PLOS ONE

RESEARCH ARTICLE

# Computational identification of key genes that may regulate gene expression reprogramming in Alzheimer's patients

**Judith A. Potashkin[1]*, Virginie Bottero[1], Jose A. Santiago[2], James P. Quinn[3]**

**1** The Cellular and Molecular Pharmacology Department, The Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL, United States of America, **2** NeuroHub Analytics, LLC, Chicago, IL, United States of America, **3** Q Regulating Systems, LLC, Gurnee, IL, United States of America

* judy.potashkin@rosalindfranklin.edu





**Citation:** Potashkin JA, Bottero V, Santiago JA, Quinn JP (2019) Computational identification of key genes that may regulate gene expression reprogramming in Alzheimer's patients. PLoS ONE 14(9): e0222921. https://doi.org/10.1371/journal.pone.0222921

**Editor:** Stephen D. Ginsberg, Nathan S Kline Institute, UNITED STATES

**Received:** July 8, 2019

**Accepted:** September 9, 2019

**Published:** September 23, 2019



**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** J.A.P. received funding from National Institute on Aging (NIA) (https://www.nia.nih.gov/) grant number R01AG062176. The NIA had not role in the design, data collection and analysis, decision to publish or preparation of this manuscript. NeuroHub Analytics, LLC. and Q Regulating Systems, LLC provided support in the form of salaries for authors J.P.Q. and J.A.S., but did not

## Abstract

The dementia epidemic is likely to expand worldwide as the aging population continues to grow. A better understanding of the molecular mechanisms that lead to dementia is expected to reveal potentially modifiable risk factors that could contribute to the development of prevention strategies. Alzheimer's disease is the most prevalent form of dementia. Currently we only partially understand some of the pathophysiological mechanisms that lead to development of the disease in aging individuals. In this study, Switch Miner software was used to identify key switch genes in the brain whose expression may lead to the development of Alzheimer's disease. The results indicate that switch genes are enriched in pathways involved in the proteasome, oxidative phosphorylation, Parkinson's disease, Huntington's disease, Alzheimer's disease and metabolism in the hippocampus and posterior cingulate cortex. Network analysis identified the krupel like factor 9 (KLF9), potassium channel tetramerization domain 2 (KCTD2), Sp1 transcription factor (SP1) and chromodomain helicase DNA binding protein 1 (CHD1) as key transcriptional regulators of switch genes in the brain of AD patients. These transcriptions factors have been implicated in conditions associated with Alzheimer's disease, including diabetes, glucocorticoid signaling, stroke, and sleep disorders. The specific pathways affected reveal potential modifiable risk factors by lifestyle changes.

## Introduction

Dementia affects over 50 million people worldwide, approximately 67% of which have Alzheimer's disease (AD) [1]. By 2050 it is predicted that as many as 152 million people may have dementia [1]. Unfortunately, there remains no cure for AD and only four drugs have been approved for treatment that manages symptoms of dementia in some patients. We currently do not completely understand the cause of AD, but there is strong data to support the involvement of the proteins amyloid and tau. In AD patients, amyloid-ß and hyperphosphorylated tau are produced abundantly in the brain. Amyloid-ß forms inter-neuronal plaques that

have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

**Competing interests:** The commercial affiliation with NeuroHub Analytics, LLC and Q Regulating Systems, LLC does not alter our adherence to all PLOS ONE policies on sharing data and materials.

disrupt cell function, whereas tau forms intra-neuronal neurofibrillary tangles that block intracellular transport. Risk factors for AD include genetics (i.e. *APOE epsilon 4* carrier), biology (i.e. aging and gender), and environmental factors (i.e. glucose and cholesterol metabolism, inflammation and oxidative stress) related to lifestyle choices or accidents (i.e. diet, exercise, smoking, education and head trauma).

Plaques and tangles may appear in the brain 18 years before the onset of symptoms [2]. Initially plaques and tangles form in the hippocampus (HIP) and entorhinal cortex (EC) of the temporal lobe, which are involved in learning and memory [3]. In addition to the HIP and EC, metabolic and pathological differences have been found in the middle temporal gyrus (MTG), posterior cingulate cortex (PCC), and superior frontal gyrus (SFG) of AD patients [4–24]. The primary visual cortex (VCX) usually does not show disease-related neurodegeneration [25, 26]. Recently, laser capture microdissection (LCM) was used to select neurons in regions of the brain affected in AD patients and healthy elderly controls. Gene expression profiling of the neurons identified changes that occur in the development and pathogenesis of AD [19, 26]. One striking finding from these studies is that AD patients have significantly reduced expression of mitochondrial transport chain genes in the PCC, MTG and HIP compared to controls [27].

The key events in AD development remain unknown, but gene expression studies on post-mortem brain tissue are expected to reveal pathways that are dysregulated in patients. Recently SWItch Miner (SWIM) software has been developed that combines gene expression data with topological properties of correlation networks to reveal major changes in cellular phenotype that are at the root of biological processes [28]. In SWIM, the Pearson correlation coefficient between the expression of two genes is used to build co-expression networks. RNA transcripts are the nodes of the network and connections between nodes are made if the expression of the genes is significantly correlated or anti-correlated. Clustering algorithms are used to identify disease modules. SWIM analysis identifies "switch genes" that may be fundamental to a disease. SWIM has been used to identify switch genes involved in the transformation of glioblastomas from stem-like to the differentiated state [29] and reprogramming in grapevine development from immature to mature growth [30]. The use of SWIM could be expanded to the study of chronic diseases in order to reveal key players that lead to disease development.

In this study, we have used SWIM to identify genes whose expression is associated with drastic changes in the brain of AD patients. The results show that the switch genes in the HIP and PCC regions of the brain that are affected in AD patients are enriched in proteasome, oxidative phosphorylation, metabolic, Parkinson's disease, Huntington's disease, and Alzheimer's disease pathways.

## Methods

### Database mining

The NCBI GEO database (https://www.ncbi.nlm.nih.gov/gds) and ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/) were searched on June 3, 2019 for studies in which gene expression data was available from laser-captured neurons in the brain of Alzheimer's patients (Fig 1). The NCBI GEO database was queried using the search terms Alzheimer's, brain, neuron and "Homo sapiens"[Organism]) for the study types expression profiling by array and expression profiling by high-throughput sequencing. 44 studies were identified, 21 were brain-specific studies and 4 had data from laser-captured neurons (GSE28146, GSE66333, GSE5281, GSE4757). The ArrayExpress database was searched using the keywords Alzheimer's, Homo sapiens, and transcription and 74 studies were identified, 27 of these studies were brain-

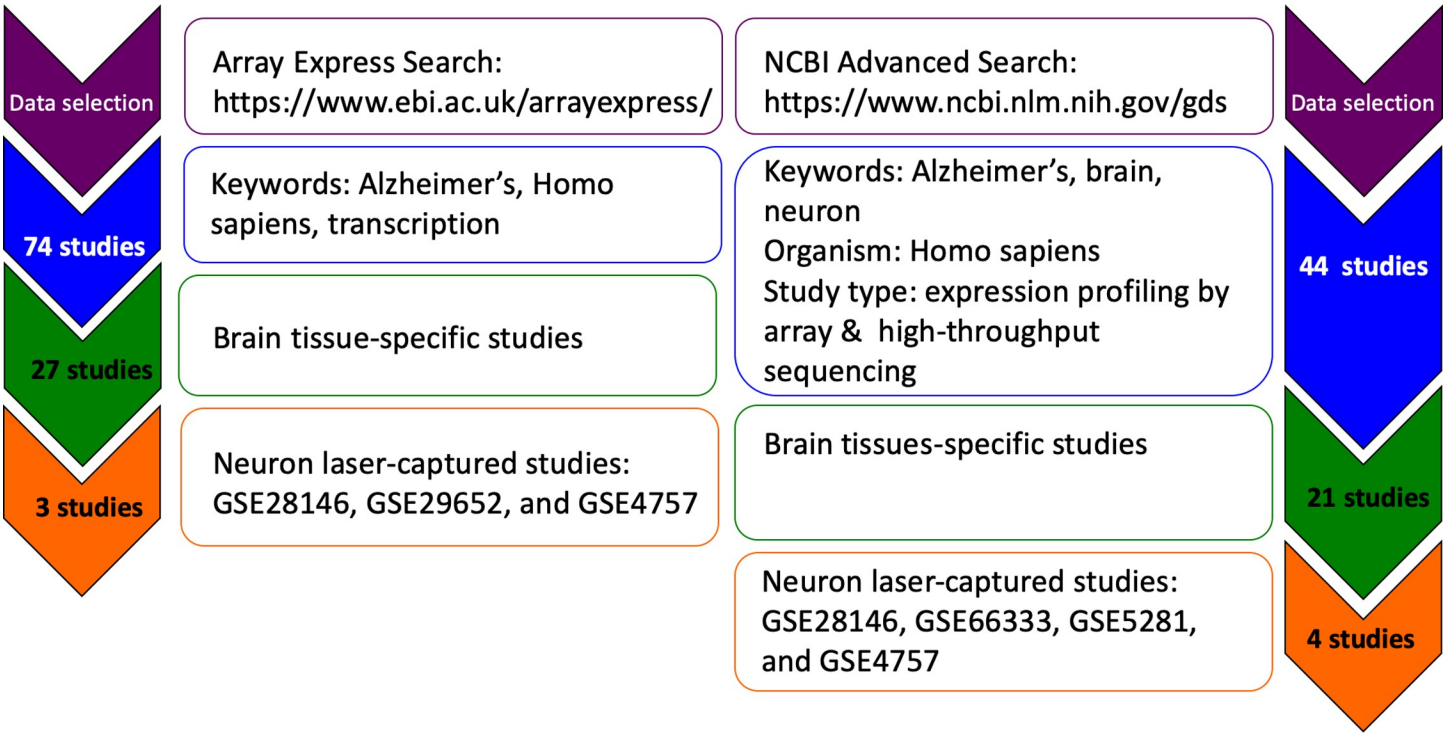


**Fig 1. Database mining.** The NCBI GEO database (https://www.ncbi.nlm.nih.gov/gds) and ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/) were searched on June 3, 2019 for studies in which gene expression data was available from laser-captured neurons in the brain of Alzheimer's patients.

https://doi.org/10.1371/journal.pone.0222921.g001

specific and 3 studies had data from laser-captured neurons (GSE28146, GSE29652, and GSE4757).

