



# Identification of Expressed, Mutated Genes and Binding Sites of DNA Associated Proteins in Alzheimer's Disease

V. Maruthi Raj<sup>2</sup>(✉) and Preenon Bagchi<sup>1,2,3</sup>

<sup>1</sup> Padmashree Institute of Management and Science, Bangalore, India

<sup>2</sup> Vasishth Academy of Advanced Studies and Research (Sarvasumana Association), Bangalore, India

maruthi1206@gmail.com

<sup>3</sup> MGM Institute of Biosciences and Technology, Aurangabad, India

**Abstract. Background:** A neurodegenerative disorder in the elderly individuals. It is a type of dementia that affects memory subsequently followed by executive dysfunction, confusion, agitation, and behavioral disturbances. Three causative genes have been associated with autosomal dominant familial AD (APP, PSEN1, and PSEN2). Identification of these genes has led to a number of animal models that have been useful to study the pathogenesis underlying AD. The primary goal of this study is to identify a set of biomarkers that maximise prediction performance in the blood and could aid in the early detection of Alzheimer's disease as well as provide valuable information about the molecular mechanisms that underpin Alzheimer's disease.

**Keywords:** Alzheimer's disease (AD) · neurodegeneration · transcriptomics · amyloid · Presenilin

## 1 Introduction

A progressive and neurodegenerative disease that destroys important mental functions as well as memory which is also called as senile dementia. This is the disorder of cognitive functions that include orientation, language, visuospatial function, judgement and predominantly memory.

It is estimated that Alzheimer's disease accounts for 60% of dementias, and it is expected to increase in triple fold by 2050. But till now, the treatments and methods to treat this is not determined to cure this in an effective manner.

Till now, drug discovery based on amyloid cascade hypothesis, neurofibrillary tangles of hyperphosphorylated tau, brain vascular pathology, neurotransmitter ailment, cell cycle dysregulation, oxidative stress and mitochondrial impairment have failed to cure dementia there are some drugs that include acetylcholinesterase inhibitors and N-methyl-D-aspartate receptor (NMDA) antagonists which perform temporary actions however, they do not change the course of illness or the rate of decline. In view of the upcoming epidemiological burden, drug repositioning might provide an elegant solution to the need

for curative treatment. Thus, identifying the proper molecular mechanisms this disease's progression is necessary for the disease modifying drug discovery. Numerous research approaches are being done and is very challenging to study the brain tissue as the cells has to be collected post mortem. Yet there is no definite early pre-mortem diagnosis for AD, and for cognitive assessment we use brain imaging methods that reveal the degradation. The degradation in cognition due to AD is irreversible, and it also conducts the immutable neural degradation that has occurred; hence, it is very much necessary to detect the AD earlier finding out the biomarkers for AD in the blood samples helps in determining the treatment for this disease.

We used transcriptomics as a global biochemical approach to identify potential biomarkers and molecular pathways in the brain, we used curated gold benchmark databases for AD to analyse human genomic and transcriptomic data to determine the involvement of genes that aid in pathogenic processes. The primary goal of this study is to identify a set of biomarkers that maximise prediction performance in the blood and could aid in the early detection of Alzheimer's disease as well as provide valuable information about the molecular mechanisms that underpin Alzheimer's disease [1–6].

### **Genes Involved**

Mutations in amyloid-beta precursor protein (APP), presenilin1, and presenilin2 are associated with early-onset forms of familial AD, whereas sporadic AD occurs in people over the age of 65 years [7–11]

### **Transcriptomics**

Transcriptomics is the bioinformatic analysis of the regulation and the expression of genes within certain tissues and developmental stage.

Generally, it is the study of all RNA molecules within a cell. RNA is copied from pieces of DNA and contains information to make the desired proteins and perform other important functions in the cell using high-throughput methods, such as microarray analysis, RNA-Seq . Transcriptomics is used to study more about how genes are turned on in different types of cells and how this may help cause certain diseases, such as Alzheimer's.

The key aims of transcriptomics are:

- to catalogue all species of transcript
- non-coding RNAs and small RNAs
- to determine the transcriptional structure of genes in terms of their start site
- 5' and 3' ends, splicing patterns and other post-transcriptional modifications and to quantify the changing

## **2 Materials and Methods**

Galaxy tutorial by James Peter Taylor, it is a web-based platform for accessible, reproducible and transparent computational biochemical research. The analysis of transcriptomics and also provide structure and organization of the genome. The sequences **SRR12786091**, **SRR12786088**, **SRR12786097** and **SRR12786097** were retrieved from

SRA database. Firstly, we will identify the quality score of our raw sequence, so the tool we used for QC is the FastQC tool, which provides a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines. It is a very popular tool used for providing an overview of basic quality control metrics for raw next generation sequencing data. In the next run Trimmomatic on each pair of forward and reverse reads, this is done to increase mapping efficiency. Next run Trimmomatic. It provides support for quality checking, filters the sequence, and provides statistical analysis of differential gene expression. Next do Mapping, it is the process of comparing each one of the reads with the reference genome to obtain one alignment, or more, between each read and the genome. Lastly conclude by HISAT2, It is a fast and sensitive alignment programme for mapping sequence reads of the human genome against the reference genome.

