Critical Role of Downregulated Synaptic Vesicle Cycles in the Pathogenesis of Alzheimer's Disease

Caryse Chang[1], Asha Daftary[2], Dhruva Gopalakrishnan[3], Inhan Lee[4]

[1] Campolindo High School, Moraga, CA
[2] Beckman High School, Irvine, CA
[3] University of Michigan, Ann Arbor, MI
[4] miRcore, Ann Arbor, MI

KEYWORDS: Alzheimer's disease, Synaptic Vesicle Cycle, Calcium Signalling

OVERVIEW. Using computational bioinformatic tools, we studied which genes are dysregulated in Alzheimer's disease patients. Using these findings, we used other tools to analyze the results of the specific dysregulated genes on molecular pathways and cross-referenced peer-reviewed research articles to understand the contribution of those dysregulated genes to the pathogenesis of Alzheimer's disease and the underlying symptoms.

SUMMARY. Alzheimer's disease (AD) is a neurodegenerative disease characterized by progressive loss of memory, functionality, and cognitive abilities, globally affecting 38.5 million every year. Despite its widespread prevalence, there are no cures for AD, only treatments that slow its progression. A deeper understanding of the molecular mechanisms driving this disease is critical to identify potential therapeutic targets. This study investigates differential gene expression in AD to unravel significantly dysregulated pathways associated with the pathology of this disease. High throughput Illumina® microarray analyses of the middle temporal gyrus biopsies of 97 AD patients and 98 non-disease controls were obtained from Gene Expression Omnibus (GEO) dataset GSE132903. Differential gene expression was analyzed using GEO2R, followed by pathway enrichment analysis of the top 250 significantly dysregulated genes through String-db. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG), we examined the synaptic vesicle cycle (hsa04721) pathway. RStudio was used for statistical analysis and heatmap visualizations. Differential RNA analysis revealed statistically significant downregulation of the synaptic vesicle cycle (SVC) in AD samples. The SVC is essential for neuronal communication, and regulates neurotransmitter release through calcium-dependent vesicle fusion. SYT1 is a crucial calcium-binding gene in synaptic vesicle fusion, and was found to be the second most significantly dysregulated gene in the dataset. Disruption of the SVC may impair neurotransmitter release, contributing to ß-amyloid (Aß) accumulation, calcium imbalance, synaptic dysfunction, and neuronal toxicity. Each of the above imbalances are associated with impaired cognitive and motor function observed in AD patients. This study demonstrates that downregulation of the SVC may contribute significantly to AD pathology, potentially supporting the development of treatments that slow disease progression.

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia, characterized by a severe decline in cognitive abilities, memory loss, and behavioral

changes such as depression, social withdrawal, and mood swings (1). Pathologically, AD leads to an accumulation of extracellular neuritic plaques comprised of Aß peptides and intracellular neurofibrillary tangles, which collectively drive neurodegeneration and the subsequent atrophy of critical brain regions, including the cerebral cortex and hippocampus (1). While advanced age, genetic predisposition, traumatic brain injuries, and environmental variables are well-established risk factors, the precise molecular mechanisms underlying AD onset and progression remain unknown (1). Current treatments only temporarily alleviate symptoms or modestly slow clinical progression, underscoring an urgent need to understand the genetic architecture of AD, to identify novel biomarkers for early diagnosis, and to develop therapeutic targets capable of addressing the root cause of the disease (2).

The synaptic vesicle cycle (SVC) is the process by which vesicles in neurons release neurotransmitters into the synaptic cleft for neural communication (2). Pathological accumulations of Aß have been shown to disrupt the SVC, stimulating increased secretion of the glutamate neurotransmitter (3). Initially, this increases neural communication, but over extended periods of time, it will damage the neuron and can cause synaptic depression, in which synapse strength weakens over time as more vesicles are released (3). During prolonged synaptic depression, the rapid exocytosis of neurotransmitters depletes the pool of available vesicles, drastically lowering the probability of subsequent vesicle release in response to action potentials and ultimately paralyzing network signaling (3). Because the structural mobilization and fusion of these vesicles are heavily regulated by precise intracellular calcium signaling, disruptions to calcium-sensing proteins within the SVC are increasingly suspected to drive early synaptic failure (4). When there are higher levels of calcium, there is a greater level of exocytosis, where the neurotransmitters are released. Conversely, lower levels of calcium lead to lower levels of exocytosis (4). However, how specific gene networks within this pathway collapse in AD patients requires further systemic investigation.

One of our key findings was the downregulation of the SYT1 gene, which triggers the binding of the vesicle to the presynaptic neuronal membrane. This gene was the second most significantly dysregulated gene in our genetic dataset. We hypothesized that downregulation of key calcium-regulated genes within the synaptic vesicle cycle disrupts presynaptic exocytosis, which will decrease neural communication and ultimately lead to a decline in cognitive abilities and neuronal damage.

To address this gap, this study aimed to identify key genes contributing to the progression of Alzheimer's disease by analyzing and comparing differences in gene expression in Alzheimer's and control patients. We used computational bioinformatics tools like String-db, KEGG, and GEO2R. Our transcriptomic analysis revealed widespread dysregulation across 17,525 genes, characterized by a collapse of vital structural components of the SVC. Crucially, we identified the SYT1 gene, which encodes a calcium sensor that triggers vesicle binding to the presynaptic membrane, as the second most significantly dysregulated gene in the dataset. Ultimately, these findings demonstrate that AD pathology is tied to the transcriptional suppression of the synaptic vesicle pathway, suggesting that stabilizing presynaptic calcium-sensing mechanisms may be an avenue for future therapeutic intervention.

RESULTS

GEO/GEO2R. We used GEO2R, a web-based tool that compares multiple groups of samples in a GEO series to analyze the GEO dataset of GSE132903. This experiment compared transcriptomic profiles from AD patient brain tissue against healthy controls to isolate statistically

significant variations. The initial analysis revealed a total of 17,525 genes that were significantly dysregulated in AD patients relative to controls. The global distribution of these alterations was visualized in a volcano plot, which demonstrated an approximately equal distribution of statistically significant downregulated and upregulated genes in AD patients compared to controls (**Figure 1**).

