Omics Comparison of Alzheimer's Neurons and Induced Neurons: Insights into Molecular Signatures

Max Van Puyvelde

December 2023

Abstract

Alzheimer's disease remains a serious health problem, necessitating the need for more in-depth knowledge of its underlying molecular mechanisms. To examine potential epigenetic and transcriptomic changes associated with AD pathology, we integrated methylation data with gene expression data in this study.
An analysis of Infinium Methylation Assay data indicated a significant difference in methylation levels between induced neurons from Alzheimer's patients and control groups. Specifically, neurons derived from Alzheimer patients displayed consistently higher methylation levels when compared to the control cohort.
A comparative analysis of microarray expression data on neurons with neurofibrillary tangles from AD patients versus those without neurofibrillary tangles showed no statistically significant differences in gene expression profiles between these two groups.

1 Introduction

Alzheimer's disease (AD) involves complex changes in cell behavior and molecular patterns. It is the most common cause of dementia, accounting for an estimated 60% to 80% of cases and affects millions worldwide [1] While omics studies have greatly improved our grasp of diseases like cancer [2], understanding AD would require a similar depth in exploring its genetic and cellular changes.

This research endeavors to uncover the genetic and cellular characteristics associated with AD pathology via a comparative analysis of Infinium methylation and microarray expression data, creating a framework for multiomics research in Alzheimer's disease.

2 Methods

This section will cover the methodologies employed for data collection and subsequent statistical analysis. The data collection involved two distinct technologies: the Infinium methylation assay and microarray assay data.

2.1 Methylation Data

This study aimed to compare the methylation patterns in induced neurons derived from individuals with Alzheimer's disease (AD) and healthy controls. The analysis utilized Infinium EPIC array data retrieved from ArrayExpress - accession E-MTAB-10344.

Data preprocessing began with handling missing values, annotating the samples, and filtering probes with insufficient calling p-values. Subsequently, a subset was selected, comprising a total of 16 samples, evenly divided between 8 AD cases and 8 controls.

Visual analysis was performed, revealing higher average methylation levels in induced neurons derived from AD patients compared to controls. Statistical validation through the Welch Two Sample t-test confirmed the significant elevation in AD samples.

The normalization process was performed by the 'dasen' function from the 'wateRmelon' package, which included dye color adjustment [3], while quality control (QC) steps involved visualizing density plots for M-values, demonstrating clear separation between methylated and unmethylated pixels.

The subsequent stage involved differential methylation analysis using the 'limma' package - a tool widely used for analyzing differential gene expression or methylation by fitting linear models and employing

moderated statistics [4]. Annotation was performed to identify CpGs situated in genic and promoter regions. Further analysis investigated Gene Ontology (GO) enrichment analysis, revealing associations with key biological processes.

2.2 Gene Expression Data

The comparative analysis handles microarray data from neurons extracted from Alzheimer's patients, focusing on samples with and without neurofibrillary tangles. The dataset was sourced from ArrayExpress - Accession: E-GEOD-4757.

Quality control checks were performed, assessing both raw and log-transformed data integrity using the 'arrayQualityMetrics' function, which generates microarray quality metrics reports for data in Bioconductor microarray data containers [5]. Following this, the raw microarray data underwent preprocessing via RMA (Robust Microarray Analysis) to normalize and summarize probe-level data into gene-level information.

Differential expression analysis involved fitting the data into a linear model, comparing gene expression differences between neurons with and without neurofibrillary tangles.

Visual representations were generated to render p-value and expression value distributions, aiming to identify genes exhibiting differential expression.

Adjusting for multiple testing was achieved using the Benjamini-Hochberg method to correct raw p-values. Finally, an enrichment analysis using Gene Ontology (GO) as performed to identify functional enrichments among differentially expressed genes.

3 Results

3.1 Methylation Levels

This analysis found significantly higher average methylation levels within the AD samples compared to the control group, as confirmed by statistical tests (t-test, $p < 0.05$), visible in Figure 1. The elevation in methylation levels indicates distinct epigenetic differences present in induced neurons obtained from AD-affected individuals.