**Quality control:** - Before any analysis or sequence alignment is done, the data quality has to be checked.

**FastQC:** -It provides few features like whether the data has many problems to be aware before further analysis. To assess the quality of the reads FastQC was to be run.

**MultiQC:** - across many samples for analyses of bioinformatics a single report is created with interactive plots using the tool MULTIQC.

**Trimmomatic:** - To increase mapping efficiency, from low-quality bases, trimming is performed. After this, to inspect the differences, re-run FastQC.

**Mapping tool:** - There are many software available from which mapping tool is process where the reads are compared with reference genome and one or more alignment will be recognized between each read & genome.

**HISAT:** - For mapping, next generation- sequence reads (HISAT-tool) is performed which a sensitive alignment.

**De novo transcript reconstruction** Following the HISAT, de novo transcript reconstruction determines transcript structures illustrated by aligned reads. This even identifies all transcripts present in sample.

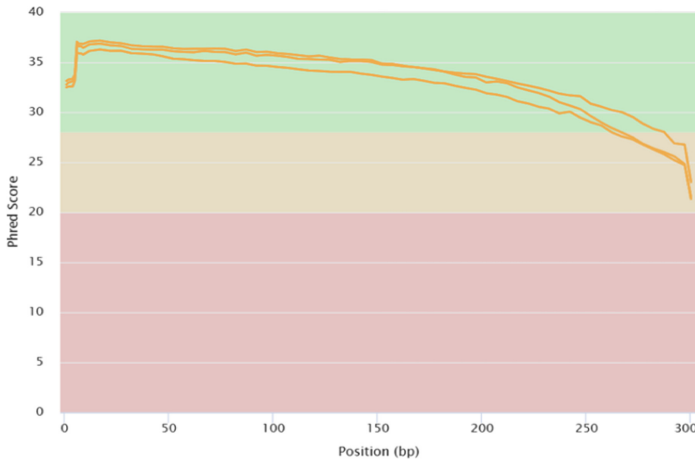
**Stringtie tool:** highly efficient assembler of RNA-seq alignment into potential transcript data.

### Transcriptome Assembly

Each 4 RNA seq libraries were analyzed/represented by transcriptome with Stringtie. Transcriptome database is made to know which transcript structures corresponds to annotated transcripts in absence of reference transcriptome.

1. **Stringtie-merge tool:** Stringtie-merge tool help to assemble transcript along with their refseq file.
2. **GFFCompare tool:** GFFCompare tool are used for the merge the stringtie and RefSeq annotation file together.

**Feature Counts tool:** Included in genomic feature, feature counts work as to quantify reads generated from RNA or DNA sequence. To designate the reads with feature for high efficiency genomic feature, blocking, chromosome hashing is to be implemented.



**Fig. 1.** FastQC: Mean Quality Scores

**DESeq:** - It is used to analyse the RNA-seq data for gene expression.

**Working of DESeq:** - DESeq2 provides a function collapse Replicates which can assist in combining the counts from technical replicates into single columns of the count matrix. The term technical replicate implies multiple sequencing runs of the same library.

### 3 Result and Discussion

**FastQC tool** - the quality score of our raw sequence, so the tool we used for QC is the FastQC tool, which provides a simple way to do some quality control checks on raw sequence data coming from high throughput sequencing pipelines (Fig. 1).

**Trimmomatic tool** - To increase mapping efficiency, from low-quality bases, trimming is performed. After this, to inspect the differences, re-run FastQC.

**Mapping** - The mapping process of comparing each one of the reads with the reference genome. We will obtain one alignment, or more, between each read and the genome. Like for any other bioinformatic task there is a lot of mapping software available.

**HISAT2 tool:** For mapping, next generation- sequence reads (HISAT-tool) is performed which a sensitive alignment.

Run HISAT2 on one forward/reverse read pair and modify the following settings (Fig. 2).

**Feature counts** – Feature Counts is a program that counts how many reads map to genomic features, such as genes, exon, promoter and genomic bins. Feature Counts function checks if reads from the same pair are adjacent to each other (this could happen when reads were for example sorted by their mapping locations), and it automatically reorders those reads that belong to the same pair but are not adjacent to each other in the input read file (Fig. 3).