FIGURE 1

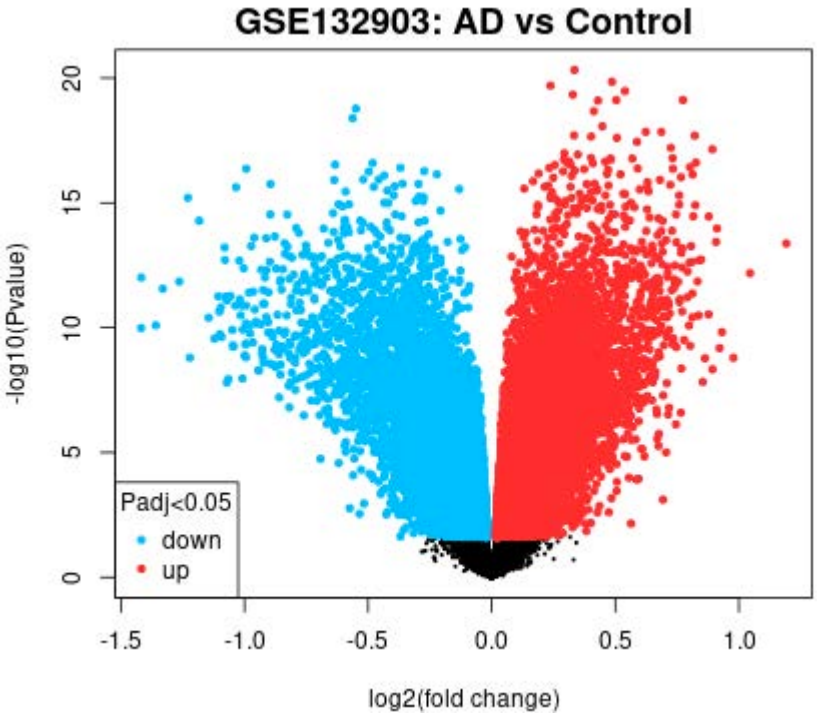
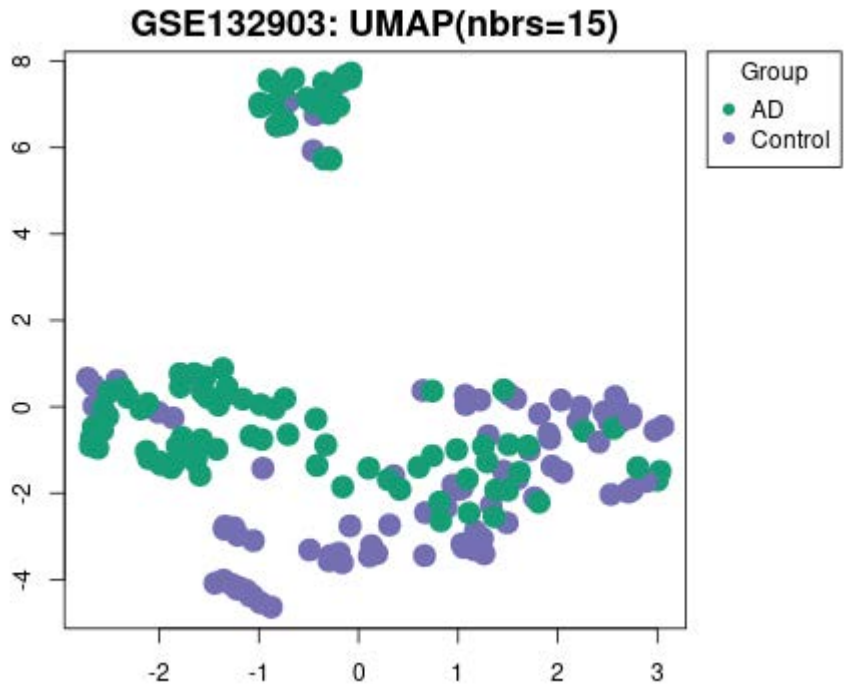


Figure 1. Volcano plot of dysregulated genes in control compared to AD patients from GEO. The x-axis of this volcano plot (the log fold change) measures the difference in gene expression in the control group compared to the Alzheimer's samples. The y-axis indicates the statistical significance of each gene; any gene with an adjusted p-value of less than 0.05 indicates statistical significance and is highlighted as blue or red. In this volcano plot, it's shown that the dataset GSE132903 has an approximately equal distribution of upregulated genes (red dots) and downregulated genes (blue dots) in control patients compared to Alzheimer's patients. Dots that are colored as black indicate not statistically significant data, as they fall out of range of the adjusted p-value of 0.05.

FIGURE 2



.

Figure 2. UMAP of control samples compared to AD samples. Each dot in the UMAP indicates a patient sample. The UMAP reveals partial separation between AD and control samples, with some overlap, indicating that while gene expression differences exist, they are not sufficient for complete group segregation. Notably, a distinct cluster enriched in AD samples suggests potential heterogeneity or subtype-specific gene expression patterns within the AD group.

It was also visualized as a UMAP to assess overall cohort distribution based on the total transcriptome (**Figure 2**). In addition, GEO2R produced a table of the genes measured in the microarray that are dysregulated between the two sample groups. This table was exported into a Google Sheet and used to find the top 250 statistically significant dysregulated genes based on adjusted p-values. These top 250 genes were input to String-db. Notably, SYT1, a gene vital to synaptic vesicle fusion in the synaptic vesicle cycle, was found to be the second most downregulated gene in the SVC (adj.p-value = 3.60E-09, logFC = -1.42) in GEO2R.

String-db. String-db is a database that helps find the best fitting pathway when a list of genes or proteins is provided. The database does so with the use of text mining. This analysis allows us to evaluate whether the observed genetic dysregulation represented randomized cellular noise or a coordinated disruption of interconnected functional systems.