## SWIM analysis to identify switch genes

Raw data from the expression arrays was imported into SWIM. The SWIM algorithm is comprised of several steps shown in Fig 2 [31]. In the pre-processing phase, genes that are not expressed or only slightly expressed are removed. In the filtering phase, the fold-change limit was set between 2–4 and genes that were not significantly expressed differently between AD patients compared to controls are removed. The False Discovery Rate method was used to correct for multiple tests [32] and then a Pearson correlation analysis was used to build a co-expression network of genes differentially expressed between AD patients and controls. In step 4, the k-means algorithm was then used to identify communities within the network [33]. To determine the number of clusters, SWIM uses Scree plot, which allows replicating the clustering many times with a new set of initial cluster centroid positions, and for each replicate the k-means algorithm performs iterations until the minimum of the sum of the squared error (SSE) function is reached. The cluster configuration with the lowest SSE values among the replicates is designated as the number of clusters. The heat cartography map is built using a clusterphobic coefficient $K\pi$, which measures external and internal node connections, and the global within-module degree $Zg$, which measures the extent each node is connected to others in its own community. When $Zg$ exceeds 5 a node it is considered a hub. The average Pearson correlation coefficient (APCC) between the expression profiles of each node and its nearest neighbors is used to build the heat cartography map. Using APCC, three types of hubs may be identified. Date hubs show low positive co-expression with their partners (low APCC), party

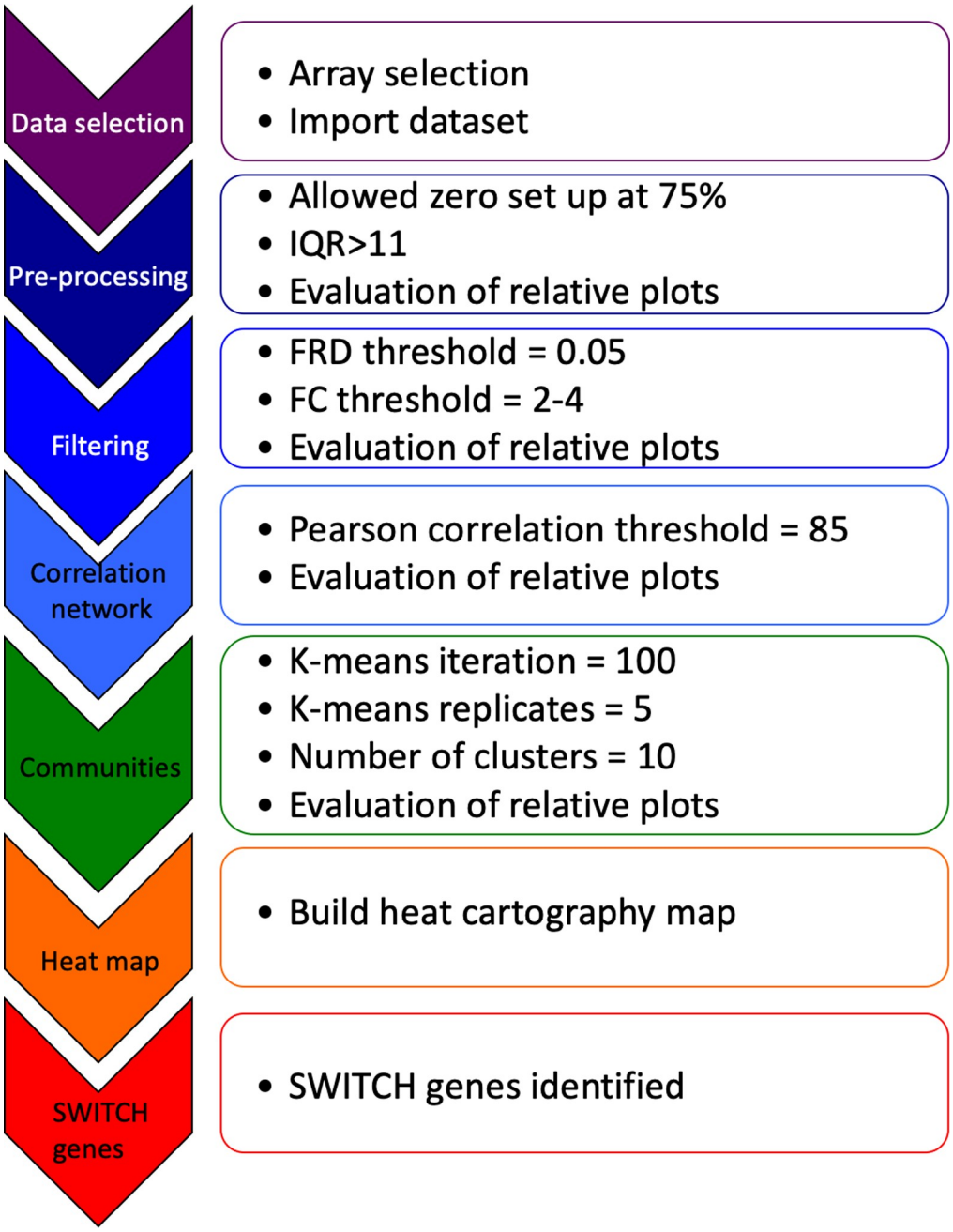


**Fig 2. SWIM analysis to identify switch genes.** Data from the expression arrays was imported into SWIM. The SWIM algorithm is comprised of the steps depicted in the figure [31]. For each brain section and each group, diseased and normal, the mean value of the gene expression is calculated for every cell in the microarray. The p-value is then calculated for each gene. Control for the false discovery rate is performed using the Benjamini and Hochberg approach. The sets of genes to analyze is further filtered by setting thresholds for the magnitude of the fold change and false discovery rate. Next, the Pearson correlation coefficient is calculated for the remaining pairs of genes and those below a threshold are eliminated. At this point 1000 to 2000 genes out of the original 54675 remain. These genes become the nodes in the network that will be analyzed.

https://doi.org/10.1371/journal.pone.0222921.g002

hubs show high positive co-expression (high APCC), and nodes that have negative APCC values are called fight-club hubs [28]. In the final step of SWIM analysis, switch genes are

identified that are a subset of the fight-club hubs that interact outside of their community. Switch genes are characterized as not being a hub in their own cluster (low $Zg < 2.5$), having many links outside their own cluster ($K\pi > 0.8$, when $K\pi$ is close to 1 most of its links are external to its own module), and having a negative average weight of incident links (APCC $< 0$) [28].

The switch genes are the set of genes that interact outside their own community, are not in local hubs and are mainly anti-correlated with their interaction partners. After the switch genes are identified, KEGG pathway analysis was conducted to see if the results imply a relationship to a disease. The random set of genes is chosen from the identified nodes from above. The analysis is performed by studying the effect on the network connectivity of removing different types of nodes by decreasing degree. The total number of nodes to be removed must be equal to the total number of switch genes and the cumulative node deletion is carried out by type (i.e., total hubs, party hubs, date hubs, fight-club hubs, switch genes, and randomly chosen nodes).

### Pathway analysis

Entrez gene identifiers from the SWIM analysis were imported into the Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) which uses singular enrichment analysis [34, 35]. The functional annotation tool was used to select Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The biological functions of KEGG charts were enriched with $p < .05$.

### Transcription factor analysis

Entrez gene identifiers from the SWIM analysis were imported into NetworkAnalyst for network analysis of transcription factors [36]. In NetworkAnalyst, we used the transcription factor and gene target data derived from the ENCODE ChIP-seq data. Transcription factor analysis in NetworkAnalyst uses the BETA Minus Algorithm in which only peak intensity signal $< 500$ and the predicted regulatory potential score $< 1$ is used. Transcription factors were ranked according to network topology measurements including degree and betweenness centrality.

## Results

In order to identify gene expression changes in the brain that may lead to the transition from a healthy aging brain to that of AD patient, we used the SWIM algorithm. The Array Express and NCBI databases were searched to identify studies that contained expression data from postmortem brain tissue of AD patients and age-matched controls (Fig 1). LCM neuron studies (GSE28146, GSE29652, GSE4757, GSE66333 and GSE5281) and several microarrays with samples from brain sections were identified. SWIM analysis was conducted on each of these datasets. In all of the brain section studies and four of the LCM studies, the p values were not sufficiently robust to complete SWIM analysis. Only GSE5281 had data with sufficiently robust p values for the analysis to be completed.

The characteristics of the participants in the GSE5281 study [19, 26, 27] are presented in Table 1. Brain samples from a total of 22–23 participants (10 AD patients and 12–13 controls, depending on the brain region). The average age of the participants was 83 years old. The controls were matched as closely as possible for age at death and mean education level. The AD patients had a Braak stage ranging from III-IV [26].

Gene expression data from laser-captured neurons from HIP, EC, MTG, PCC, SFG and VCX was imported into SWIM. The data for EC using a linear fold-change of 3 is presented in Fig 3. The samples that were filtered out in step 2 of the analysis are depicted as grey bars in Fig 3A, whereas those that are retained for further analysis are shown in red. In Fig 3B the correlation communities are identified. The fight-club hubs are depicted in R4 in blue and are

**Table 1. Characteristics of study participants.**

| GSE5281 | Entorhinal | cortex | |
|---|---|---|---|
| Variable | HC | AD | *P* Value |
| Number of participant | 13 | 10 | |
| Age, mean (SD) | 80.3 (9.2) | 85.6 (6.3) | 0.13 |
| Female/male, No (% male) | 3/10 (76) | 6/4 (40) | 0.07 |
| GSE5281 | Hippocampus | | |
| Variable | HC | AD | *P* Value |
| Number of participant | 13 | 10 | |
| Age, mean (SD) | 79.6 (9.4) | 77.8 (5.7) | 0.6 |
| Female/male, No (% male) | 3/10 (76) | 4/6 (60) | 0.38 |
| GSE5281 | Mid | temporal | gyrus |
| Variable | HC | AD | *P* Value |
| Number of participant | 12 | 10 | |
| Age, mean (SD) | 80.1 (9.8) | 79.1 (6.4) | 0.76 |
| Female/male, No (% male) | 4/8 (67) | 6/10 (63) | 0.82 |
| GSE5281 | Posterior | cingulate | cortex |
| Variable | HC | AD | *P* Value |
| Number of participant | 13 | 10 | |
| Age, mean (SD) | 79.8 (9.4) | 77.5 (6.2) | 0.56 |
| Female/male, No (% male) | 4/9 (69) | 3/6 (67) | 0.90 |
| GSE5281 | Superior | frontal | gyrus |
| Variable | HC | AD | *P* Value |
| Number of participant | 13 | 10 | |
| Age, mean (SD) | 79.3 (10.2) | 79.2 (7.5) | 0.97 |
| Female/male, No (% male) | 4/7 (64) | 10/13 (67) | 0.69 |
| GSE5281 | Primary | visual | cortex |
| Variable | HC | AD | *P* Value |
| Number of participant | 13 | 10 | |
| Age, mean (SD) | 79.9 (7) | 80.2 (6.7) | 0.38 |
| Female/male, No (% male) | 3/9 (75) | 8/11 (58) | 0.33 |

https://doi.org/10.1371/journal.pone.0222921.t001

negatively correlated in expression with their interaction partner. A heat map of the expression of the switch genes, step 5, is shown in Fig 3C. After the switch genes are identified, the data is analyzed further to assess robustness. The data in Fig 3D indicates that fight-club hubs differ from date and party hubs and the switch genes are significantly different that random. The switch genes identified in the EC are listed in S1 Table.

