Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.53555/jxtvch70

IDENTIFICATION OF HIPPOCAMPUS-RELATED HUB GENES AND PATHWAYS IN ALZHEIMER'S DISEASE USING DIFFERENTIAL GENE EXPRESSION AND NETWORK ANALYSIS

Muhammad Afzal[1], Muhammad Farrukh Tahir[2*], Anam Khalid[3*], Abdullah Sethar[4], Muhammad Ali[5], Muhammad Abbas[6], Sanobar Gull[7], Hidayatullah Soomro[8], Tamseel Saleem[9], Aatif Masood Ahmad Khan[10]

[1]Department of Biochemistry, Faisalabad Medical University, Faisalabad, Pakistan.
[2*]Department of Biochemistry, University of Jhang, Jhang. Pakistan.
[3*]Department of Biochemistry, University of Jhang, Jhang, Pakistan.
[4]Livestock Breeding Services Authority (LBSA) Sindh, Livestock & Fisheries Department, Government of Sindh, Pakistan.
[5]Punjab Ranjers Teaching Hospital, Lahore, Pakistan.
[6]1st affiliated Hospital of Hebei North University Hebei Province, China.
[7]Department of Biochemistry, Faculty of Life Science, Govt College University Faisalabad.
[8]Department of Veterinary Medicine, Faculty of Veterinary Sciences SBBUVAS Sakrand, Pakistan.
[9]Department of Biochemistry and Physiology, Faculty of Animal Husbandry and Veterinary Sciences (AHVS), Sindh Agriculture University, Tandojam, Pakistan.
[10]Livestock and Dairy Development Department, Quetta, Balochistan, Pakistan.

***Corresponding Author(s):** Muhammad Farrukh Tahir & Anam Khalid
*Email(s): drfarrukh@uoj.edu.pk & anamkhalidrai@gmail.com

Abstract
Alzheimer's Disease (AD) is a complex neurodegenerative disorder characterized by cognitive decline and memory loss, with the hippocampus being significantly affected. This study aimed to identify differentially expressed genes (DEGs) in the hippocampus of AD patients using the GSE48350 dataset. Through rigorous preprocessing and quality control, we identified 250 top DEGs, including 189 upregulated and 61 downregulated genes. Protein-protein interaction (PPI) network analysis revealed SRC, BTK, and CD80 as key hub genes, indicating their central roles in AD pathology. Functional enrichment analysis of these hub genes highlighted several significantly overrepresented pathways, including Thiamine metabolism, Regulation of lipolysis in adipocytes, Axon guidance, Influenza A, Calcium signaling, and Thermogenesis. These findings underscore the critical involvement of synaptic function, neuroinflammation, and immune modulation in AD and suggest potential metabolic and neurodevelopmental links to the disease. This study provides new insights into the molecular mechanisms of AD and identifies potential therapeutic targets for future research.

Keywords: Alzheimer's Disease; GSE48350 dataset; Hub genes

Introduction

Alzheimer's disease (AD) is a devastating neurodegenerative disorder and the most common cause of dementia, accounting for 60-80% of all dementia cases worldwide [1-3]. The prevalence of AD is rising sharply with the aging population, making it one of the most critical public health challenges of our time [4]. AD is characterized by a progressive decline in cognitive function, including memory loss, impaired judgment, and changes in behavior, which severely affect the quality of life of patients and place an enormous burden on caregivers and healthcare systems [5].

The pathological hallmarks of AD include the extracellular accumulation of amyloid-beta (Aβ) plaques, the intracellular formation of neurofibrillary tangles (NFTs) composed of hyperphosphorylated tau protein, and significant neuronal loss, particularly in brain regions associated with memory and learning [6]. Among these regions, the hippocampus is critically involved in the early stages of AD and exhibits some of the most profound pathological changes [7]. The deterioration of hippocampal function is strongly correlated with the cognitive deficits observed in AD, underscoring the importance of this brain region in the disease's progression [8, 9].

Despite extensive research efforts, the precise molecular mechanisms driving AD pathogenesis remain elusive, and there are currently no disease-modifying therapies available. The heterogeneity of AD and the complex interplay between genetic, epigenetic, and environmental factors contribute to the difficulty in understanding its underlying biology. Therefore, identifying key molecular players and pathways involved in AD is crucial for the development of effective diagnostic markers and therapeutic targets.