FIGURE 3A

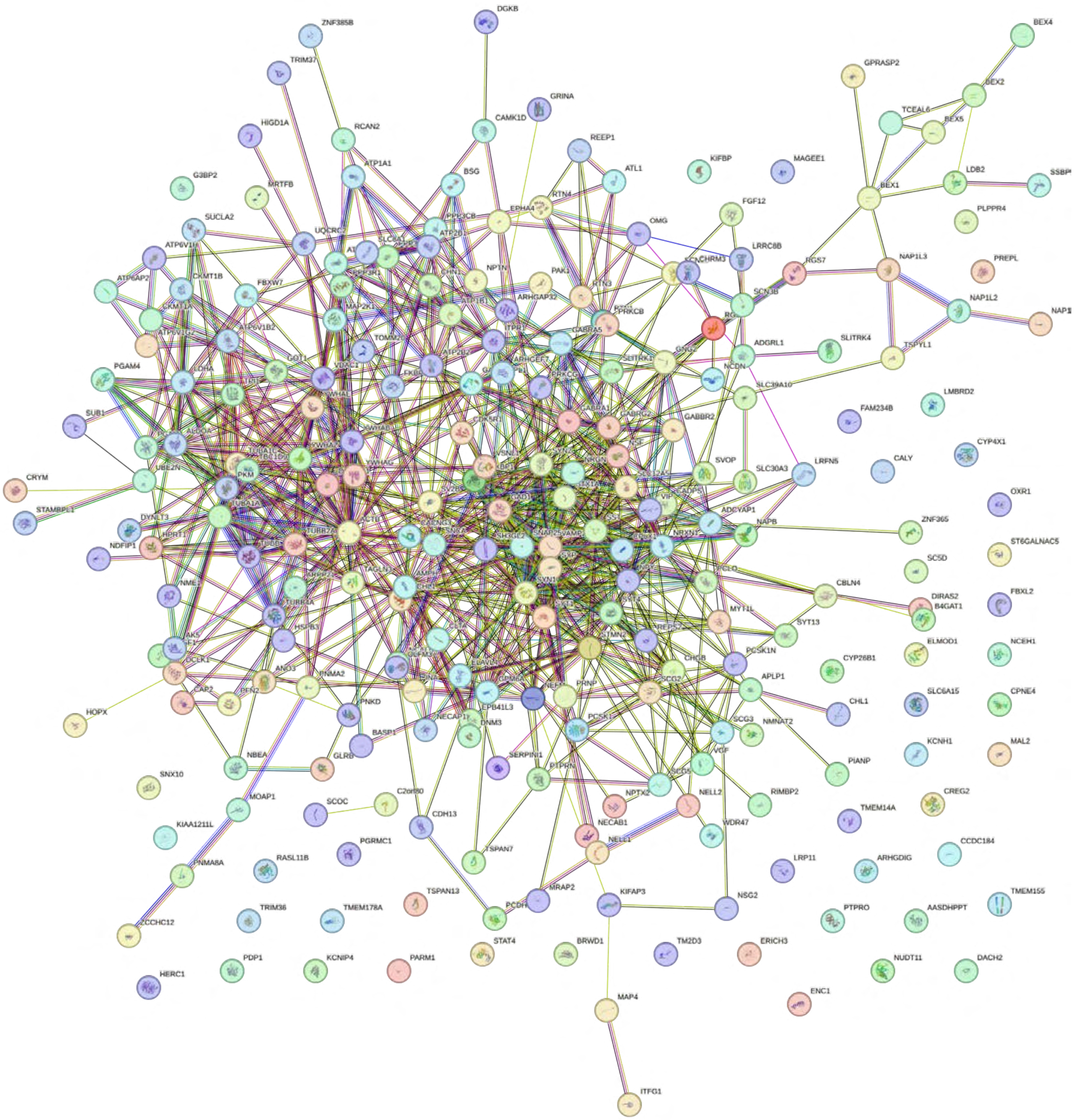

ZNF385B
DGKB
BEX4
TRIM37
GPRASP2
BEX2
GRINA
HIGD1A
RCAN2
CAMK1D
TCEAL6
BEX5
REEP1
KIFBP
MAGEE1
ATP1A1
ATL1
G3BP2
MRTFB
BSG
BEX1
LDB2
FGF12
PLPPR4
SUCLA2
EPHA4
RTN4
OMG
UQCRC2
LRRC8B
PREPL
NAP1L3
RGS7
NAP1L2
SLITRK4
ADGRL1
TSPYL1
PGAM4
SLC39A10
LMBRD2
FAM234B
SUB1
GABBR2
CYP4X1
SVOP
SLC30A3
LRFN5
CALY
CRYM
STAMBPL1
OXR1
ZNF365
ADCYAP1
NAPB
ST6GALNAC5
SC5D
NDFIP1
CBLN4
FBXL2
DIRAS2
B4GAT1
SYT13
ELMOD1
NCEH1
MYT1L
PCSK1N
CYP26B1
APLP1
SLC6A15
CPNE4
HOPX
CHL1
SCG3
NMNAT2
KCNH1
MAL2
NBEA
PIANP
SNX10
MOAP1
SCOC
C2orf80
CDH13
PTPRN
NPTX2
NELL2
RIMBP2
CREG2
TMEM14A
KIAA1211L
WDR47
NECAB1
CCDC184
PGRMC1
NELL1
RASL11B
PNMA8A
TSPAN7
LRP11
ARHGDIG
TSPAN13
MRAP2
KIFAP3
NSG2
TMEM155
TRIM36
TMEM178A
PTPRO
AASDHPPT
ZCCHC12
STAT4
BRWD1
TM2D3
ERICH3
PDP1
KCNIP4
PARM1
NUDT11
DACH2
HERC1
ENC1
MAP4
ITFG1
SERPINI1
BASP1
GLRB
DNM3
EPB41L3
NECAP1
PNKD
PNMA2
ANO3
CAP2
PFN2

FIGURE 3B

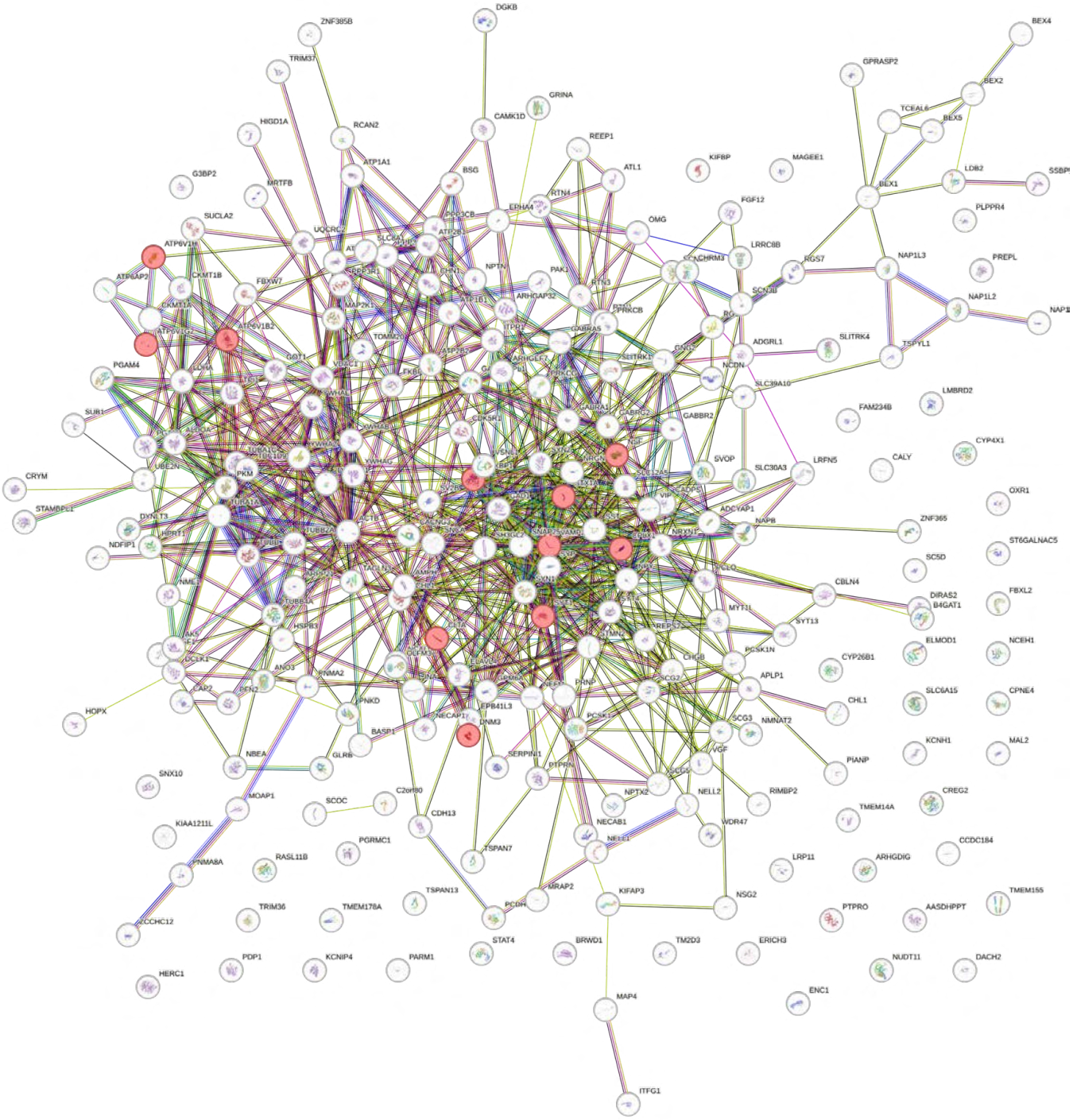


Figure 3. (a) String-db network of all 250 dysregulated genes. An image of the STRING network resulting from the top 250 dysregulated and statistically significant genes in the

GSE132903 dataset. The database indicated that there were over 843 interactions between proteins. Furthermore, the colors of the lines connecting nodes align with the type of interaction taking place, and a greater thickness in these lines also indicate a higher confidence in the protein's interactions. Having more lines come from a single node also demonstrates that the protein is part of a significant biological process. **(b) String-db network highlighting genes involved in the synaptic vesicle cycle** This string-db map highlights in red the proteins involved within the synaptic vesicle cycle that were input from GEO.

