0.0026	<i>FBLN1</i>	-1.8663	0.007021	-0.33266417	0.0118
	<i>GABRA1</i>	2.4678	0.0303	1.2669	0.0001	<i>SOX21</i>	-1.9109	0.012743	-0.39442583	0.0059
<i>OXR1</i>	1.7054	0.0308	0.6058	0.0051	<i>SMOX</i>	-1.4258	0.015775	-0.41022292	0.0495	
Common genes in Hippocampus	<i>PRKCQ-AS1</i>	3.6348	0.00005	0.4930	0.0337	<i>GLS2</i>	-2.4687	0.00005	-0.4732	0.0467
	<i>TRIM16L</i>	0.9125	0.00005	0.7303	0.0295	<i>SLC22A24</i>	-2.4552	0.0006	-0.3981	0.0043
	<i>PLK5</i>	1.8366	0.0001	0.9223	0.0271	<i>PDXDC1</i>	-4.8043	0.001	-0.2931	0.0144
	<i>CNOT1</i>	2.4604	0.0007	0.3935	0.0049	<i>MYH16</i>	-0.7303	0.0010	-0.7485	0.0449
	<i>LINC00648</i>	3.7324	0.001	0.3031	0.0224	<i>ZNF704</i>	-3.4464	0.0011	-0.3069	0.0272
	<i>APOL4</i>	1.6363	0.0010	0.4707	0.0166	<i>DEFB108B</i>	-1.9448	0.0045	-0.7843	0.0343
	<i>APBB1IP</i>	2.6487	0.0011	1.0084	0.0081	<i>CDC42SE2</i>	-1.8370	0.0083	-0.4003	0.0001
	<i>LINC00700</i>	1.9264	0.0014	0.5462	0.0453	<i>DNAH14</i>	-0.5588	0.019	-0.6766	0.0079
	<i>STRAP</i>	2.3408	0.0018	1.0526	0.0248	<i>DNAJC13</i>	-2.1886	0.0205	-0.4092	0.0464
	<i>NUMA1</i>	2.1248	0.0020	0.7587	0.0088	<i>KCNK10</i>	-0.4884	0.0226	-1.1787	0.0216
Common genes in STG and MTG	<i>C9orf3</i>	4.61003	0.00005	1.364899	0.01602	<i>CRYBB2P1</i>	-4.42879	0.00005	-0.81855005	0.03695
	<i>CPNE3</i>	3.3581	0.00005	1.1969213	0.0243	<i>FAM208A</i>	-4.10653	0.00005	-0.4882346	0.00118
	<i>GLUL</i>	2.82662	0.00005	1.393713	0.0218	<i>KNTC1</i>	-3.7896	0.00005	-1.6199	0.0109
	<i>IKKBK</i>	2.82657	0.00005	1.240875	0.0020	<i>NAT9</i>	-4.2662	0.00005	-0.6446	0.0109
	<i>PFKFB3</i>	3.89407	0.00005	2.198218	4.49E-07	<i>NRG4</i>	-5.5145	0.00005	-1.1152	0.0405
	<i>SLC11A2</i>	2.16711	0.00005	1.957787	0.000037	<i>SETD5</i>	-2.676	0.00005	-0.8030	0.0046
	<i>SPP1</i>	3.2241	0.00005	0.9801302	0.0206	<i>TMCC1</i>	-2.63472	0.00005	-1.098779	0.008445
	<i>USP40</i>	2.10273	0.00005	1.4242945	0.00124	<i>TNKS</i>	-2.05614	0.0002	-0.8721588	0.027795
	<i>ZCCHC7</i>	5.67435	0.00005	1.0294421	0.00326	<i>ZBTB45</i>	-4.02681	0.0003	-1.16852	0.024
<i>ZNF528</i>	3.51179	0.00005	0.8404265	0.0287	<i>NAA38</i>	-3.54314	0.00035	-0.92694	0.0019665	

Table 4. The top significant common genes between RNA-seq and microarray. According to cut off P value < 0.05 and $FC \geq 1.23$. The common and top significant Up- and Down-regulated genes in GSE53697 (sampling from BA9 which part of the dorsolateral prefrontal cortex (DL-PFC)), GSE67333 (sampling from hippocampi) and GSE57152 (sampling from superior temporalis gyrus) with microarray results which consist of submeta analysis of hippocampus, expression profile of frontal cortex and medial temporal gyrus, have been shown.

Studying on the RNA-seq data in AD and on the three parts of brain regions (part of the dorsolateral prefrontal cortex, hippocampus and superior temporalis gyrus) demonstrated the most important DEGs. The difference in brain regions sampling and low replication on the RNA-seq data rather than the microarray results can be a contributing factor in reducing the common genes between these two techniques. Although, microarray meta-analysis is still a widely used method in the context of gene expression, but the analysis of RNA-seq data beside to the microarray results can lead to heighten the accuracy of these results³¹. Due to the lack of sufficient RNA-seq data for all brain regions on Alzheimer's disease, we did our study on three brain regions and compared the outputs with hippocampus submeta analysis, the expression profiles of frontal cortex and medial temporal gyrus. Eventually, the genes which were Up- or Down-regulated in this brain regions have been reported.

Our analysis reports have shown the significant pathways between Up-regulated genes already linked to AD: i) The ECM-receptors and their ligands as a molecular function, play a significant role in neuronal development and synaptic activity²⁵. The ECM receptors and cell adhesion molecules (CAMs) participate in cellular interactions, which are involved in major neurodegenerative diseases such as AD. Also, ECM proteins were up-regulated in AD and during the aging³²; hence, we demonstrated these pathways with their related DE genes which are Up-regulated in AD. ii) The CAM pathway, which has a central role in the neuronal cell adhesion and has a critical function for the synaptic formation and blood-brain barrier integrity for neurotransmission. In AD, loss of the synaptic pathway is the strongest impairment that has been done. In addition, a different expression of cell adhesion genes mainly seen in AD and Parkinson disease^{33,34}.

