A project report submitted on

Gene expression analysis for brain derived neurotrophic factors (BDNF) and their role in Alzheimer's disease

Under the supervision of

Dr. Tiratha Raj Singh

And

Dr. Saurabh Srivastava

Submitted by

Saishta Shree (161511)

Bachelors Of Technology In Bioinformatics

Department Of Biotechnology And Bioinformatics,

Jaypee University Of Information Technology,

Waknaghat, Solan 173234, Himachal Pradesh, India

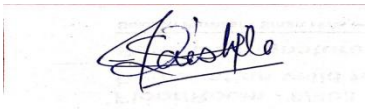
ACKNOWLEDGEMENT

It is our privilege to express our sincerest regards to **Dr. Tiratha Raj Singh and Dr. Saurabh Srivastava**, for his valuable inputs, able guidance, encouragement, wholehearted cooperation, constructive criticism, helpful information, practical advice and unceasing ideas which have helped me tremendously at all time.

We deeply express our sincere thanks to our Head of Department Prof. Dr. Sudhir Kumar for encouraging and allowing us to present the project on the topic "**Gene Expression Analysis For Brain Derived Neurotrophic Factors (Bdnf) And Their Role In Alzheimer's Disease"** at our department premises for the partial fulfilment of the requirements leading to the award of B. Tech degree.

DECLARATION

I hereby declare that the project report entitled "**Gene Expression Analysis For Brain-Derived Neurotrophic Factors (BDNF) And Their Role In Alzheimer's Disease"** submitted at "Jaypee University of Information Technology, Waknaghat, India" is an authentic record of my work carried out under the supervision of Dr.Tiratha Raj Singh. I have not submitted this work elsewhere for any other degree or diploma.



(Signature of the Student)

Saishta Shree (161511)

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Dr. Tiratha Raj Singh

Associate Professor

Department of Biotechnology and Bioinformatics,

Jaypee University of Information Technology, Waknaghat, India

Date:

CERTIFICATE

This is to certify that project report entitled "**Gene Expression Analysis For Brain-Derived Neurotrophic Factors (BDNF) And Their Role In Alzheimer's Disease**", submitted by Ms. **Saishta Shree** is in its partial fulfillment for the award of degree of Bachelor of Technology in Bioinformatics to Jaypee University of Information Technology Waknaghat, Solan (H.P.), India is an authentic record of candidate's work carried out by her under my supervision.

This work has not been submitted partially or fully to any other university or institution to achieve any award or any other degree.

Dr. Tiratha Raj Singh

Associate Professor,

Department of Biotechnology and Bioinformatics,

Jaypee University of Information Technology,

Waknaghat, Distt. Solan (173234), Himachal Pradesh, India

Dr. Saurabh Srivastava

Assistant Professor,

Department of Mathematics,

Jaypee University of Information Technology,

Waknaghat, Distt. Solan (173234), Himachal Pradesh, India

TABLE OF CONTENTS

ABSTRACT

Diminished articulation of mind determined neurotrophic thing (BDNF) has a critical limit in the prognosis of Alzheimer's issue (AD), that is portrayed by means of the development of plaques including Abeta and neurofibrillary tangles made out of hyperphosphorylated tau protein. A developing collection of proof shows a capacity defensive impact of BDNF towards Aβ-hastened neurotoxicity in AD, However, the immediate remedial effect of BDNF inserted on tauopathy in AD stays to be mounted. In this report, we saw that the BDNF degree was diminished in the immunizer and cerebrum of AD patients. Here we show tremendous neuroprotective results of entorhinal BDNF the board in creature models of Alzheimer's illness, with the augmentation of helpful gifts into the declining hippocampus. BDNF quality transportation while managed after infection beginning in amyloid-transgenic mice, turns around the neural connection misfortune along with incomplete standardization of atypical quality articulation, hence, upgrading cell flagging and recuperating of picking up information on and memory. These outcomes happen freely of results on amyloid plaque load. In matured rodents, BDNF implantation switches subjective decay, improves age-related bothers in quality articulation and reestablishes cell flagging. In adult rodents and primates, BDNF forestalls injury instigated death toll of entorhinal cortical neurons. In matured primates, BDNF turns around neuronal decay and improves age-related intellectual impedance. All things considered, those discoveries demonstrate that BDNF applies broad ensuring results on fundamental neuronal hardware worried in Alzheimer's sickness, performing through amyloid-autonomous systems. In this manner, making BDNF a prevalent detail in potential solution for Alzheimer's issue.

1 INTRODUCTION

Alzheimer's issue (AD) is a neurodegenerative illness that reasons degeneration, or misfortune, of neurons inside the cerebrum, for the most part in the wide region of the cerebral cortex and hippocampus. The distortion is first recognized in the psyche tissue that incorporates the frontal and transient projections, and afterward gradually progress to different regions of the neocortex (as appeared in Figure 1). Alzheimer's infection is related with the assortment of insoluble sorts of amyloid-β (Aβ) in plaques in extracellular territories, notwithstanding inside the parcels of veins, and accumulation of the microtubule protein tau in neurofibrillary tangles in neurons. Aβ is determined with the guide of the proteolytic cleavage of amyloid forerunner protein (APP) with the guide of a convoluted hover of family members of catalysts (γ-secretases and β-secretases), which comprise of presenilin 1 (PS1; encoded by methods for PSEN1) and PS2 (encoded by utilizing PSEN2).

The normal length of pollution is 8–10 years, however the clinical suggestive stages are gone before by means of preclini-cal and prodromal degrees that usually enhance over numerous years. Irregular Alzheimer's infection is the most extreme com-mon type and has a middle period of beginning of 80 years. The key thought process is the inability to clear Aβ peptide from the cerebrum tissue. Be that as it may, co-morbidities alongside cerebro-vascular scatter and hippocampal sclerosis are visit at this age, which entangles examination and control-ment. An own family ancestry of influenced close to companion and kids isn't surprising in irregular illness, be that as it may, a little extent (<1%) of victims have autosomal prevailing acquired Alzheimer's issue (DIAD); this shape has an early time of beginning (mean time of ~45 years). In this subgroup, patho-genetic transformations in the qualities encoding APP, PS1, and PS2 are found, which thought process overproduction or arrangement of an atypical state of Aβ. In most extreme clinical regards, the irregular and familial kinds of Alzheimer's sickness are practically identical, which incorporate the charge of infirmity movement and biomarker profiles [1].