On analyzing the top 250 most dysregulated genes in AD patients on String-db, and specifying our organism as homo-sapiens, the database indicated over 843 interactions (edges) between proteins (nodes) (**Figure 3A**). There was an average node degree of 6.8 and an average clustering coefficient of 0.391, indicating moderate connectivity among proteins. The observed number of edges greatly exceeded the 240 interactions expected by chance, and the PPI enrichment p-value ($<1.0\times10^{-16}$) indicates that the proteins are significantly more interconnected than expected for a random set of genes, confirming shared biological functions or pathways (**Figure 3A**). Several of the key Gene Ontology (GO) processes that had many genes from this network were vesicle-mediated transport in synapse, chemical synaptic transmission, trans-synaptic signaling, synaptic vesicle cycling, synaptic signaling, and several other biological processes related to neuronal transportation, health, and communication. Each of these processes are crucially related to cognitive development. Dysregulation in these processes may be consistent with Alzheimer's disease pathology. Differential RNA analysis showed that the synaptic vesicle cycle (SVC) (hsa04721) had the most statistically significant downregulation (**Figure 3B**).

FIGURE 4

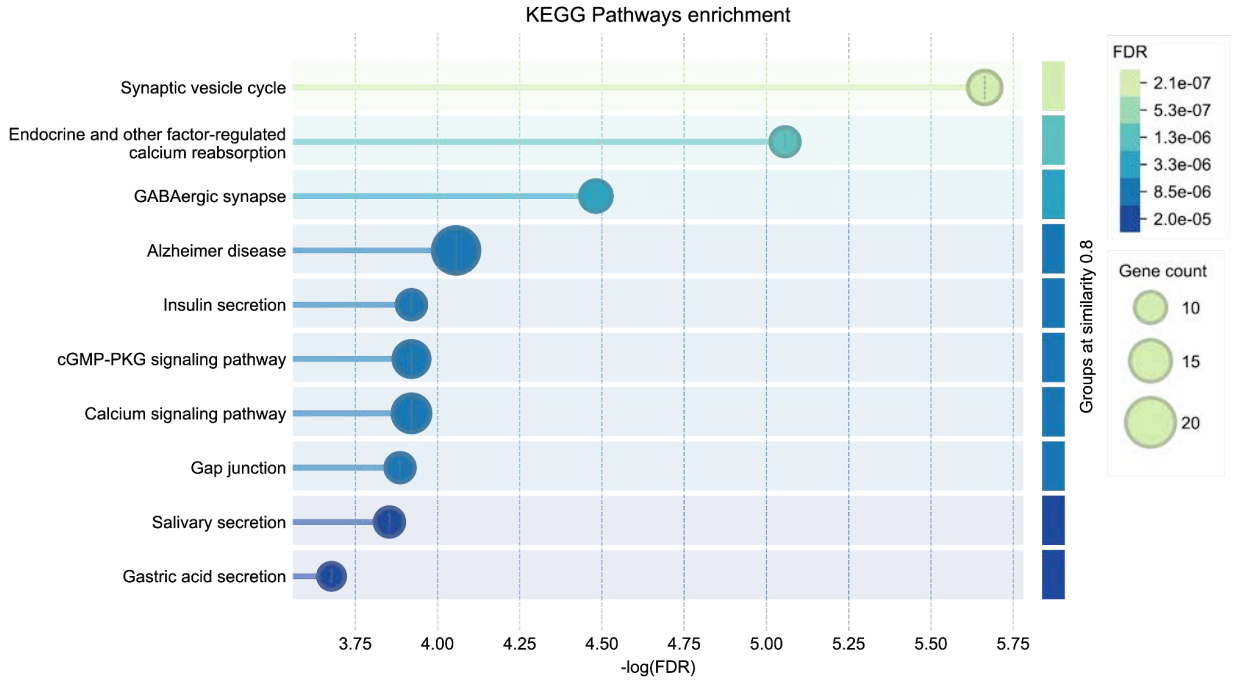


Figure 4. KEGG enrichment pathways found in String-db. The left side's y-axis indicates the top 10 KEGG pathways found to involve genes significantly dysregulated in AD through the GEO data. The right side clumps together different KEGG pathways that are related based on a similarity rate of 0.8 or higher, and the x-axis signifies the -log(FDR) of each pathway. The color of each pathway indicates the false discovery rate. Only the top 10 KEGG pathways were included and the pathways were sorted based on highest -log(FDR).

The synaptic vesicle cycle is the biological process in which synaptic vesicles are loaded with neurotransmitters, move to an active zone, and then undergo endocytosis before repeating the process again. The other top 10 KEGG pathways are shown below in order of false discovery rate (**Table 1**, **Figure 4**).

TABLE 1

term ID	term description	observed gene count	background gene count	strength	signal	false discovery rate	matching proteins in your network
hsa04721	Synaptic vesicle cycle	11	72	1.08E+00	1.29	2.17E-06	STX1A,CLTA,SYT1,ATP6V1B2,ATP6V1G2,CPLX1,DNM3,ATP6V1H,STXBP1,NSF,VAMP2
hsa04961	Endocrine & other factor-regulated calcium reabsorption	9	51	1.15E+00	1.2	8.78E-06	CLTA,PRKCG,DNM3,ATP2B2,ATP1B1,SLC8A1,ATP2B1,ATP1A1,PRKCB
hsa04727	GABAergic synapse	10	85	9.70E-01	1.01	3.30E-05	GABBR2,PRKCG,GNG2,GAD1,NSF,GABRA5,SLC12A5,GABRA1,GABARAPL1,PRKCB
hsa05010	Alzheimer disease	18	354	6.10E-01	0.76	8.78E-05	PPP3R1,CHRM3,TUBB4A,UQCRC2,MAP2K1,ITPR1,CDK5R1,RTN4,TUBB,RTN3,TUBB2A,PPP3CB,PPP3CA,VDAC1,TUBA1A,ATP2A2,TUBA1C,SNCA
hsa04020	Calcium signaling pathway	13	191	0.73	0.81	0.00012	PPP3R1,CHRM3,PRKCG,ITPR1,ATP2B2,PPP3CB,PPP3CA,VDAC1,SLC8A1,ATP2B1,ATP2A2,CAMK1D,PRKCB

hsa04911	Insulin secretion	9	82	0.94	0.9	0.00012	STX1A,CHRM3,PRKCG,PCLO,ATP1B1,VAMP2,ATP1A1,ADCYAP1,PRKCB
hsa04022	cGMP-PKG signaling pathway	12	163	0.77	0.82	0.00012	PPP3R1,MAP2K1,ITPR1,ATP2B2,ATP1B1,PPP3CB,PPP3CA,VDAC1,SLC8A1,ATP2B1,ATP2A2,ATP1A1
hsa04540	Gap junction	9	87	0.91	0.88	0.00013	PRKCG,TUBB4A,MAP2K1,ITPR1,TUBB,TUBB2A,TUBA1A,TUBA1C,PRKCB
hsa04970	Salivary secretion	9	89	0.9	0.87	0.00014	CHRM3,PRKCG,ITPR1,ATP2B2,ATP1B1,ATP2B1,VAMP2,ATP1A1,PRKCB
hsa04971	Gastric acid secretion	8	71	0.95	0.85	0.00021	CHRM3,PRKCG,SST,ITPR1,ATP1B1,ATP1A1,ACTB,PRKCB

Table 1. Table from String-db indicating the top 10 KEGG pathways. Top 10 KEGG pathways that involve many of the input genes from GEO2R that were found to be differentially regulated. The KEGG pathways are sorted based on lowest false discovery rate (FDR), with the synaptic vesicle cycle having the lowest FDR of 2.17E-06. Each row also indicates the dysregulated genes input from GEO2R involved in that pathway.

