# Diversity study of small open reading frames (sORFs) of healthy and in Alzheimer's Disease brain

Saloe Bispo, Fabio Passetti

*Laboratory of Functional Genomics and Bioinformatics, Oswaldo Cruz Institute (IOC), Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, RJ, Brazil. Laboratory of Gene Expression Regulation, Carlos Chagas Institute (ICC), Oswaldo Cruz Foundation (Fiocruz), Curitiba, PR, Brazil.*

**Abstract**

The aging of the world population is associated with the increased frequency of people diagnosed with dementias. These are responsible for the greatest burden of neurodegenerative diseases, with Alzheimer's representing approximately 60-70% of dementia cases. Computational approaches that integrate genome, transcriptome and proteome data have been developed by our research group to study human transcripts and their polypeptide products in an area known as proteogenomics. The discovery of small open reading frames (sORFs) in gene and protein databases called small ORF encoded polypeptides (sORF-encoded polypeptides) or SEPs has revealed a fundamental shortcoming in our knowledge of protein-coding genes. Some of these new sORFs have crucial biological roles in cells and organisms, which motivates the search for new sORFs. In this study, we are developing a proteogenomics approach for the identification of proteoforms in human and mouse focusing on SEPs. Thus, predictions of SEPs will be incorporated into the human and murine proteoform repository SpliceProt maintained by our group. Currently, we are using public shotgun proteomics data from healthy and AD affected brain samples, searching for new SEPs expressed under such conditions. Preliminary data have shown the presence of previously undetected microproteins in proteomics data from AD, derived from sORFs (lncRNAs and antisense).

Funding: CNPq and ICC