

What Do Machines Tell us About Dementia? Machine Learning Applied to Aging, Dementia and Traumatic Brain Injury Study

Denis Arthur Pinheiro Moura (✉ drmouradap@gmail.com)

Universidade Federal de Pernambuco <https://orcid.org/0000-0002-7462-5455>

Joao Ricardo Mendes de Oliveira

Universidade Federal de Pernambuco

Research Article

Keywords: Transcriptomics, Artificial Intelligence, Genomic Data Science, Classification model

Posted Date: September 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-840907/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Dementia, a syndrome characterized by the progressive deterioration of memory and cognition, arises from different pathologies, with Alzheimer's Disease (AD) its most common cause. Patterns of gene expression during dementia of different etiologies may function as generalist biomarkers of the condition. We used RNA-Seq data from the Allen Dementia and Traumatic Brain Injury Study (ADTBI) to identify differentially expressed genes in brains with dementia. Machine Learning algorithms Decision Trees (DT) and Random Forest (RF) were used to create models to identify dementia samples based on their gene expression profile. Importance analyses were conducted to identify the most relevant genes in each classification model. A total of 1629 differentially expressed (DE) genes were found in brains with the condition. Gene PAN3-AS1 was the only DE gene across more than three brain regions. The artificial intelligence models were capable of identifying correctly up to 92.85% of dementia samples. Our analyses provide interesting insights regarding using brain-specific gene expression profiles as biomarkers of dementia, identifying genes possibly involved with dementia, and guiding future studies in prediction and early identification of the syndrome.

Introduction

Dementia is one of the most challenging issues in aging populations. Its prevalence is between 5–7% of the elderly, affecting around 50 million people worldwide (Prince et al, 2013). Dementia is a clinical syndrome caused by different etiologies, including complex, multigenic, and even epigenetic causes. Alzheimer's Disease (AD), vascular dementia, Lewy body disease, frontotemporal dementia, and Creutzfeldt-Jakob disease are among its most common causes. Traumatic brain injury and alcohol consumption are also frequent causes of the condition, especially in younger people (Harvey, Skelton-Robinson and Rossor, 2003; Shively et al, 2012).

Machine Learning (ML) computes probabilities and classifies data based on statistics. This method has been proved invaluable for disease classification based on images, with growing precision and predictive power (Rizk-Jackson et al, 2011; Liu et al, 2015). ML is currently used to aid the diagnostics of heart and liver diseases, diabetes, dengue, and many others (Fatima and Pasha, 2017). RNA-Seq is a high-throughput technology capable of assessing transcriptomic profiles (Courtney et al, 2010). The combination of RNA-Seq with ML models can help us identify potential biomarkers of diseases, and assess the expression of these genes to predict the disease development (Singireddy et al, 2015).

Changes in gene expression caused by neurodegenerative diseases can highlight the pathogenic mechanisms (Arneson et al, 2018). Identifying these changes may help early predict developing diseases by the identification of gene expression patterns related to the pathologies. And to further understand the brain molecular networks involved in the process of memory and dementia, we investigated the transcriptome patterns of healthy and dementia RNA-Seq data from the Aging, Dementia, and Traumatic Brain Injury Study (ADTBI, Miller, et al, 2017). We used the set of differentially expressed (DE) genes from

dementia samples to build ML models capable of predicting the condition. We were capable of identifying DE genes during dementia and that can be used as biomarkers of the condition.

Methods

Data Acquisition

RNA-Seq data of human brains with and without dementia were extracted from the ADTBI (aging.brain-map.org). This database contains data of 50282 transcripts of 107 brain transcriptomes. Of the dementia samples, thirty were from AD, and 21 from different etiologies (vascular, multiple etiologies, others, and unknown causes), and 56 samples are from non-dementia brains. The parietal cortex (PC), temporal cortex (TC), forebrain white matter (FWM), and hippocampus are the four sampled areas in the ADTBI and included in this analysis.

Data Cleansing and Treatment

R language was used to clean and treat the datasets. Data from three files were merged to create the final dataset (Donor information, Gene Expression Matrix, and Column Samples). We used the column DSM IV (clinical diagnosis) from within the donor information table to classify samples. We considered all sources of dementia as just "Dementia", and the remaining samples as "No Dementia". Non-normalized Fragments per Kilobase of sequence per million mapped reads (FPKM) data were loaded and DESeq2 library (Love, 2017) to extract differential expression data of the genes. Python module Pandas was used to further filter the DE genes by false discovery rate (FDR) adjusted p-values, considering significant DE genes only those with p-adjusted below 0.05.

Disease modeling

To check whether the dementia samples would cluster together, we have drawn a heatmap with the 1000 lowest false discovery rate-adjusted p-values for each data stratification.

Transcriptome patterns of DE genes were used as input for the classification models Decision Tree (DT) using R library rpart (Therneau and Atkinson, 2019), and Random Forest (RF), using R library randomForest (Liaw and Wiener, 2002).