The data for HIP using a linear fold-change of 3 is presented in Fig 4. The samples that are retained for further analysis are depicted in red in Fig 4A, the correlation communities are identified in Fig 4B and the fight-club hubs are depicted in R4 in blue. A heat map of the expression of the switch genes, is shown in Fig 4C. The data indicates that fight-club hubs differ from date and party hubs and the switch genes are significantly different that random, Fig 4D. The switch genes identified in the HIP are listed in S2 Table.

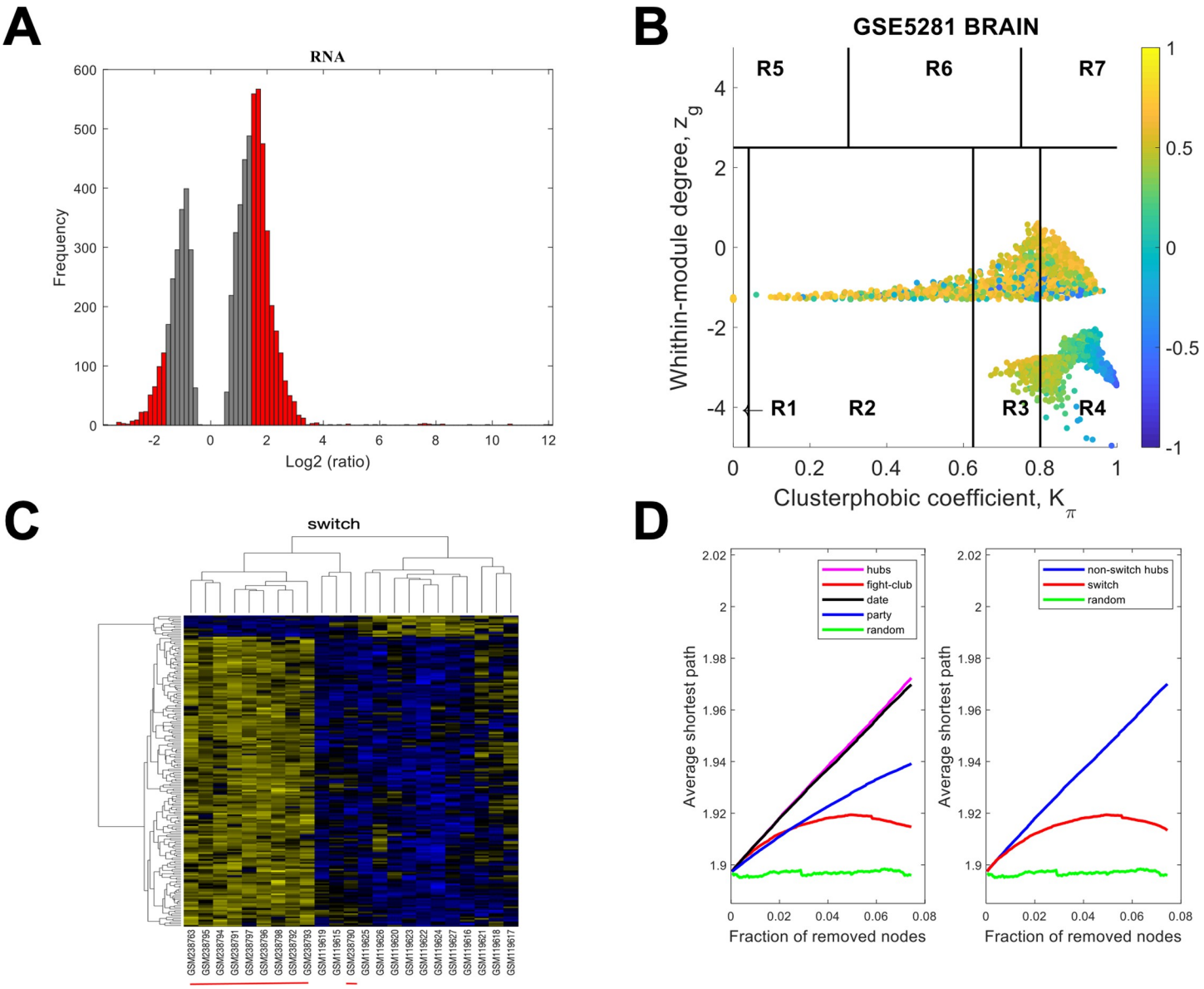


**Fig 3. SWIM analysis of the entorhinal cortex. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the EC. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the entorhinal cortex correlation network. The plane is identified by two parameters: *Zg* (within-module degree) and *Kπ* (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High *Zg* values correspond to nodes that are hubs within their module (local hubs), whereas low *Zg* values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low *Zg* and by high *Kπ* values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the entorhinal cortex. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the entorhinal cortex. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g003

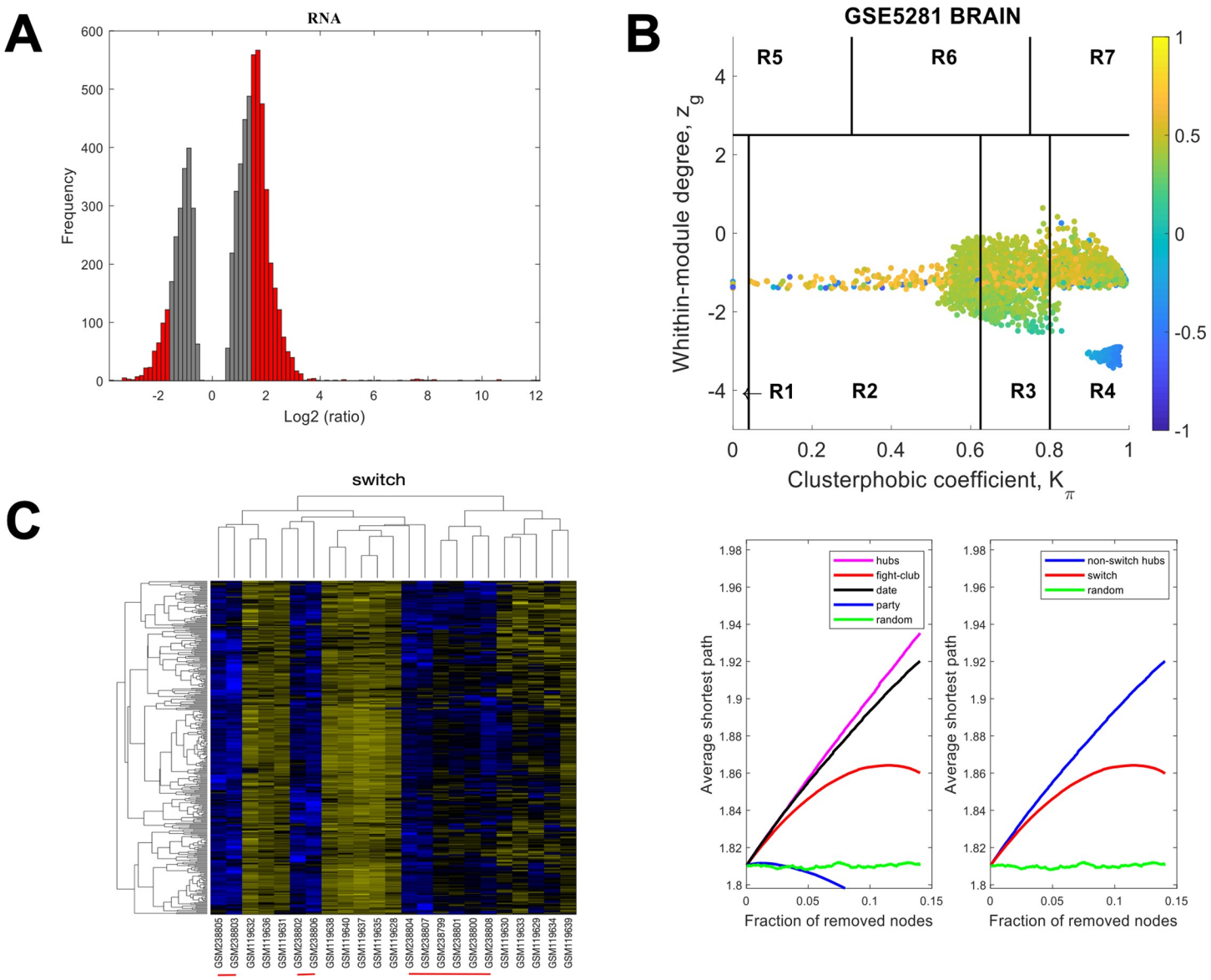


**Fig 4. SWIM analysis of the hippocampus. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the HIP. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the HIP. correlation network. The plane is identified by two parameters: *Zg* (within-module degree) and *Kπ* (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High *Zg* values correspond to nodes that are hubs within their module (local hubs), whereas low *Zg* values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low *Zg* and by high *Kπ* values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the HIP. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the HIP. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g004

The data for MTG using a linear fold-change of 4 is presented in Fig 5. The samples that are retained are depicted in red, Fig 5A and the correlation communities are identified in Fig 5B