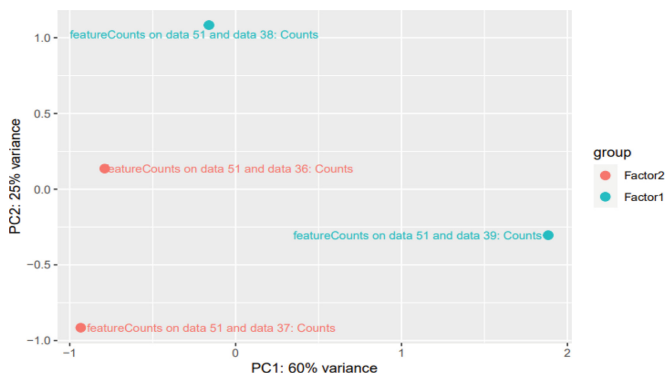
48622	113	chr10	221318	60	255M31S	=	221318	-286	GCCAAC
48622	177	chr10	221318	60	255M31S	=	221318	-286	GCCAAC
1175915	113	chr10	221381	60	192M31S	=	221381	-223	GCCTGT
1175915	177	chr10	221381	60	192M31S	=	221381	-223	GCCTGT
1282489	385	chr10	341497	0	12556M	chr11	92976630	0	TGCAG
984239	113	chr10	437972	60	173M15S	=	437972	0	AAGTTC
984239	177	chr10	437972	60	173M15S	=	437972	0	AAGTTC
48952	65	chr10	441497	60	275234M24S	=	441497	285	TGTGTG
48952	129	chr10	441497	60	275234M24S	=	441497	-285	TGTGTG
274521	65	chr10	441497	60	275106M1D116M	=	441497	250	AGGGTC
274521	129	chr10	441497	60	275106M1D116M	=	441497	-250	AGGGTC
308671	65	chr10	441497	60	275133M	=	441497	160	TGTGTG
308671	129	chr10	441497	60	275133M	=	441497	-160	TGTGTG
314166	65	chr10	441497	60	275234M23S	=	441497	284	TGCATA
314166	129	chr10	441497	60	275234M23S	=	441497	-284	TGCATA
344405	65	chr10	441497	60	275106M1D102M	=	441497	236	GCGCTC
344405	129	chr10	441497	60	275106M1D102M	=	441497	-236	GCGCTC
359469	65	chr10	441497	60	275234M7S	=	441497	268	ACGGTG
359469	129	chr10	441497	60	275234M7S	=	441497	-268	ACGGTG
475198	65	chr10	441497	60	275234M7S	=	441497	268	GCTCTG
475198	129	chr10	441497	60	275234M7S	=	441497	-268	GCTCTG
691646	65	chr10	441497	60	275234M3S	=	441497	0	GCTCTG
691646	129	chr10	441497	60	275234M3S	=	441497	0	GCTCTG
764768	65	chr10	441497	60	275106M1D110M	=	441497	244	CAGTTT

**Fig. 2.** HISAT2 tool for mapping

Status	HISAT2 on data 16 and data 15: aligned reads (BAM)
Status	HISAT2 on data 16 and data 15: aligned reads (BAM)
Assigned	52
Unassigned_Unmapped	1418971
Unassigned_Read_Type	0
Unassigned_Singleton	0
Unassigned_MappingQuality	0
Unassigned_Chimera	1419858
Unassigned_FragmentLength	0
Unassigned_Duplicate	0
Unassigned_MultiMapping	19661
Unassigned_Secondary	0
Unassigned_NonSplit	0
Unassigned_NoFeatures	77
Unassigned_Overlapping_Length	0
Unassigned_Ambiguity	51

**Fig. 3.** Feature counts

**DESeq2-** It is a tool for differential gene expression analysis of RNA-seq data. It is a new version of DESeq and can detect more differentially expressed genes (DEGs) than DESeq2. However, it also seems to allow more false positives. The DESeq2 package is designed for normalization, visualization, and differential analysis of high- dimensional count data.



**Fig. 4.** Principal component analysis

It makes use of empirical Bayes techniques to estimate priors for log fold change and dispersion, and to calculate posterior estimates for these quantities.

**PCA (principal component analysis)** - Principal component analysis, or PCA, is a statistical procedure that allows you to summarize the information content in large data tables by means of a smaller set of “*summary indices*” that can be more easily visualized and analyzed. The underlying data can be measurements describing properties of production samples, chemical compounds or reactions, process time points of a continuous process, batches from a batch process, biological individuals or trials of a DOE-protocol (Fig. 4).

**Heat map** - Heat map is a data visualization technique that shows magnitude of a phenomenon as color in two dimensions. The variation in color may be by hue or intensity, giving obvious visual cues to the reader about how the phenomenon is clustered or varies over space. It helps to visualize density (Fig. 5).

**Dispersion estimates** - The dispersion is a parameter describing how much the variance deviates from the mean. The Poisson distribution is sometimes said to be a special case of the NB distribution, when dispersion=1 and thereby mean = variance. So, to answer your question: no, the dispersion is not the variance of your gene (Fig. 6).

**Histogram** - Histogram is a chart that plots the distribution of a numeric variable’s values as a series of bars. Each bar typically covers a range of numeric values called a bin or class; a bar’s height indicates the frequency of data points with a value within the corresponding bin (Fig. 7).

**MA plot** - An MA plot visualizes the relationships between the log ratio and mean values of two variables. An MA-plot is a plot of log-intensity ratios (M-values) versus log-intensity averages (A-values). The “M” refers to minus in the log scale. The log ratios of the two measurements are plotted on the vertical (y) axis (Fig. 8).