The significant pathway results between Down-regulated meta-analysis genes have demonstrated which Down-regulated genes are involved in AD i.e.: Oxidative phosphorylation (OXPHOS), Parkinson's disease,

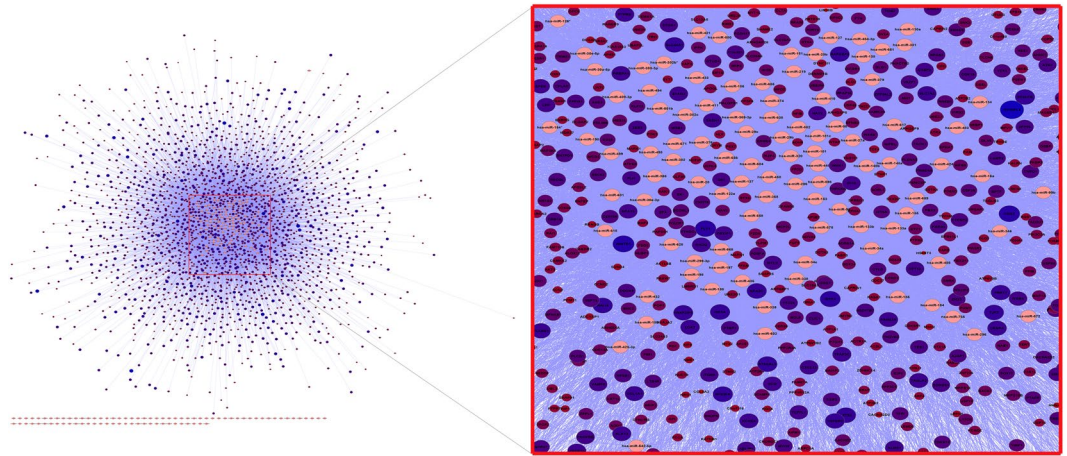


Figure 8. The regulatory subnetwork in AD that showed the DE miRNAs and their target genes. The differential expression has been shown by node color gradient. The pink nodes represent miRNAs, but a gradient red to blue colors showed Down-regulated and Up-regulated target genes respectively.

Synaptic vesicle cycle (SVC), Alzheimer's disease, Epithelial cell signaling in *Helicobacter pylori* (HPI) infection and Huntington's disease were the top involved pathways. In recent studies, this subject has been confirmed that alterations in the function of Oxidative phosphorylation (OXPHOS) have involved in the pathogenesis of AD. Also, the toxic effect of A β and tau protein on the OXPHOS cause the decreased neuronal survival³⁵. Therefore, alterations in function OXPHOS can increase the risk of AD.

The Epithelial cell signaling in *Helicobacter pylori* infection are other pathways associated with AD. Based on AD different subtypes-levels of A β , tau hyperphosphorylation, ubiquitination and *etc.*, - inflammation and immune processes has displayed the different roles since Alzheimer disease patients which use the nonsteroidal anti-inflammatory drugs (NSAIDs) reduce the risk of Alzheimer disease³⁶. Therefore, immune and inflammatory pathways play an important role in modulators of AD³⁶. It can also be inferred that epithelial cell signaling in *Helicobacter pylori* infection has an effect on AD development it is noteworthy that hyperphosphorylation of tau protein is associated with defects in neurons or the loss them³⁷. Infection with *Helicobacter pylori* (*H. Pylori*) which is a gram negative bacterium, is related to the AD³⁷. Hyperphosphorylation of tau protein is one of the events seen in the AD patient's brains³⁷. Studies on the AD patients and normal people showed the significant *H. Pylori* and anti-*H. Pylori* IgG antibodies have been observed in cerebrospinal fluid (CSF) and their serum^{38,39}. *H. Pylori* induced the tau phosphorylation in the AD special sites³⁶ so, it is effective in occurring AD. Here we showed the involvement genes in each pathway with the expression value.

Alzheimer's disease (AD), Parkinson's disease (PD) and Huntington's disease (HD) are important neurodegenerative disorders, and their involved genes mostly overlap. It is noteworthy that there are factors in common between AD and PD. A common and major neuropathological alteration, for instance, a decline in locus coeruleus (LC) noradrenergic neurons occur in AD and PD, that leads to progression of both of these disorders⁴⁰. In addition, AD and PD have different clinical and pathological features but display the overlap molecular mechanisms⁴¹. Huntington's disease (HD) is dominantly inherited and occurs in the 4th to 5th decade of life, but the AD is an aging disease and happens in older age. Albeit, it is mentioned that HD increases the risk of AD among elderly HD patients and demonstrated the co-occurrence of both of them⁴². The role of SVC in neurodegenerative diseases is specified and any changes in this pathway contributed in the development of diseases^{43,44}. In about the TCA cycle pathway, any alteration in (TCA) cycle enzymes of mitochondria and changes in its activity may be critical to increase the risk of AD. Therefore, studying on TCA cycle will improve our understanding the mechanisms and will be effective for AD patients⁴⁵. Therefore, our results categorized the genes that are specific or common for each or both of them. likewise, For all of the significant pathways, we categorized the involvement genes (Up- and Down-regulated) that will be crucial for continuous researchers and other types of detection or diagnose in AD patients. In the second part of our analysis, we focused on the miRNAs and target genes that are most important in AD. Centiscape results according to the degree and betweenness of our network were miR-335 and miR-30a-5p which had the most score and were the regulators in AD network. The expression of both of them decreased so, subnetworks were constructed with near neighbor nodes of miR-30a-5p and miR-335, which were the hubs of AD network analysis. miR-30a-5p with 356 and miR-335 with 397 neighbor nodes were the most interacted nodes respectively.