The first model considered all DE genes as classification features. Feature selection was made using all genes whose attribute importance was superior to 3 defined by the random.forest.importance method from randomForest library. This step was done for each brain region, and new RF models were created using only these selected genes.

For the RF models, the number of trees (*ntree*) and the number of variables at each split (*mtry*) were adjusted based on the Out-of-bag error, and the proximity was calculated and visualized using multidimensional scaling.

Model metrics analysis

For each created model, a confusion matrix was calculated, containing the total number of true positives (correctly identified dementia samples), true negatives (correctly identified non-dementia samples), false positives (non-dementia samples classified as dementia), and false negatives (dementia samples classified as non-dementia). From these confusion matrices, the accuracy, precision, recall, and F1 score metrics were calculated for each model. Accuracy is calculated as the ratio of correctly predicted observations (True positives + True Negatives) over the total number of observations. Precision was calculated as the ratio of true positives over the total number of positive observations and relates to a low false-positive rate. Recall, also known as sensitivity, is the ratio of true positives overall real positives (true positives and false negatives). And finally, the F1 Score measures the test's accuracy based on the average between precision and recall. All metrics were added to a single comparison data frame.

Data Visualization

R library EnhancedVolcano (Blighe, Rana, and Lewis, 2020) was used to draw the volcano plots showing DE genes in each data stratification. R library VennDiagram (Chen, 2018) was used to plot a venn diagram of the four datasets showing DE genes shared between the brain regions. R library ggplot2 was used to create and visualize DT trees, the Out-of-bag errors, and multidimensional scaling plots. R library volcanoPlot was used to show the significant DE genes.

Functional Enrichment and Protein Network Analysis

STRING (Szklarczyk et al, 2019) database API was accessed using the lists of DE genes for each stratification as input. First, each gene was mapped with a non-ambiguous identifier, using "get_string_ids" method, returning only the best matches for each gene name. The network images were generated using the "network" method, with nodes connected based on confidence, and with a score of at least 750, without secondary nodes. Finally, the method "function_annotation" was used for the functional enrichment of the networks, considering all annotated functions shared by at least 5 nodes based on Gene Ontology (The Gene Ontology Consortium, 2019), Reactome (Jassal et al, 2020) or SMART (Letunic e Bork, 2017). The script to access the API can be found in the author's repository (github.com/mouradap/StringAPI).

Results

Data Clustering

Our data clustering analysis revealed that taking the 1000 genes with the lowest adjusted p-values, the dementia samples would group closer than with any other non-dementia sample (Fig. 1). This clustering is an important indicator that the models are able to find patterns and correctly classify the samples.

Differential Expression of all samples

We found differential expression patterns in brains with dementia, with a total of 1629 DE genes between non-dementia samples. There were a total of 1550 overexpressed genes across the brain samples, while only 79 were under expressed. Of these, 382 DE genes were found in the TC, 377 in the PC, 376 in the FWM, and 502 in the hippocampus. Figure 2 presents the amount of DE genes through a volcano plot in the FWM as an example. Only genes with false discovery rate-adjusted p-values below 0.05 were considered significantly DE (green and gray in the figure).

A Venn diagram was drawn to identify DE genes in all brain regions (Fig. 3). This revealed a total of 24 DE genes shared by at least 2 brain regions. Except for under expressed genes *TNFRSF21* in the PC (-3.37-fold) and *OR4C5* in the FWM (-3.44 fold), all genes were overexpressed by at least 2.74-fold in the dementia samples. Gene *PAN3-AS1* was the only DE gene shared by three brain areas (hippocampus, FWM, and PC), and overexpressed by over 3-fold.

Machine Learning Models

All DT based models had precision over 80%, with Hippocampus model with 100% precision, and the lowest false-negative rate (Fig. 4A). The Whole-Brain model had a precision of 93.10% (Fig. 4B). The TC had the highest F1 Score between all models (80%) (Fig. 4C).

The RF models overall did not score as well as the DT models. The Hippocampus model created with feature selection had the highest Recall (sensitivity) between all models (77.27%), with the lowest number of mislabeling of true dementia samples. These comparisons are presented in Table 1.

Feature selection Genes

The attribute importance for feature selection revealed several genes whose patterns could be particularly useful in the classification of dementia samples. Models for each brain region had different numbers of selected features. The TC had 9 chosen genes (*HDAC11*, *CNGB3*, *TRIM56*, *TBC1D13*, *PDLIM7*, *PPP1CB*, *ADGRB3*, *GET4*, and *TYK2*, in ascending order of importance), and the FWM had 7 chosen genes (*ARL10*, *OBSCN*, *STX7*, *VAPA*, *CRHR1*, *SAR1AP3*, and *MARCH2*). The hippocampus and the PC both had 3 genes selected based on the attribute importance: *TMEM106B*, *PCCB*, and *ZNF358* for PC, and *FPGT*, *TNNI3K*, *CRB2*, and *ST6GALNAC2* for the hippocampus.