**Volcano plots** - volcano plots are commonly used to display the results of RNA-seq or other omics experiments. A volcano plot is a type of scatterplot that shows statistical significance (P value) versus magnitude of change (fold change). It enables quick visual identification of genes with large fold changes that are also statistically significant. These may be the most biologically significant genes. In a volcano plot, the most upregulated

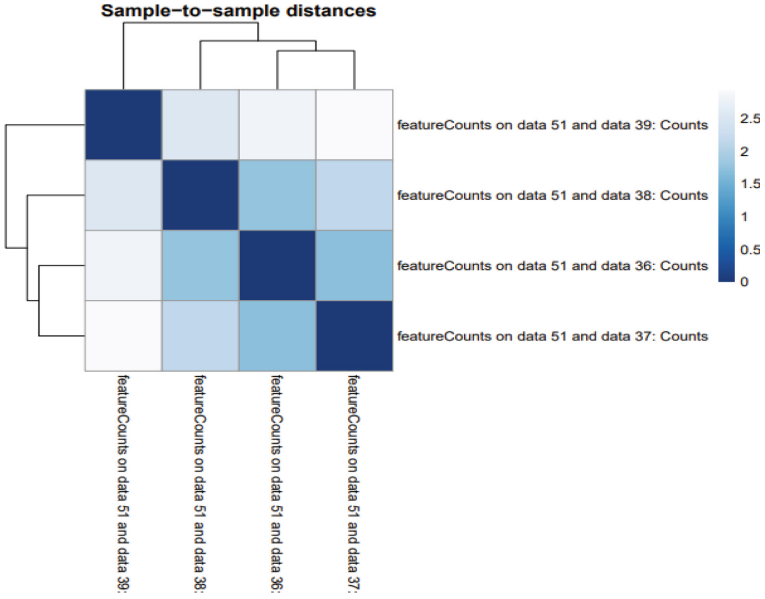


Fig. 5. Heat map

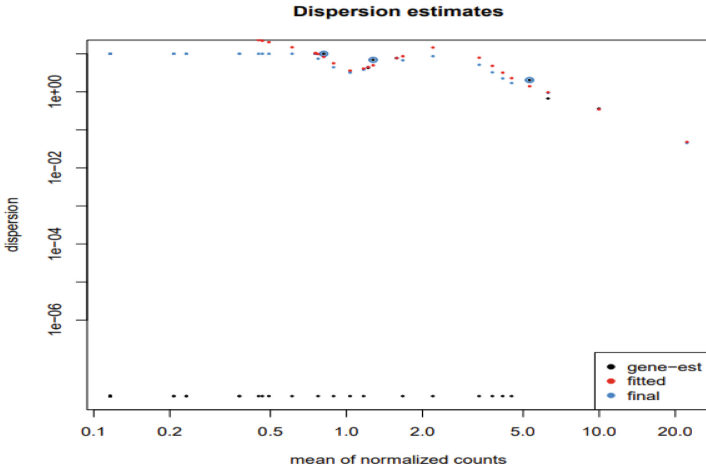
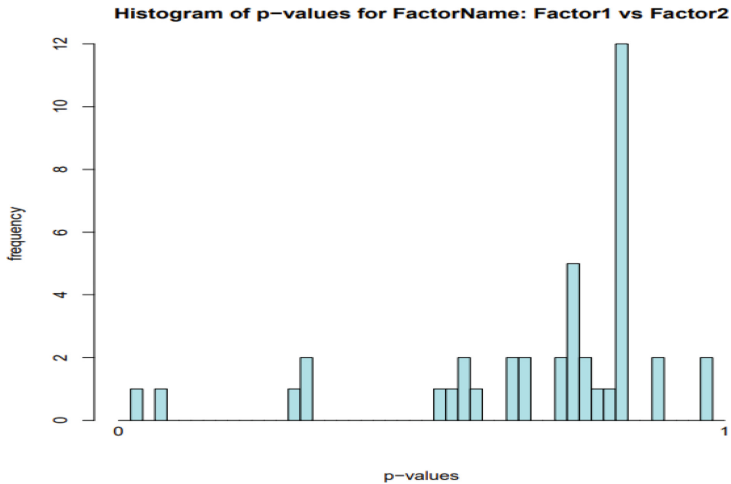
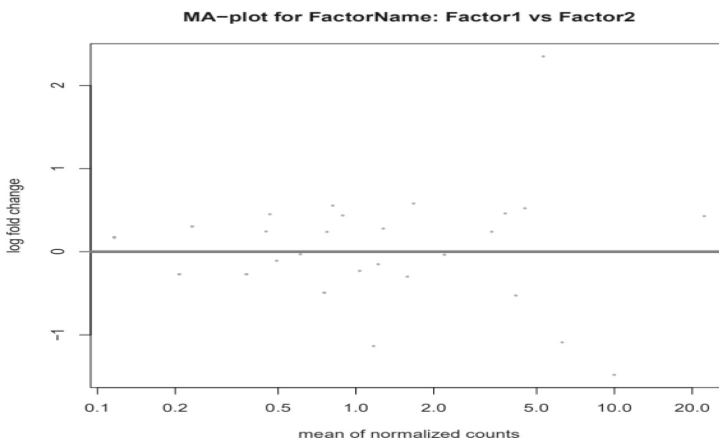


Fig. 6. Dispersion estimates

genes are towards the right, the most downregulated genes are towards the left, and the most statistically significant genes are towards the top (Fig. 9).