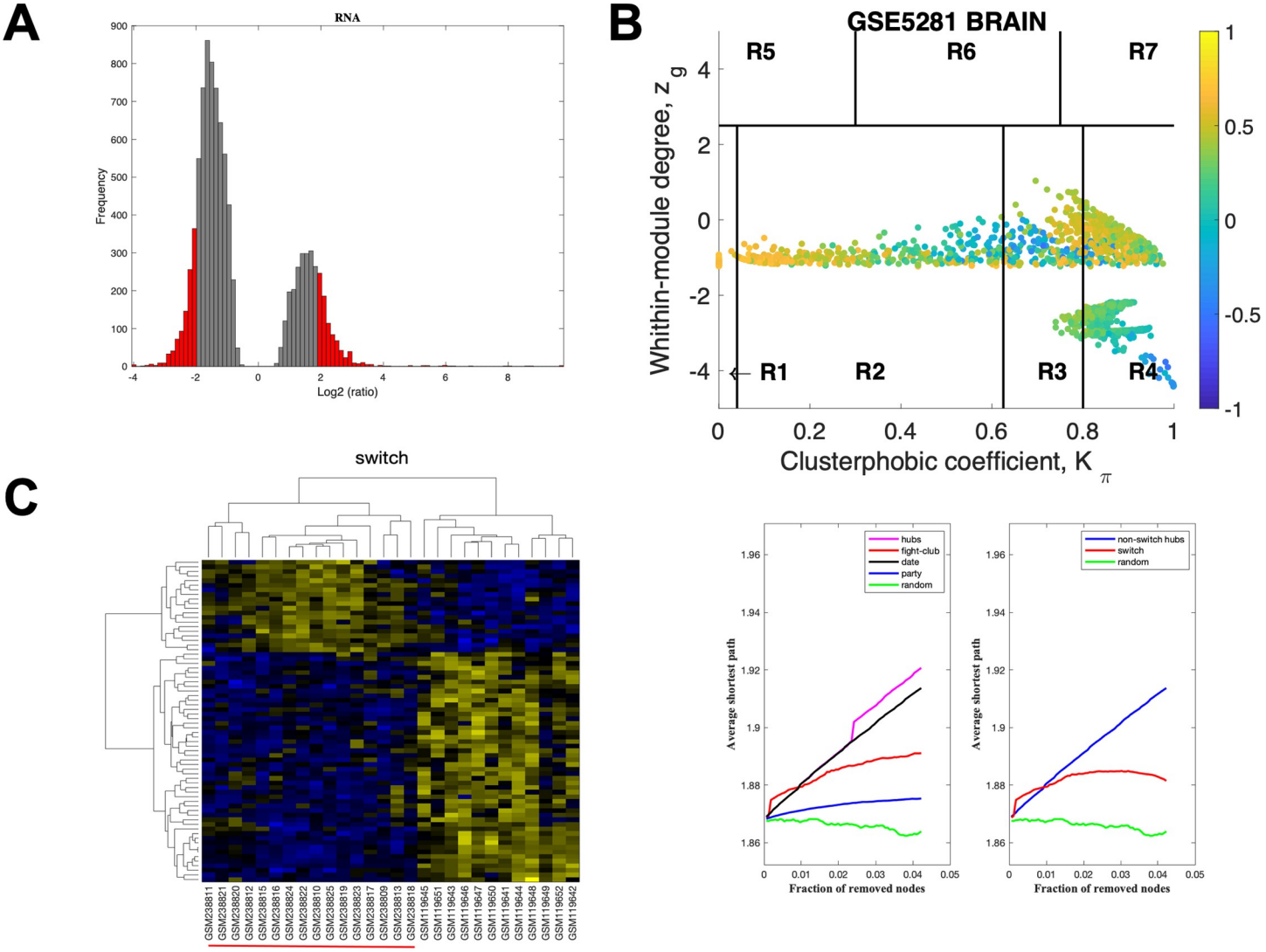


**Fig 5. SWIM analysis of the mid temporal gyrus. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the MTG. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the MTG. correlation network. The plane is identified by two parameters: *Zg* (within-module degree) and *Kπ* (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High *Zg* values correspond to nodes that are hubs within their module (local hubs), whereas low *Zg* values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low *Zg* and by high *Kπ* values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the MTG. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the MTG. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g005

with the fight-club hubs depicted in R4 in blue. A heat map of the expression of the switch genes is presented in Fig 5C. The data indicates that fight-club hubs differ from date and party hubs and the switch genes are significantly different that random, Fig 5D. The switch genes identified in the MTG are listed in S3 Table.

The data for PCC using a linear fold-change of 4 is presented in Fig 6. The samples that are retained for further analysis are depicted in red, Fig 6A, the correlation communities are identified in Fig 6B with the fight-club hubs depicted in R4 in blue. A heat map of the expression of the switch genes, step 5, is shown in Fig 6C. The data indicates that fight-club hubs differ from date and party hubs and the switch genes are significantly different that random, Fig 6D. The switch genes identified in the PCC are listed in S4 Table.

For SFG, initially the fold-change was set at 3 and SWIM was not able to identify any switch genes. We then set the linear fold-change of 2.5 and the data obtained from SWIM analysis is presented in Fig 7. The samples that are retained for further analysis are depicted in red, Fig 7A, the correlation communities are identified in Fig 7B and very few fight-club hubs were found. A heat map of the expression of the switch genes, step 5, is shown in Fig 7C. The data indicates that fight-club hubs differ from date and party hubs, but the switch genes differ only slightly from random, Fig 7D. The switch genes identified in the SFG are listed in S5 Table.

For VCX, initially the fold-change was set at 3 and SWIM was not able to identify any switch genes. We then set the linear fold-change at 2 and the data obtained from SWIM analysis is presented in Fig 8. The samples that are retained for further analysis are depicted in red in Fig 8A and the correlation communities are identified in Fig 8B. The fight-club hubs are depicted in R4 in blue. A heat map of the expression of the switch genes, step 5, is shown in Fig 8C. The data indicates that fight-club hubs do not differ measurably from date and party hubs, but the switch genes are significantly different than random, Fig 8D. The switch genes identified in the VCX are listed in S6 Table.

A Venn diagram analysis and UpSetR plot analysis of the switch genes identified in each brain region is shown in Fig 9 and S7 Table. The order of brain regions which have the largest number of switch genes to the least number is HIP>PCC>EC>MTG>VCX>SFG. The regions that had the largest number of unique switch genes to that which had the least is HIP>EC>PCC>MTG>VCX>SFG. The PCC shares 53 switch genes with the HIP (S7 Table). Interestingly, the EC shares only 1 switch gene with HIP and one with the PCC. Most of the EC switch genes are unique for that region of the brain.

Pathway analysis of the switch genes was performed in order to identify functions. In the HIP the majority of the disrupted pathways were involved with metabolism, specifically glutamine, glutamate, steroid, arginine, pyruvate and amino acids metabolism (Fig 10). Changes in gene expression involved with oxidative phosphorylation, RNA transport and the spliceosome are enriched in switch genes. The switch genes of the HIP are also enriched in Parkinson's, Alzheimer's and Huntington's disease pathways. The PCC switch genes are also enriched in metabolic and Parkinson's, Alzheimer's and Huntington's disease pathways (Fig 11). The switch genes dysregulated pathways shared in the PCC and HIP are the proteasome, oxidative phosphorylation and metabolism (S1 Fig and S7 Table). The switch genes in the EC, MTG, VCX and SFG are not enriched in any particular pathways.

In order to identify key transcriptional regulators of the switch genes from the different brain regions, a transcription factor analysis was performed using NetworkAnalyst [36]. Network analysis was performed using the brain regions with the greater number of switch genes, the HIP and PCC. Network analysis revealed that switch genes identified in the HIP region were regulated by the transcription factors, krupel like factor 9 (KLF9), and potassium channel tetramerization domain 2 (KCTD2), whereas those from the PCC region were regulated by KLF9, Sp1 transcription factor (SP1), and chromodomain helicase DNA binding protein 1

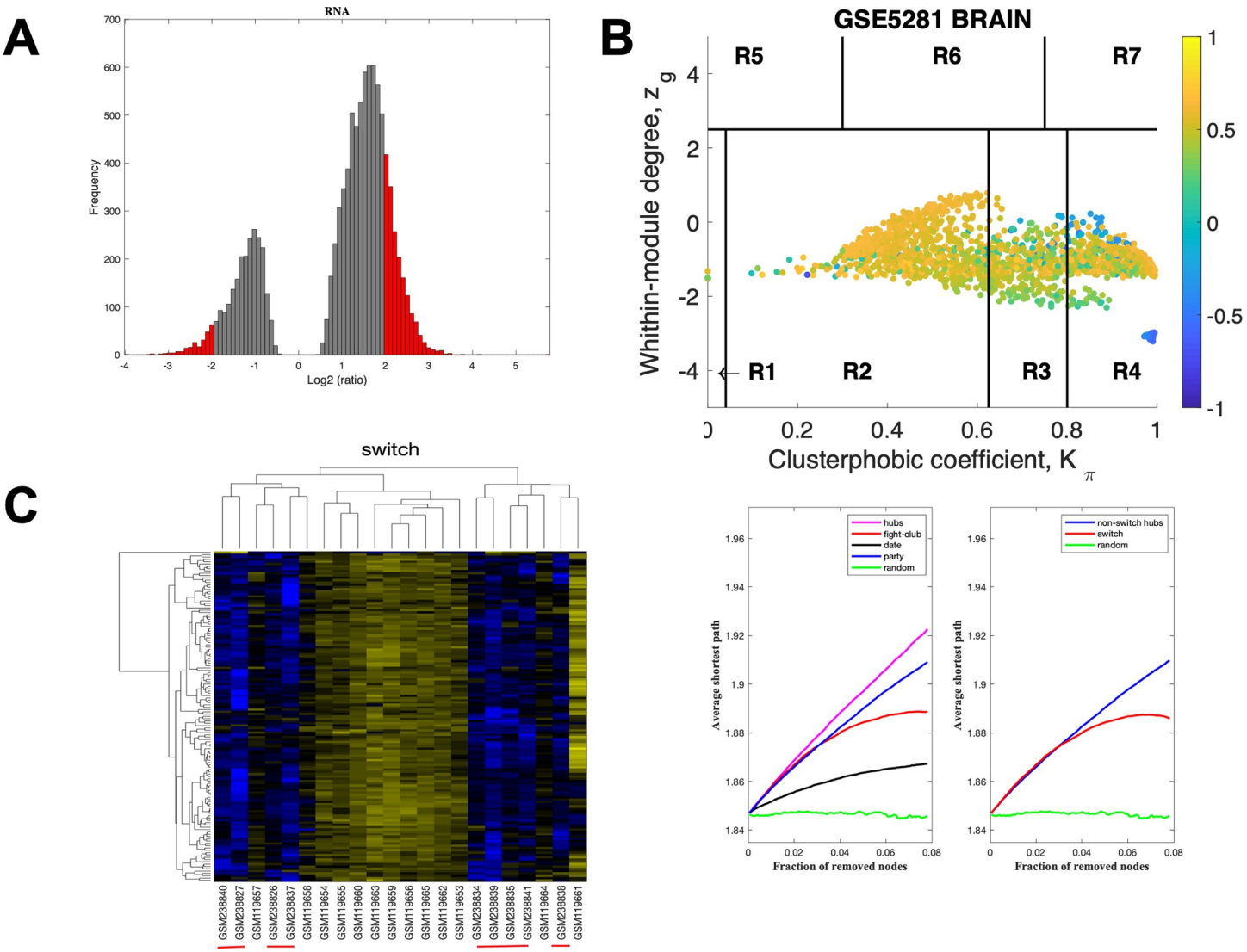


**Fig 6. SWIM analysis of the posterior cingulate. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the PC. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the PC. correlation network. The plane is identified by two parameters: *Zg* (within-module degree) and *Kπ* (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High *Zg* values correspond to nodes that are hubs within their module (local hubs), whereas low *Zg* values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low *Zg* and by high *Kπ* values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the PC. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the PC. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g006

(CHD1). As noted above, KLF9 was shared between the HIP and PCC brain regions (S3 and S4 Figs).