Advancements in high-throughput technologies, such as microarrays and next-generation sequencing, have enabled researchers to investigate gene expression profiles on a genome-wide scale [10]. The GSE48350 dataset, generated from post-mortem brain tissues of AD patients, provides a valuable opportunity to study the transcriptional changes associated with AD in specific brain regions [11]. By focusing on the hippocampus, a region particularly vulnerable to AD-related pathology, we aimed to uncover differentially expressed genes (DEGs) that may play crucial roles in the onset and progression of the disease.

In this study, we performed a comprehensive bioinformatics analysis of the GSE48350 dataset to identify hippocampus-related DEGs and potential hub genes in AD patients. Through integrative approaches, including functional enrichment analysis, protein-protein interaction (PPI) network construction, and hub gene identification, we sought to elucidate the molecular mechanisms underlying AD pathogenesis. Our findings not only enhance our understanding of the complex molecular landscape of AD but also highlight specific genes that could serve as novel biomarkers or therapeutic targets for this debilitating disease.

Methodology

Data acquisition

The gene expression dataset GSE48350 was obtained from the Gene Expression Omnibus (GEO) database [12], a publicly accessible repository of high-throughput gene expression data. The dataset consists of post-mortem hippocampal tissue samples from AD patients and age-matched non-demented control individuals. The dataset includes microarray-based gene expression profiles generated using the platform GPL570, which measures the expression levels of thousands of genes across the genome (Figure 1).

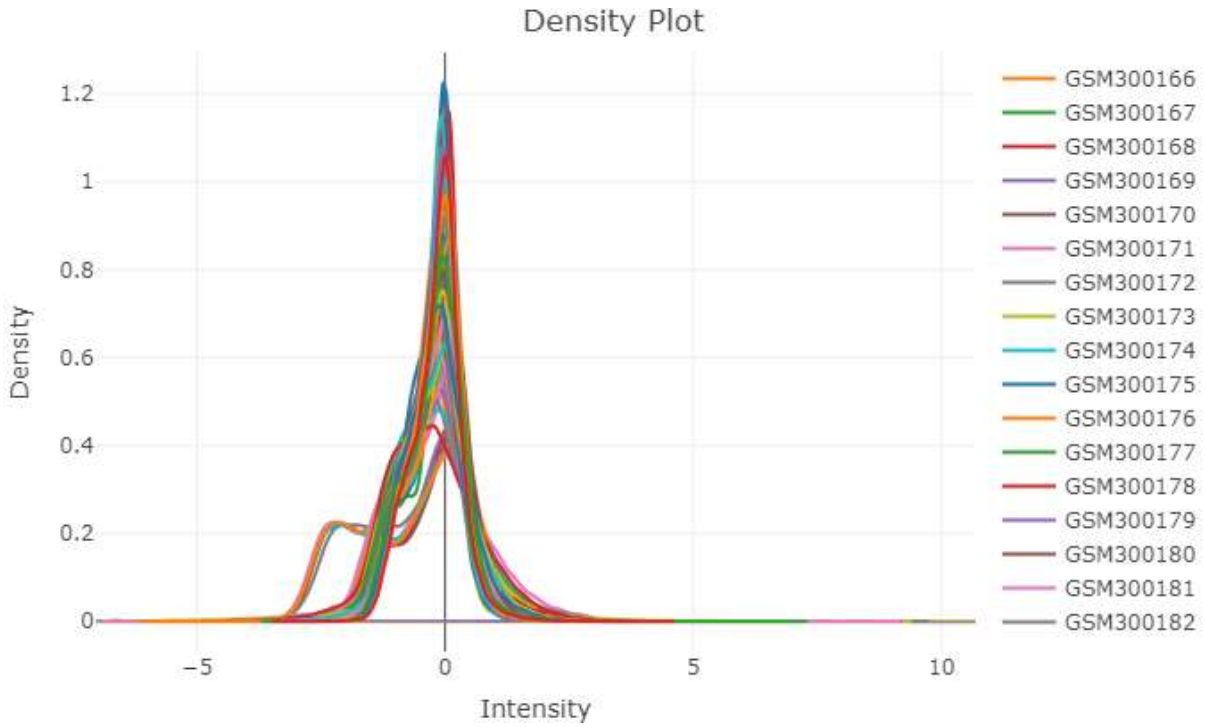

Figure 1: Expression density graph of all the AD and control samples in GSE48350 dataset.