Protein network analyzes

All DE gene lists were capable of recreating protein functional interaction networks from the STRING database. The most connected network was created from DE genes from the hippocampus, with 45 direct nodes connected to the central *UBA52* gene, with 7 adjacent protein clusters (Fig. 5).

Functional enrichment

The most-reported process in Gene Ontology in all sample stratifications was the Cellular Process, shared by around 78% of DE genes. Biological Regulation was the second most reported process, shared by 60%

of the genes. In Reactome pathways, Signal Transduction was the most present process, followed by Immune System, Metabolism and Protein Metabolism (present in 19%, 12%, 12%, and 10% of the genes, respectively). Specific pathways change in each brain region, with relatively fewer participating genes in each. 'Asparagine-bound N-glycosylation', 'Signal Transduction Diseases', and 'MAPK Family Signaling Cascades' belong in the pathways of the Hippocampus DE genes. The complete list of functionally enriched DE genes is available in the supplementary material (Supplementary 1–8).

Discussion

The continuous growth of human elderly populations makes understanding, preventing, and treating age-related conditions a priority in research and policies for public health. This study attempted to assess dementia's neuroanatomical expression profile shared throughout multiple etiologies. Our most efficient models were able to correctly identify up to 92.85% of dementia samples. These models prove that there are differential expression patterns in brains with some type of dementia. While these models might not be used as dementia predicting tools, they contain important information about genes in which a differential expression profile correlates to brains with dementia. This is consistent with a recent comparative transcriptome study between AD, Frontotemporal Dementia, and Huntington's Disease (Annese et al, 2018; Stopa et al, 2018). These models contain important information regarding DE genes during dementia, pointing to potential biomarkers for the condition.

ML models can predict dementia in RNA-Seq samples. This indicates that we can utilize brain-specific gene expression profiles as disease biomarkers. Changes in TC, PC, and FWM are sufficient to be modeled and used to compare and classify dementia. Resultant genes from the importance analysis can become potential candidate genes for the development of a gene expression panel that could work as a type of dementia indicator.

Our results show that differences in transcriptome patterns between healthy and dementia brains were most significant in the Hippocampus. This was expected since there is strong evidence of hippocampus involvement in the origin and development of dementia. Hippocampus volume loss and shape change is a common feature to distinguish healthy and Alzheimer's Disease dementia brains (Wang et al, 2003). This phenotype can also be a potential predictor of future dementia (Apostolova et al, 2010). In addition to the atrophy, molecular changes in the dementia hippocampus are also reported. The presence of α -synuclein, of ubiquitin-immunoreactive inclusions, and reduced choline acetyltransferase are molecular aspects of the hippocampus in dementia of different etiologies (Galvin et al, 1999).

Most DE genes are involved in trivial cellular processes, such as metabolism and protein regulation. Proteins involved in the immune response are also overexpressed in this condition, pointing to the role of the innate immune system in neurodegeneration. This hypothesis is reported by Richards and colleagues (2016), which states that neurodegenerative disease is the result of chronic activation of an innate surveillance pathway.

DE genes shared by the brain regions during dementia are linked to collagen metabolism, tyrosin kinases, adaptive immunity, protein digestion and absorption, and osteoclast differentiation (Sup. Table 1). Genes *SYK* and *TYK2* participate in most of these molecular functions and are differentially expressed in Temporal Cortex, Parietal Cortex, and Temporal Cortex and Hippocampus, respectively. Indeed, *SYK* appears to be directly involved in the susceptibility of vascular dementia (Kim, Kong, and Lee, 2013). In addition, *TYK2* participates in the *TYK2/STAT3* signaling pathway that mediates β -amyloid-induced neuronal cell death (Wan et al, 2010). In the Decision Tree classification model, overexpression of *TYK2* (over 7-fold) was the most important feature distinguishing dementia and no dementia brains in our study.

Other DE genes shared by more than one region are reported in the context of dementia. *TNFRSF21* is known as death receptor 6. This gene is differentially expressed in the Parietal Cortex and Temporal Cortex in dementia brains and is involved in neuronal apoptosis by binding to β -amyloid precursor protein and recruiting Caspase 6 (Nikolaev et al, 2009). Gene *TIMM23B* codifies an inner mitochondrial membrane protein, which is also associated with App plaques in AD (Heinemeyer et al, 2019).

RNA genes and pseudogenes were also important features in distinguishing dementia brains in our models. These have little to no physiological information available yet. Between these genes, some have an association with neurodegenerative conditions. *RPS27AP4* is a pseudogene differentially expressed in both Temporal Cortex and Hippocampus is possibly involved with schizophrenia (Xu et al, 2013). Lastly, *MIR181C* gene is DE in both Parietal Cortex and Hippocampus, and is involved with microglia-mediated apoptosis, increased amyloid-beta plasma levels with aging, and cognitive impairment (Fang et al, 2017; Crespo, Atienza e Cantero, 2019).