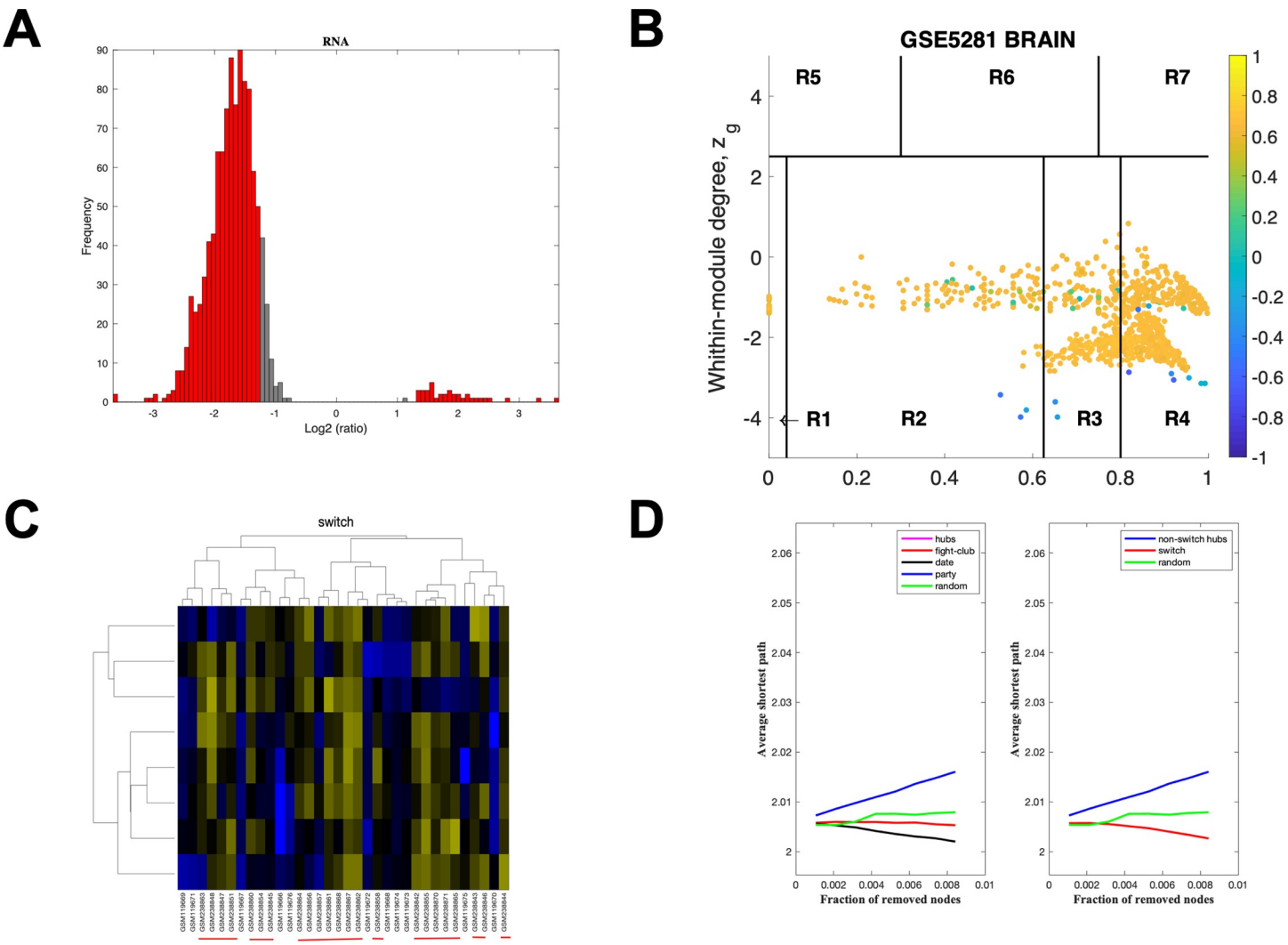


**Fig 7. SWIM analysis of the superior frontal gyrus. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the SFG. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the SFG. correlation network. The plane is identified by two parameters: *Zg* (within-module degree) and *Kπ* (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High *Zg* values correspond to nodes that are hubs within their module (local hubs), whereas low *Zg* values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low *Zg* and by high *Kπ* values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the SFG. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the SFG. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g007

## Discussion

In this study we used SWIM analysis to identify key genes in regions of the brain known to show metabolic and pathological differences in AD patients compared to healthy aging

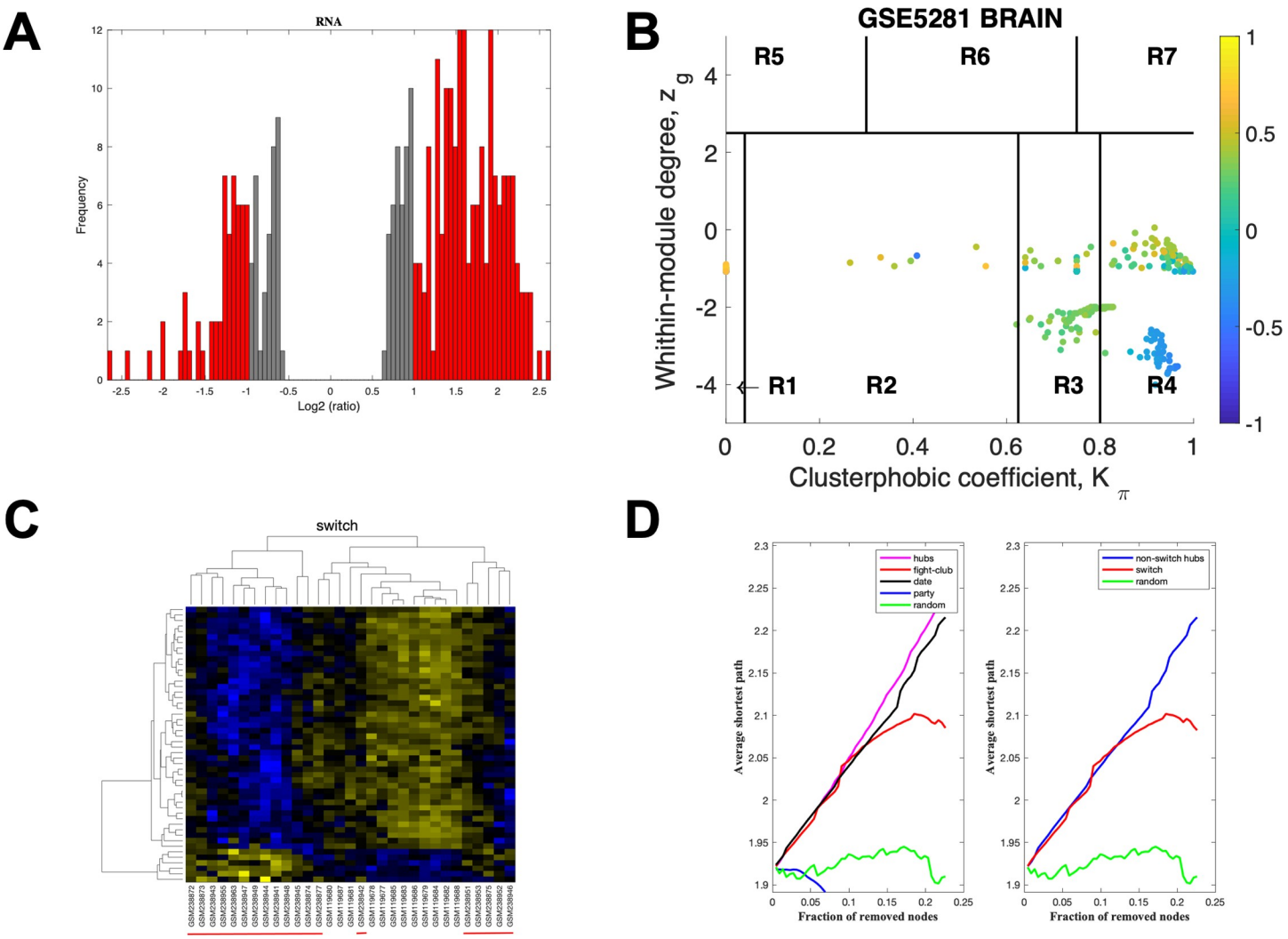


**Fig 8. SWIM analysis of the primary visual cortex. A.** Distribution of the fold-change values for GSE5281 brain microarray gene expression data from the VCX. The x-axis represents the fold-change value (log2 of the fold-change) that is the ratio of the average expression data in AD patients compared to the average expression data in normal controls computed for protein-coding and non-coding RNAs. The y-axis represents the frequency of the obtained fold-change values. The grey bars represent the fold-change values associated with protein-coding and non-coding RNAs that will be discarded according to the selected threshold. The red bars represent the fold-change values associated with protein-coding and non-coding RNAs that were retained for further analysis. **B.** Heat cartography map for GSE5281 brain data from the VCX. correlation network. The plane is identified by two parameters: $Zg$ (within-module degree) and $K\pi$ (clusterphobic coefficient) and it is divided into seven regions each defining a specific node role (R1-R7). High $Zg$ values correspond to nodes that are hubs within their module (local hubs), whereas low $Zg$ values correspond to nodes with few connections within their module (non-hubs within their communities, but they could be hubs in the network). Each node is colored according to its average Pearson Correlation coefficient (APCC) value. Yellow nodes are party and date hubs, which are positively correlated in expression with their interaction partners. Blue nodes are the fight-club hubs, which have an average negative correlation in expression with their interaction partners. Blue nodes falling in the region R4 are the switch genes, which are characterized by low $Zg$ and by high $K\pi$ values and are connected mainly outside their module. **C.** Dendrogram and heat map for switch genes in GSE5281 brain microarray gene expression data from the VCX. The expression profiles of switch genes (including protein-coding and non-coding RNAs) are clustered according to rows (switch genes) and columns (samples) of the switch genes expression data (biclustering). The colors represent different expression levels that increase from blue to yellow. The red line under the x axis labels denotes AD samples. **D.** Robustness for the GSE5281 brain correlation network from the VCX. The x-axis represents the cumulative fraction of removed nodes, while the y-axis represents the average shortest path. The shortest path between two nodes is the minimum number of consecutive edges connecting them. Each curve corresponds to the variation of the average shortest path of the correlation network as function of the removal of nodes specified by the colors of each curve.

https://doi.org/10.1371/journal.pone.0222921.g008

individuals including the HIP, PCC, EC, MTG, and SFG. For comparison we also analyzed gene expression data from the VCX, which usually does not show disease-related neurodegeneration. Transcription data from laser-captured neurons was interrogated to identify switch