Preprocessing and quality control

Before analysis, the raw gene expression data were subjected to preprocessing steps to ensure data quality and consistency. The preprocessing involved:

Normalization: The raw expression data were normalized using the Robust Multi-array Average (RMA) method to correct for technical variations and ensure comparability across samples.

Background correction: Non-specific background signals were subtracted from the raw data to enhance the accuracy of the gene expression measurements.

Log transformation: The expression values were log2-transformed to stabilize variance and improve the normality of the data distribution.

Batch effect correction: If necessary, batch effects due to technical differences across experimental runs were corrected using the ComBat method from the sva package in R.

Identification of DEGs

To identify genes that were differentially expressed between AD patients and control samples, we performed differential expression analysis using the limma package in R. The analysis involved:

Design matrix construction: A design matrix was constructed to model the relationship between the gene expression levels and the disease status (AD vs. control).

Linear model fitting: A linear model was fitted for each gene to assess the differential expression between the two groups.

Statistical testing: The empirical Bayes method was applied to compute moderated t-statistics and corresponding p-values for each gene.

Multiple testing corrections: To control the false discovery rate (FDR), p-values were adjusted using the Benjamini-Hochberg method. Genes with an adjusted p-value < 0.05 and an absolute log2 fold change > 1 were considered significantly differentially expressed.

Protein-Protein Interaction (PPI) network construction

To explore the interactions between the identified DEGs, a protein-protein interaction (PPI) network was constructed using the STRING database. The steps included:

Network construction: The DEGs were input into the STRING database to retrieve known and predicted PPIs with a confidence score threshold of 0.7 (high confidence).

Network visualization: The resulting PPI network was visualized using Cytoscape software, allowing for the identification of highly connected nodes (hub genes) and network modules.

Hub gene identification

Hub genes, which are central to the PPI network and may play critical roles in AD pathogenesis, were identified using network analysis tool (Cytohubba) within Cytoscape. Specifically:

Degree centrality analysis: The degree of each node (number of connections) was calculated, and genes with the highest degrees were considered potential hub genes.

Functional enrichment analysis

To gain insights into the biological significance of the identified DEGs, we performed functional enrichment analysis using the DAVID tool.

Statistical analysis

All statistical analyses were performed using R software (version 2.0). Results were considered statistically significant at $p < 0.05$ unless otherwise specified. Visualizations of the results, including volcano plots, heatmaps, and PPI networks, were generated using the ggplot2 and ComplexHeatmap packages in R.

Results

Identification of DEGs

After rigorous preprocessing and quality control of the GSE48350 dataset, differential expression analysis was conducted to identify genes that were significantly dysregulated in the hippocampus of AD patients compared to age-matched controls. A total of top 250 DEGs were identified, with 189 genes upregulated and 61 genes downregulated in AD samples (Figure 2). The selection of these top DEGs was based on an adjusted p-value < 0.05 and an absolute log2 fold change > 1, ensuring robust and biologically relevant results (Figure 2).

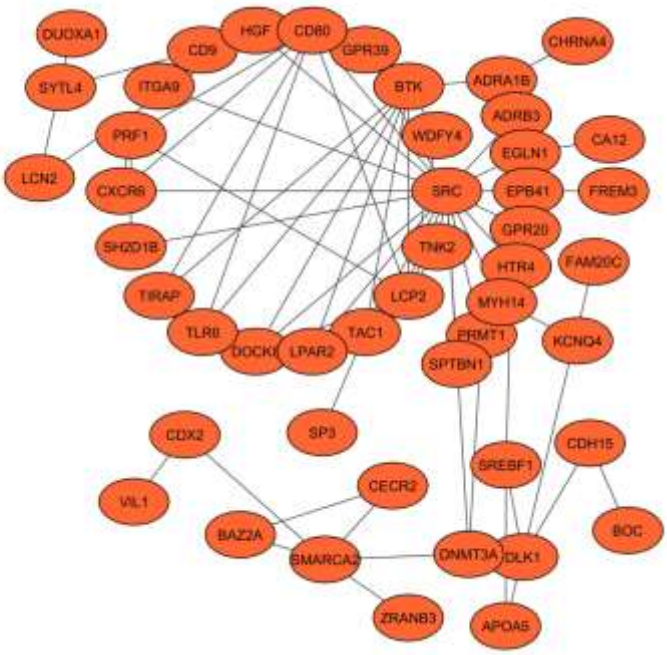

Figure 2: A PPI network of the DEGs identified in GSE48350 dataset.