KEGG. To understand the physical molecular effects of these genetic changes, we used Kyoto Encyclopedia of Genes and Genomes (KEGG) to analyze the synaptic vesicle cycle biological process (hsa04721) (**Figure 5**). Out of the 125 total genes that comprise the human SVC pathway, 17 were present and highly dysregulated within our filtered dataset, yielding a high pathway enrichment strength of 1.03 and an overall network signal of 1.75 (indicating the harmonic mean between the observed/expected ratio and -log(FDR)). The SVC also had a false discovery rate (FDR) of 2.17×10^{-6} (**Table 1**), which is less than our p-value of 0.05. Genes identified as dysregulated in GEO2R and STRING are highlighted in yellow (**Figure 5**).

FIGURE 5

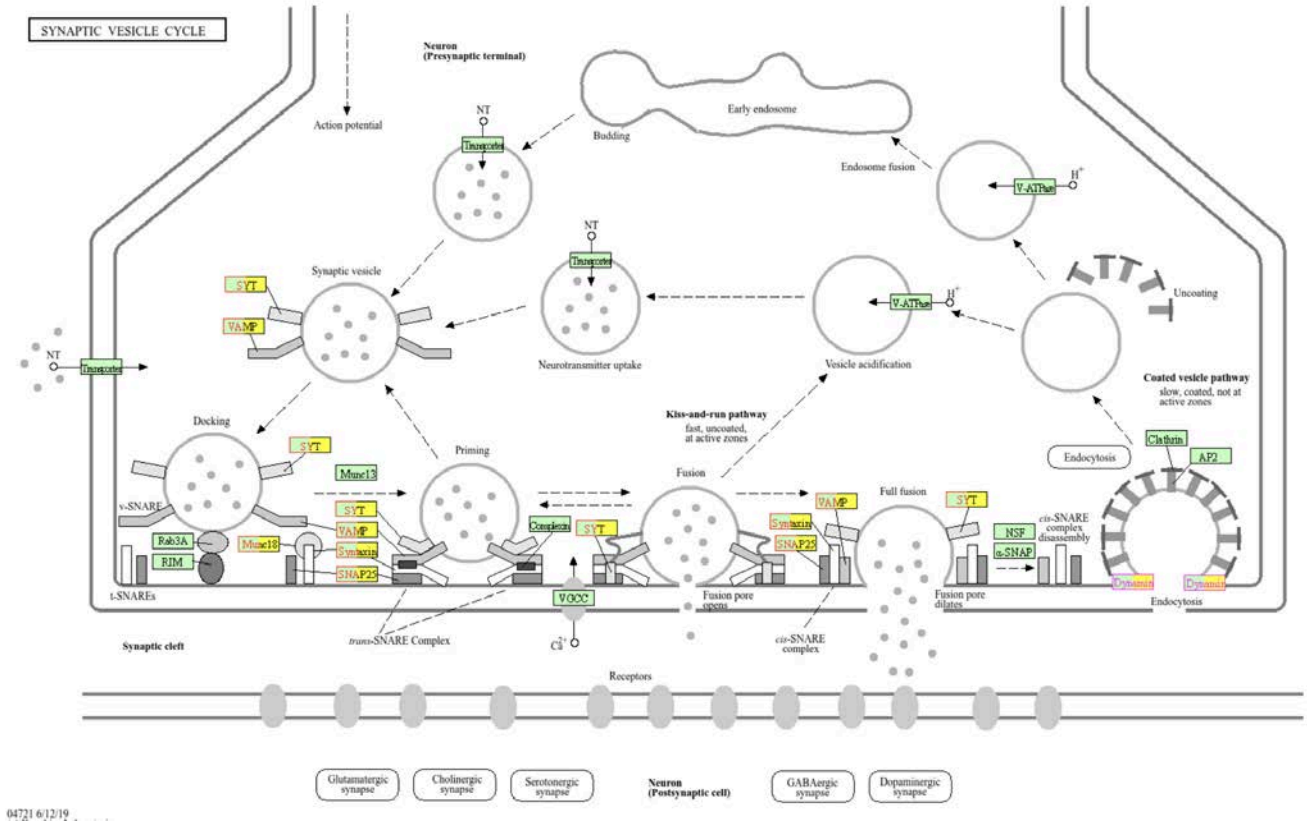


Figure 5. Synaptic Vesicle Cycle's KEGG pathway diagram. This image of the synaptic vesicle cycle's KEGG diagram highlights genes identified as dysregulated by String-DB and GEO2R. These dysregulated genes indirectly alter the state of future proteins that thus allow the vesicle to release from the neuron and exocytosis the neurotransmitters.

The SVC's typical function is to regulate the release and uptake of vesicles during synaptic signaling so that neurotransmitters can be released between neurons. The dysregulated proteins were found to cluster within critical stages of neurotransmitter loading, active zone docking, exocytosis, and subsequent endocytic recycling (Figure 5). The primary structural drivers within this pathway were all verified to be downregulated in the initial differential expression analysis, including DNM3, CLTA, VAMP2, SYT1, CPLX1, STX1A, STXBP1, NSF, ATP6V1B2, ATP6V1H, and ATP6V1G2 (**Table 1**). This suppression of these genes indicates cascading failure that disrupts presynaptic vesicle release and subsequent neurotransmitter exocytosis.

Heatmap. The heatmap provides a comprehensive view of gene expression differences between Alzheimer's disease (AD) and control (CTL) samples (**Figure 6**). The heatmap was generated from z-score normalized expression values, and displays distinct patterns of dysregulated genes between samples of Alzheimer's and control patients using the pheatmap()

function from the pheatmap package. Blue indicates lower expression in the corresponding gene for that patient while red indicates increased expression for that gene (upregulation). The differences and clusters of color highlight differences in gene expression profiles between the two groups. The tree diagram at the top of the heatmap shows that the samples segregate into two primary clusters, with AD (on the right) and control (on the left) samples grouping separately. This clustering indicates a substantial difference in gene expression profiles between the two groups. Most AD samples cluster together, demonstrating consistency in their gene expression patterns. A few samples show partial overlap between the groups, suggesting possible individual variability or intermediate gene expression profiles.