Based on our expression analysis on miR-30a-5p and miRNA-335 they have been decreased. The expression of miRNAs are different in each part of a brain. Mark N. Ziats *et al.* demonstrated in 2014 that miRNAs in the dorsolateral prefrontal cortex differentially expressed rather than other parts of the brain, such as the cerebellum, hippocampus and other regions of the prefrontal cortex⁴⁶. On the other hand, in our study, the miRNA microarray data set had been sampling from the parietal lobe of postmortem of Alzheimerian patient's brains⁴⁷ so the different types of expression are perhaps because of the sampling regions. miR-30a-5p and miR-335 are one of the miRNAs that are involved in the AD, and their expression decreased in this brain region of AD patients. Wang-Xia Wang *et al.* at 2011 studied the miRNA expression profile in gray and white matter showed miR-335 and miR-30a-5p have been Down-regulated in a gray matter of AD patient⁴⁸. Also, it is worthwhile to say miRNA

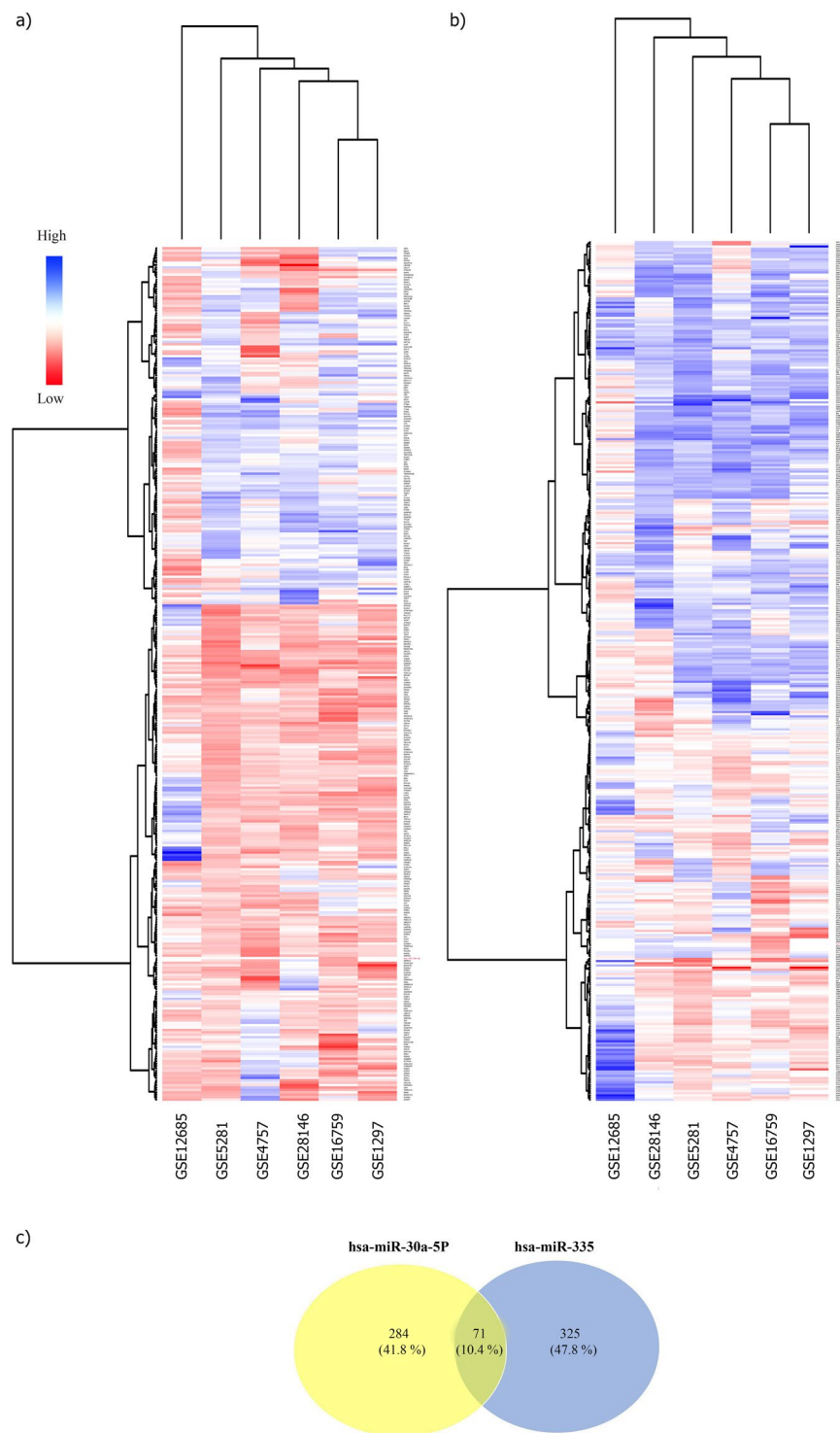


Figure 9. The cluster heat map of miR-30a-5p and miR-335 with their Venn diagram. Differentially expressed microRNAs (DEmiRs) hub in network and genes that were targeted of it. The red color showed the low expression value, and the blue color showed higher expression value. **(a)** The miR-30a-5p with expression profile of its target genes. **(b)** The miR-335 with expression profiles of its target genes. **(c)** This panel showed the overlap between miR-30a-5p and miR-335 target genes; 71 genes are common between both miRNAs.

often could be co-expressed with its targets²¹. In AD the main cause of the synaptic dysfunction is the aggregation of amyloid- β ($A\beta$) peptides in the neuronal system which leads to a disarrangement, and cell loss⁴⁹. The KEGG pathway analysis of miRNA30a-5p and its first neighboring nodes have been sorted by the Term P value corrected with step down showed that Long-term potentiation (LTP) is the significant pathway for all the near neighbor targets but about miR-335 we could not find any significant pathway.