Moreover, the analysis identified statistically significant increases in methylation levels in both promoter and genic regions of the genome. These differential methylation patterns suggest alterations in specific genomic areas associated with gene regulation and functional sequences. The significance of these observations was underscored by their association with

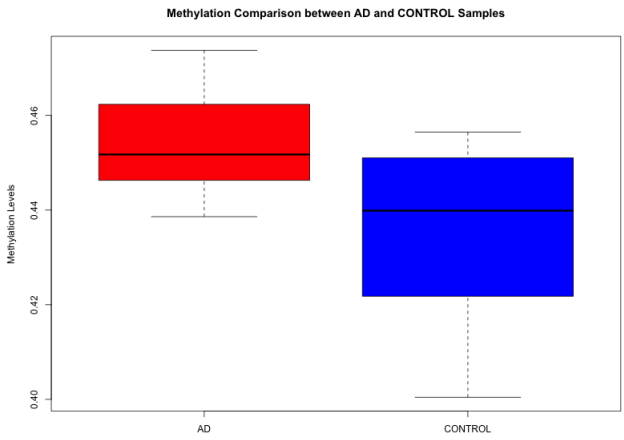

Figure 1: Boxplot showing the difference in methylation levels between AD and control samples.

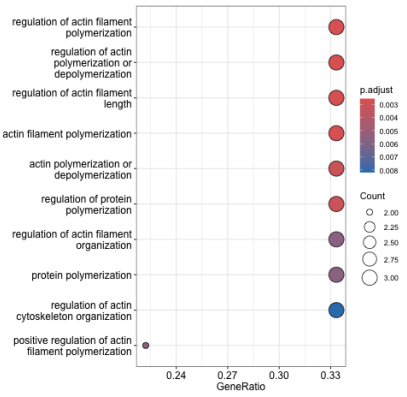

Figure 2: Dotplot showing the biological pathways, belonging to the top statistically significant methylated genes in CpG sites of genic regions in AD induced neurons.

biological pathways identified through Gene Ontology (GO) enrichment analysis.

The enriched biological pathways associated with the elevated methylation levels in genic regions were primarily linked to actin regulation, protein polymerization, and cytoskeleton organization, as seen in Figure 2. On the other hand, the differential methylation observed in promotor regions was associated with biological pathways involving aerobic electron transport chains and ATP synthesis, as seen in Figure 3

These findings imply distinct epigenetic signatures, suggesting the presence of clear variations in methylation profiles in fibroblast-induced neurons from AD patients.

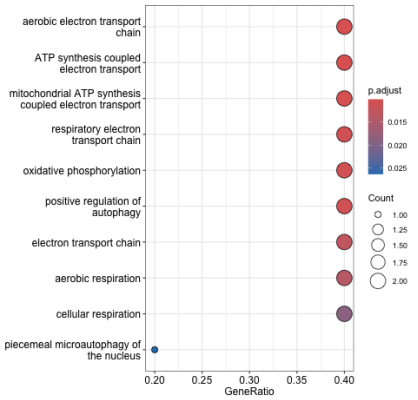

Figure 3: Dotplot showing the biological pathways, belonging to the top statistically significant methylated genes in CpG sites of promotor regions in AD induced neurons.

logFC	P.Value	adj.P.Val	Gene	Chrom	Pos	Abs_diff_meth
2.334274	0.00000001278517	0.01080174	RFXAP	13	37,394,589	0.29651977
-1.305930	0.00000221373054	0.63305194	HTRA1	10	124,231,341	0.15087011
1.416394	0.00000451869831	0.72487695	FHOD3	18	34,337,838	0.21901726
-1.098934	0.00000536130158	0.72487695	TRIM8	10	104,403,707	0.06484409

Figure 4: Table showing the top statistically significant methylated genes.

Top statistically significant methylated gene RFXAP, as seen in Figure 4 is primarily recognized for its role in the Major Histocompatibility Complex (MHC) class II gene regulation as part of the RFX transcription factor complex. The gene has been found significantly downregulated in cases of alzheimer's disease, affirming these findings [6].

3.2 Gene Expression Profiles

The microarray data analysis aimed to identify differential gene expression patterns between neurons with and without neurofibrillary tangles (NFTs) in Alzheimer's disease (AD) patients. Quality control measures confirmed successful normalization of the data using robust microarray analysis (RMA).

Despite observing trends in gene expression via various visual representations such as histograms, volcano, and MA plots, no statistically significant differentially expressed genes were identified after adjusting for multiple testing using the Benjamini-Hochberg method, as seen in Figures 5 and 6

Array annotation verification showed consistency with processed data, but the absence of statistically significant genes limited the meaningful interpretation of enrichment analysis using Gene Ontology (GO).