Although the deep brain tissues used in this study are not accessible in exploratory examinations, this effort was important to reveal several transcriptomic aspects of the syndrome. The predictive power of these models reveals key molecules that may be potentially involved in the development and progress of the condition. Furthermore, we demonstrated that ML algorithms can be used to classify disease samples based on transcriptome data. Future studies in the field can allow the development of dementia early detection tools, promoting healthier aging for the populations.

Declarations

Ethical Statement

Ethics approval and consent to participate

Not Applicable

Consent for publication

Not Applicable

Availability of data and materials

All data used in the production of this paper is publicly available at the Allen Insitute's Aging Dementia and Traumatic Brain Injury Study website.

Code Availability

All data and scripts underlying this article are available in GitHub repository rnaSeqDementia at <https://github.com/mouradap/rnaSeqDementia>.

Competing interests

Conflict of Interest The authors declare that they have no conflict of interest.

Funding

This project was funded by the Brazilian funding agency FACEPE (IBPG-0465-2.02/17).

Authors' contributions

Denis Moura was responsible for all the data operations and analytics in this paper, and João Oliveira contributed to the ideation and support throughout the entire project.

Acknowledgements

We would like to acknowledge the Allen Institute for Brain Science for the public availability of their datasets, which allowed this and many other research projects in the advancement of neurogenetics.

References

1. Annese, A., Manzari, C., Lionetti, C., Picardi, E., Horner, D. S., Chiara, M., Caratozzolo, M. F., Tullo, A., Fosso, B., Pesole, G., & D'Erchia, A. M. (2018). Whole transcriptome profiling of Late-Onset Alzheimer's Disease patients provides insights into the molecular changes involved in the disease. *Scientific Reports*, 8(1), 4282. <https://doi.org/10.1038/s41598-018-22701-2>
2. Apostolova, L. G., Mosconi, L., Thompson, P. M., Green, A. E., Hwang, K. S., Ramirez, A., Mistur, R., Tsui, W. H., & de Leon, M. J. (2010). Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. *Neurobiology of Aging*, 31(7), 1077–1088. <https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2008.08.008>
3. Arneson, D., Zhang, Y., Yang, X., & Narayanan, M. (2018). Shared mechanisms among neurodegenerative diseases: from genetic factors to gene networks. *Journal of Genetics*, 97(3), 795–806. <https://pubmed.ncbi.nlm.nih.gov/30027910>Courtney, E., Kornfeld, S., Janitz, K., & Janitz, M. (2010). Transcriptome profiling in neurodegenerative disease. *Journal of Neuroscience Methods*, 193(2), 189–202. <https://doi.org/10.1016/j.jneumeth.2010.08.018>

4. Blighe, K. Rana, S. and Lewis, M. (2020). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. *R package version 1.6.0*.
<https://github.com/kevinblighe/EnhancedVolcano>
5. Chen, H. (2018). VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.6.20. <https://CRAN.R-project.org/package=VennDiagram>
6. Fang, C., Li, Q., Min, G., Liu, M., Cui, J., Sun, J., & Li, L. (2017). MicroRNA-181c Ameliorates Cognitive Impairment Induced by Chronic Cerebral Hypoperfusion in Rats. *Molecular Neurobiology*, 54(10), 8370–8385. <https://doi.org/10.1007/s12035-016-0268-6>
7. Fatima, M. and Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *Journal of Intelligent Learning Systems and Applications*, 9(01), 1.
8. Galvin, J. E., Uryu, K., Lee, V. M. Y., & Trojanowski, J. Q. (1999). Axon pathology in Parkinson's disease and Lewy body dementia hippocampus contains α -, β -, and γ -synuclein. *Proceedings of the National Academy of Sciences of the United States of America*, 96(23), 13450–13455. <https://doi.org/10.1073/pnas.96.23.13450>
9. Gene Ontology Consortium (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1), D330–D338.
10. Harvey, R. J., Skelton-Robinson, M., & Rossor, M. N. (2003). The prevalence and causes of dementia in people under the age of 65 years. *Journal of Neurology, Neurosurgery, and Psychiatry*, 74(9), 1206–1209. <https://doi.org/10.1136/jnnp.74.9.1206>
11. Heinemeyer, T., Stemmet, M., Bardien, S., & Neethling, A. (2019). Underappreciated Roles of the Translocase of the Outer and Inner Mitochondrial Membrane Protein Complexes in Human Disease. *DNA and Cell Biology*, 38(1), 23–40.
12. Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., & Haw, R. (2020). The reactome pathway knowledgebase. *Nucleic Acids Research*, 48(D1), D498–D503.
13. Kim, Y., Kong, M., & Lee, C. (2013). Association of intronic sequence variant in the gene encoding spleen tyrosine kinase with susceptibility to vascular dementia. *The World Journal of Biological Psychiatry*, 14(3), 220–226.
14. Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*, 46(D1), D493–D496.
15. Liaw, A. and Wiener, M. (2002). Classification and Regression by randomForest. *R News* 2(3), 18–22.
16. Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Feng, D., & Fulham, M. J. (2015). Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Transactions on Bio-Medical Engineering*, 62(4), 1132–1140. <https://doi.org/10.1109/TBME.2014.2372011>
17. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>