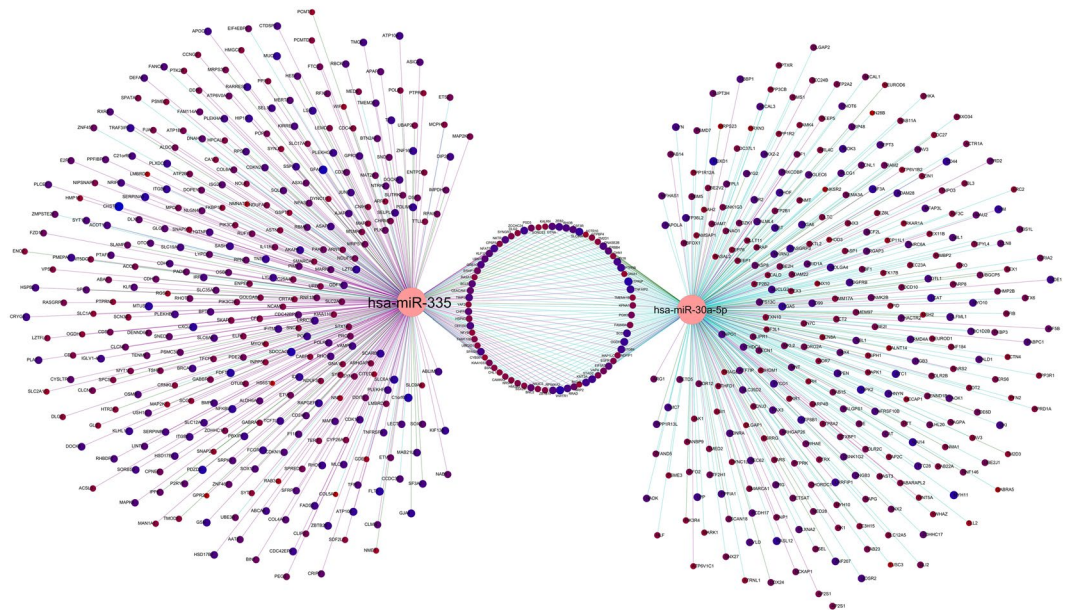


Figure 10. The near neighbor nodes subnetwork of miR-30a-5p and miR-335. This subnetwork represented the near neighbor target genes of miR-30a-5p and miR-335 with their common genes. The node size and color showed the expression value (blue color; big node showed the higher expression value and red color; small node showed the low expression value), and the edges color demonstrated the source of each target gene prediction. TargetScan: indigo, MicroCosm V5: green, mirTarbase: Violet.

miRNA-30a-5p near neighbor nodes			
GO ID	GO Term	Term P value Corrected with Bonferroni step down	Associated Genes Found
KEGG:04720	Long-term potentiation	0.00756605	<i>CAMK2B, CAMK4, GRIA1, GRIA2, MAPK1, PPP3CB, PPP3R1, RPS6KA2</i>
miRNA-335 near neighbor nodes			
KEGG:00100	Steroid biosynthesis	0.105889168	<i>CEL, FDFT1, HSD17B7, SQLE</i>

Table 5. The significant pathways of miR-30a-5p and miR-335 near the neighbor target genes in AD. For miR-30a-5p, Long-term potentiation is the pathway that near neighbor target genes interfered in that with the significant P value. But for miR-335, it is not a significant pathway. Here we showed the target genes are involved in each pathway.

The KEGG pathway analysis on the miRNA 30a-5p near neighbor target genes demonstrated that among all of them, *CAMK2B, CAMK4, GRIA1, GRIA2, MAPK1, PPP3CB, PPP3R1, RPS6KA2* genes are significantly involved in LTP. Studying in the hippocampal LTP demonstrated its association with learning and memory. So, in AD LTP is flawed because of accumulation of A β and other fragments⁵⁰. In neuroscience, LTP is the chemical transmission function and synaptic activity between two neurons⁵¹. The expression of *CAMK2B* and *RPS6KA2* genes increased, but *CAMK4, GRIA1, GRIA2, MAPK1, PPP3CB, PPP3R1* genes decreased in our meta-analysis study on AD.

The present study prepared and categorized genes and miRNAs in AD based on available microarray and RNA-seq data sets. Therefore, the reliable sources, which have been proposed in this study, are the most important ones. Hence, we faced with some inevitable technical limitation. Some of these limitations which are mentioned here: Microarray is one of the most powerful technique for analysis of gene or protein expression simultaneously; but, this technology has several limitations. The hybridization of probes to bind to a target sequence and specificity of them with great importance, but cross-hybridization in this context will diminish the specificity. Albeit, determining the four levels of hybridization specificity in the context of microarray hybridization could be effective but low hybridization specificity in each of which four level might be occurred. Therefore, specificity of hybridization in microarray is one of the critical limitations of this technique^{52,53}. Using microarray data sets, which were obtained from multiple studies in different conditions, increased the heterogeneity in our study. In addition, the low sample size and unknown characteristic details (for instance, gender, ethnic, age, etc.), different quality and amount of RNA are the major challenge in microarray analysis. However, we tried to minimize the effects of this limitation where we doing late integration meta-analysis and utilizing appropriate methods for analyzing data. Quality control and normalization is the critical content in microarray analysis and will decrease the false data in microarray analysis. As already, mentioned microarray is one of the most powerful techniques to evaluate the gene or protein expression. Also, according to differentially expressed genes in the microarray,