FIGURE 6

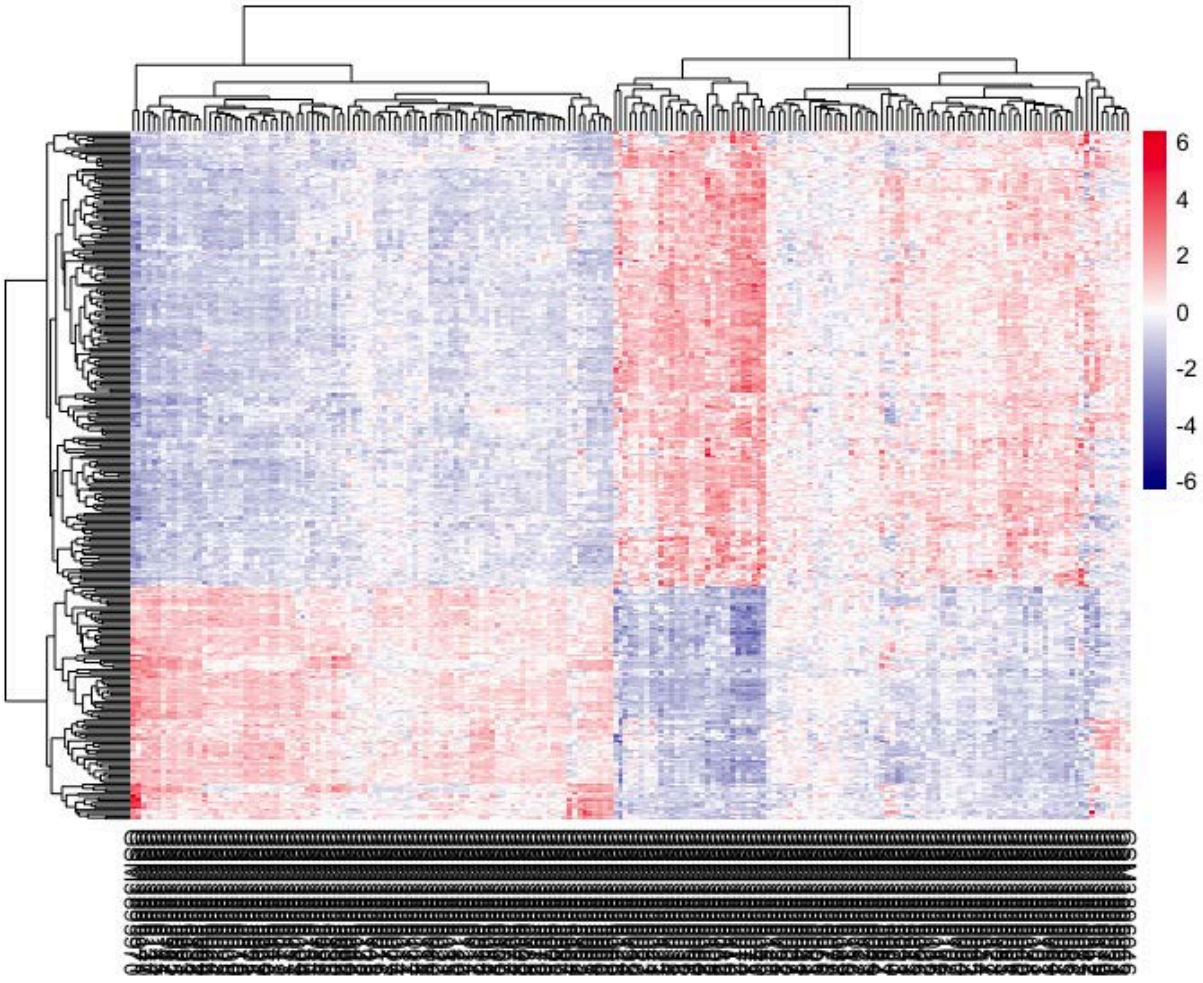


Figure 6. Heatmap comparing control samples to AD samples. This heatmap, produced using R Studio, shows the contrast and clustering of dysregulated genes between control samples and Alzheimer's disease samples. Alzheimer's samples indicated both strong upregulation (red clusters) and downregulation (blue clusters) of genes compared to normally expressed genes in control samples.

The heatmap only displays sample names and group annotations without individual gene labels since there are many genes. Additionally, clustering dendrograms help visualize relationships between samples and co-expression patterns among genes.

The heatmap reveals several distinct clusters of genes with differential expression between AD and control samples. A large cluster of genes shows consistent upregulation (red) in AD samples compared to controls, suggesting their potential involvement in pathways like apoptosis and the extracellular matrix (ECM), which both increase with AD (7). Another prominent cluster displays downregulation (blue) in AD samples, which may indicate genes involved in pathways like the synaptic vesicle cycle, which decreases neural signaling during disease progression (4). Additionally, some gene clusters show variable expression patterns across samples, reflecting individual variability or potential subtypes within the AD cohort. Genes with the highest differential expression (intense red or blue) are tightly clustered, indicating strong co-regulation and potential involvement in shared biological processes.

The clustering and distinct expression patterns observed in the heatmap confirm significant differences between Alzheimer's disease and control samples. The separated sample clusters and consistent gene expression patterns support the reliability of the data and identify potential biomarkers for Alzheimer's disease diagnosis and progression monitoring.

DISCUSSION

The results of this study highlight the prominent downregulation of the synaptic vesicle cycle (SVC), the importance of calcium homeostasis, and the role of the synaptotagmin-1 (SYT1) gene in Alzheimer's disease (AD) pathology. The findings of this study allow for a better understanding of the symptoms of AD and how cognitive decline is caused, in part, by impaired neurotransmitter release. By analyzing transcriptomic data via GEO and GEO2R, we visualized the dysregulation of approximately 17,525 genes between AD patients and healthy controls.

It is worth noting that the global transcriptomic profiles visualized as a UMAP demonstrated weak separation and overlapping clusters between the AD and control cohorts (**Figure 1**). This lack of distinct global clustering is common in bulk tissue transcriptomics of neurodegenerative disorders, often driven by cellular heterogeneity (e.g., varying ratios of neurons, microglia, and astrocytes) and the clinical spectrum of disease progression across patient samples. Despite this limited global variance, utilizing this dataset remains justified. While UMAP evaluates the entire transcriptome simultaneously, our later analysis relied on the statistical strength of GEO2R, which successfully isolated highly significant, differentially expressed genes. By filtering these specific differentially expressed genes through STRING-db and KEGG pathway analyses, we successfully extracted a robust, biologically meaningful signal from the background noise, specifically mapping a concerted collapse of the synaptic vesicle pathway.

Our study has several limitations. We relied only on gene expression data without confirming corresponding changes at the protein level. Since changes at the RNA level do not always directly translate to protein abundance or function, future studies should use protein assays or functional experiments to assess the consequence of gene expression changes. In addition, the samples were obtained from brains frozen at autopsy at the Banner Sun Health Research Institute Brain and Body Donation Program. There may be discrepancies in the DNA expression profiling from non-living tissue compared to living tissue.

Functional enrichment using String-db and KEGG pathway analysis revealed that a significant portion of the downregulated differentially expressed genes are very important in the

SVC, which governs the storage and release of neurotransmitters. These downregulated genes do not merely act in isolation; rather, they propagate a molecular chain reaction across the synaptic membrane (**Figure 5**). When these core structural proteins are suppressed, synaptic vesicles lose their capacity to efficiently dock, prime, and fuse with the presynaptic membrane during calcium-dependent exocytosis.

Our analysis identified several key components of this pathway that were significantly downregulated. The KEGG pathway analysis of the downregulated synaptic vesicle cycle highlighted how the downregulated genes interact. Furthermore, our String-db analysis (**Figure 4**) demonstrated that the genes within the network were connected to GO processes such as synaptic signaling, chemical synaptic transmission, and synaptic vesicle cycling. The analysis of both the KEGG pathway and of the STRING-DB network underscore the connections between the downregulated synaptic vesicle cycle, calcium regulation, and genes such as SYT1.

Several of the key downregulated genes that appeared in the synaptic vesicle cycle were SYT, VAMP, Syntaxin, SNAP25, and Dynamin, highlighted as yellow boxes (**Figure 5**). Below, each gene is defined in their function and how they are dysregulated in the synaptic vesicle cycle.

