Association of stochastic epigenetic mutations with age and Alzheimer's disease: meta-analysis

Tinatin Mekvabishvili

Master's thesis submitted to the Faculty of Natural Sciences and Medicine of the Ilia State University, according to the requirements for the Degree of Master of Science in Applied Genetics

Program: Applied Genetics

Supervisor: Vincenzo Lagani, PhD

Ilia State University

Tbilisi, 2022

Table of Content:

Abbreviations

SEM – Stochastic Epigenetic mutation

AD – Alzheimer's disease

DNAm – DNA methylation

STG – Superior Temporal Gyrus

MTG - Middle Temporal Gyrus

ITG - Inferior Temporal Gyrus

DMR – Differentially Methylated Region

OXT – OXT (Oxytocin/Neurophysin I Prepropeptide) is a Protein Coding gene

CHRNB1 - Cholinergic Receptor Nicotinic Beta 1 Subunit

RHBDF2 - Rhomboid 5 Homolog 2

ADAM1 - ADAM Metallopeptidase Domain 1A (Pseudogene)

M - Methylated

U - Unmethylated

Abstract

As modern technologies develop a huge amount of publicly available data is deposited through internet resources. Using Gene Expression Omnibus database, it became possible to have hands-on gene expression data (Edgar, Domrachev, and Lash 2002). It allows us to track gene expression patterns from patients with different conditions, including Alzheimer's disease.

In our work, we downloaded eight publicly available datasets which contained DNA methylation levels and different phenotypic characteristics of Alzheimer's Diagnosed patients and healthy controls. We preprocessed the datasets and performed statistical analysis with the R Statistical Software and the packages/libraries hosted in the Bioconductor repository (v1.14.0; Marini et al., 2020). Our main goal was to find aberrant DNA mutation loci (loci of stochastic epigenetic mutations - SEMs) and find a correlation between the accumulation of SEMs, AD, and age.

Preprocessed datasets were analyzed individually. Different statistical methods were applied for the analysis: logistic regression, correlation test, and t-test. Effect sizes generated from individual dataset analysis were pooled and a random effect model meta-analysis was performed. Logistic-regression meta-analysis model used to predict Alzheimer's based on age and a significant accumulation of SEMs, showed a weak correlation between AD and age with a p-value equal to 0.06. In contrast with it, correlation test analysis showed no association between stochastic epigenetic mutations and age. As for our analysis Logistic regression can be considered as a more robust method as it allows us to explore the individual effects of age and AD by removing possible mutual effects.