**Fig. 7.** Histogram



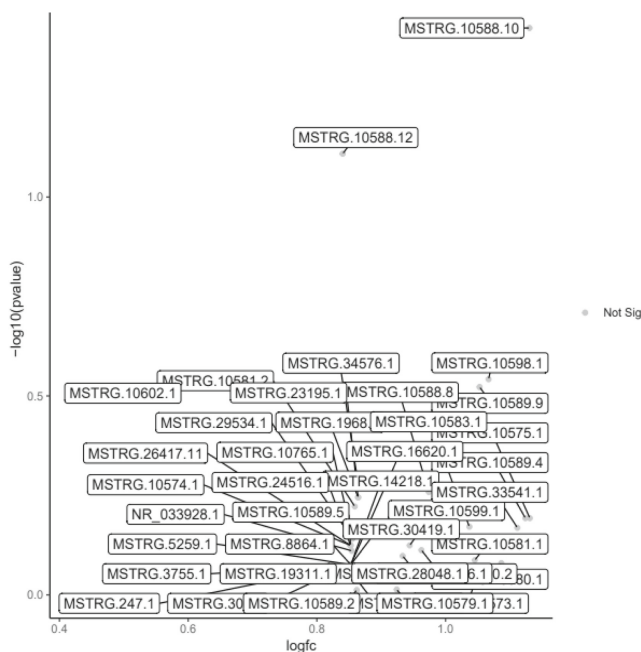
**Fig. 8.** MA plot

## 4 Drug Design [12-16]

### Neuroprotective Herbs for the Management of Alzheimer's Disease

1. **Ashwagandha** (*Withania somnifera*)-. ergostane-type steroidal lactones, withanolides A-Y, dehydrowithanolide-R, withasomniferin-A, withasomidienone, withasomniferols A-C, withaferin A, withanone
2. **Brahmi** (*Bacopa monnieri*)- saponins, bacopasides III, IV, V, bacosides A and B, bacosaponins A, B, C, D, E, and F, alkaloids, sterols, betulic acid, polyphenols
3. **Cat's Claw** (*Uncaria tomentosa*)- oxindole alkaloids, polyphenols (flavonoids, proanthocyanidins, and tannins), glycosides, pentacyclic alkaloids, and sterols





**Fig. 9.** Volcano plot

4. **Ginkgo Biloba**-flavonoid glycosides, terpene lactones, and ginkgolic acids
5. **Gotu Kola** (*Centella asiatica*)- (asiaticosides, asiatic acid, madecassoside, and madasiatic acid
6. **Lion's Mane** (*Hericium erinaceus*)- hericenones and erinacines
7. **Saffron** (*Crocus sativus*)- safranal, a carbox-aldehyde
8. **Shankhpushpi** (*Convolvulus pluricaulis*)- triterpenoids, flavonol glycosides, anthocyanins, and steroids
9. **Turmeric** (*Curcuma longa*)- curcumin, demethoxycurcumin (DMC), bis-demethoxycurcumin (BDMC), and cyclocurcumin
10. **Moringa oleifera**- alkaloids, saponins, tannins, steroids, phenolic acids, glucosinolates, flavonoids, and terpenes (Tables 1, 2, 3, and 4)

Compound demethoxycurcumin (DMC) has a docking score of  $-4084$  kcal/mol and has the most 3 interactions with the protein **NR\_033928.1**. Hence this compound can be taken as ligand for the disease- **ALZHEIMER'S DISEASE**.

**Table 1.** Phytochemical compounds and their SMILE sequence

<i>Plant name</i>	<i>Phytocompound</i>	<i>SMILES</i>
<b>Turmeric</b>	<i>Curcumin</i>	<chem>COC1=C(C=CC(=C1)C=CC(=O)CC(=O)C=CC2=CC(=C(C=C2)O)OC)O</chem>
	<i>demethoxycurcumin (DMC)</i>	<chem>COC1=C(C=CC(=C1)C=CC(=O)CC(=O)C=CC2=CC=C(C=C2)O)O</chem>
	<i>bisdemethoxycurcumin</i>	<chem>C1=CC(=CC=C1C=CC(=O)CC(=O)C=CC2=CC=C(C=C2)O)O</chem>
	<i>cyclocurcumin</i>	<chem>COC1=C(C=CC(=C1)C=CC2=CC(=O)CC(O2)C3=CC(=C(C=C3)O)OC)O</chem>
<b>Saffron</b>	<i>safranal</i>	<chem>CC1=C(C(CC=C1)(C)C)C=O</chem>
	<i>carboxaldehyde</i>	<chem>C1=CC(=C2C=COC2=C1C=O)F</chem>
<b>Cat's Claw</b> ( <i>Uncaria tomentosa</i> )	<i>proanthocyanidin</i>	<chem>COC1=C(C=C(C=C1O)C2C(CC3=C(O2)C(=C(C=C3O)O)C4C(C(OC5=CC(=CC(=C45)O)O)C6=CC=C(C=C6O)O)O)O)O</chem>
	<i>Flavonoid</i>	<chem>CC1(C(CC2=C(O1)C=C(C3=C2OC(CC3=O)C4=CC=C(C=C4)O)O)O)C</chem>
<b>Brahmi</b> ( <i>Bacopa monnieri</i> )	<i>saponin</i>	<chem>CC1(C2CCC3(C(C2(CCC1OC4C(C(C(CO4)OC5C(C(C(CO5)O)O)OC6C(C(C(C(O6)CO)O)O)OC7C(C(C(C(O7)CO)O)O)OC8C(C(C(C(O8)CO)O)O)O)C)CCC9C3(CC(C2(C9C(C(C2)C)C=O)CO1)O)C)C</chem>
	<i>Bacopaside IV</i>	<chem>CC(=CC1CC(C2C3CCC4C5(CCC(C(C5CCC4(C36CC2(O1)OC6)C)(C)OC7C(C(C(CO7)O)OC8C(C(C(C(O8)CO)O)O)O)C)(C)O)C</chem>
	<i>Bacopaside A</i>	<chem>CC(=CCCC(C)(C1C2CCC3C(C2(CC1=O)C)(CCC4C3(CCC(C4(C)C)OC5C(C(C(C(O5)CO)OC6C(C(C(CO6)O)O)O)O)CO)C)O)C</chem>
	<i>Bacopasaponin C</i>	<chem>CC(=CC1COC23CC4(CO2)C(C3C1(C)O)CCC5C4(CCC6C5(CCC(C6(C)C)OC7C(C(C(CO7)O)OC8C(C(C(C(O8)CO)O)O)OC9C(C(C(O9)CO)O)O)C)C</chem>
	<i>betulic acid</i>	<chem>CC(=C)C1CCC2(C1C3CCC4C5(CCC(C(C5CCC4(C3(CC2)C)C)(C)O)C)C(=O)O</chem>