18. Miller, J. A., Guillozet-Bongaarts, A., Gibbons, L. E., Postupna, N., Renz, A., Beller, A. E., Sunkin, S. M., Ng, L., Rose, S. E., & Smith, K. A. (2017). Neuropathological and transcriptomic characteristics of the aged brain. *Elife*, 6, e31126.
19. Nikolaev, A., McLaughlin, T., O'Leary, D. D. M., & Tessier-Lavigne, M. (2009). APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature*, 457(7232), 981–989. <https://doi.org/10.1038/nature07767>
20. Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, 9(1), 63-75.e2. <https://doi.org/10.1016/j.jalz.2012.11.007>
21. Richards, R. I., Robertson, S. A., O'Keefe, L. V, Fornarino, D., Scott, A., Lardelli, M., & Baune, B. T. (2016). The Enemy within: Innate Surveillance-Mediated Cell Death, the Common Mechanism of Neurodegenerative Disease. *Frontiers in Neuroscience*, 10, 193. <https://doi.org/10.3389/fnins.2016.00193>
22. Rizk-Jackson, A., Stoffers, D., Sheldon, S., Kuperman, J., Dale, A., Goldstein, J., Corey-Bloom, J., Poldrack, R. A., & Aron, A. R. (2011). Evaluating imaging biomarkers for neurodegeneration in pre-symptomatic Huntington's disease using machine learning techniques. *NeuroImage*, 56(2), 788–796. <https://doi.org/10.1016/j.neuroimage.2010.04.273>
23. Shively, S., Scher, A. I., Perl, D. P., & Diaz-Arrastia, R. (2012). Dementia resulting from traumatic brain injury: what is the pathology? *Archives of Neurology*, 69(10), 1245–1251. <https://doi.org/10.1001/archneurol.2011.3747>
24. Singireddy, S., Alkhateeb, A., Rezaeian, I., Rueda, L., Cavallo-Medved, D., & Porter, L. (2015). Identifying differentially expressed transcripts associated with prostate cancer progression using RNA-Seq and machine learning techniques. *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 1–5.
25. Stopa, E. G., Tanis, K. Q., Miller, M. C., Nikonova, E. V, Podtelezchnikov, A. A., Finney, E. M., Stone, D. J., Camargo, L. M., Parker, L., Verma, A., Baird, A., Donahue, J. E., Torabi, T., Eliceiri, B. P., Silverberg, G. D., & Johanson, C. E. (2018). Comparative transcriptomics of choroid plexus in Alzheimer's disease, frontotemporal dementia and Huntington's disease: implications for CSF homeostasis. *Fluids and Barriers of the CNS*, 15(1), 18. <https://doi.org/10.1186/s12987-018-0102-9>
26. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., & Bork, P. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1), D607–D613.
27. Therneau, T. and Atkinson, B. (2019). rpart: Recursive Partitioning and Regression Trees. *R package version 4.1-15*. <https://CRAN.R-project.org/package=rpart>
28. Vijayarani, S., & Dhayanand, S. (2015). Liver disease prediction using SVM and Naïve Bayes algorithms. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 4(4), 816–820.

29. Wan, J., Fu, A. K. Y., Ip, F. C. F., Ng, H.-K., Hugon, J., Page, G., Wang, J. H., Lai, K.-O., Wu, Z., & Ip, N. Y. (2010). Tyk2/STAT3 signaling mediates beta-amyloid-induced neuronal cell death: implications in Alzheimer's disease. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(20), 6873–6881. <https://doi.org/10.1523/JNEUROSCI.0519-10.2010>
30. Wang, L., Swank, J. S., Glick, I. E., Gado, M. H., Miller, M. I., Morris, J. C., & Csernansky, J. G. (2003). Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. *NeuroImage*, 20(2), 667–682. [https://doi.org/https://doi.org/10.1016/S1053-8119\(03\)00361-6](https://doi.org/https://doi.org/10.1016/S1053-8119(03)00361-6)
31. Xu, C., Aragam, N., Li, X., Villa, E. C., Wang, L., Briones, D., Petty, L., Posada, Y., Arana, T. B., Cruz, G., Mao, C., Camarillo, C., Su, B. Bin, Escamilla, M. A., & Wang, K. (2013). BCL9 and C9orf5 are associated with negative symptoms in schizophrenia: meta-analysis of two genome-wide association studies. *PloS One*, 8(1), e51674–e51674. <https://doi.org/10.1371/journal.pone.0051674>

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Supplementary Table

Supplementary Table 1 is not available with this version.