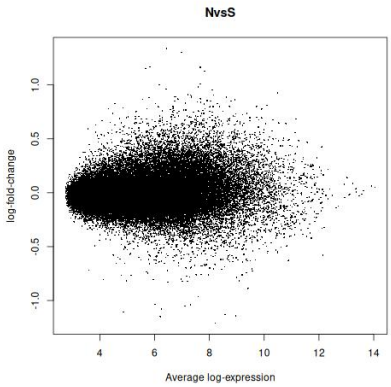

Figure 5: MA plot showing average log expression (x-axis) against log fold change (y-axis).

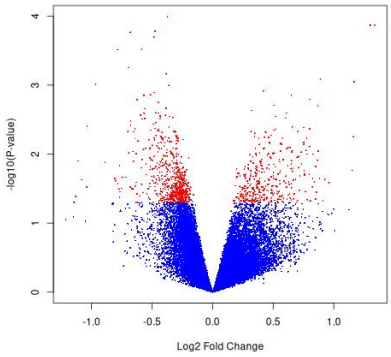

Figure 6: MA plot showing average log expression (x-axis) against log fold change (y-axis). P-values ¡ 0.05 are marked red.

In conclusion, while the analysis revealed observable gene expression trends, it did not yield statistically significant differences between neurons with and without NFTs in AD patients. Caution is warranted in interpreting the biological implications of these findings.

4 Discussion

Methylation upregulations in genic regions were linked to crucial pathways governing actin regulation, protein polymerization, and cytoskeleton organization. Similarly, changes in promoter regions were associated with pathways involving aerobic electron transport chains and ATP synthesis. These findings hint at a potential correlation between upregulated

methylation profiles and disruptions in fundamental cellular mechanisms implicated in AD development and progression.

Highlighting the significance of specific methylated genes, the study underscored the downregulation of RFXAP, a gene known for its involvement in MHC class II gene regulation. This aligns with prior research, strengthening the potential association of identified genes with AD pathogenesis and suggesting their role as potential biomarkers or targets for further investigation.

Contrarily, the analysis of gene expression in neurons with neurofibrillary tangles (NFTs) did not yield statistically significant differences. However, due to the rather low number of samples, further research into gene expression differences is encouraged.

Future research should further investigate the implications of these methylation alterations, validating potential biomarkers, and exploring multi-omics approaches to comprehensively unravel the intricate molecular mechanisms underlying AD pathogenesis.

5 Conclusion

This study identified distinct methylation patterns in induced neurons of AD patients, notably higher than controls in genic and promoter regions. These alterations were linked to critical cellular pathways, suggesting potential disruptions in AD progression. Specific methylated genes, like RFXAP, hinted at their role in AD pathogenesis, possibly as biomarkers. However, gene expression analysis in neurons with NFTs did not reveal statistically significant differences in gene expression among neurons with NFTs, possibly due to limited sample sizes.

6 Acknowledgements

I would like to thank the lecturers and supervisors of the course *Applied High Throughput Analysis* at *Ghent University* for their mentorship. Performed analyses are largely based on their foundation.

References

[1] D. Gibbs and T. H. Barker, "A tattoo on my brain: A neurologist's personal battle against alzheimer's disease," *Cambridge University Press*, 2023.

[2] T. Das, G. Andrieux, M. Ahmed, and S. Chakraborty, "Integration of online omics-data resources for cancer research," *Frontiers in Genetics*, vol. 11, 2020.

[3] R. Pidsley, C. C. Y Wong, M. Volta, K. Lunnon, J. Mill, and L. C. Schalkwyk, "A data-driven approach to preprocessing illumina 450k methylation array data," *BMC Genomics*, vol. 14, p. 293, May 2013.

[4] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, pp. e47–e47, 01 2015.

[5] A. Kauffmann, R. Gentleman, and W. Huber, "arrayQualityMetrics–a bioconductor package for quality assessment of microarray data," *Bioinformatics*, vol. 25, pp. 415–416, Dec. 2008.

[6] J. Liu, L. Jiao, X. Zhong, W. Yao, K. Du, S. Lu, Y. Wu, T. Ma, J. Tong, M. Xu, W. Jiang, Y. Wang, M. He, W. Xin, and M. Liu, "Platelet activating factor receptor exaggerates Microglia-Mediated microenvironment by IL10-STAT3 signaling: A novel potential biomarker and target for diagnosis and treatment of alzheimer's disease," *Front Aging Neurosci*, vol. 14, p. 856628, Apr. 2022.