(continued)

**Table 1.** (continued)

<b>Lion's Mane</b> ( <i>Hericiu m erinaceus</i> )	<i>Hericenone A</i>	<chem>CC(=CC(=O)CC(=CCC1=C(C=C2C(=C1O)COC2=O)OC)C)C</chem>
	<i>Erinacine G</i>	<chem>CC(C)C(=O)CCC1(CCC2(C(C1=O)CC=C3C4C2OC5C4(C(C(C3O))(CO5)O)O)O)C</chem>
	<i>Hericenone G</i>	<chem>CCCCCCCCCCCCCCCCC(=O)OCC1=CC(=C2CCC(OC2=C1C=O)(C)CC(=O)C=C(C)C)OC</chem>
	<i>Hericenone f</i>	<chem>CCCCCCCCCCCCCCCCC(=O)OCC1=CC(=C2CCC(OC2=C1C=O)(C)CC(=O)C=C(C)C)OC</chem>
	<i>Erinacine E</i>	<chem>CC(C)C1=C2C3CC=C4C5C(C3(CCC2(CC1)C)C)OC6C5(C(C4O)(CO6)O)O</chem>
<b>Ashwaga ndha</b> ( <i>Withani a sommifer a</i> )	<i>withanolide A</i>	<chem>CC1=C(C(=O)OC(C1)C(C)C2CCC3C2(CCC4C3C5C(O5)C6(C4(C(=O)C=CC6)C)O)C)O)C</chem>
	<i>ergostane</i>	<chem>CC(C)C(C)CCC(C)C1CCC2C1(CCC3C2CCC4C3(CCCC4)C)C</chem>
	<i>Withanone</i>	<chem>CC1=C(C(=O)OC(C1)C(C)C2(CCC3C2(CCC4C3C5C(O5)C6(C4(C(=O)C=CC6)C)O)C)O)C</chem>
	<i>withaferin A</i>	<chem>CC1=C(C(=O)OC(C1)C(C)C2CCC3C2(CCC4C3CC5C6(C4(C(=O)C=CC6)C)O5)C)CO</chem>
	<i>Withanolide b</i>	<chem>CC1=C(C(=O)OC(C1)C(C)C2CCC3C2(CCC4C3C5C(O5)C6(C4(C(=O)C=CC6)C)O)C)C</chem>
<b>Gotu Kola</b> ( <i>Centella asiatica</i> )	<i>Asiaticoside</i>	<chem>CC1CCC2(CCC3(C(=CCC4C3(CCC5C4(CC(C(C5(C)CO)O)O)C)C2C1C)C(=O)OC6C(C(C(C(O6)COC7C(C(C(C(O7)CO)OC8C(C(C(C(O8)C)O)O)O)O)O)O)O)O)O)O</chem>
	<i>asiatic acid</i>	<chem>CC1CCC2(CCC3(C(=CCC4C3(CCC5C4(CC(C(C5(C)CO)O)O)C)C2C1C)C(=O)O</chem>
	<i>madecassoside</i>	<chem>CC1CCC2(CCC3(C(=CCC4C3(CC(C5C4(CC(C(C5(C)CO)O)O)O)C)C2C1C)C(=O)OC6C(C(C(C(O6)COC7C(C(C(C(O7)CO)OC8C(C(C(C(O8)C)O)O)O)O)O)O)O)O)O</chem>