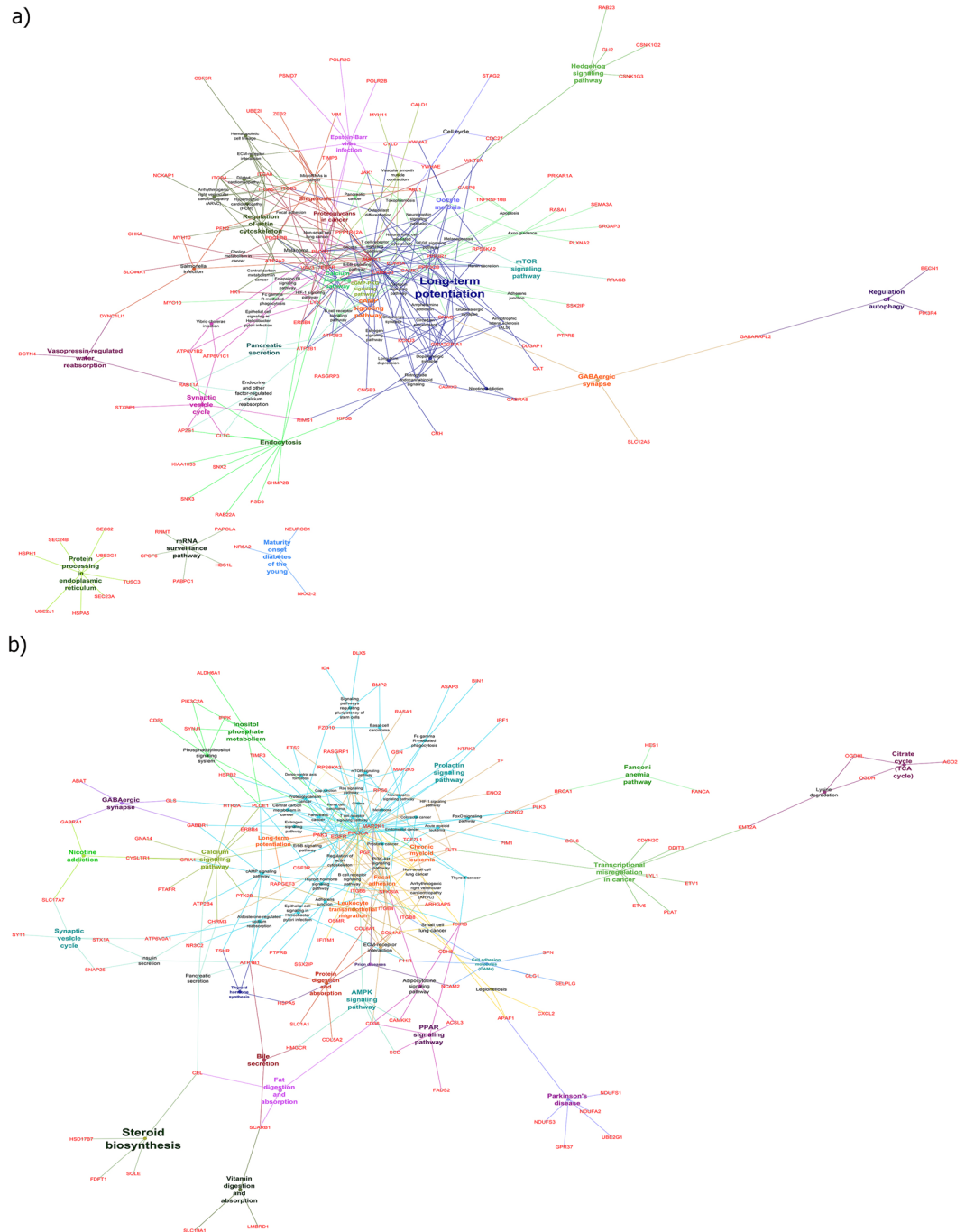


Figure 11. Functional analysis of target genes (a) miR-30a-5p (b) miR-335 in KEGG pathway by ClueGo software in Cytoscape 3.2.1. The results of centiscap analysis elicited from the gene regulatory network and then the target genes of those hub nodes transferred into CluGo and were grouped with it as a functional cluster. Each node represents a KEGG pathways process, and their associated genes are represented as dots. All the nodes and their colors showed the functional group to which they belong. Edge of this functional analysis demonstrated the term-term interactions or term-genes interactions. The enrichment significant term showed by the size of each node.

usually we have a widespread differentially expressed genes, for instance, gene expression with very low expression value but insignificant or modest DEGs with significant values. Filtering microarray data is an essential factor are mostly assigned to cut off. In about RNA sequencing technique, we have a limitation in the context of the depth of sequencing and the number of replications that are important factors for demonstrating the reliable DEGs. Meanwhile, due to the cost of this method, considering the depth of sequencing and the number of replications in the appropriate condition is very difficult and usually, one of them is being investigated. On the other hands, the different sequencing methods in RNA-seq analyzing produced various DEGs output. Thus, All of them

should be considered⁵⁴. Nevertheless, RNA-Seq is a new technique and under development and despite the cost, widely used in the context of gene expression⁵⁵ and when it studied with microarray, they can be complemented with each other⁵⁶.

In conclusion, our finding suggests that the different expressions of genes and miRNAs are one of the most important variables in AD. Bioinformatic Analysis could help us to find the most important genes, miRNA, miRNA-mRNA interactions and their related pathways. Hence, these should be probed in further studies for better understanding of the gene regulatory network, molecular mechanisms of AD, developing new therapeutic approaches, future studying of miRNA function and regulation and their potential as diagnostic biomarkers for AD. However, this project is a bioinformatics analysis study based on the high throughput data and is not derived in our laboratory, but a large number of experimental studies confirm that the pathways and genes which were involved in AD are supported.

Methods

We have detected miRNA and their target genes in AD as well as the Gene Ontology and their signaling pathways. First, we detected differentially expressed genes (DEGs) by meta-analyzing six gene microarray data sets, for the differentially expressed miRNAs (DEmiRs), a miRNA expression profile could be detected. Then using the Cytoscape 3.2.1⁵⁷ software, DEGs and DEmiRs in order to visualize and draw the miRNA-gene interaction network. Furthermore, we calculated the active hubs and their immediate neighbors in our miRNA-gene network; we obtained the potential active miRNAs and target genes in AD. Last but not least, using the ClueGO v2.2.5 plugin⁵⁸ in Cytoscape 3.2.1, we detected the pathways of our top miRNA-target genes and therefore, the significant pathways involved in AD (Fig. 2) were revealed. The summary of the overall study process has been shown in the diagram supplementary Figure S1.

Microarray analysis and availability. Seven microarray data sets of human AD up to 31 December 2016 according to our criteria (Fig. 1) have been selected that are available in the public repository: NCBI Gene Expression Omnibus (GEO): GSE1297, GSE4757, GSE5281, GSE12685, GSE28146, GSE16759, in which GSE16759 was used for its miRNA expression profile^{47,59–64} and totally 264 samples (113 control and 151 cases) were analyzed in this study which their details have been mentioned in Table 1. Brain samples of each data sets were collected from Research Institute of Alzheimer's disease in the USA; so, descriptions and categorization of the donor samples details, including the mean age, gender, and brain regions were listed in supplementary Table S9. These seven data sets were independently generated using different protocols. The quality control and normalization of our array data have been done with affy, plier and piano packages in R^{65–67}. Their output has been reported in supplementary Figure S2. Analyzing AD has been done in two groups control and AD patients using the GEO2R tool, to detect differentially expressed genes (DEGs) and differentially expressed miRNAs (DEmiRs)⁶⁸.

Differentially expressed miRNAs in AD. High-throughput techniques to investigate miRNA expression in AD have thus far rarely been used; we found only one miRNA microarray data set in GEO (GSE16759), that studied both miRNA and mRNA expression in AD patients and controls⁴⁷.