Figures

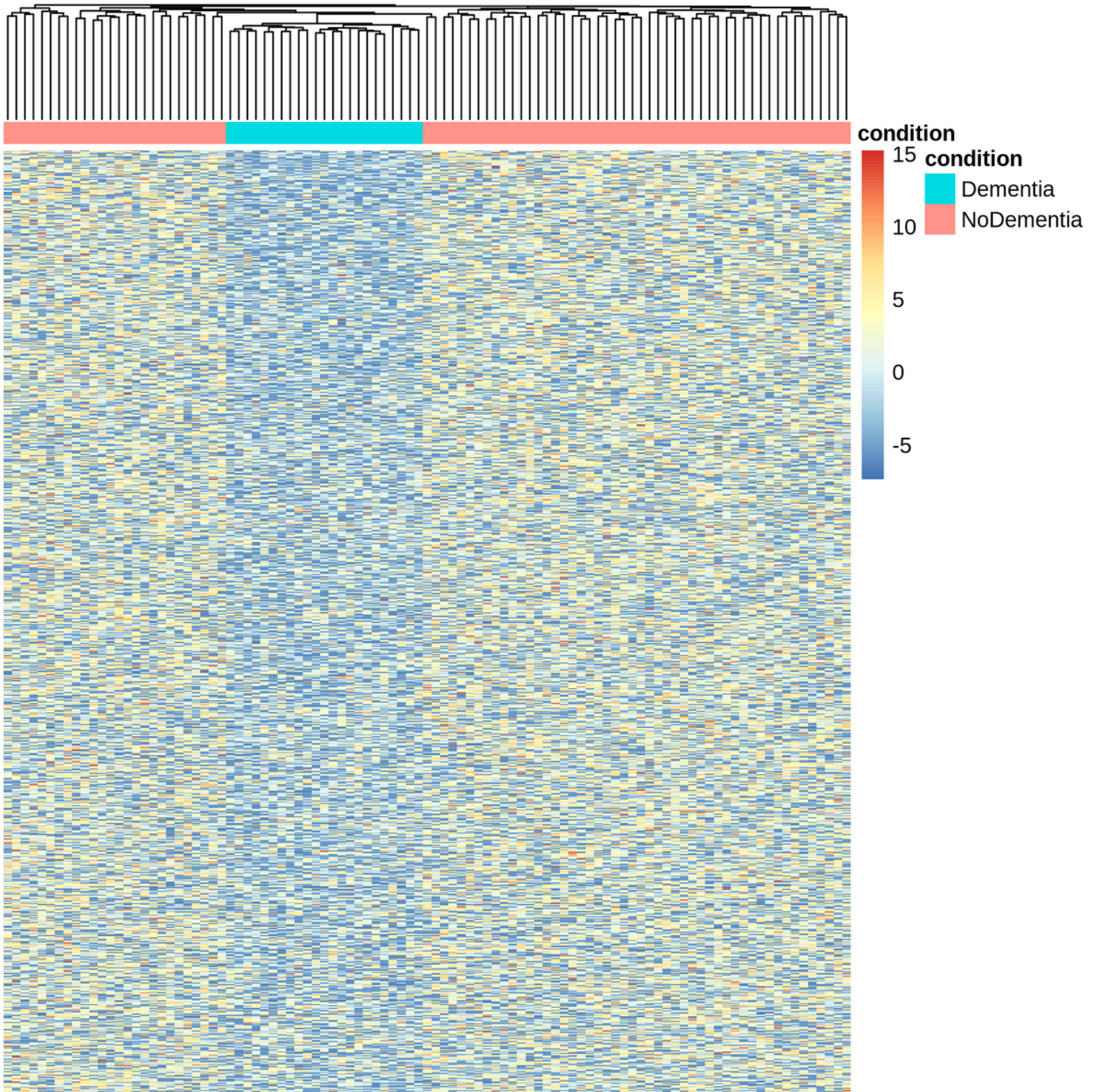
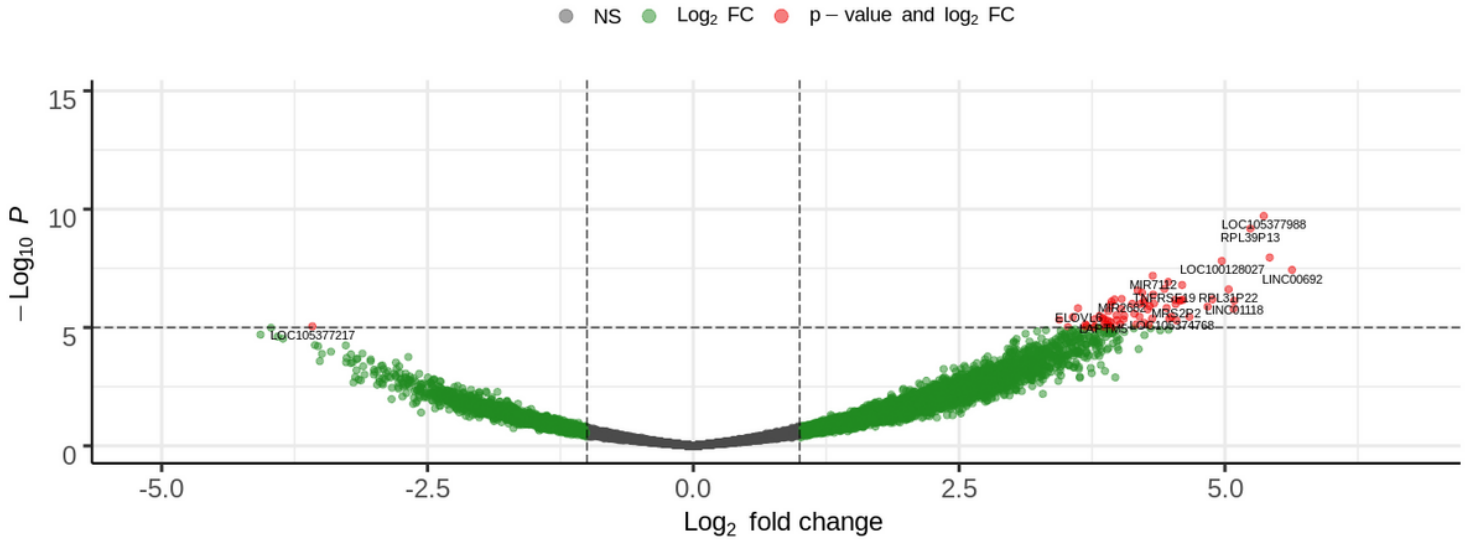