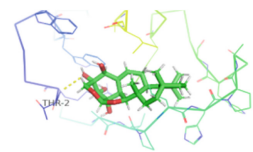
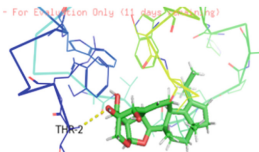
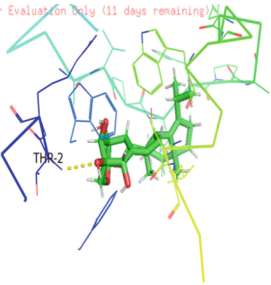
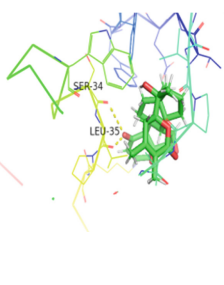
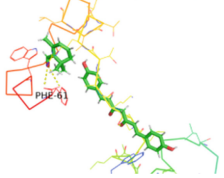
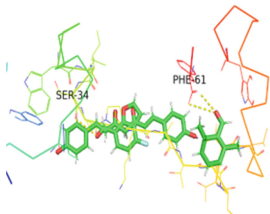
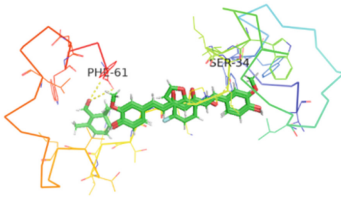
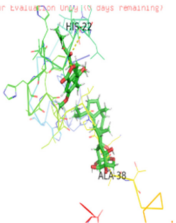
**Table 2.** Molinspiration

Phytochemical compounds	miLogP	TPSA	natoms	nW	noN	nOHNH	Nviolations	Nroth	Volume
<i>CurcuminJLLK</i>	2.30	93.07	27	368.38	6	2	0	8	332.18
<i>demethoxycurcumin (DMC)</i>	2.48	83.83	25	338.36	5	2	0	7	306.64
<i>bisdemethoxycurcumin</i>	2.67	74.60	23	308.33	4	2	0	6	281.09
<i>Cyclocurcumin</i>	3.03	85.23	27	368.38	6	2	0	5	328.17
<i>Safranal</i>	2.95	17.07	11	150.22	1	0	0	1	158.60
<i>carboxaldehyde</i>	2.12	30.21	12	164.13	2	0	0	1	133.52
<i>proanthocyanidin</i>	3.05	209.75	43	592.55	12	9	3	4	493.20
<i>Flavonoid</i>	2.79	96.22	26	356.37	6	3	0	1	311.66
<i>Saponin</i>	-2.34	422.06	85	1223.36	27	15	3	14	1086.35
<i>Bacopaside IV</i>	3.21	197.00	54	766.97	13	7	3	6	715.53
<i>Bacopaside A</i>	2.54	215.83	54	768.98	13	8	3	10	729.72
<i>Bascopasaponin C</i>	2.79	255.92	63	899.08	17	9	3	9	822.81
<i>betulic acid</i>	7.04	57.53	33	456.71	3	2	1	2	472.04
<i>Hericenone A</i>	3.58	72.84	24	330.38	5	1	0	6	308.64
<i>Erinacine G</i>	0.35	133.52	33	464.56	8	4	0	4	424.71
<i>Hericenone G</i>	9.56	78.92	43	598.87	6	0	2	24	619.87
<i>Hericenone f</i>	9.32	78.92	41	570.81	6	0	2	22	586.26
<i>Erinacine E</i>	2.27	99.38	31	432.56	6	4	0	1	403.32
<i>withanolide A</i>	4.15	96.36	34	470.61	6	2	0	2	441.81

**Table 3.** Interacting amino acids and docking scores

Plant name	Phytocompound	Interacting aminoacids	No of interactions	Docking score	Docking
<i>Turmeric</i>	<i>Curcumin</i>	LEU-35 THR-2 ALA-89 THR-39	4	-3964 Kcal/mol	Yes
<i>Turmeric</i>	<i>demethoxycurcumin (DMC)</i>	ALA-38 PHE-61 LYS-37	3	-4084 Kcal/mol	Yes
<i>Turmeric</i>	<i>Bisdemethoxycurcumin</i>				
<i>Turmeric</i>	<i>Cyclocurcumin</i>				
<i>Saffron</i>	<i>Safranal</i>	PHE-61	1	-2058 Kcal/mol	Yes
<i>Saffron</i>	<i>Carboxaldehyde</i>	THR-2 ASP-4	2	-2410 Kcal/mol	Yes
<i>Uncaria tomentosa</i>	<i>Flavonoid</i>	ARG -3 HIS-22 LEU-22	3	-3812 kcal/mol	Yes
<i>Hericium erinaceus</i>	<i>Hericenone A</i>	HIS -22	1	-4182 Kcal/mol	Yes
<i>Hericium erinaceus</i>	<i>Erinacine G</i>	SER-34	1	-4528 Kcal/mol	Yes
<i>Hericium erinaceus</i>	<i>Erinacine E</i>	THR -2	1	-4374 Kcal/mol	Yes
<i>Withania somnifera</i>	<i>withanolide A</i>	LEU-35 SER-34	2	-4470 Kcal/mol	Yes