Meta-analysis. Different gene expression profiles may demonstrate the varied differentially expressed (DE) genes to obtain accurate gene expressions⁶⁹; hence, we used a meta-analysis in our study. In this study, we implemented the meta-analysis by using the R package RobustRankAggreg⁷⁰. The scores based on the P value were calculated for all seven data sets, genes and miRNA with P-values less than 0.05 and fold change ≥ 1.23 were considered as DEGs. In our study, we performed a sub meta-analysis on Male and Female with four GSEs (GSE16759 because of the low sample size and GSE4757 because the unknown gender have been excluded). On the other hand, the expression profile of six brain regions and sub meta-analysis on the hippocampus and entorhinal cortex have been done to comprise the effect of brain regions on DEGs. Then region specific genes for each brain region in AD have analyzed.

Studying of RNA-seq data. According to our inclusion and exclusion criteria, by searching the GEO and SRA (<https://www.ncbi.nlm.nih.gov/sra>), all available gene expression data in the context of AD have been collected (supplementary Figure S3). The GSE53697 by the Illumina HiSeq. 2500, GSE67333 by the Illumina HiSeq. 2000 platforms and GSE57152 by AB 5500xl Genetic Analyzer platforms respectively have been selected for further analysis. The quality control (FastQC) on the RNA-seq (GSE53697, GSE67333, and GSE57152) short reads have been done (supplementary Figure S4)^{71,72}; hence, samples with low quality, even after trimming were omitted. Totally 20 control samples and 21 patient samples have been selected for RNA-seq data analysis. The whole descriptions of each RNA-seq donor sample detail were categorized in supplementary Table S10. Therefore, for comparing the expression profile of RNA-seq (GSE67333 and GSE57152), TopHat2 (version 2.0.8)⁷³ was utilized to align the RNA short reads to the reference human genome (hg38). Then individual transcripts were assembled with cufflinks and for generating the RNA-seq DEGs, Cuffdiff, which calculates the expression and finds significant changes in samples have been done^{74,75}. Also, the available analyzed expression profile of GSE53697 was used which is in the supplementary files¹².

Identification of miRNA-target genes. To construct the Gene Regulatory network (GRN) for miRNA-target genes based on gene expression and DEmiRs, all of them were integrated and visualized in Cytoscape 3.2.1. In this regard, we used cyTargetLinker plugging in Cytoscape 3.2.1 to draw the miRNA-target genes network¹³. After that, an organic layout was applied on the network, all of them except Homo sapiens filtered. Furthermore, node size and color have been done on the miRNA-target genes network based on logFC for all of them.

Identification of potential active miRNA-target genes in AD. Traced networks often have a complex structure, even if it is distilled into the smaller network. For solving these problems and analyzing the network easily, Cytoscape 3.2.1 conspicuously helps us. By centiscape plugging in that we could find the hub nodes in miRNA-target genes network. In this analysis, we calculate degree-closeness and degree-betweenness for all nodes, edges and detected the potential active miRNA node. Then we continued the study¹⁴. In order to determine the validated miRNA target genes as compared to predicted ones that were studied in this analysis, miRTarBase⁷⁶, TarBase v.8⁷⁷ and miRWalk 2.0^{78,79} databases have been utilized.

Enriched gene ontology and pathways analysis. To identify the biological processes and their pathways in the miRNA-target genes and gender specific genes; also, molecular function of each region specific genes, the ClueGO v2.2.5 plugin of Cytoscape 3.2.1 was used⁵⁸. In the advanced statistical option in ClueGo v2.2.5 plugin, Two-sided hypergeometric test to calculate the importance of each term was selected and Bonferroni step-down, and Kappa score = 0.4 were used for P value correction. In this part of our study, we detect the KEGG pathways of all genes that were interacted by hub node. In our previous study in the context of network biology, we utilized ClueGO v2.2.4 plugin for visualizing the functionally related gene by network and showed the significant results based on Bonferroni step down correction and kappa score threshold⁸⁰.

Clustering gene expression data. Clustering of DEGs was done using the R packages gplots, RcolorBrewer which we have visualized it^{81,82}. In this part of our study, we used hierarchical clustering, which is a powerful method for analyzing high throughput expression data. R calculated the similarity between genes in each data sets and showed the expression value by colors then clustering them.