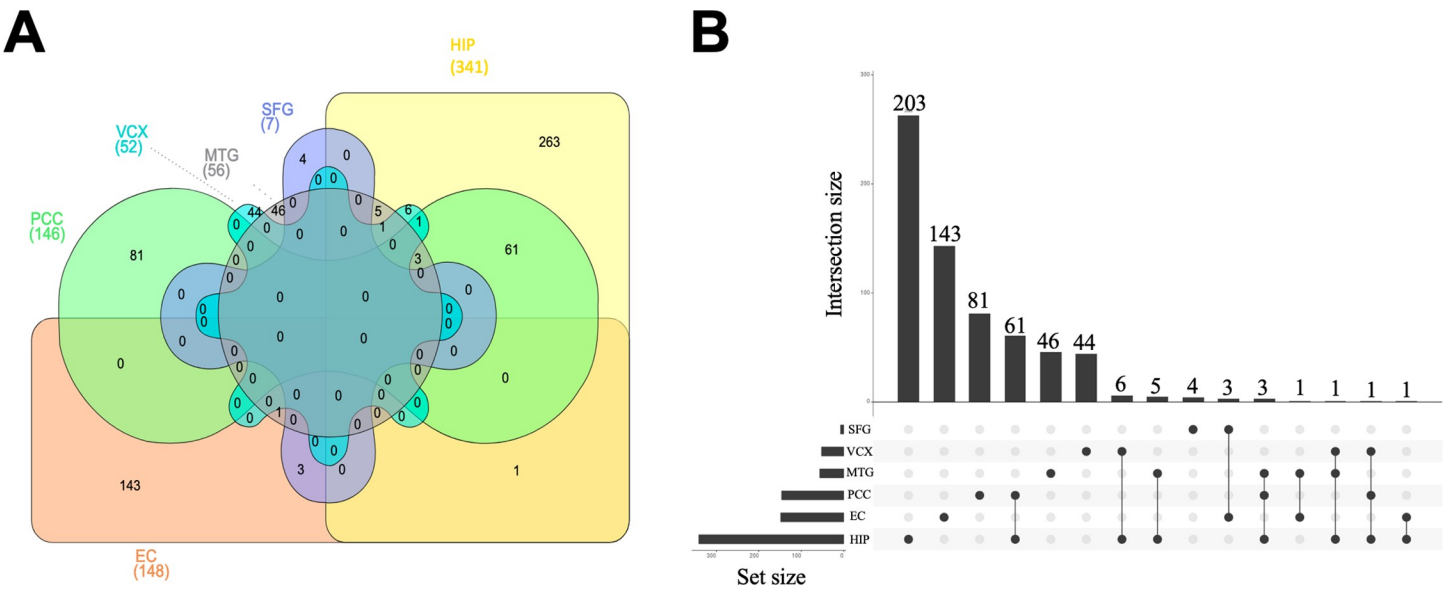


**Fig 9. Venn diagram and UpsetR plot of switch genes from different brain regions.** A. The Venn diagram was created using http://www.interactivenn.net/. The genes symbols were imported for the different brain areas. EC: Entorhino cortex, HIP: Hippocampus, MTG: Mid Temporal Gyrus, SFG: Superior Frontal Gyrus, PCC Posterior Cingulate, VCX: Primary Visual Cortex. B. The UpSetR plot was created as described [37]. The horizontal bars with labels at the lower left of the panel represent the six data sets that were included in the Venn diagram, with the length of each bar displaying the total set size. The dot pattern to the right shows the intersections between the sets. The vertical bars at the top show the size of the corresponding intersection, ranked by decreasing set size, where a gray dot indicates an empty set and a single black dot indicates no intersection with another set.

https://doi.org/10.1371/journal.pone.0222921.g009

genes. The results indicate that changes in gene expression in both the HIP and PCC may alter brain function by disrupting metabolism, oxidative phosphorylation, and the proteasome (S1 Fig).

Previous studies have shown that the PCC and HIP are affected in AD patients [12]. The PCC shows a reduction in glucose metabolism in early AD and has the largest abnormal positron emission tomography scans of cognitively normal late-middle age individuals who carry the *APOE epsilon 4* allele [24, 38, 39]. The HIP shows neurofibrillary tangles in AD patients [12, 14–16]. In addition, energy metabolism genes showed lower expression levels in the PCC and HIP in AD patients compared to controls [27].

In contrast to our study, a previous study that analyzed gene expression changes using the same microarray data found that cellular physiological processes, transport, metabolism and cellular localization were pathways affected across most brain regions including the PCC, HIP, EC, and MTG [26]. The unique aspect of the SWIM algorithm that most likely explains the difference in the results is that it includes fight-club hubs that are negatively correlated with their interaction partners. Therefore, although there is a dysregulation of gene expression that leads to pathways affected in multiple brain regions in AD patients, our data indicate that key switch gene changes that alter significant pathways are present mainly in in the PCC and HIP.

Our results also showed that switch genes in the EC, MTG, VCX and SFG are not enriched in any particular pathways. This suggests that these brain regions may be more resistant to key switch events that cause neurodegeneration compared to the PCC and HIP regions. Similar to our findings, the earlier study that analyzed the GSE5281 data found that the SFG and VCX areas, which are affected in later stages of AD, are relatively neuroprotected and capable of resisting disease pathology [26].

Caberlotto and colleagues used network analysis of AD-related genes to conclude that metabolism-associated processes including insulin and fatty acid metabolism underlie the

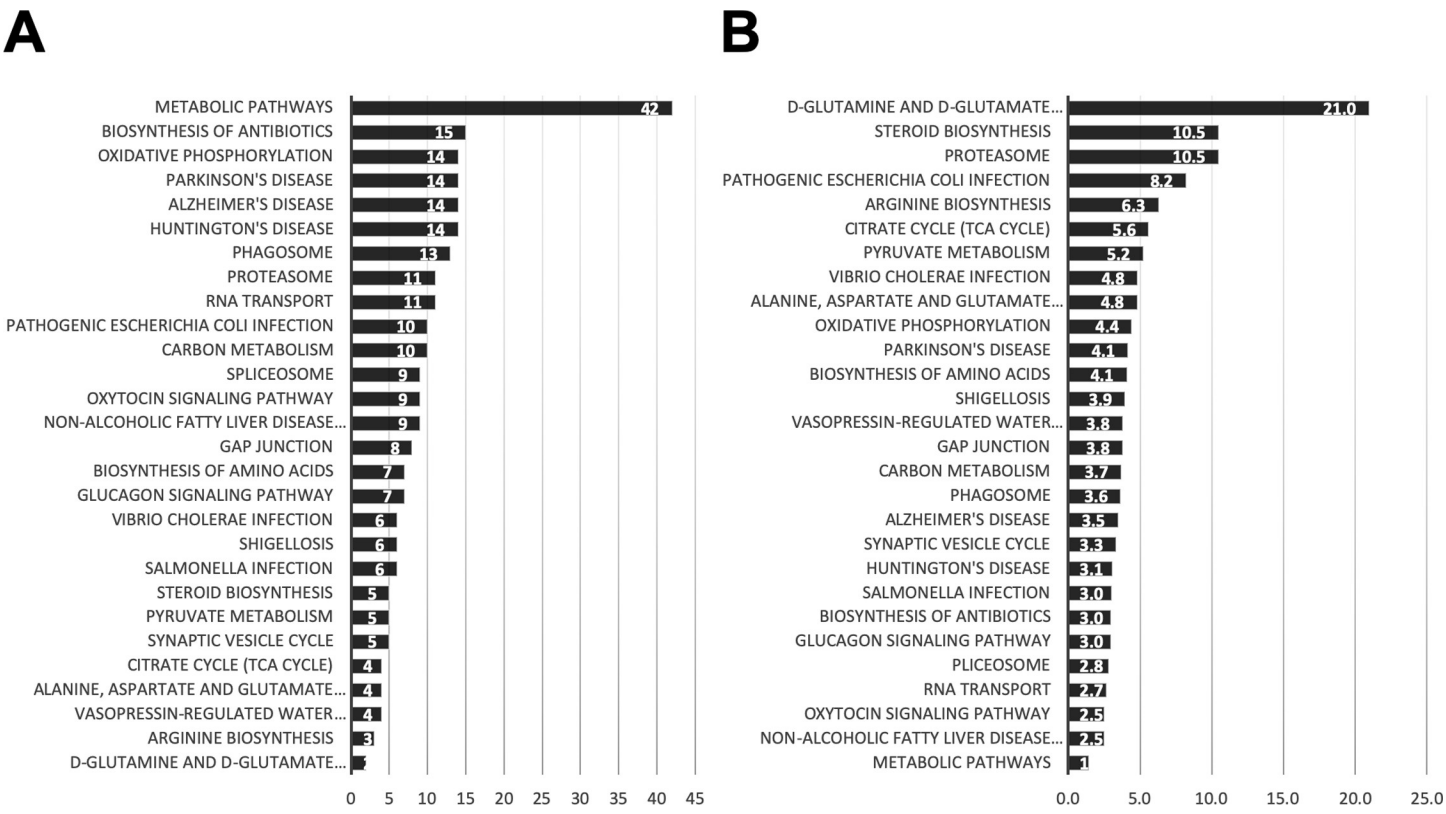


**Fig 10. Hippocampus pathway enrichment analysis.** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of swim genes of the HIP was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID). The gene count for each pathway is represented in A, whereas B represents the fold enrichment.

https://doi.org/10.1371/journal.pone.0222921.g010

development of AD [40]. In this study, seed genes associated with AD were obtained from the same transcriptomic data we used (microarray GSE5281). In addition, they used single nucleotide polymorphism data from AD, molecular targets of AD drugs and AD genes present in the Online Mendelian Inheritance in Man database. We compared the switch genes identified in our study to the AD-related seed genes identified by Caberlotto and colleagues and the results indicate that many of the switch genes were the same as the seed genes, especially in the HIP and PCC strongly suggesting that dysregulation of metabolic processes are key events important to the development of AD (S2 Fig and S8 Table).

Network analysis of the switch genes in the HIP and PCC brain regions identified several transcription factors relevant to the pathogenesis of AD. For example, network analysis identified KLF9 and KCTD2 as the main regulatory transcription factors of the HIP switch genes. Recently, Cui and colleagues demonstrated that KLF9 promotes the expression of peroxisome proliferator-activated receptor γ coactivator 1α (PGC1 α) resulting in hepatic gluconeogenesis and suggested that KLF9 may be responsible for the glucocorticoid therapy-induced diabetes [41]. In this context, diabetes has been extensively associated with an increased risk for AD [42]. Moreover, glucocorticoid overexposure has been associated with cognitive decline, amyloid beta misprocessing and ultimately, the development of AD [43, 44]. Given the involvement of KLF9 in glucose homeostasis and glucocorticoid signaling, its potential as a therapeutic target for AD warrants further investigation.

