Identifying network biomarkers for Alzheimer's Disease using single cell RNA sequencing data.

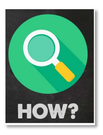
Ioannis Aslanis, Marios Krokidis, Georgios N. Dimitrakopoulos, Aristidis G. Vrahatis

Department of Informatics, Ionian University, Greece

https://bihelab.di.ionio.gr

Abstract

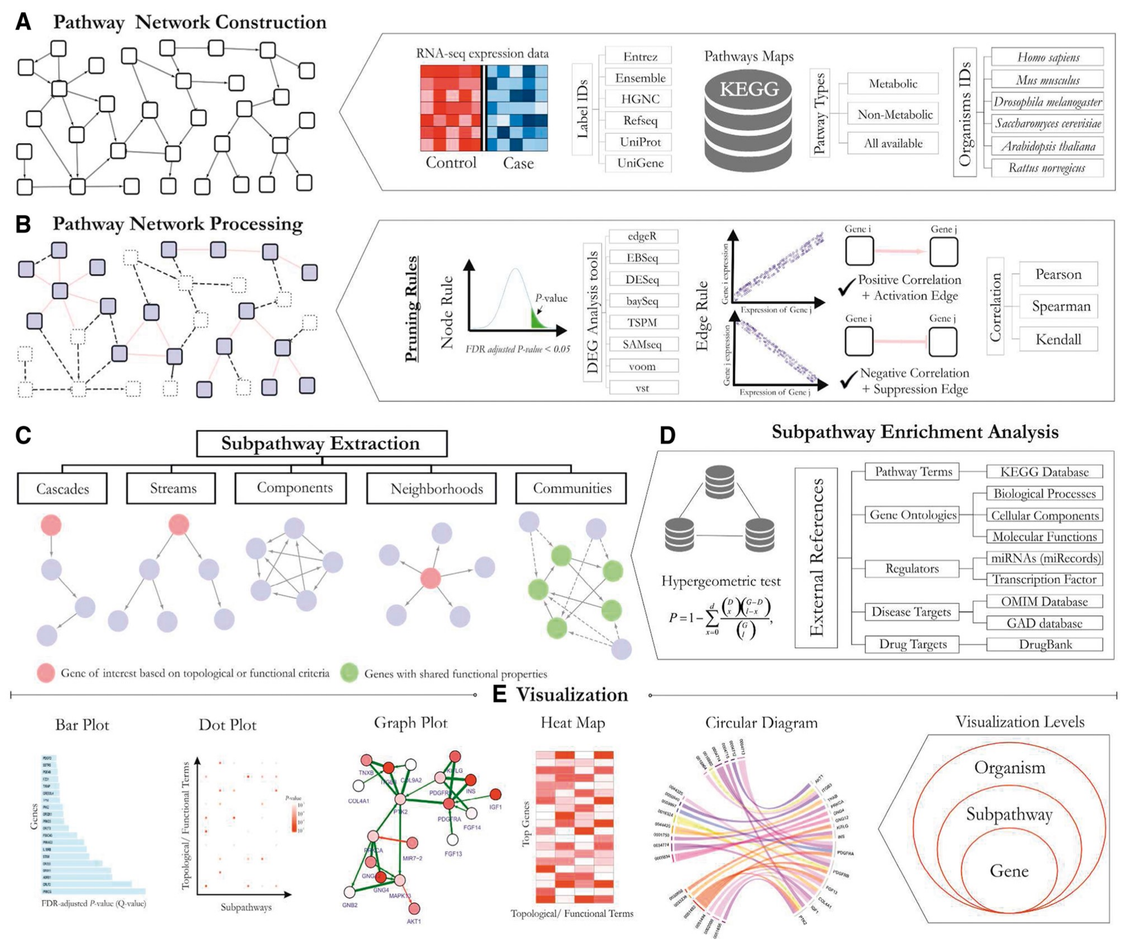
A network-based systems biology R workflow that extracts highly expressed subpathways within a pathway network as recorded by an Alzheimer's Disease - based single-cell RNA-seq experiment.



The most active subpathways are isolated by pruning the gene-gene interaction network based on rules derived from Differential Expression analysis from the AD-based scRNA-seq data in conjunction with biological knowledge extracted from the pathway network and the networks topology.

The nature and complexity of Alzheimer's disease has made it a very prominent field to approach through the lens and tools of network biology while the technology of single-cell RNA-seq promises new insights in the extent and complexity cell behaviour.