Figure 1

Temporal Cortex Differentially Expressed genes in Dementia

Volcano plot

EnhancedVolcano



Total = 50281 variables

Figure 2

enhanced volcano plot of DE genes in the temporal cortex

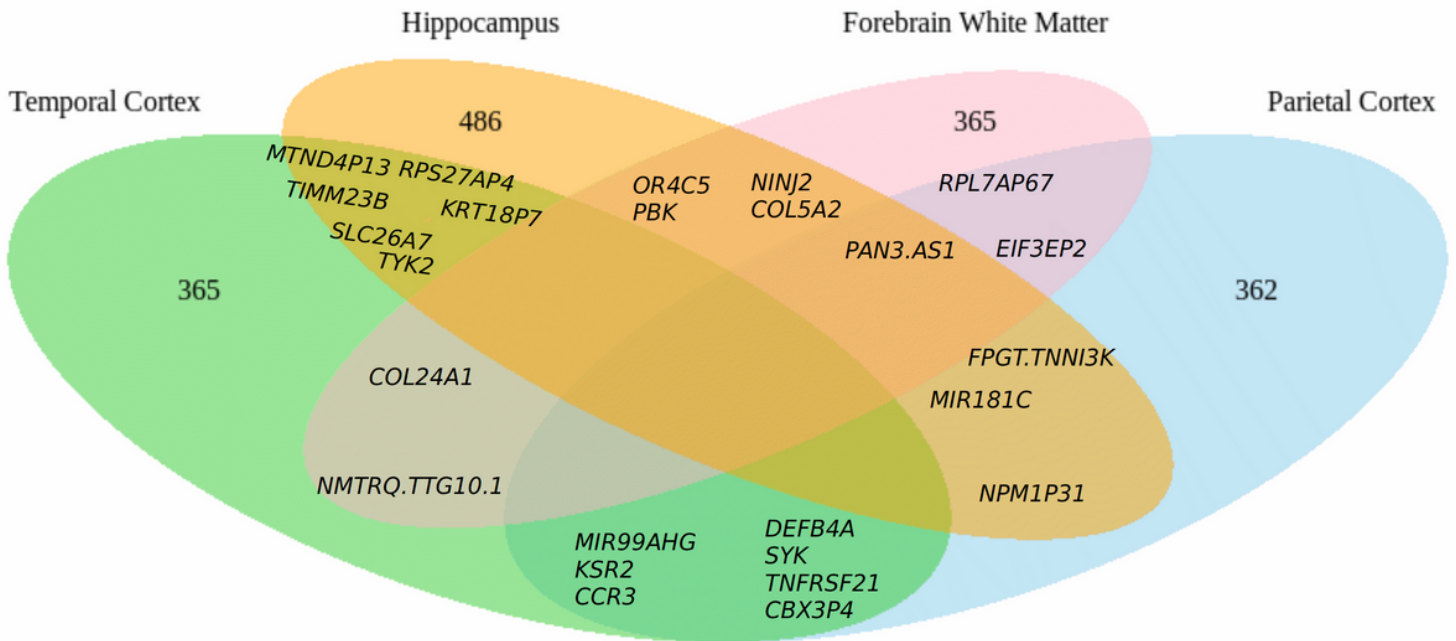


Figure 3

Venn Diagram of differentially expressed genes shared by brain areas

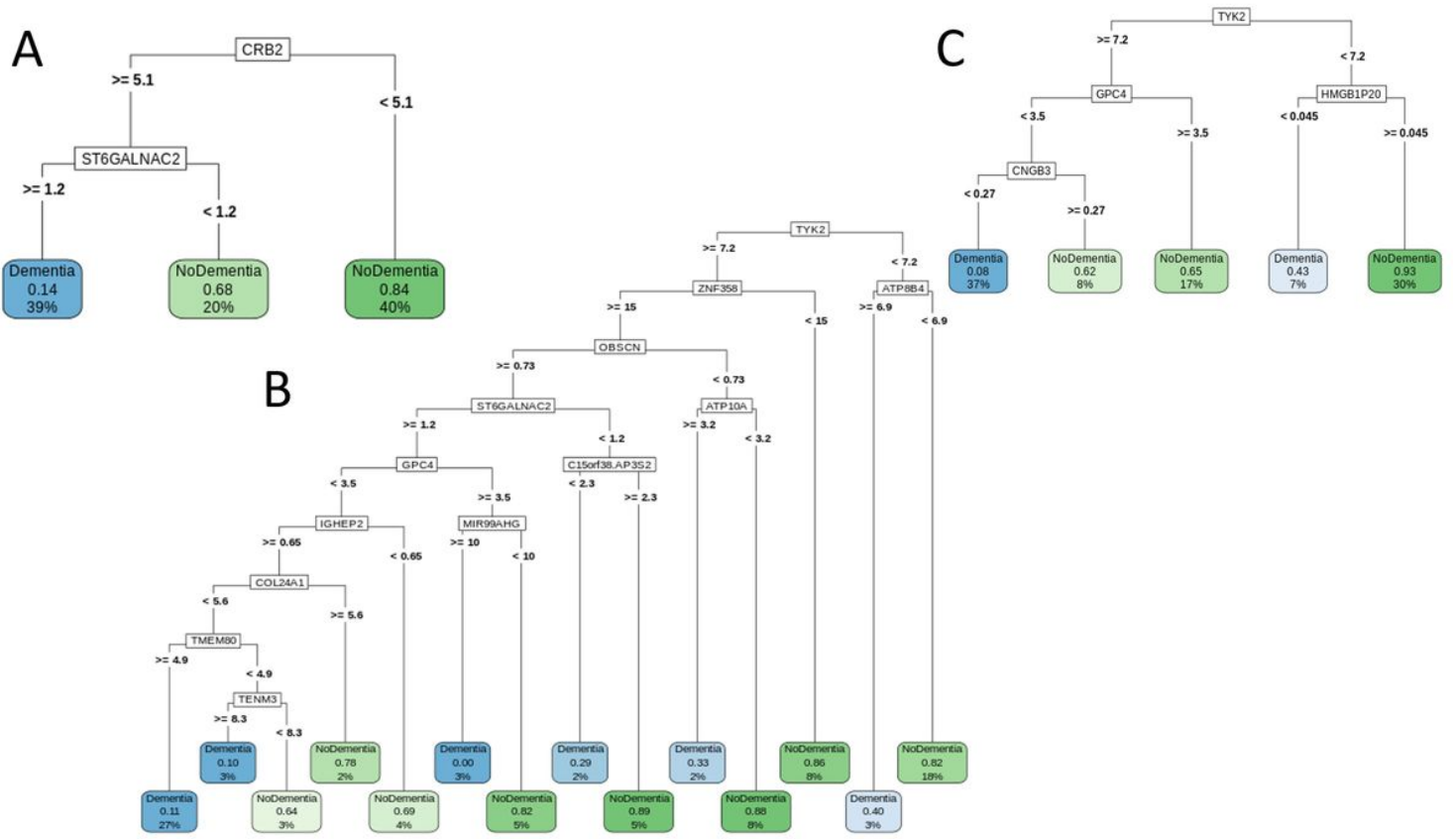


Figure 4

Decision tree based dementia classification models

- [Table1ModelComparisonmetrics.xlsx](#)
- [Supplementarymaterial1allDETCxfunctionalEnrichment.txt](#)
- [Supplementarymaterial2allDETCxNetwork.png](#)
- [Supplementarymaterial3allDEPCxfunctionalEnrichment.txt](#)
- [Supplementarymaterial4allDEPCxNetwork.png](#)
- [Supplementarymaterial5allDEHipfunctionalEnrichment.txt](#)
- [Supplementarymaterial6allDEHipNetwork.png](#)
- [Supplementarymaterial7allDEFWMfunctionalEnrichment.txt](#)
- [Supplmentarymaterial8allDEFWMNetwork.png](#)