SYT (synaptotagmins) are a family of membrane proteins that detect Ca^{2+} to allow fast, synchronous neurotransmitter release. In our GEO2R analysis, we found that SYT1 is downregulated in AD (logFC = -1.42), meaning that there are less proteins expressed and produced to detect the calcium that signals for neurotransmitter release, slowing down neuron signalling. This analysis aligns with published literature that explores the importance of targeting the synaptic vesicle cycle in order to treat AD patients.

VAMP (vesicle-associated membrane proteins), SNAP25, and syntaxin are membrane proteins that mediate vesicle fusion with target membranes by forming a complex together. This complex brings together the vesicles and the plasma membranes, which are then used for exocytosis of neurotransmitters. When these genes are downregulated (VAMP logFC = -0.982, SNAP25 logFC = -0.707, Syntaxin logFC = -1.26), the neuron is unable to efficiently expel the neurotransmitters at their normal level with exocytosis.

Dynamin is a protein that acts as a “scissor” that severs budding vesicles during endocytosis, a key role during synaptic signaling to release the neurotransmitters (5,6). Since it’s downregulated (logFC = -0.742), this prevents as many vesicles from being severed during exocytosis, disrupting the synaptic homeostasis.

These findings are consistent with prior studies that emphasize synaptic dysfunction as a central feature of AD pathology (8, 9). Similar to previous research demonstrating impaired synaptic signaling and vesicle trafficking, our results show that neuronal communication is disrupted by reduced vesicle fusion efficiency, impaired calcium sensing, and defective vesicle recycling (10). The disruption of the SVC directly impacts the trafficking of the amyloid precursor protein (APP), which resides within synaptic membranes. Under physiological conditions, APP is cleaved normally; however, when vesicle cycling stalls, Aß accumulates abnormally, forming toxic oligomers in the brain (5,6). This extracellular Aß accumulation triggers an aberrant intracellular calcium influx, which further destabilizes the calcium-sensing capabilities of proteins like SYT1. The resulting deficit in neurotransmitter release leads to progressive synaptic pruning and a breakdown in neural networks, ultimately manifesting as the cognitive decline and memory loss characteristic of AD (8, 9, 11, 12).

Given its profound downregulation and central role in calcium sensing, SYT1 represents a compelling focal point for AD research. While its distinct expression pattern highlights it as a

potential pathological biomarker, extracting brain tissue from live patients to measure transcript levels is clinically unfeasible and dangerous. Therefore, future translational research must explore less invasive diagnostic avenues, such as evaluating SYT1 protein levels or associated synaptic degradation fragments within patient cerebrospinal fluid (CSF) or blood-derived exosomes. Therapeutic strategies should move beyond targeting amyloid clearance alone, focusing instead on stabilizing downstream calcium homeostasis and restoring synaptic vesicle fusion efficiency. Integrating these transcriptomic findings with proteomic profiling and functional electrophysiological studies will be essential to fully map how these microscopic membrane disruptions translate into macroscopic cognitive failure.

MATERIALS AND METHODS

GEO & GEO2R. Data was collected from the Gene Expression Omnibus (GEO), a database of genomic data, through the dataset GSE132903, which measures RNA expression from the middle temporal gyrus (5,6). This dataset was completed through a study by the Translational Genomic Research Institute and was last updated in January of 2024. The aim for this study was to identify specific genes, pathways, and molecular networks that contribute to AD pathology through expression profiling using microarray and by analyzing transcriptome changes in the middle temporal gyrus of AD patients. Findings were verified using qPCR and independent datasets. Braak stage, MMSE scores, and tangle density were also used to explore the relationships between gene expression and clinical measures.

Within this dataset, RNA expression of the middle temporal gyrus was analyzed from 97 AD patients and 98 non-disease controls. By using GEO2R, an application within GEO, differential gene expression was analyzed and visualized in a volcano plot (**Figure 1**) and UMAP (**Figure 2**). In the volcano plots, the fold change was set as the x-axis and the p-value was set as the y-axis. The red and blue dots in the volcano plot indicate a significantly dysregulated gene (adjusted p-value < 0.05) between Alzheimer's patients and control patients, with blue indicating downregulation and red indicating upregulation in AD patients compared to controls (**Figure 1**).

String-db. The top 250 statistically significant genes that were identified as dysregulated using GEO and GEO2R were input into STRING, a database of known and predicted protein-protein interactions. The results were filtered to only include homo-sapiens databases. Each circle, or node, represents a protein. The different colored lines connecting nodes indicate different types of interactions between the relevant nodes, such as how the interactions were determined or how they may interact with each other. A greater thickness of lines indicates a higher confidence in the proteins' interactions. Greater thickness of lines from each node also indicates the protein has greater connectivity with other proteins and may be part of a significant biological process.

KEGG. The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to visualize the interaction of proteins in interested pathways. Specifically, we used the KEGG pathway for the synaptic vesicle cycle (hsa04721). Using the search box, we found dysregulated genes (**Table 1**) and highlighted them as yellow to indicate those from String-db involved within that pathway (**Figure 5**).

R. R was used to generate a hierarchical clustering heatmap (**Figure 6**) to visualize the differential gene expression patterns between Alzheimer's (AD) and control samples. RStudio's

pheatmap, methods, stats, and utils package were used for analysis. Gene expression was obtained from the GEO dataset (file GSE132903.top.table.tsv) and loaded using the read function. To prepare the data, we removed non-numeric columns and only kept the gene expression values (adj.P.val, p.value, t, B, and logFC). The ID column represents the gene identifiers and was converted into row names using "row.names(cid)=colnames(data_CTL_AD)." We assigned the AD group as AD and control group as CTL. The code used can be found in Table S1 of the Appendix.

REFERENCES

[1] Breijyeh, Zeinab and Rowaida Karaman. "Comprehensive Review on Alzheimer's Disease: Causes and Treatment." *Molecules*, vol. 25, no. 24, Dec. 2020, p. 5789, https://doi.org/10.3390/molecules25245789.

[2] Ovsepian, S. V., et al. "Synaptic Vesicle Cycle and Amyloid β: Biting the Hand That Feeds." *Alzheimer's & Dementia*, vol. 14, no. 4, Feb. 2018, pp. 502–513, https://doi.org/10.1016/j.jalz.2018.01.011.

[3] Fagiani, Francesca, et al. "Amyloid-β and Synaptic Vesicle Dynamics: A Cacophonic Orchestra." *Journal of Alzheimer's Disease*, vol. 72, no. 1, Sep. 2019, pp. 1–14, https://doi.org/10.3233/JAD-190771.

[4] Chang, Shuwen, et al. "Synaptotagmin-1 Drives Synchronous Ca2+-Triggered Fusion by C2B-Domain-Mediated Synaptic-Vesicle-Membrane Attachment." *Nature Neuroscience*, vol. 21, no. 1, Dec. 2017, pp. 33–40, https://doi.org/10.1038/s41593-017-0037-5.

[5] "GEO Accession Viewer: GSE132903." National Center for Biotechnology Information, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132903. Accessed 25 May 2026.

[6] Piras, Ignazio S., et al. "Transcriptome changes in the Alzheimer's disease middle temporal gyrus: Importance of RNA metabolism and mitochondria-associated membrane genes." *Journal of Alzheimer's Disease*, vol. 70, no. 3, 3 Aug. 2019, pp. 691–713, https://doi.org/10.3233/jad-181113.