PPI network construction and hub gene identification

To further understand the interactions among the DEGs, a protein-protein interaction (PPI) network was constructed using the STRING database. The network analysis revealed several highly connected nodes, indicating potential hub genes that play central roles in the network's integrity and function.

Using degree centrality and Molecular Complex Detection (MCODE) algorithms, SRC, BTK, and CD80 were identified as the top hub genes within the PPI network (Figure 3). These genes exhibited the highest degree of connectivity, suggesting their critical involvement in the molecular mechanisms of AD.

SRC (Proto-Oncogene Tyrosine-Protein Kinase Src): SRC is a non-receptor tyrosine kinase that regulates various cellular processes, including cell growth, differentiation, and survival. Its role in AD has been associated with synaptic function and plasticity, as well as the regulation of amyloid-beta production [13].

BTK (Bruton Tyrosine Kinase): BTK is a key enzyme in the signaling pathways of the immune system, particularly in B cell development and function. The involvement of BTK in neuroinflammation suggests a potential link to the chronic inflammatory response observed in AD [14].

CD80 (Cluster of Differentiation 80): CD80 is a co-stimulatory molecule expressed on antigen-presenting cells and plays a crucial role in the activation of T cells. The identification of CD80 as a hub gene highlights the importance of immune modulation in the progression of AD [15].

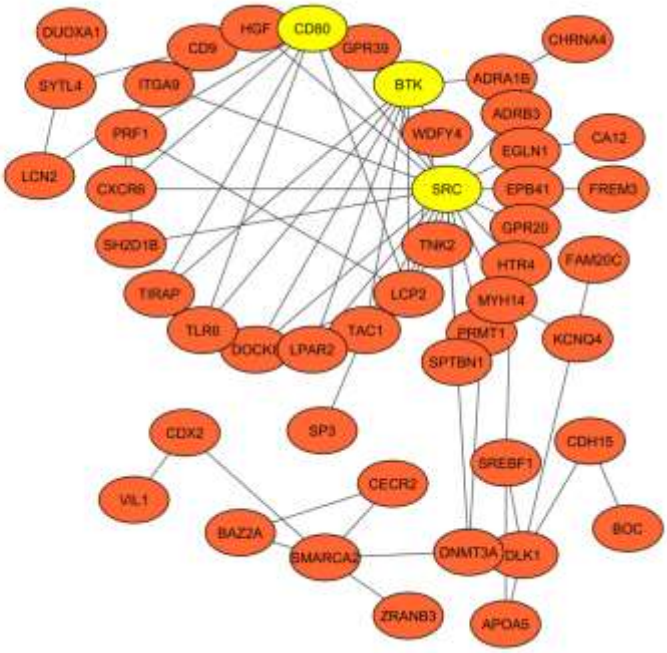

Figure 3: A PPI network of the DEGs highlighting hub genes in GSE48350 dataset.

Functional enrichment analysis

The dot plot in Figure 4depicts the results of a pathway enrichment analysis of hub genes, highlighting several key signaling pathways that are significantly overrepresented in the dataset. Among the most enriched pathways are Thiamine metabolism and Regulation of lipolysis in adipocytes, both of which exhibit high fold enrichment values, indicating a strong overrepresentation in the gene list (Figure 4). This suggests these pathways might play critical roles in the context of the study, potentially linking metabolic processes to the condition under investigation. Additionally, pathways such as Axon guidance and Influenza A also show substantial enrichment, suggesting their involvement in neurobiological processes or immune responses (Figure 4). Despite having lower fold enrichment, pathways like Calcium signaling and Thermogenesis still display statistical significance, indicating they may contribute in more specific or nuanced ways to the overall biological context (Figure 4).. Collectively, these insights provide valuable directions for further exploration, particularly in understanding the underlying mechanisms of AD, and may help identify potential therapeutic targets.