As a confusion element, Alzheimer's affliction stocks numerous attributes with various molecularly characterized neuro-degenerative maladies, comprehensive of Parkinson's issue and the frontotemporal dementias [2]. One would conceivably, hence, question whether Alzheimer's infection is an inescapable a piece of ordinary maturing or whether it is a discrete issue framework [3].

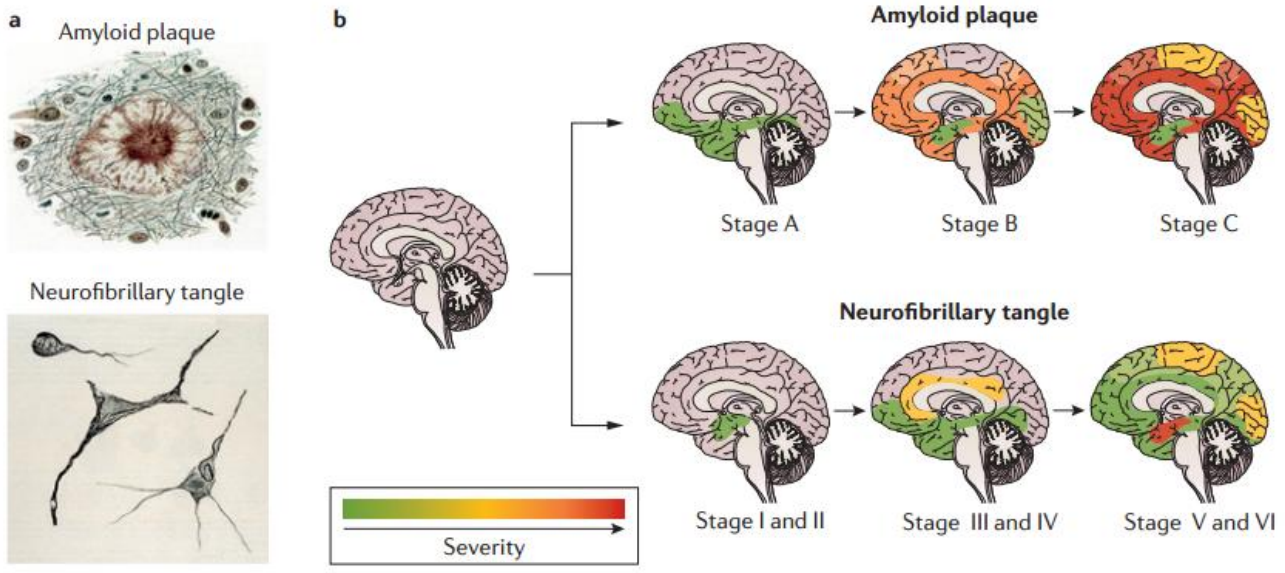


Figure 1:AD, (Aβ) sworn explanation goes before neurofibrillary and neuritic conflict joined by an unmistakable explanation in the facade and brief projections, hippocampus and limbic device (pinnacle line). Little regularly, distress appears to rise out of different districts of the cerebral neocortex with near saving of the hippocampus. Tangles and neuritic decrease start inside ordinary transient projections and hippocampus, and grade by grade grow to various areas of the neocortex (back line). Nearness of atomic imaging procedures for Aβ and tau, broad dispersal of depressed person modifications will get managable to real ‑time in vivo assessment and could now not be dependent upon analyzation ages as depicted here.

1.1 How does Alzheimer's disease impact the cerebrum?

The mind ordinarily therapists somewhat in energizing getting old be that as it may, hugely, does now no more lose nerve cell in tremendous not numeral. While in AD the mischief is tremendous, the indistinguishable numeral of nerve cell hinder working, there's an absence of associations with different neurons, and bite the dust. Alzheimer's upset procedures basic to to neurocyte and their structure, along with discussion, assimilation, and fix. From the begining, AD regularly demolishes neurons and their relationship in elements of the brain worried in recollection, which incorporates the entothorax and limbic brain. It later effects area inside the cerebral cortex at risk for decision making, language, thinking, and direct. At long last, a wide scope of districts of the cerebrum are hurt. After

some time, a person with AD a tiny bit at a time loses his convenience to live and trademark independently. Over the long haul, the ailment initiated death toll (as shown in Figure 2).

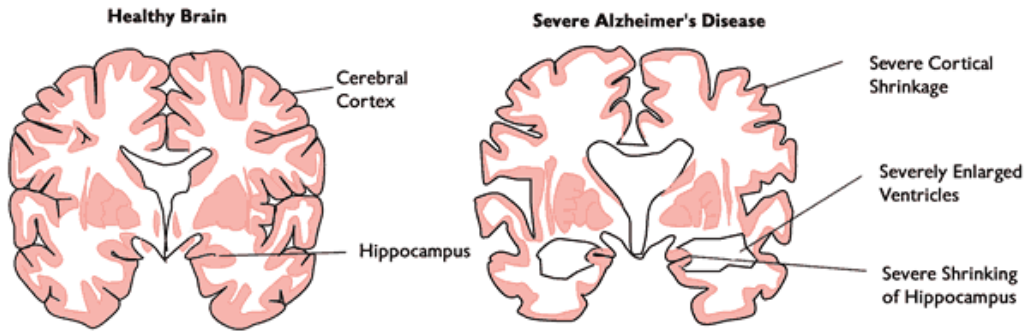


Figure 2. Examination of an ordinary matured cerebrum (left) and the mind of an individual with Alzheimer's (correct). Attributes that different the two are brought up.

Advertisement is portrayed by mental decay and loss of neurons in unequivocal cerebrum locale. Late disclosures have proposed a consideration of psyche construed BDNF in the focalisation of AD. BDNF is an exogenous protein connected with the upkeep of neuronal cutoff, synaptic flexibility and assistant uprightness in the grown-up mind.

1.2 Brain Derived Neurotrophic Factor

BDNF plays a vital trademark in directing picking up information on and memory. BDNF is a neurotrophin that has a place with an own group of proteins that advance the endurance, gifts, and improvement of neurons. The declaration of the BDNF quality might be chosen inside the cortex, hippocampus, and basal forebrain districts which may be critical for memory, picking up information on, and higher intellectual trademark. BDNF improves neurogenesis and neurotransmission all through the neurotransmitters, advances synaptic development, and balances synaptic versatility. BDNF besides initiates hippocampal extensive timespan potentiation, it's fundamental for memory development. Weinstein et al. Discovered that better fringe BDNF degrees secure more seasoned grown-ups toward AD [4]. By having BDNF levels better by utilizing one far reaching deviation, the danger for AD or dementia got reduced with the guide of 33% [5].