In addition to KLF9, KCTD2 was another key transcription factor regulating the HIP switch genes. Genome wide association studies identified KCTD2 as a shared susceptibility

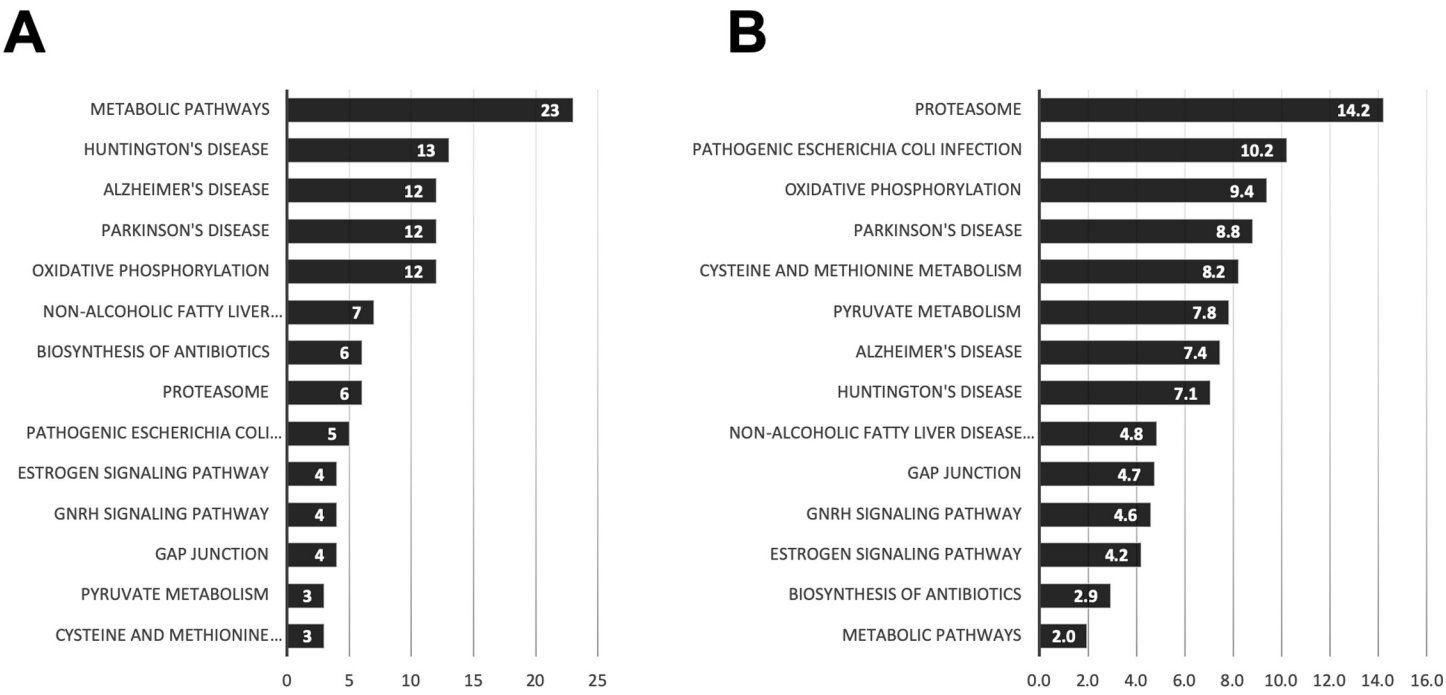


**Fig 11. Posterior cingulate cortex (PCC) pathway enrichment analysis.** KEGG pathway enrichment analysis of swim genes of the posterior cingulate cortex was performed using DAVID. The gene counts for each pathway is represented in A, whereas B represents the fold enrichment.

https://doi.org/10.1371/journal.pone.0222921.g011

gene between AD and ischemic stroke [45, 46]. Interestingly, KCTD2 may play a role in sleep regulation [47, 48] and sleep disturbances have been linked to the development of AD [49].

Similarly, network analysis identified KLF9, SP1 and CHD1 as central regulators of PCC switch genes. Dysregulation of SP1 in AD has been documented in several studies. For instance, SP1 mRNA was upregulated in brains of both human and transgenic AD model mice [50]. Inhibition of SP1 function in a transgenic AD model mice increased memory deficits suggesting that it may be a useful therapeutic target [51]. Another important transcription factor, CHD1, is involved in TDP-43 mediated neurodegeneration [52]. Recently, Chd1 has been found to play a role in learning and memory in mice [53]. Furthermore, Chd1 knockdown in mouse embryonic stem cells mimicked high fat diet and aging-induced gene expression changes [54]. Collectively, the transcription factors identified in this study are involved in processes related to the pathogenesis of AD and thus may be important therapeutic targets.

There is a potential caveat that should be kept in mind when interpreting the results from this study. Although several possible gene expression datasets were identified, only one study achieved the high stringent p-values required for the SWIM analysis. Therefore, the results presented herein may be specific for this dataset and not of AD in general. Nonetheless, the pathways and transcription factors identified in this study have been associated with AD by other investigations. Future studies will seek to confirm the validity of these findings in an independent microarray.

## Conclusions

This study provides novel insights into the key switch events that occur in the HIP and PCC involved in the transformation from a healthy aging brain to that of an AD patient. The majority of the pathways in the HIP and PCC that are altered in AD patients are involved with metabolism including disruption of glutamine, glutamate, steroid, arginine, pyruvate and

amino acids metabolism. In addition, some of the transcriptional regulators of the switch genes are involved in glucose homeostasis, glucocorticoid signaling, sleep regulation, and memory. Targeting these transcription factors may provide novel therapeutics for AD.

## Supporting information

**S1 Fig. Genes common in HIP and PCC pathway enrichment analysis.** KEGG pathway enrichment analysis of swim genes common between the HIP and the PCC were performed using DAVID. The gene counts for each pathway is represented in A whereas B represent the fold enrichment.
(TIFF)

**S2 Fig. Venn diagram between SWIM genes and Caberlotto et al seed genes.** The Venn diagrams were created using http://www.interactivenn.net/. The orange sets represent the SWIM genes whereas the green sets represent the seed genes from Caberlotto et al study.
(TIF)

**S3 Fig. Transcription factor analysis of the switch genes identified in the hippocampus.** Network analysis of HIP switch genes was performed using NetworkAnalyst. Transcription factor data was derived from the ENCODE ChIP-seq database. Transcription factors (blue rectangles) and switch genes (pink circles) are ranked according to network topology measurements, degree and betweenness centrality. Transcription factors with the highest values of degree and betweenness centrality measurements are enclosed in the yellow oval. Gray lines represent protein-protein interactions. Network analysis was performed on June 2019.
(TIF)

**S4 Fig. Transcription factor analysis of the switch genes identified in the posterior cingulate cortex.** Network analysis of posterior cingulate cortex switch genes was performed using NetworkAnalyst. Transcription factor data was derived from the ENCODE ChIP-seq database. Transcription factors (blue rectangles) and switch genes (pink circles) are ranked according to network topology measurements, degree and betweenness centrality. Transcription factors with the highest values of degree and betweenness centrality measurements are enclosed in the yellow oval. Gray lines represent protein-protein interactions. Network analysis was performed on June 2019.
(TIF)

**S1 Table. Enthorinal cortex SWIM genes.**
(XLSX)

**S2 Table. Hippocampus SWIM genes.**
(XLSX)

**S3 Table. Mid temporal gyrus SWIM genes.**
(XLSX)

**S4 Table. Posterior cingulate cortex SWIM genes.**
(XLSX)

**S5 Table. Superior frontal gyrus SWIM genes.**
(XLSX)

**S6 Table. Primary visual cortex SWIM genes.**
(XLSX)

**S7 Table. HIP and PCC SWIM genes venn diagram.**
(XLSX)

**S8 Table. Hippocampus pathway enrichment analysis.**
(XLSX)

**S9 Table. Posterior cingulate cortex pathway enrichment analysis.**
(XLSX)

**S10 Table. HIP/PCC pathway enrichment analysis.**
(XLSX)

## Acknowledgments

This study was funded by the National Institute on Aging (NIA) grant number R01AG062176 to JAP. The NIA had no role in the design, data collection and analysis, decision to publish or preparation of this manuscript.

Q Regulating Systems, LLC and NeuroHub Analytics, LLC provided support in the form of salaries for authors J.P.Q. and J.A.S., but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

**Conceptualization:** Judith A. Potashkin.

**Formal analysis:** Virginie Bottero, Jose A. Santiago, James P. Quinn.

**Funding acquisition:** Judith A. Potashkin.

**Writing – original draft:** Judith A. Potashkin, Virginie Bottero, Jose A. Santiago, James P. Quinn.

**Writing – review & editing:** Judith A. Potashkin, Virginie Bottero, Jose A. Santiago, James P. Quinn.

## References

1. Patterson C. World Alzheimer Report 2018. The state of the art of dementia research: New frontiers London, UK: 2018.
2. Rajan KB, Wilson RS, Weuve J, Barnes LL, Evans DA. Cognitive impairment 18 years before clinical diagnosis of Alzheimer disease dementia. Neurology. 2015; 85(10):898–904. https://doi.org/10.1212/WNL.0000000000001774 PMID: 26109713; PubMed Central PMCID: PMC4560057.
3. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. Cold Spring Harb Perspect Med. 2011; 1(1):a006189. https://doi.org/10.1101/cshperspect.a006189 PMID: 22229116; PubMed Central PMCID: PMC3234452.
4. Angelie E, Bonmartin A, Boudraa A, Gonnaud PM, Mallet JJ, Sappey-Marinier D. Regional differences and metabolic changes in normal aging of the human brain: proton MR spectroscopic imaging study. AJNR Am J Neuroradiol. 2001; 22(1):119–27. PMID: 11158897.
5. Beach TG, Walker R, McGeer EG. Patterns of gliosis in Alzheimer's disease and aging cerebrum. Glia. 1989; 2(6):420–36. https://doi.org/10.1002/glia.440020605 PMID: 2531723.
6. Bliss TV, Collingridge GL. A synaptic model of memory: long-term potentiation in the hippocampus. Nature. 1993; 361(6407):31–9. https://doi.org/10.1038/361031a0 PMID: 8421494.
7. Bobinski M, de Leon MJ, Convit A, De Santi S, Wegiel J, Tarshish CY, et al. MRI of entorhinal cortex in mild Alzheimer's disease. Lancet. 1999; 353(9146):38–40. https://doi.org/10.1016/s0140-6736(05)74869-8 PMID: 10023955.
8. Bouras C, Hof PR, Giannakopoulos P, Michel JP, Morrison JH. Regional distribution of neurofibrillary tangles and senile plaques in the cerebral cortex of elderly patients: a quantitative evaluation of a one-

year autopsy population from a geriatric hospital. Cereb Cortex. 1994; 4(2):138–50. https://doi.org/10.1093/cercor/4.2.138 PMID: 8038565.

9. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 1991; 82(4):239–59. https://doi.org/10.1007/bf00308809 PMID: 1759558.
10. Braak H, Braak E. The human entorhinal cortex: normal morphology and lamina-specific pathology in various diseases. Neurosci Res. 1992; 15(1–2):6–31. https://doi.org/10.1016/0168-0102(92)90014-4 PMID: 1336586.
11. Davies L, Wolska B, Hilbich C, Multhaup G, Martins R, Simms G, et al. A4 amyloid protein deposition and the diagnosis of Alzheimer's disease: prevalence in aged brains determined by immunocytochemistry compared with conventional neuropathologic techniques. Neurology. 1988; 38(11):1688–93. https://doi.org/10.1212/wnl.38.11.1688 PMID: 3054625.
12. de Leon MJ, George AE, Stylopoulos LA, Smith G, Miller DC. Early marker for Alzheimer's disease: the atrophic hippocampus. Lancet. 1989; 2(8664):672–3. https://doi.org/10.1016/s0140-6736(89)90911-2 PMID: 2570916.
13. Du AT, Schuff N, Zhu XP, Jagust WJ, Miller BL, Reed BR, et al. Atrophy rates of entorhinal cortex in AD and normal aging. Neurology. 2003; 60(3):481–6. https://doi.org/10.1212/01.wnl.0000044400.11317.ec PMID: 12578931; PubMed Central PMCID: PMC1851672.
14. Fox NC, Warrington EK, Stevens JM, Rossor MN. Atrophy of the hippocampal formation in early familial Alzheimer's disease. A longitudinal MRI study of at-risk members of a family with an amyloid precursor protein 717Val-Gly mutation. Ann N Y Acad Sci. 1996; 777:226–32. https://doi.org/10.1111/j.1749-6632.1996.tb34423.x PMID: 8624089.
15. Frisoni GB, Laakso MP, Beltramello A, Geroldi C, Bianchetti A, Soininen H, et al. Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. Neurology. 1999; 52 (1):91–100. https://doi.org/10.1212/wnl.52.1.91 PMID: 9921854.
16. Hyman BT, Van Hoesen GW, Damasio AR, Barnes CL. Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. Science. 1984; 225(4667):1168–70. https://doi.org/10.1126/science.6474172 PMID: 6474172.
17. Ibanez V, Pietrini P, Alexander GE, Furey ML, Teichberg D, Rajapakse JC, et al. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. Neurology. 1998; 50(6):1585–93. https://doi.org/10.1212/wnl.50.6.1585 PMID: 9633698.
18. Jack CR Jr., Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. Neurology. 1998; 51(4):993–9. https://doi.org/10.1212/wnl.51.4.993 PMID: 9781519; PubMed Central PMCID: PMC2768817.
19. Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Walker DG, et al. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol Genomics. 2007; 28(3):311–22. https://doi.org/10.1152/physiolgenomics.00208.2006 PMID: 17077275; PubMed Central PMCID: PMC2259385.
20. Mielke R, Herholz K, Grond M, Kessler J, Heiss WD. Clinical deterioration in probable Alzheimer's disease correlates with progressive metabolic impairment of association areas. Dementia. 1994; 5(1):36–41. PMID: 8156085.
21. Morris JC, Price JL. Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease. J Mol Neurosci. 2001; 17(2):101–18. PMID: 11816784.
22. Price JL, Davis PB, Morris JC, White DL. The distribution of tangles, plaques and related immunohistochemical markers in healthy aging and Alzheimer's disease. Neurobiol Aging. 1991; 12(4):295–312. https://doi.org/10.1016/0197-4580(91)90006-6 PMID: 1961359.
23. Rogers J, Morrison JH. Quantitative morphology and regional and laminar distributions of senile plaques in Alzheimer's disease. J Neurosci. 1985; 5(10):2801–8. PMID: 4045553.
24. Small GW, Ercoli LM, Silverman DH, Huang SC, Komo S, Bookheimer SY, et al. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. Proc Natl Acad Sci U S A. 2000; 97(11):6037–42. https://doi.org/10.1073/pnas.090106797 PMID: 10811879; PubMed Central PMCID: PMC18554.
25. Metsaars WP, Hauw JJ, van Welsem ME, Duyckaerts C. A grading system of Alzheimer disease lesions in neocortical areas. Neurobiol Aging. 2003; 24(4):563–72. https://doi.org/10.1016/s0197-4580(02)00134-3 PMID: 12714113.
26. Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Ramsey K, et al. Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. Physiol Genomics. 2008; 33(2):240–56. https://doi.org/10.1152/physiolgenomics.00242.2007 PMID: 18270320; PubMed Central PMCID: PMC2826117.

27. Liang WS, Reiman EM, Valla J, Dunckley T, Beach TG, Grover A, et al. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci U S A. 2008; 105(11):4441–6. https://doi.org/10.1073/pnas.0709259105 PMID: 18332434; PubMed Central PMCID: PMC2393743.

28. Fiscon G, Conte F, Farina L, Paci P. Network-Based Approaches to Explore Complex Biological Systems towards Network Medicine. Genes (Basel). 2018; 9(9). https://doi.org/10.3390/genes9090437 PMID: 30200360; PubMed Central PMCID: PMC6162385.

29. Fiscon G, Conte F, Licursi V, Nasi S, Paci P. Computational identification of specific genes for glioblastoma stem-like cells identity. Sci Rep. 2018; 8(1):7769. https://doi.org/10.1038/s41598-018-26081-5 PMID: 29773872; PubMed Central PMCID: PMC5958093.

30. Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, et al. Integrated network analysis identifies fight-club nodes as a class of hubs encompassing key putative switch genes that induce major transcriptome reprogramming during grapevine development. Plant Cell. 2014; 26(12):4617–35. https://doi.org/10.1105/tpc.114.133710 PMID: 25490918; PubMed Central PMCID: PMC4311215.

31. Paci P, Colombo T, Fiscon G, Gurtner A, Pavesi G, Farina L. SWIM: a computational tool to unveiling crucial nodes in complex biological networks. Sci Rep. 2017; 7:44797. https://doi.org/10.1038/srep44797 PMID: 28317894; PubMed Central PMCID: PMC5357943.

32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995:289–300.

33. Hartigan J, Wong M. Algorithm AS 136: A k-means clustering algorithm. J R Stat Soc Ser B Stat Methodol. 1979:100–8.

34. Dennis G Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003; 4(5):P3. PMID: 12734009.

35. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007; 35(Web Server issue):W169–75. https://doi.org/10.1093/nar/gkm415 PMID: 17576678; PubMed Central PMCID: PMC1933169.

36. Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat Protoc. 2015; 10(6):823–44. https://doi.org/10.1038/nprot.2015.052 PMID: 25950236.

37. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph. 2014; 20(12):1983–92. Epub 2015/09/12. https://doi.org/10.1109/TVCG.2014.2346248 PMID: 26356912; PubMed Central PMCID: PMC4720993.

38. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. Ann Neurol. 1997; 42(1):85–94. https://doi.org/10.1002/ana.410420114 PMID: 9225689.

39. Reiman EM, Chen K, Alexander GE, Caselli RJ, Bandy D, Osborne D, et al. Correlations between apolipoprotein E epsilon4 gene dose and brain-imaging measurements of regional hypometabolism. Proc Natl Acad Sci U S A. 2005; 102(23):8299–302. https://doi.org/10.1073/pnas.0500579102 PMID: 15932949; PubMed Central PMCID: PMC1149416.

40. Caberlotto L, Lauria M, Nguyen TP, Scotti M. The central role of AMP-kinase and energy homeostasis impairment in Alzheimer's disease: a multifactor network analysis. PLoS One. 2013; 8(11):e78919. https://doi.org/10.1371/journal.pone.0078919 PMID: 24265728; PubMed Central PMCID: PMC3827084.

41. Cui A, Fan H, Zhang Y, Zhang Y, Niu D, Liu S, et al. Dexamethasone-induced Kruppel-like factor 9 expression promotes hepatic gluconeogenesis and hyperglycemia. J Clin Invest. 2019; 130:2266–78. https://doi.org/10.1172/JCI66062 PMID: 31033478; PubMed Central PMCID: PMC6546458.

42. de la Monte SM, Wands JR. Alzheimer's disease is type 3 diabetes-evidence reviewed. J Diabetes Sci Technol. 2008; 2(6):1101–13. https://doi.org/10.1177/193229680800200619 PMID: 19885299; PubMed Central PMCID: PMC2769828.

43. Canet G, Chevallier N, Zussy C, Desrumaux C, Givalois L. Central Role of Glucocorticoid Receptors in Alzheimer's Disease and Depression. Front Neurosci. 2018; 12:739. https://doi.org/10.3389/fnins.2018.00739 PMID: 30459541; PubMed Central PMCID: PMC6232776.

44. Ouanes S, Popp J. High Cortisol and the Risk of Dementia and Alzheimer's Disease: A Review of the Literature. Front Aging Neurosci. 2019; 11:43. https://doi.org/10.3389/fnagi.2019.00043 PMID: 30881301; PubMed Central PMCID: PMC6405479.

45. Boada M, Antunez C, Ramirez-Lorca R, DeStefano AL, Gonzalez-Perez A, Gayan J, et al. ATP5H/KCTD2 locus is associated with Alzheimer's disease risk. Mol Psychiatry. 2014; 19(6):682–7. https://doi.org/10.1038/mp.2013.86 PMID: 23857120; PubMed Central PMCID: PMC4031637.

46. Traylor M, Adib-Samii P, Harold D, Alzheimer's Disease Neuroimaging I, International Stroke Genetics Consortium UKYLSDNAr, Dichgans M, et al. Shared genetic contribution to Ischaemic Stroke and Alzheimer's Disease. Ann Neurol. 2016; 79(5):739–47. https://doi.org/10.1002/ana.24621 PMID: 26913989; PubMed Central PMCID: PMC4864940.

47. Li Q, Kellner DA, Hatch HAM, Yumita T, Sanchez S, Machold RP, et al. Conserved properties of Drosophila Insomniac link sleep regulation and synaptic function. PLoS Genet. 2017; 13(5):e1006815. https://doi.org/10.1371/journal.pgen.1006815 PMID: 28558011; PubMed Central PMCID: PMC5469494.

48. Pirone L, Smaldone G, Esposito C, Balasco N, Petoukhov MV, Spilotros A, et al. Proteins involved in sleep homeostasis: Biophysical characterization of INC and its partners. Biochimie. 2016; 131:106–14. https://doi.org/10.1016/j.biochi.2016.09.013 PMID: 27678190.

49. Ju YE, Lucey BP, Holtzman DM. Sleep and Alzheimer disease pathology—a bidirectional relationship. Nat Rev Neurol. 2014; 10(2):115–9. https://doi.org/10.1038/nrneurol.2013.269 PMID: 24366271; PubMed Central PMCID: PMC3979317.

50. Citron BA, Dennis JS, Zeitlin RS, Echeverria V. Transcription factor Sp1 dysregulation in Alzheimer's disease. J Neurosci Res. 2008; 86(11):2499–504. https://doi.org/10.1002/jnr.21695 PMID: 18449948.

51. Citron BA, Saykally JN, Cao C, Dennis JS, Runfeldt M, Arendash GW. Transcription factor Sp1 inhibition, memory, and cytokines in a mouse model of Alzheimer's disease. Am J Neurodegener Dis. 2015; 4(2):40–8. PMID: 26807343; PubMed Central PMCID: PMC4700125.

52. Berson A, Sartoris A, Nativio R, Van Deerlin V, Toledo JB, Porta S, et al. TDP-43 Promotes Neurodegeneration by Impairing Chromatin Remodeling. Curr Biol. 2017; 27(23):3579–90 e6. https://doi.org/10.1016/j.cub.2017.10.024 PMID: 29153328; PubMed Central PMCID: PMC5720388.

53. Schoberleitner I, Mutti A, Sah A, Wille A, Gimeno-Valiente F, Piatti P, et al. Role for Chromatin Remodeling Factor Chd1 in Learning and Memory. Front Mol Neurosci. 2019; 12:3. https://doi.org/10.3389/fnmol.2019.00003 PMID: 30728766; PubMed Central PMCID: PMC6351481.

54. Green CD, Huang Y, Dou X, Yang L, Liu Y, Han JJ. Impact of Dietary Interventions on Noncoding RNA Networks and mRNAs Encoding Chromatin-Related Factors. Cell Rep. 2017; 18(12):2957–68. https://doi.org/10.1016/j.celrep.2017.03.001 PMID: 28329687.