Methodology



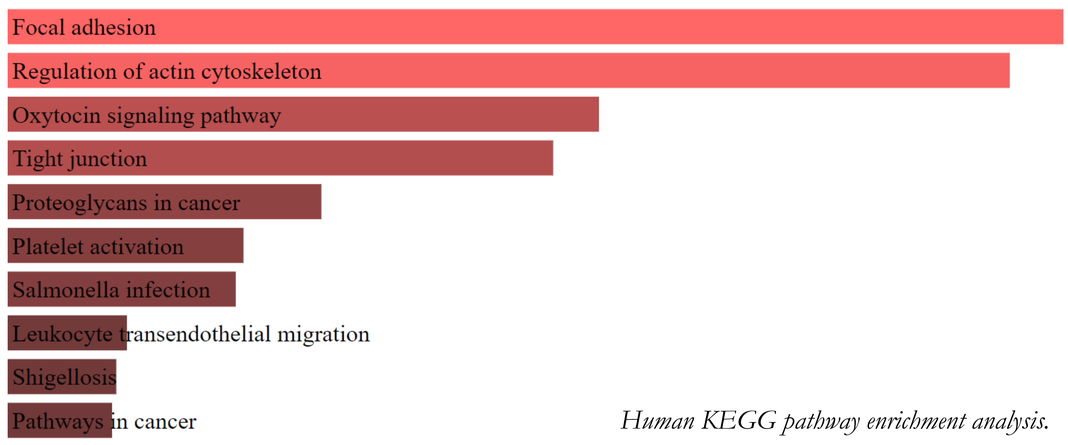
We create a gene to gene network by converting all available mouse KEGG signaling pathway maps to a pathway network while preserving its topological attributes. We keep the genes that are both in our AD-based dataset and the pathway network. In order to locate the genes with statistically significant Differential Expression (DE) we use the well-established DE tool limmatrend on the gene expression of our sc-RNA-seq data pruning gene nodes with statistical significance bellow our threshold. We then proceed to estimate the Pearson Correlation Coefficient between all the adjacent remaining genes in order to highlight further the interactions that agree with KEGGs pathway map prior biological knowledge. We end up with a final network where any subset of connected nodes contains statistically significant expressed genes. From this we extract the most topologically interconnected subpathways.

Experimental Analysis

The data used for this experiment is real high dimensional biomedical datasets with single cell expression profiles by high throughput sequencing (data are available in NCBI Gene Expression Omnibus, accession: GSE150934). The experimental data include purified cells obtained from the hippocampus of two-, 6-, and 9-month-old 5xFAD mice and a 3-monthold wild-type mouse. Approximately 1000-1500 cells were obtained from each dissected brain [3].

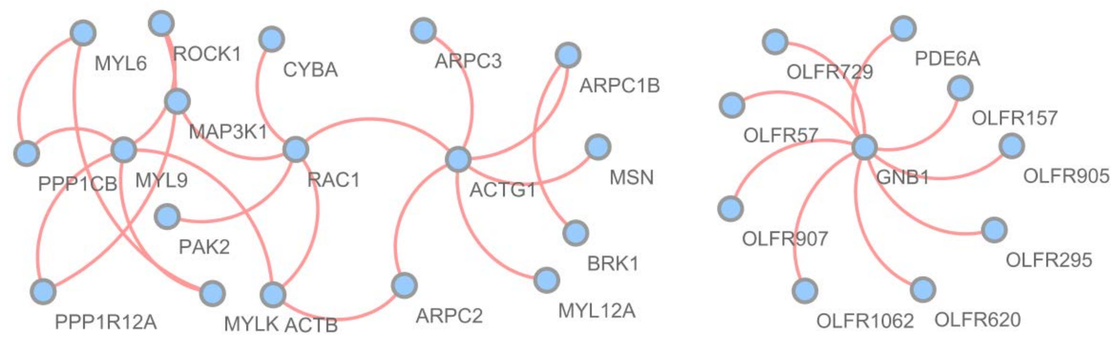
Results

We use KEGG and Gene Ontology enrichment analysis to reveal the functional implications our extracted gene modules.



Human KEGG pathway enrichment analysis.

Gene Ontology enrichment analysis.



Conclusion

Our proposed computational method identified several disease-perturbed subpathways after the integration of real gene expression data in the pathway network. These can be regarded as potential biomarkers revealing new gene modules that play an important role in AD and thus opening new windows in the understanding and treatment of the disease.

ACKNOWLEDGEMENTS

This research is funded by the European Union and Greece (Partnership Agreement for the Development Framework 2014-2020) under the Regional Operational Programme Ionian Islands 2014-2020, project title: "Study of drug protocols with biomarkers that define the evolution of non-genetic neurodegenerative diseases - NEUROPHARMA", project number: 5016117.

url: https://bihelab.di.ionio.gr/

References

1. Andrews, T. S., & Hemberg, M. (2017). Identifying cell populations with scRNASeq. *Molecular aspects of medicine*.
2. Lawlor, N., et al. (2017). Single-cell transcriptomes identify human islet cell signatures and reveal cell-type–specific expression changes in type 2 diabetes. *Genome research*, *27*(2), 208-222.
3. Choi H, Choi Y, Lee EJ, et al. Hippocampal glucose uptake as a surrogate of metabolic change of microglia in Alzheimer's disease. J Neuroinflammation. 2021;18(1):190. Published 2021 Aug 31. doi:10.1186/s12974-021-02244-6

Identifying network biomarkers for Alzheimer's Disease using single cell RNA sequencing data.