1.2.1 Mechanism of action

BDNF ties with two receptors at the floor of cells which could answer to the expansion perspective, TrkB (expressed "Track B") and the LNGFR (Low-proclivity nerve blast factor receptor) otherwise called p75.

1.2.2 Tropomyosin receptor kinase B (TrkB)

TrkB is a receptor for BDNF. The TrkB receptor is encoded by method of the NTRK2 quality and TrkB is an individual from a receptor hover of family members of tyrosine kinases that comprises of TrkA and TrkC. TrkB autophosphorylation is reliant upon ligand-exact alliance with BDNF, a comprehensively communicated side interest subordinate hypochondriac thing that manages versatility and is unregulated after hypoxic harm. The actuation of the BDNF-TrkB pathway is basic in the improvement of brisk term memory and the blast of neurons.

1.2.3 Low-affinity nerve growth factor receptor (LNGFR)

The situation of the inverse BDNF and p75 receptors is less clear. The TrkB receptor connects with BDNF in a ligand-exact, and all neurotrophins interface with the p75 receptor. When p75 receptor is enacted, it brings about initiation of the NFkB receptor. LNGFR may likewise sign a cell to bite the dust by means of apoptosis as opposed to endurance pathways in cells communicating the p75 receptor without Trk receptors.

1.2.4 Common SNPs in the BDNF gene

BDNF has various known unmarried nucleotide polymorphisms (SNP) alongside, rs6265, rsC270T, rs7103411, rs2030324, rs2203877, rs2049045 and rs7124442. Starting at 2008, rs6265 is the most extreme remember SNP of the BDNF quality.

1.2.5 Val66Met

Val66Met is exact to human A piont transformation in the coding grouping is guanine to adenine which switches at work 196 as outcomes in an amino corrosive change to valine to methionine

substitute at codon sixty six, Val66Met, that is inside the prodomain of BDNF. Val66Met (a missense change on the codon sixty six) dissimilarity of the BDNF quality mastermind cognizance to AD and AD object show decreased mRNA and protein degrees of BDNF inside the serum and cerebrum in assessment with fit more established guideline [5].

We planned to identify the differentially expressed genes in BDNF using available datasets in GEO, which can further be used as potential biomarkers for AD. The dataset of microarray for differential expressed gene was used for our analysis.

1.2.6 Microarray

Microarray is a technique which is used to detect the expression of thousands of genes. Principle behind the DNA microarray is nucleic acid hybridization[2]. The differentially expression data generated in microarray analysis can be deposited in the freely accessible database such as Gene Expression Omnibus (GEO) [6].

1.2.7 Gene Expression Omnibus

Gene Expression Omnibus is a public repository for nucleotide sequence data obtained by DNA microarray and sequencing methods. Gene expression data related to BDNF AND AD was selected from GEO for our analysis. Biomarkers for bdnf are important owning to their clinical importance for accurate diagnosis. We studied differential gene expression in normal individuals and those suffering from Alzheimer's Diease AD. Peripheral tissue like blood was the best choice for the analysis as it is difficult to obtain brain samples for analysis.

The microarray analysis for the dataset identified for differentially expressed genes in BDNF AND AD was carried out using R language. The output of the analysis was further represented as a volcano plot of genes.

2 MATERIALS AND METHODS

The work flow of methodology for following analysis is given below (Figure 3).

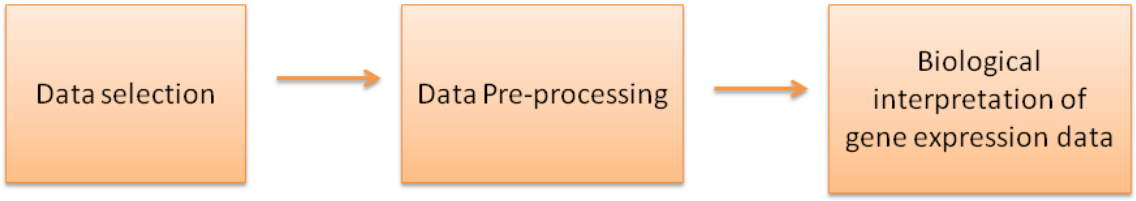


Figure 3. ***Methodology flow chart for microarray data analysis*****.**

2.1 Dataset selection

Microarray-based quality articulation information of subjects with "Advertisement and BDNF" were acquired from Gene Expression Omnibus (GSE). The qualified investigations were looked with the catchphrases "Alzheimer's Disease" and "BDNF". The creature discovered was Homo sapiens, Mus musculus and Rattus norvegicus cluster type as "Articulation profiling by exhibit". Subsequently, just Mus musculus was chosen as a delegate dataset for the investigation. Crude test level information (CEL records) that concentrated on quality articulation profiling in the limbic brain and entorhinal cortex gatherings of BDNF and AD rewarded ojects were gathered. Data on covariates, including age, genotype, treatment, and clump impact, was required for this investigation. The CEL records for GSE14522 were recovered and further broke down utilizing R Affy. Various examples were dissected utilizing various stages and just one of them was utilized.

The features of GSE14522 were as follows:

- Platform: GPL 1261 and GPL1355
- Number of sample : 53
- Sample groups: BDNF, GFP, Sham lesion
- PMID: 19198615

2.2 Data pre-processing

Data processing was performed using R programming language. The downloaded raw CEL files were loaded into different R package like Affy, Limma, Dplyr, Mouse4302.db, and Calibrate available on bioconductor. The details about these packages are as follows :

- Affy: Affy is an R package of functions and classes for the analysis of oligonucleotide array manufactured by Affymetrix. It allows the user to normalize the probe intensity data [7].
- Limma: Limma is a R group used for data assessment, direct models and differential explanation for microarray data [4].
- Dplyr: Dplyr is a R group used to switch and summarize data in plain association. The course of action of limits that grant data control like isolating for lines, picking express segments, re-mentioning lines and including new portions makes Dplyr capable for data assessment [5].
- Mouse4302.db: Mouse4302.db is an R package used for array annotation data assembled using data from public repositories [8].
- Calibrate: Adjustment is a R bundle utilized for drawing aligned scales with tick blemishes on (non-symmetrical) variable vectors in scatterplots and biplot. It likewise gives some capacity to multivariate examination like the chief facilitate investigation [3].