References

- Zhao, Y., Tan, W., Sheng, W. & Li, X. Identification of Biomarkers Associated With Alzheimer's Disease by Bioinformatics Analysis. *American journal of Alzheimer's disease and other dementias* **31**, 163–168 (2016).
- Eckert, A., Schulz, K. L., Rhein, V. & Götz, J. Convergence of amyloid- β and tau pathologies on mitochondria *in vivo*. *Molecular neurobiology* **41**, 107–114 (2010).
- Jiang, W. *et al.* Identification of active transcription factor and miRNA regulatory pathways in Alzheimer's disease. *Bioinformatics* **29**, 2596–2602 (2013).
- Maciotta, S., Meregalli, M. & Torrente, Y. The involvement of microRNAs in neurodegenerative diseases. *7*, 265 (2013).
- Qin, L. *et al.* Computational characterization of osteoporosis associated SNPs and genes identified by genome-wide association studies. *PLoS one* **11**, e0150070 (2016).
- Cheng, L., Quek, C. Y., Sun, X., Bellingham, S. A. & Hill, A. F. The detection of microRNA associated with Alzheimer's disease in biological fluids using next-generation sequencing technologies. **4**, 150 (2013).
- Femminella, G. D., Ferrara, N. & Rengo, G. The emerging role of microRNAs in Alzheimer's disease. *Frontiers in physiology* **6**, 40 (2015).
- Goodall, E. F., Heath, P. R., Bandmann, O., Kirby, J. & Shaw, P. J. Neuronal dark matter: the emerging role of microRNAs in neurodegeneration. *Frontiers in Cellular Neuroscience* **7**, 178 (2013).
- Kumar, S. & Reddy, P. H. Are circulating microRNAs peripheral biomarkers for Alzheimer's disease? *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **1862**, 1617–1627 (2016).
- Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401–1414 (2007).
- Magistri, M., Velmeshv, D., Makhmutova, M. & Faghihi, M. A. Transcriptomics Profiling of Alzheimer's Disease Reveal Neurovascular Defects, Altered Amyloid-beta Homeostasis, and Deregulated Expression of Long Noncoding RNAs. *Journal of Alzheimer's disease: JAD* **48**, 647–665 (2015).
- Scheckel, C. *et al.* Regulatory consequences of neuronal ELAV-like protein binding to coding and non-coding RNAs in human brain. *eLife* **5**, 19 (2016).
- Kutmon, M., Kelder, T., Mandaviya, P., Evelo, C. T. & Coort, S. L. CyTargetLinker: a cytoscape app to integrate regulatory interactions in network analysis. *PLoS one* **8**, e82160 (2013).
- Scardoni, G., Petterlini, M. & Laudanna, C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* **25**, 2857–2859 (2009).
- Oliveros, J. C. *An interactive tool for comparing lists with Venn's diagrams*, <http://bioinfogp.cnb.csic.es/tools/venny/index.html> (2007–2015).
- Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research* **27**, 29–34 (1999).
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C. & Kanehisa, M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**, W182–185 (2007).
- Kotera, M., Yamanishi, Y., Moriya, Y., Kanehisa, M. & Goto, S. GENIES: gene network inference engine based on supervised analysis. *Nucleic Acids Res* **40**, W162–167 (2012).
- Cheng, L., Quek, C., Sun, X., Bellingham, S. & Hill, A. The detection of microRNA associated with Alzheimer's disease in biological fluids using next-generation sequencing technologies. *Frontiers in Genetics* **4**, 150 (2013).
- Lee, K. *et al.* Replenishment of microRNA-188-5p restores the synaptic and cognitive deficits in 5XFAD Mouse Model of Alzheimer's Disease. *Scientific Reports* **6**, 34433 (2016).
- Grasso, M., Piscopo, P., Confaloni, A. & Denti, M. A. Circulating miRNAs as biomarkers for neurodegenerative disorders. *Molecules* **19**, 6891–6910 (2014).
- Berchtold, N. C. *et al.* Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proceedings of the National Academy of Sciences* **105**, 15605–15610 (2008).
- Ellegren, H. & Parsch, J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* **8**, 689–698 (2007).
- Shi, L., Zhang, Z. & Su, B. Sex Biased Gene Expression Profiling of Human Brains at Major Developmental Stages. **6**, 21181 (2016).
- Li, R. & Singh, M. Sex differences in cognitive impairment and Alzheimer's disease. *Frontiers in neuroendocrinology* **35**, 385–403 (2014).
- Viña, J. & Lloret, A. Why women have more Alzheimer's disease than men: gender and mitochondrial toxicity of amyloid- β peptide. *Journal of Alzheimer's disease* **20**, 527–533 (2010).
- Grimm, A., Lim, Y.-A., Mensah-Nyagan, A. G., Götz, J. & Eckert, A. Alzheimer's disease, oestrogen and mitochondria: an ambiguous relationship. *Molecular neurobiology* **46**, 151–160 (2012).
- Ferrer, I. *et al.* Olfactory Receptors in Non-Chemosensory Organs: The Nervous System in Health and Disease. *Frontiers in Aging Neuroscience* **8**, 163 (2016).
- Wolozin, B. & Apicco, D. RNA binding proteins and the genesis of neurodegenerative diseases. *Advances in experimental medicine and biology* **822**, 11–15 (2015).