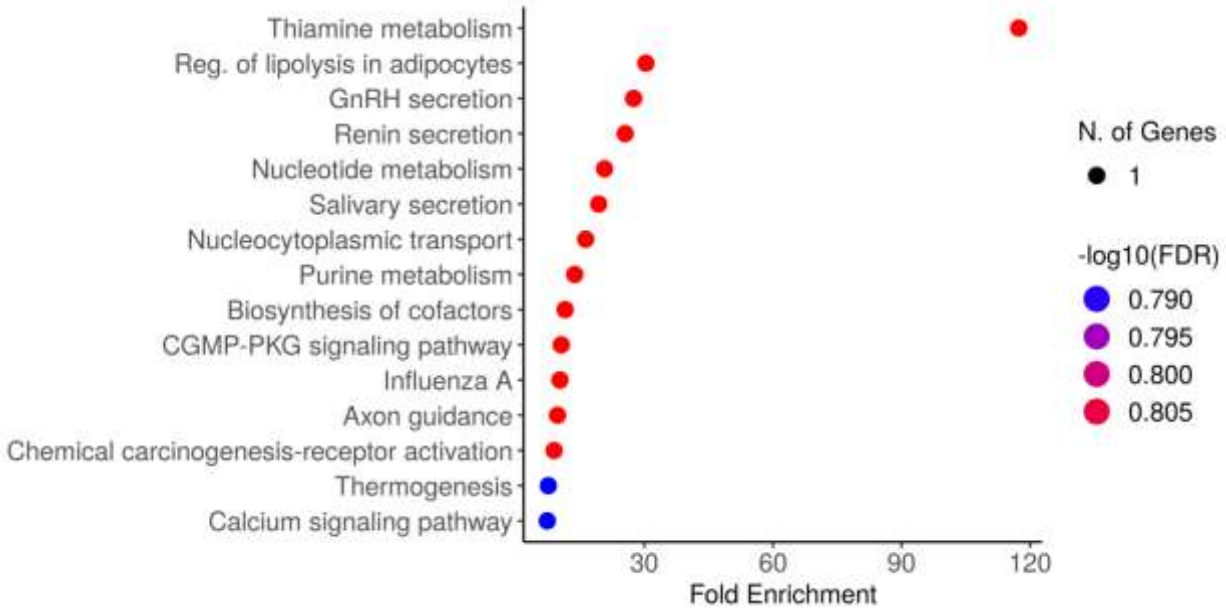

Figure 4: Hub genes-associated pathways.

Discussion

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder characterized by cognitive decline, memory loss, and behavioral changes [16]. It is the most common form of dementia and is primarily associated with the accumulation of amyloid-beta plaques and neurofibrillary tangles in the brain [17]. The hippocampus, a critical region for memory formation and spatial navigation, is significantly affected early in the disease process, leading to its pronounced involvement in AD pathology [18]. The complex pathogenesis of AD involves a combination of genetic, environmental, and lifestyle factors, with emerging evidence highlighting the role of neuroinflammation and dysregulated signaling pathways.

In this study, we analyzed the GSE48350 dataset to identify differentially expressed genes (DEGs) in the hippocampus of Alzheimer's Disease (AD) patients. Rigorous preprocessing and quality control resulted in the identification of 250 top DEGs, with 189 genes upregulated and 61 genes downregulated in AD samples, based on an adjusted p-value < 0.05 and an absolute log2 fold change > 1. This robust selection of DEGs provides a comprehensive view of gene expression alterations associated with AD.

The protein-protein interaction (PPI) network analysis of these DEGs revealed several highly connected nodes, identifying SRC, BTK, and CD80 as key hub genes. SRC, a non-receptor tyrosine kinase, plays a crucial role in regulating synaptic function and plasticity, and its involvement in AD is consistent with previous research linking SRC activity to synaptic dysfunction and amyloid-beta regulation [19]. Similarly, BTK, an enzyme integral to immune signaling pathways, aligns with studies highlighting neuroinflammation as a significant contributor to AD pathology [20]. The identification of CD80, a co-stimulatory molecule crucial for T cell activation, underscores the importance of immune modulation in AD progression [21].

Functional enrichment analysis of the hub genes revealed several significantly overrepresented pathways, including Thiamine metabolism, Regulation of lipolysis in adipocytes, Axon guidance, Influenza A, Calcium signaling, and Thermogenesis. The enrichment of pathways related to Thiamine metabolism and lipid regulation supports the hypothesis that metabolic disturbances are closely linked to AD. The involvement of pathways such as Axon guidance and Influenza A suggests potential links between neurodevelopmental processes and immune responses in AD. While pathways like Calcium signaling and Thermogenesis exhibited lower fold enrichment, their statistical significance indicates their potential contribution to AD pathology.