The expression data was read in R using 'ReadAffy' function, which extracted the data from CEL files. Boxplots were utilized to distinguish any exception tests that were in this manner expelled. The datasets were standardized utilizing the Robust Multi-cluster Average (RMA) work in the affy R bundle. After grouping of these samples, we considered only BDNF and GFP for the analysis of relative gene expression. A design matrix for the selected samples was created to fit the linear model by combining all levels as determined by the groups made using the "as.factor" function. After designing a matrix, a "Top250" table of probesets and expression values was created. Probesets were first mapped to Gene IDs utilizing the AnnotationDbi bundle to comment on documents. Probesets that mapped to various qualities were expelled, and for any qualities that mapped to numerous probesets, just the probeset that had the biggest total assessed impact size was kept.

2.3 Quality control

Quality control of microarray data begins with the visual examination of microarray images. The data analysis software packages can be used to make plots (for example of background signal, average intensity values and percentage of genes above background) to help identify arrays, reporters or samples.

2.4 Biological interpretation of gene expression data

Many of the strategies for visualisation and interpreting microarray data can also be used for RNA-seq experiments. Some common ways of visualising and deciphering gene expression facts.

3 RESULTS AND DISCUSSION

The holder plots in observe four and parent five are showing the microarray realities going before and after standardization, separately. In recognize four; the level dark follows speaking to the middle quality articulation cost for each example have been particular. In Figure five, the even dark follows speaking to the middle quality articulation cost show up on a straightforwardly line after which information after standardization were fit for additional assessment.

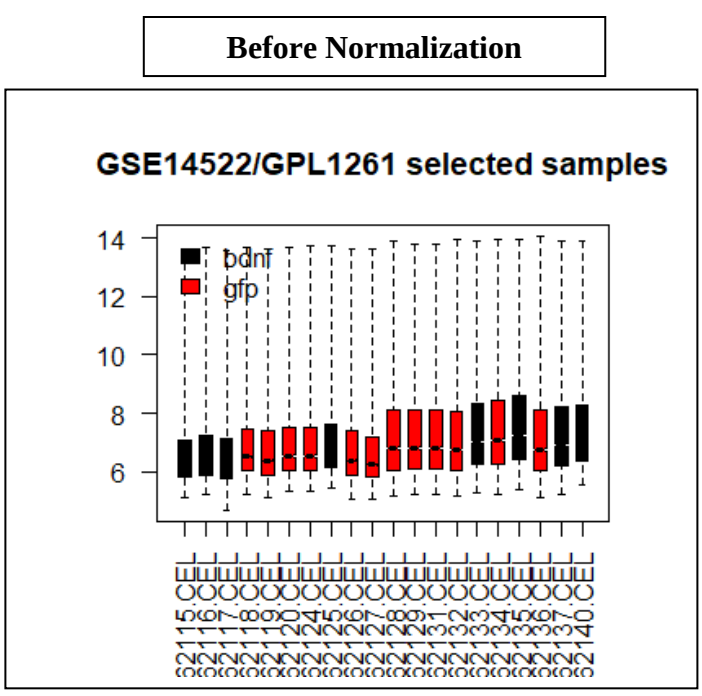


Figure 4: ***Box plot for the quality articulation information before standardization. Articulation esteems were resolved utilizing the Affy bundle in R programming. The dark bar demonstrates the middle worth.***

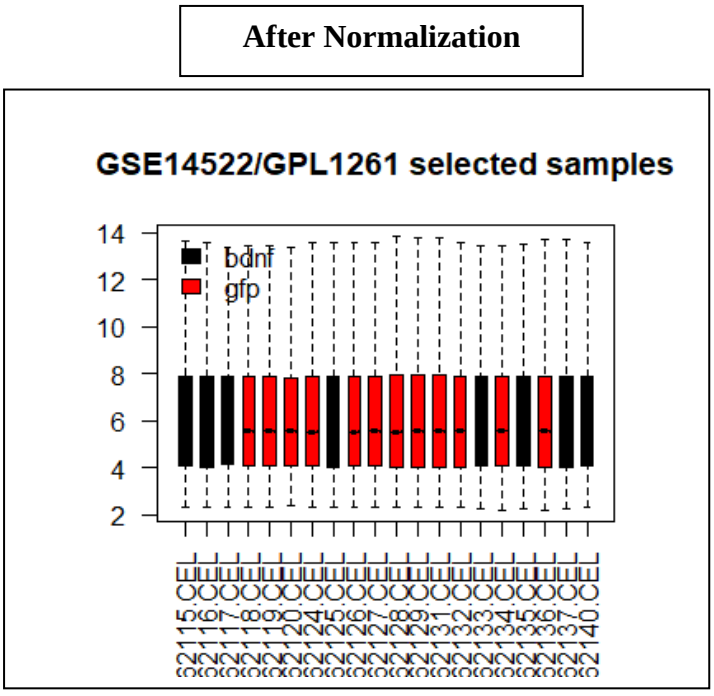


Figure 5: ***Box plot for the quality articulation information following standardization. Articulation esteems were resolved utilizing the Affy bundle in R programming, trailed by standardization utilizing the hearty multiarray normal calculation The dark bar shows the middle worth.***

3.1.2 Retrieved gene symbols associated with probe Ids.

In figure 6, we extracted the gene symbols using Mouse4302.db package in R software corresponding to probe Ids from which we listed the top 20 genes out of 250.