[7] Kumari, Sneha, et al. "Apoptosis in Alzheimer's Disease: Insight into the Signaling Pathways and Therapeutic Avenues." *Apoptosis*, vol. 28, 26 Apr. 2023, https://doi.org/10.1007/s10495-023-01848-y.

[8] Yang, Ling, et al. "Therapeutic Role of Neural Stem Cells in Neurological Diseases." *Frontiers in Bioengineering and Biotechnology*, vol. 12, Mar. 2024, pp. 1329712, https://doi.org/10.3389/fbioe.2024.1329712.

[9] Öhrfelt, Annika, et al. "The Pre-Synaptic Vesicle Protein Synaptotagmin Is a Novel Biomarker for Alzheimer's Disease." *Alzheimer's Research & Therapy*, vol. 8, Oct. 2016, pp. 41, https://doi.org/10.1186/s13195-016-0208-8

[10] Subramanian, Jaichandar “How Do Amyloid Pathology and Aberrant Neuronal Activity Disrupt Plasticity and Memory in Alzheimer’s Disease?” *The Neuroscientist*, vol. 32, no. 2, Jan. 2026, pp. 112-125, https://doi.org/10.1177/10738584251414384.

[11] Wesson, Daniel W., et al. “Mechanisms of Neural and Behavioral Dysfunction in Alzheimer’s Disease.” *Molecular Neurobiology*, vol. 43, no. 3, Mar. 2011, pp. 163–179, https://doi.org/10.1007/s12035-011-8177-1.

[12] De Strooper, Bart and Zetterberg, Henrik. “Tracking the Turning Point in Alzheimer’s Disease.” *Science*, vol. 392, no. 6797, Apr. 2026, pp. 468–469, https://doi.org/10.1126/science.aeb6987.

SUPPLEMENTAL TABLE

```
data <- read.table(file="GSE132903_series_matrix.txt",row.names = 1,header = T,
comment.char = "!")

one <- data[,1]
summary(one)
boxplot(one)

# obtain sample ids in each group from GEO2R process
AD <- c("GSM3896032", "GSM3896033", "GSM3896034", "GSM3896035", "GSM3896036",
"GSM3896037", "GSM3896038", "GSM3896039", "GSM3896040", "GSM3896041",
"GSM3896042", "GSM3896043", "GSM3896044", "GSM3896045", "GSM3896046",
"GSM3896047", "GSM3896048", "GSM3896049", "GSM3896050", "GSM3896051",
"GSM3896052", "GSM3896053", "GSM3896054", "GSM3896055", "GSM3896056",
"GSM3896057", "GSM3896058", "GSM3896059", "GSM3896060", "GSM3896061",
"GSM3896062", "GSM3896063", "GSM3896064", "GSM3896065", "GSM3896066",
"GSM3896067", "GSM3896068", "GSM3896069", "GSM3896070", "GSM3896071",
"GSM3896072", "GSM3896073", "GSM3896074", "GSM3896075", "GSM3896076",
"GSM3896077", "GSM3896078", "GSM3896079", "GSM3896080", "GSM3896081",
"GSM3896082", "GSM3896083", "GSM3896084", "GSM3896085", "GSM3896086",
"GSM3896087", "GSM3896088", "GSM3896089", "GSM3896090", "GSM3896091",
"GSM3896092", "GSM3896093", "GSM3896094", "GSM3896095", "GSM3896096",
"GSM3896097", "GSM3896098", "GSM3896099", "GSM3896100", "GSM3896101",
"GSM3896102", "GSM3896103", "GSM3896104", "GSM3896105", "GSM3896106",
"GSM3896107", "GSM3896108", "GSM3896109", "GSM3896110", "GSM3896111",
"GSM3896116", "GSM3896117", "GSM3896118", "GSM3896119", "GSM3896120",
 "GSM3896121", "GSM3896122", "GSM3896123", "GSM3896124", "GSM3896125",
"GSM3896127", "GSM3896128", "GSM3896129", "GSM3896130", "GSM3896131",
"GSM3896132", "GSM3896133")

CTL <- c("GSM3895951", "GSM3895952", "GSM3895953", "GSM3895954", "GSM3895955",
"GSM3895956", "GSM3895957", "GSM3895958", "GSM3895959", "GSM3895960",
"GSM3895961", "GSM3895962", "GSM3895963", "GSM3895964", "GSM3895965",
"GSM3895966", "GSM3895967", "GSM3895968", "GSM3895969", "GSM3895970",
"GSM3895971", "GSM3895972", "GSM3895973", "GSM3895974", "GSM3895975",
"GSM3895976", "GSM3895977", "GSM3895978", "GSM3895979", "GSM3895980",
"GSM3895981", "GSM3895982", "GSM3895983", "GSM3895984", "GSM3895985",
"GSM3895986", "GSM3895987", "GSM3895988", "GSM3895989", "GSM3895990",
"GSM3895991", "GSM3895992", "GSM3895993", "GSM3895994", "GSM3895995",
"GSM3895996", "GSM3895997", "GSM3895998", "GSM3895999", "GSM3896000",
"GSM3896001", "GSM3896002", "GSM3896003", "GSM3896004", "GSM3896005",
"GSM3896006", "GSM3896007", "GSM3896008", "GSM3896009", "GSM3896010",
"GSM3896011", "GSM3896012", "GSM3896013", "GSM3896014", "GSM3896015",
"GSM3896016", "GSM3896017", "GSM3896018", "GSM3896019", "GSM3896020",
```

```
"GSM3896021", "GSM3896022", "GSM3896023", "GSM3896024", "GSM3896025",
"GSM3896026", "GSM3896027", "GSM3896028", "GSM3896029", "GSM3896030",
"GSM3896031", "GSM3896112", "GSM3896113", "GSM3896114", "GSM3896115",
"GSM3896126", "GSM3896134", "GSM3896135", "GSM3896136", "GSM3896137",
"GSM3896138", "GSM3896139", "GSM3896140", "GSM3896141", "GSM3896142",
"GSM3896143", "GSM3896144", "GSM3896145")

# Combine both control and AD
data_CTL_AD <- data[,c(CTL,AD)]

# read GEO2R results
# top nrow genes: read.table for nrows rows
fulltopdiff <- read.table(file="GSE132903.top.table.tsv",header = T, sep = "\t", nrows=300) #top
300 genes

# fulltopdiff$ID is the column with column name ID
# select data only with the fulltopdiff$ID
selecteddata <- data_CTL_AD[fulltopdiff$ID,]

install.packages('pheatmap')
pheatmap(selecteddata, scale="row",
        color=colorRampPalette(c("darkblue", "white", "red2"))(256),
        border_color = NA,
        show_rownames = F)

# column id cluster
group <- as.factor(c(rep("CTL", 21),rep("AD", 21)))
cid <- data.frame(group)
row.names(cid)=colnames(data_CTL_AD)

pheatmap(selecteddata,scale="row",
        color=colorRampPalette(c("darkblue", "white", "red2"))(256),
        border_color = NA,
        show_rownames = F,
        annotation_col = cid)
```

Table S1. Code used in RStudio to produce the pheatmap (Figure 6). Code was written in R using local files downloaded from GEO2R (eg. GSE132903_series_matrix.txt). The sample ids were found in GEO when selecting the group samples. Packages used include pheatmap, methods, stats, and utils packages.