Conclusion

Overall, these findings provide valuable insights into the molecular mechanisms underlying AD, highlighting the critical roles of synaptic function, neuroinflammation, and immune modulation. The enrichment of metabolic and neurodevelopmental pathways further emphasizes the complexity of AD and offers new directions for future research. Further validation of these results could pave the way for identifying novel therapeutic targets and improving our understanding of AD's etiology and progression.

Conflict of interest

None

Acknowledgement

None

References

[1] Birajdar SV, Mulchandani M, Mazahir F and Yadav AK. Dementia and neurodegenerative disorder: An introduction. In: editors. Nanomedicine-based approaches for the treatment of dementia. Elsevier; 2023. p. 1-36.
[2] Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, Cummings J and van der Flier WM. Alzheimer's disease. The Lancet 2021; 397: 1577-1590.
[3] Penke B, Szűcs M and Bogár F. New pathways identify novel drug targets for the prevention and treatment of Alzheimer's disease. International journal of molecular sciences 2023; 24: 5383.
[4] Revi M. Alzheimer's disease therapeutic approaches. GeNeDis 2018: Genetics and Neurodegeneration 2020; 105-116.
[5] Islam Bu, Jabir NR and Tabrez S. The role of mitochondrial defects and oxidative stress in Alzheimer's disease. Journal of Drug Targeting 2019; 27: 932-942.
[6] Sosa-Ortiz AL, Acosta-Castillo I and Prince MJ. Epidemiology of dementias and Alzheimer's disease. Archives of medical research 2012; 43: 600-608.
[7] Mufson EJ, Mahady L, Waters D, Counts SE, Perez SE, DeKosky S, Ginsberg SD, Ikonomovic MD, Scheff S and Binder L. Hippocampal plasticity during the progression of Alzheimer's disease. Neuroscience 2015; 309: 51-67.
[8] Mu Y and Gage FH. Adult hippocampal neurogenesis and its role in Alzheimer's disease. Molecular neurodegeneration 2011; 6: 1-9.
[9] Duyckaerts C and Dickson D. Neuropathology of Alzheimer's disease and its variants. Neurodegeneration: the molecular pathology of dementia and movement disorders 2011; 62-91.
[10] Morrison JH and Hof PR. Selective vulnerability of corticocortical and hippocampal circuits in aging and Alzheimer's disease. Progress in brain research 2002; 136: 467-486.
[11] Braak H and Del Tredici K. Neuroanatomy and pathology of sporadic Alzheimer's disease. 2014;
[12] Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau W-C, Ledoux P, Rudnev D, Lash AE, Fujibuchi W and Edgar R. NCBI GEO: mining millions of expression profiles—database and tools. Nucleic acids research 2005; 33: D562-D566.
[13] de Pins B, Mendes T, Giralt A and Girault J-A. The non-receptor tyrosine kinase Pyk2 in brain function and neurological and psychiatric diseases. Frontiers in Synaptic Neuroscience 2021; 13: 749001.
[14] McDonald C, Xanthopoulos C and Kostareli E. The role of Bruton's tyrosine kinase in the immune system and disease. Immunology 2021; 164: 722-736.
[15] Mir MA and Agrewala JN. Influence of CD80 and CD86 co-stimulation in the modulation of the activation of antigen presenting cells. Current Immunology Reviews 2007; 3: 160-169.
[16] Buchake V, Muthal A, Saudagar R and Bachhav R. A neurodegenerative disorder-Alzheimer disease: A treatise. Research Journal of Pharmacology and Pharmacodynamics 2010; 2: 268-273.

[17] Trojsi F, Christidi F, Migliaccio R, Santamaría-García H and Santangelo G. Behavioural and cognitive changes in neurodegenerative diseases and brain injury. Behavioural neurology 2018; 2018:
[18] Cummings JL. Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis. Neurobiology of aging 2000; 21: 845-861.
[19] Marcello E, Epis R, Saraceno C and Di Luca M. Synaptic dysfunction in Alzheimer's disease. Synaptic Plasticity: Dynamics, Development and Disease 2012; 573-601.
[20] Albertini C. Development of new chemical entities to target neuroinflammatory pathways. 2022;
[21] Grosche L, Knippertz I, König C, Royzman D, Wild AB, Zinser E, Sticht H, Muller YA, Steinkasserer A and Lechmann M. The CD83 molecule–an important immune checkpoint. Frontiers in Immunology 2020; 11: 721.