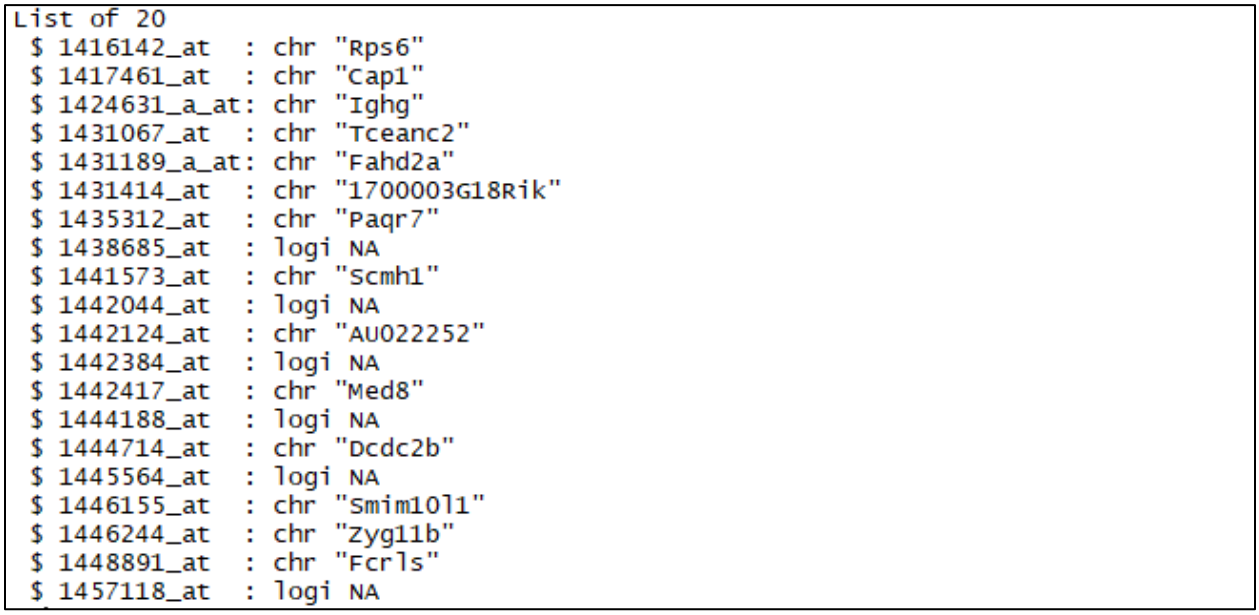


Figure 6. ***List of map probe attributes with there associated gene symbols were extracted using Mouse4302.db package in R software.***

3.1.3 Addition of gene symbols in the original gene expression dataset

In figure 7, we added annotation data to gene expression dataset for our further analysis.

	SYMBOL	ID	logFC	AveExpr	t	P.Value	adj.P.Val	B
1416142_at	Rps6	1416142_at	-0.7355592	7.776196	-6.798351	9.603225e-07	0.02165575	1.6186072
1417461_at	Cap1	1417461_at	2.6467248	10.167493	4.340459	2.826996e-04	0.55663515	-0.9737364
1424631_a_at	Ighg	1424631_a_at	3.3369746	7.098682	4.965158	6.347454e-05	0.27139201	-0.2274303
1431067_at	Tceanc2	1431067_at	-1.0732151	5.247144	-4.864221	8.069952e-05	0.30330243	-0.3444039
1431189_a_at	Fahd2a	1431189_a_at	0.2458326	6.492271	4.597841	1.524826e-04	0.42981978	-0.6600074
1431414_at	1700003G18Rik	1431414_at	-0.2935089	3.269198	-5.316777	2.766032e-05	0.15593853	0.1681192
1435312_at	Paqr7	1435312_at	-0.8990398	6.669100	-4.947517	6.619171e-05	0.27139201	-0.2477675
1438685_at	<NA>	1438685_at	-0.3621151	8.496603	-5.622828	1.354158e-05	0.12214773	0.4966798
1441573_at	Scmh1	1441573_at	-0.6572601	4.143303	-4.707607	1.172645e-04	0.39446977	-0.5287797
1442044_at	<NA>	1442044_at	-0.9749295	5.380354	-5.759518	9.872939e-06	0.11131986	0.6385696
1442124_at	AU022252	1442124_at	-0.9326505	6.841191	-5.353056	2.540320e-05	0.15593853	0.2078433
1442384_at	<NA>	1442384_at	-1.2120372	4.362384	-6.813604	9.289543e-07	0.02165575	1.6317264
1442417_at	Med8	1442417_at	-0.4166520	5.096730	-5.040780	5.304826e-05	0.26583663	-0.1407758
1444188_at	<NA>	1444188_at	-0.4870005	3.819511	-4.377560	2.586068e-04	0.55663515	-0.9280212
1444714_at	Dcdc2b	1444714_at	-0.2943911	4.515978	-5.823287	8.525602e-06	0.11131986	0.7037314
1445564_at	<NA>	1445564_at	-1.0616608	4.258858	-4.437908	2.237356e-04	0.55663515	-0.8540039
1446155_at	Smim10l1	1446155_at	-0.6340721	3.527299	-4.632357	1.403886e-04	0.42211100	-0.6185689
1446244_at	Zyg11b	1446244_at	-0.9539897	8.084431	-4.689509	1.224491e-04	0.39446977	-0.5503044
1448891_at	Fcrls	1448891_at	-1.0439134	7.222387	-5.501355	1.796038e-05	0.13500521	0.3680605
1457118_at	<NA>	1457118_at	0.3207445	6.252312	4.402598	2.435213e-04	0.55663515	-0.8972589

Figure 7: ***List of Added Annotation data to differential gene expression results like(ID, logFC, AveExpr, t-value, P.Value, adj. P.Val, B, etc)***

3.1.4 Created a volcano plot highlighting significant genes

A spring of gushing lava plot is such a disperse plot that suggests statistical significance (P-value) as opposed to the magnitude of exchange (fold exchange). It allows short visible identity of genes with large fold changes that are additionally statistically enormous. These genes can be the maximum biologically giant . In a volcano plot, the maximum upregulated genes are closer to the proper, the maximum downregulated are closer to the left. The genes are coloured in the event that they bypass the thresholds log fold change, blue are upregulated and orange are downregulated genes (Figure 8).