**Table 4.** Structures of phytochemical compounds.

 <p>Molecular docking of Hericenone A (green stick) into a protein binding pocket. The residue THR-2 is highlighted in blue and shown with dashed lines indicating interactions with the ligand.</p>	 <p>Molecular docking of Deoxymethylcurcumin (green stick) into a protein binding pocket. The residue THR-2 is highlighted in blue and shown with dashed lines indicating interactions. A red watermark reads: "- For Evaluation Only (11 days remaining)".</p>
HERICENONE A	DEOXYMETHYLCURCUMINE
 <p>Molecular docking of Hericenone (green stick) into a protein binding pocket. The residue THR-2 is highlighted in blue and shown with dashed lines indicating interactions. A red watermark reads: "For Evaluation Only (11 days remaining)".</p>	 <p>Molecular docking of Withanolide A (green stick) into a protein binding pocket. Residues SER-34 and LEU-35 are highlighted in blue and shown with dashed lines indicating interactions.</p>
HERICENONE	WITHANOLIDE A
 <p>Molecular docking of Bisdeoxycurcumin (green stick) into a protein binding pocket. The residue PHE-61 is highlighted in red and shown with dashed lines indicating interactions.</p>	 <p>Molecular docking of Carboxaldehyde (green stick) into a protein binding pocket. Residues SER-34 and PHE-61 are highlighted in blue and shown with dashed lines indicating interactions.</p>
BISDEOXYCURCUMIN	CARBOXALDEHYDE
 <p>Molecular docking of Curcumin (green stick) into a protein binding pocket. Residues PHE-61 and SER-34 are highlighted in blue and shown with dashed lines indicating interactions.</p>	 <p>Molecular docking of Hericenone A (green stick) into a protein binding pocket. Residues HIS-27 and LEU-38 are highlighted in blue and shown with dashed lines indicating interactions. A red watermark reads: "No License File - For Evaluation Only (11 days remaining)".</p>
CURCUMIN	HERICENONE A

## 5 Conclusion

The whole genome sequence of Alzheimer disease was retrieved and based on the Feature count tool and DESeq2 tool the PCA expressed genes and heat map expressed genes were detected. Further, the expressed genes were located using the volcano plot.

As per docking studies compound *demethoxycurcumin* (DMC) has docking score of -4084 Kcal/mol with 3 interaction with protein. Further, receptor ligand binding assay studies can be performed to prove its efficacy in treating neuroblastoma.

## References

1. J. A. Miller, S. Horvath, and D. H. Geschwind, "Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 28, pp. 12698–12703, 2010.
2. Palmqvist S, Zetterberg H, Mattsson N, et al.: Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. *Neurology*. 2015; 85(14): 1240–9
3. Albert MS, DeKosky ST, Dickson D, et al.: The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011; 7(3): 270–9
4. Singh B, Parsaik AK, Mielke MM, et al.: Association of mediterranean diet with mild cognitive impairment and Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimers Dis*. 2014; 39(2): 271–82.
5. Shen J and Kelleher RJ III: The presenilin hypothesis of Alzheimer's disease: Evidence for a loss-of-function pathogenic mechanism. *Proc Natl Acad Sci USA* 104: 403-409, 2017.
6. X. Zhou and D. P. Tuck, "MSVM-RFE: extensions of SVMRFE for multiclass gene selection on DNA microarray data," *Bioinformatics*, vol. 23, no. 9, pp. 1106–1114, 2007
7. M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000
8. Levy-Lahad E, Tsuang D, Bird TD. Recent advances in the genetics of Alzheimer's disease. *J Geriatr Psychiatry Neurol* 1998;11(2):42–54.
9. Boschert U, Merlo-Pich E, Higgins G, Roses AD, Catsicas S. Apolipoprotein E expression by neurons surviving excitotoxic stress. *Neurobiol Dis* 1999;6(6):508–514.
10. Mahley RW, Weisgraber KH, Huang Y. Apolipoprotein E4: a causative factor and therapeutic target in neuropathology, including Alzheimer's disease. *Proc Natl Acad Sci U S A* 2006;103(15):5644–5651.
11. Braak H, Del Tredici K: Reply: the early pathological process in sporadic Alzheimer's disease. *Acta Neuropathol*. 2013; 126(4): 615–8.
12. Gregory, J.; Vengalasetti, Y.V.; Bredesen, D.E.; Rao, R.V. Neuroprotective Herbs for the Management of Alzheimer's Disease. *Biomolecules* 2021, 11, 543.
13. Khan, A.; Tania, M.; Liu, R.; Rahman, M.M. *Hericium erinaceus*: An edible mushroom with medicinal values. *J. Complement. Integr. Med.* 2013, 10
14. Khan, A.; Tania, M.; Liu, R.; Rahman, M.M. *Hericium erinaceus*: An edible mushroom with medicinal values. *J. Complement. Integr. Med.* 2013, 10
15. Akram, M.; Nawaz, A. Effects of medicinal plants on Alzheimer's disease and memory deficits. *Neural Regen Res*. 2017, 12, 660–670
16. F. Fahrenholz, "Alpha-secretase as a therapeutic target," *Current Alzheimer Research*, vol. 4, no. 4, pp. 412–417, 2007.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