30. Maziuk, B., Ballance, H. I. & Wolozin, B. Dysregulation of RNA binding protein aggregation in neurodegenerative disorders. *Frontiers in molecular neuroscience* **10**, 89 (2017).
31. Trost, B. *et al.* Concordance between RNA-sequencing data and DNA microarray data in transcriptome analysis of proliferative and quiescent fibroblasts. *Royal Society open science* **2**, 150402 (2015).
32. Berezin, V., Walmod, P. S., Filippov, M. & Dityatev, A. Targeting of ECM molecules and their metabolizing enzymes and receptors for the treatment of CNS diseases. *Prog Brain Res* **214**, 353–388 (2014).
33. Liu, G. *et al.* Cell adhesion molecules contribute to Alzheimer's disease: multiple pathway analyses of two genome-wide association studies. *Journal of neurochemistry* **120**, 190–198 (2012).
34. Ramanan, V. K. & Saykin, A. J. Pathways to neurodegeneration: mechanistic insights from GWAS in Alzheimer's disease, Parkinson's disease, and related disorders. *Am J Neurodegener Dis* **2**, 145–175 (2013).
35. Biffi, A. *et al.* Genetic variation of oxidative phosphorylation genes in stroke and Alzheimer's disease. *Neurobiology of aging* **35**, 1956–e1951–1956, e1958 (2014).
36. Wyss-Coray, T. Inflammation in Alzheimer disease: driving force, bystander or beneficial response? *Nature medicine* **12**, 1005–1015 (2006).
37. Wang, X.-L. *et al.* Helicobacter pylori filtrate induces Alzheimer-like tau hyperphosphorylation by activating glycogen synthase kinase-3 β . *Journal of Alzheimer's Disease* **43**, 153–165 (2015).
38. Malaguarnera, M. *et al.* Helicobacter pylori and Alzheimer's disease: a possible link. *European Journal of Internal Medicine* **15**, 381–386 (2004).
39. Kountouras, J. *et al.* Increased cerebrospinal fluid Helicobacter pylori antibody in Alzheimer's disease. *International Journal of Neuroscience* **119**, 765–777 (2009).
40. Szot, P. Common factors among Alzheimer's disease, Parkinson's disease, and epilepsy: possible role of the noradrenergic nervous system. *Epilepsia* **53**, 61–66 (2012).
41. Caruana, M., Cauchi, R. & Vassallo, N. Putative role of red wine polyphenols against brain pathology in Alzheimer's and Parkinson's disease. *Frontiers in Nutrition* **3**, 31 (2016).
42. Davis, M. Y., Keene, C. D., Jayadev, S. & Bird, T. The co-occurrence of Alzheimer's disease and Huntington's disease: a neuropathological study of 15 elderly Huntington's disease subjects. *Journal of Huntington's disease* **3**, 209–217 (2014).
43. Esposito, G., Ana Clara, F. & Verstreken, P. Synaptic vesicle trafficking and Parkinson's disease. *Developmental neurobiology* **72**, 134–144 (2012).
44. Gylys, K. H. *et al.* Synaptic changes in Alzheimer's disease: increased amyloid- β and gliosis in surviving terminals is accompanied by decreased PSD-95 fluorescence. *The American journal of pathology* **165**, 1809–1817 (2004).
45. Bubber, P., Haroutunian, V., Fisch, G., Blass, J. P. & Gibson, G. E. Mitochondrial abnormalities in Alzheimer brain: mechanistic implications. *Annals of neurology* **57**, 695–703 (2005).
46. Ziats, M. N. & Rennert, O. M. Identification of differentially expressed microRNAs across the developing human brain. *Molecular psychiatry* **19**, 848–852 (2014).
47. Nunez-Iglesias, J., Liu, C.-C., Morgan, T. E., Finch, C. E. & Zhou, X. J. Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. *PLoS one* **5**, e8898 (2010).
48. Wang, W.-X., Huang, Q., Hu, Y., Stromberg, A. J. & Nelson, P. T. Patterns of microRNA expression in normal and early Alzheimer's disease human temporal cortex: white matter versus gray matter. *Acta neuropathologica* **121**, 193–205 (2011).
49. Koch, G. *et al.* Impaired LTP-but not LTD-like cortical plasticity in Alzheimer's disease patients. *Journal of Alzheimer's Disease* **31**, 593–599 (2012).
50. Chen, Q. S., Kagan, B. L., Hirakura, Y. & Xie, C. W. Impairment of hippocampal long-term potentiation by Alzheimer amyloid β -peptides. *Journal of neuroscience research* **60**, 65–72 (2000).
51. Cooke, S. & Bliss, T. Plasticity in the human central nervous system. *Brain* **129**, 1659–1673 (2006).
52. Koltai, H. & Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic acids research* **36**, 2395–2405 (2008).
53. Draghici, S., Khatri, P., Eklund, A. C. & Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *TRENDS in Genetics* **22**, 101–109 (2006).
54. Rapaport, F. *et al.* Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome biology* **14**, 3158 (2013).
55. Finotello, F. & Di Camillo, B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. *Briefings in functional genomics* **14**, 130–142 (2015).
56. Kogenaru, S., Yan, Q., Guo, Y. & Wang, N. RNA-seq and microarray complement each other in transcriptome profiling. *BMC genomics* **13**, 629 (2012).
57. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498–2504 (2003).
58. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
59. Blalock, E. M., Buechel, H. M., Popovic, J., Geddes, J. W. & Landfield, P. W. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *Journal of chemical neuroanatomy* **42**, 118–126 (2011).
60. Blalock, E. M. *et al.* Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proceedings of the National Academy of Sciences* **101**, 2173–2178 (2004).
61. Dunckley, T. *et al.* Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. *Neurobiology of aging* **27**, 1359–1371 (2006).
62. Williams, C. *et al.* Transcriptome analysis of synaptoneurosomes identifies neuroplasticity genes overexpressed in incipient Alzheimer's disease. *PLoS one* **4**, e4936 (2009).
63. Liang, W. S. *et al.* Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiological genomics* **28**, 311–322 (2007).
64. Liang, W. S. *et al.* Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *Proceedings of the National Academy of Sciences* **105**, 4441–4446 (2008).
65. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
66. {PICR}, A. I. a. C. J. M. a. plier: Implements the Affymetrix PLIER algorithm. *R package version 1.46.0* (2017).
67. Våremo, L., Nielsen, J. & Nookaew, I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Research* **41**, 4378–4391 (2013).
68. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research* **41**, D991–D995 (2013).
69. Ramasamy, A., Mondry, A., Holmes, C. C. & Altman, D. G. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med* **5**, e184 (2008).
70. Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**, 573–580 (2012).

71. Andrews, S. FastQC: a quality control tool for high throughput sequence data. Babraham Institute, Cambridge, UK. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (2010).
72. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic acids research* **44**, W3–W10 (2016).
73. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**, R36 (2013).
74. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**, 511–515 (2010).
75. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562–578 (2012).
76. Chou, C. H. *et al.* miRTarBase update 2018: a resource for experimentally validated microRNA–target interactions. *Nucleic Acids Res* **46**, D296–d302, <https://doi.org/10.1093/nar/gkx1067> (2018).
77. Karagkouni, D. *et al.* DIANA-TarBasev8: a decade-long collection of experimentally supported miRNA–gene interactions. *Nucleic Acids Research* **46**, D239–D245, <https://doi.org/10.1093/nar/gkx1141> (2018).
78. Dweep, H., Sticht, C., Pandey, P. & Gretz, N. miRWalk–database: prediction of possible miRNA binding sites by “walking” the genes of three genomes. *Journal of biomedical informatics* **44**, 839–847 (2011).
79. Dweep, H. & Gretz, N. miRWalk2.0: a comprehensive atlas of microRNA–target interactions. *Nature methods* **12**, 697, <https://doi.org/10.1038/nmeth.3485> (2015).
80. Bahramali, G., Goliaei, B. & Minuchehr, Z. A network biology approach to understanding the importance of chameleon proteins in human physiology and pathology. **49**, 303–315 (2017).
81. Warnes, G. R. *et al.* gplots: Various R programming tools for plotting data. *R package version 2*, 1 (2009).
82. Neuwirth, E. RColorBrewer: ColorBrewer palettes. *R package version 1* (2011).

Acknowledgements

We would like to thank Mr. Usef Moradifard and Mr. Usef Hoseinbeyki for financial support of this project. This research did not receive any grants from funding agencies in the public, commercial, or not-for-profit sectors and was conducted using the authors’ personal funds. We are particularly grateful for the assistance given by the Bioinformatics laboratory of National Institute of Genetic Engineering and Biotechnology; also, from Dr. Pardis Minuchehr for reviewing the manuscript for editorial needs.

Author Contributions

S.H.M. and M.H.B. were Masters Degree students at NIGEB and contributed equally to the work.; S.H.M. and M.H.B. and Z.M. conceived and designed the experiments, S.H.M. and M.H.B. performed the experiments, downloaded and analyzed the data; S.H.M., M.H.B., Z.M. and S.M.G. wrote the manuscript all authors read and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-20959-0>.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018