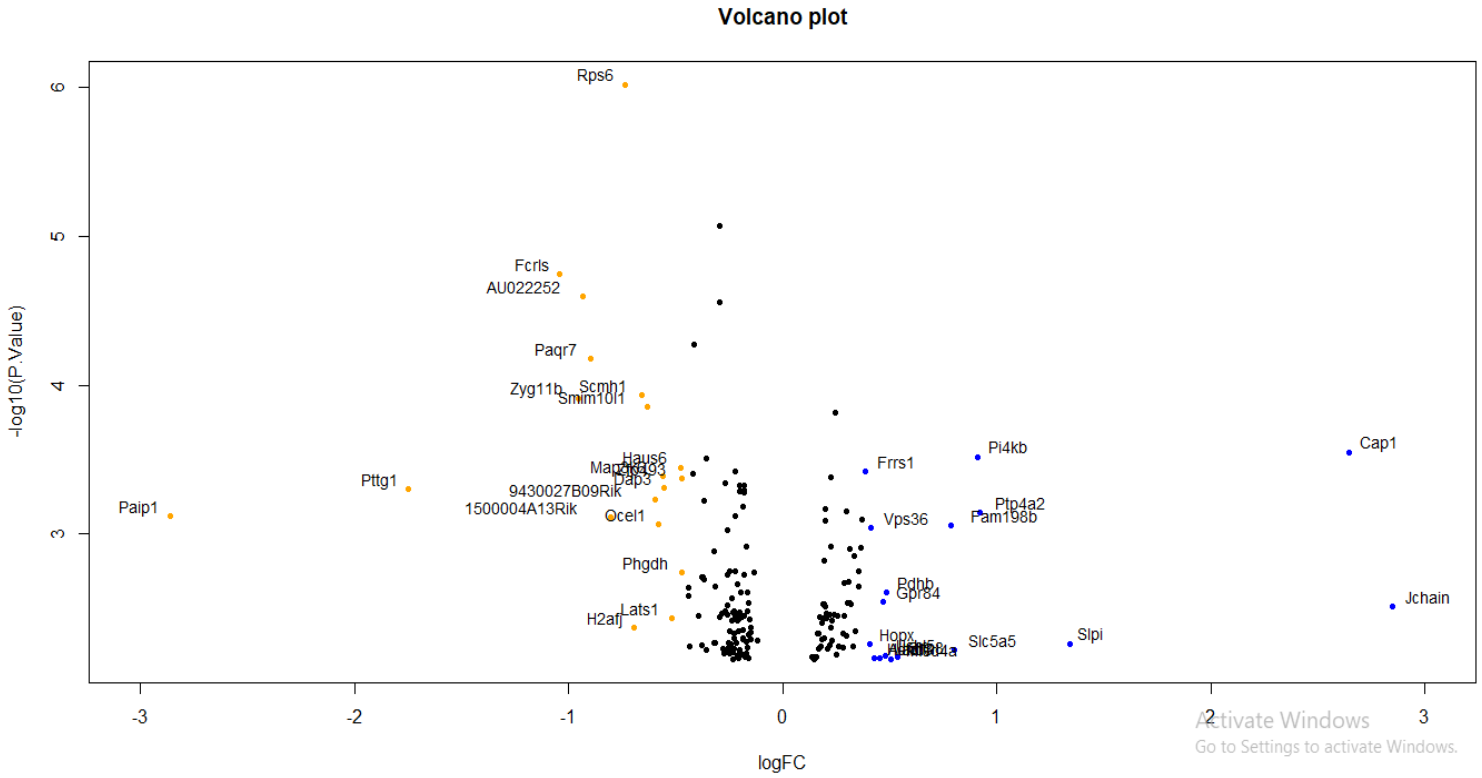


Figure 8 ***Volcano plot highlighting significant genes.***

Table 1: Rundown of up-controlled quality with their Uniprot Accession, Function, Cellular Component.

Gene Symbol	Full name	Uniprot Accession	Function
Cap1	Adenylyl cyclase-associated protein	P40124	Directly regulates filament dynamics
Jchain	Immunoglobulin J chain	P01592	link two monomer units of either IgM or IgA.
Pi4kb	Phosphatidylinositol 4-kinase beta	E9Q8A3	regulate Golgi disintegration/reorganization during mitosis, possibly via its phosphorylation
Ptp4a2	TYR_PHOSPHATASE_2 domain-containing protein	Q3UXF9	Protein tyrosine phosphatase which stimulates progression from G1 into S phase during mitosis
Pam148b	No data available	No data available	No data available
Slpi	Antileukoproteinase	P97430	Acid-stable proteinase inhibitor with strong affinities for trypsin, chymotrypsin, elastase, and cathepsin
Sic5a5	No data available	No data available	No data available

Frrs1	Ferric-chelate reductase 1	Q8K385	Ferric-chelate reductases reduce Fe3+ to Fe2+ before its transport from the endosome to the cytoplasm.
Vps36	Vacuolar protein-sorting-associated protein 36	Q91XD6	Component of the ESCRT-II complex
Pdhb	Pyruvate dehydrogenase E1 component subunit beta,	Q9D051	pyruvate dehydrogenase complex catalyzes
Gpr84	G-protein coupled receptor 84	Q8CIM5	Receptor for medium-chain free fatty acid
Hopx	Homeodomain-only protein	Q8R1H0	Atypical homeodomain protein which does not bind DNA and is required to modulate cardiac growth and development.

Table 2: Once-over of down-oversaw quality with their Uniprot Accession, Function Cellular Component.

Gene Symbol	Full name	Uniprot Accession	Function
Rps6	40S ribosomal protein S6	P62754	ribosomal protein
Fcrls	Fc receptor-like S	Q91YK7	Scavenger receptor activity
AU022252	Uncharacterized protein C1orf50	Q5EBG8	protein binding
Paqr7	progestin receptor alpha	Q80ZE4	Plasma membrane progesterone (P4) receptor coupled to G proteins
Zyg11b	zyg-11 homolog B	Q3UFS0	Serves as substrate adapter subunit in the E3 ubiquitin ligase complex ZYG11B-CUL2-Elongin BC
Scmh1	Polycomb protein SCMH1	Q8K214	Polycomb protein SCMH1
Paip1	Polyadenylate-binding protein-interacting protein 1	Q8VE62	Its stimulatory activity on translation is mediated via its action on PABPC1.

Lats1	Serine/threonine-protein kinase	Q8BYR2	Negative regulator of YAP1 in the Hippo signaling pathway that plays a pivotal role in organ size control and tumor suppression by restricting proliferation and promoting apoptosis.
H2afj	Histone H2A.J	Q8R1M2	Core component of nucleosome.
Zfp93	Zfp93 protein	Q6P7V4	involved in transcriptional regulation
Dap3	28S ribosomal protein S29	G3X9M0	Involved in mediating interferon-gamma-induced cell death
Phgdh	D-3-phosphoglycerate dehydrogenase	Q61753	Catalyzes the reversible oxidation of 3-phospho-D-glycerate to 3-phosphonooxypyruvate

4 APPENDIX

```
library(affy)
library(limma)
library(dplyr)
library(mouse4302.db)
file <- ReadAffy(celfile.path ="C:\\Users\\HP\\Desktop\\mouse",  compress = TRUE)
file
#Checking the Class of the file object
class(file)
# Dimensions
dim(file)
# Structure
str(file)
# Information about the platform
annotation(file)
# Information about the features
head(featureData(file1))
#normalize the data, for instance the rma() function.
f <- rma(file)
```

```
# Check the new object after normalization

class(f)

pData(f)

annotation(file)

head(featureData(f))

dim(f)

#group names for all samples in a series

gsms <- paste0("000111XXX101111X1101010XX0")

gsms

sml <- c()

for (i in 1:nchar(gsms)) { sml[i] <- substr(gsms,i,i) }

# eliminate samples marked as "X"

sel <- which(sml != "X")

sml <- sml[sel]

f <- f[ ,sel]

file=file[,sel]

test2=exprs(f)

test3 = log(test2, 2)
```

```
# set up the data and proceed with analysis

sml <- paste("G", sml, sep="")    # set group names

fl <- as.factor(sml)

f$description <- fl

labels <- c("bdnf","gfp")

pData(file)

#boxplot without normalisation

title <- paste ("GSE14522", '/', "GPL1261", " selected samples", sep ='')

boxplot(file, boxwex=0.6, notch=T, outline=FALSE, las=2, col=fl)

legend("topleft", labels, fill=palette(), bty="n")

#Boxplot  with normalisation

title <- paste ("GSE14522", '/', "GPL1261", " selected samples", sep ='')

boxplot(test2, boxwex=0.6, notch=T, outline=FALSE, las=2, col=fl)

legend("topleft", labels, fill=palette(), bty="n")

# Create the design matrix

design <- model.matrix(~ 0+ f$description)

colnames(design) <- levels(fl)
```

```
#Fit the model using the design matrix and contrast matrices

fit <- lmFit(f, design)

cont.matrix <- makeContrasts(G1-G0, levels=design)

cont.matrix

fit2 <- contrasts.fit(fit, cont.matrix)

# Moderation of standard errors using empirical Bayes for model fit

fit3 <- eBayes(fit2, 0.01)

s <- decideTests(tG2)

#top table

tG2 <- topTable(fit3, adjust="fdr", sort.by="B",  number = 250, resort = "logFC", genelist =
rownames(fit3))

tG2 <- tG2[order(tG2$ID), ]

#genes

sym <- mget(tG2$ID, mouse4302SYMBOL, ifnotfound = NA)

tG2ann <- cbind(SYMBOL = unlist(sym), tG2, stringsAsFactors = FALSE)

a=write.table(tG2ann, file="C:\\Users\\HP\\Desktop\\t.txt", row.names=F, sep="\t")
```

```
#Removal of gene duplication
g <- tG2ann[!duplicated(tG2ann$SYMBOL),,drop=FALSE]
str(g)
head(g)
#volcanplot
library(calibrate)
summary(g)
threshold.high <- sort(g$logFC, decreasing = TRUE)[20]
threshold.low <- sort(g$logFC, decreasing = FALSE)[20]
with(g, plot(logFC, -log10(P.Value), pch=20,main="Volcano plot", xlim=c(-3,3)))
with(subset(g, logFC < threshold.low),points(logFC, -log10(P.Value), pch=20, col="orange"))
with(subset(g, logFC > threshold.high),points(logFC, -log10(P.Value), pch=20, col="blue"))
with(subset(g, logFC > threshold.high),textxy(logFC, -log10(P.Value), labs=SYMBOL, cex=0.9))
with(subset(g, logFC < threshold.low),textxy(logFC, -log10(P.Value), labs=SYMBOL, cex=0.9))
```

5 APPENDIX

Genes	Area	Paper	Function	GO - Molecular Function	GO - Biological Process	GO - Cellular component
Rps6	S6K1 expression is upregulated in the brains of AD patients	https://www.jneurosci.org/content/35/41/14042.short	ribosomal protein	RNA binding, structural constituent of ribosome, protein binding, protein kinase binding	activation-induced cell death of T cells, erythrocyte development, G1/S transition of mitotic cell cycle, gastrulation, glucose homeostasis, mammalian oogenesis stage,mitotic cell cycle, mitotic cell cycle checkpoint, negative regulation of apoptotic process, rRNA processing, translation.	Cytosol, Endoplasmic reticulum, Nucleus, Other locations (cell, cytoplasm, dendrite, ribosome, polysome, small ribosomal subunit)
Fcrls	microglial functions in response to stress and AD pathology	https://www.sciencedirect.com/science/article/pii/S2352289518300079	Scavenger receptor activity	coreceptor activity	No data available	Membrane
AU022252	Could not find this gene related reaseach paper	n		identical protein binding, protein binding	No data available	No data available
Paqr7		NF-	Plasma membrane progesterone (P4) receptor coupled to G	signaling receptor activity, steroid binding,steroid hormone receptor activity	multicellular organism development, oogenesis, response to steroid hormone	Cell membrane, protein

Scmh1		NF-	Polycomb group (PcG) multiprotein complexes	protein binding	anterior/posterior pattern specification, chromatin remodeling, gene silencing, negative regulation of transcription, DNA-templated	Nucleus
Paip1		https://link.springer.com/article/10.1007/s11064-006-9117-8	Its stimulatory activity on translation is mediated via its action on PABPC1.	RNA binding, translation activator activity, protein binding	regulation of translational initiation, translational initiation, mRNA stabilization, positive regulation of translation, regulation of translation	cytoplasm, cytosol
Phgdh		https://sci-hub.tw/10.1016/j.bmc.2012.12.008	Catalyzes the reversible oxidation of 3-phospho-D-glycerate to 3-phosphonooxypyruvate	NAD binding, phosphoglycerate dehydrogenase activity	cellular amino acid metabolic process, G1 to G0 transition, gamma-aminobutyric acid metabolic process , glial cell development, neurogenesis, neural tube development, neuron projection development, regulation of gene expression, spinal cord development, glycine metabolic process	myelin sheath, cytosol, extracellular exosome
Lats1		https://sci-hub.tw/10.1016/j.bmc.2012.12.008	Negative regulator of YAP1 in the Hippo signaling pathway that	nucleotide binding, magnesium ion binding, protein kinase activity,	cell division, cellular protein localization, cytoplasmic sequestering of protein, G1/S	Cytoskeleton, cytoplasm, microtubule organizing center, cytosol, microtubule organizing center

					process, sister chromatid segregation,hormone -mediated signaling pathway, protein phosphorylation, regulation of organ growth	
H2afj		https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3388733/	Core component of nucleosome.	DNA binding, protein heterodimerization activity	chromatin organization, chromatin silencing,	Nucleus, Chromosome
Zfp93		https://www.sciencedirect.com/science/article/pii/S0012160607000735X	involved in transcriptional regulation.	metal ion binding, nucleic acid binding	regulation of transcription, DNA-templated	intracellular, nucleus
Dap3		https://sci-hub.tw/10.1109/biocas.2016.7833793	Involved in mediating interferon-gamma-induced cell death	RNA binding, structural constituent of ribosome, protein binding, GTP binding	apoptotic process	nucleoplasm, mitochondrion, mitochondrial ribosome,
Cap1		NF-	Directly regulates filament dynamics	actin binding, adenylate cyclase binding	cell morphogenesis, ameboidal-type cell migration, receptor-mediated endocytosis, cytoskeleton organization, establishment or maintenance of cell polarity	extracellular region, cytoplasm, plasma membrane, focal adhesion, membrane

Pi4kb		https://www.sciencedirect.com/science/article/abs/pii/0925443994900930	regulate Golgi disintegration/reorganization during mitosis, possibly via its phosphorylation	14-3-3 protein binding, 1-phosphatidylinositol 4-kinase activity, ATP binding,phosphatidylinositol kinase activity	phosphatidylinositol-mediated signaling, phosphatidylinositol phosphorylation	Cytosol, Endoplasmic reticulum, Golgi apparatus, Mitochondrion, Plasma Membrane, cytoplasm, membrane
Ptp4a2		NF-	Protein tyrosine phosphatase which stimulates progression from G1 into S phase during mitosis	protein tyrosine phosphatase activity, prenylated protein tyrosine phosphatase activity, protein tyrosine/serine/threonine phosphatase activity, hydrolase activity, phosphatase activity	protein dephosphorylation, dephosphorylation, peptidyl-tyrosine dephosphorylation, post-translational protein modification	nucleus, cytoplasm, endosome, early endosome, cytosol
Slpi		https://www.sciencedirect.com/science/article/pii/S1525260140000314	Acid-stable proteinase inhibitor with strong affinities for trypsin, chymotrypsin, elastase, and cathepsin G	DNA binding, endopeptidase inhibitor activity, enzyme binding, mRNA binding, serine-type endopeptidase inhibitor activity	antibacterial humoral response, immune response, innate immune response, negative regulation of protein binding, negative regulation of viral genome replication	Extracellular region or secreted, Golgi apparatus
Frrs1		https://sci-hub.tw/10.1002/prca.201400149	Ferric-chelate reductases reduce Fe3+ to Fe2+ before its transport from the endosome to the cytoplasm.	ferric-chelate reductase activity, metal ion binding	oxidation-reduction process	integral component of membrane

Vps36		https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3413662/	Component of the ESCRT-II complex	phosphatidylinositol-3-phosphate binding, protein C-terminus binding, ubiquitin binding	endosomal transport, protein transport to vacuole involved in ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway	Cytosol, Endosome, Lysosome, Nucleus
Pdhb		NF-	pyruvate dehydrogenase complex catalyzes	pyruvate dehydrogenase (acetyl-transferring) activity, pyruvate dehydrogenase (NAD+) activity	acetyl-CoA biosynthetic process from pyruvate, glucose metabolic process, mitochondrial acetyl-CoA biosynthetic process from pyruvate, tricarboxylic acid cycle	Mitochondrion, Nucleus, pyruvate dehydrogenase complex
Gpr84	GPR84 deficiency reduces microgliosis	https://www.sciencedirect.com/science/article/pii/S0889159115000136	Receptor for medium-chain free fatty acid	G protein-coupled peptide receptor activity, urotensin II receptor activity	neuropeptide signaling pathway	Plasma Membrane, receptor complex
Hopx		NF-	Atypical homeodomain protein which does not bind DNA and is required to modulate cardiac growth $ development.	DNA binding	regulation of heart contraction, regulation of protein binding, regulation of transcription by RNA polymerase II, trophectodermal cell differentiation	Nucleus, cytoplasm

Jchain		NF-	link two monomer units of either IgM or IgA.	antigen binding, IgA binding, immunoglobulin receptor binding, protein binding, bridging, protein homodimerization activity, single-stranded DNA binding	adaptive immune response, antibacterial humoral response, glomerular filtration, humoral immune response, innate immune response, protein-containing complex assembly	Extracellular region or secreted

6 CONCLUSION

The Alzheimer's infirmity is a raising neurodegenerative issue that impacts an enormous number of people reliably and can't be reestablished with no issue. In this study the information assume that reduced BDNF in the serum of patients will provoke Alzheimer's infirmity. In the final findings of our analysis, we have find 29 genes with their respective functions, gene ontology, nucleotide sequence, protein sequence, and associated lncRNA if any. With the assistance of these qualities we can do wet lab concentrate for additional examination. This computational analysis provides specific and robust information from a set of thousands of genes involved in gene expression studies. It is believed that the narrowed results upto 29 could be targeted experimentally to verify and provide few specific biomarkers for further investigations at various levels.

7 REFERENCES

[1] A. Kumar and T. R. Singh, "A New Decision Tree to Solve the Puzzle of Alzheimer's Disease Pathogenesis Through Standard Diagnosis Scoring System," *Interdiscip. Sci. Comput. Life Sci.*, vol. 9, no. 1, pp. 107–115, 2017.

[2] P. P. Panigrahi and T. R. Singh, "Computational studies on Alzheimer's disease associated pathways and regulatory patterns using microarray gene expression and network data: Revealed association with aging and other diseases," *J. Theor. Biol.*, vol. 334, pp. 109–121, 2013.

[3] Shukla R., Singh T.R. (2019). Virtual screening, pharmacokinetics, molecular dynamics and binding free energy analysis for small natural molecules against cyclin-dependent kinase 5 for Alzheimer's disease. Journal of Biomolecular Structure and Dynamics, 38(1), 248-262.

[4] Ashwani Kumar, Ankush Bansal, Tiratha Raj Singh (2019). ABCD: Alzheimers disease Biomarkers Comprehensive Database. 3 Biotech, 9 (10), 351-357

[5] Jan Graffelman, "No Title," *2019-10-01 17:10:02 UTC CRAN*, 2019. [Online]. Available: https://www.rdocumentation.org/packages/calibrate/versions/1.7.5.

[6] M. E. Ritchie *et al.*, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Res.*, vol. 43, no. 7, pp. e47–e47, Jan. 2015.

[7] H. Wickham and R. François, *dplyr: A Grammar of Data Manipulation*. 2014.

[8] *et al.*, "Genes Selection Comparative Study in Microarray Data Analysis," *Bioinformation*, vol. 9, no. 20, pp. 1019–1022, 2013.

[9] L. Gautier, L. Cope, B. M. Bolstad, and R. A. Irizarry, "affy—analysis of Affymetrix GeneChip data at the probe level," *Bioinformatics*, vol. 20, no. 3, pp. 307–315, Feb. 2004.

[10] Marc Carlson, "No Title," *mouse4302.db: Affymetrix Mouse Genome 430 2.0 Array annotation data (chip mouse4302). R package version 3.2.3.* [Online]. Available: http://bioconductor.org/packages/release/data/annotation/html/mouse4302.db.html.

[11] Ashwani Kumar, Tiratha Raj Singh (2017). Analysis for Biological Network Properties of Alzheimers Disease Associated Gene Set by Enrichment and Topological Examinations. International Journal of Bioinformatics Research and Applications, 13 (3), 214-222.