

Parallel integration of WGS–RNA-seq and Lipidomics-Metabolomics layers identifies robust and clinically complementary Alzheimer’s Disease clusters in the ROSMAP study

Margherita Squillario¹, Maurizio Fedriga¹, Francesco Reggiani¹, Matteo Pardini^{1 2}, Michele Piana³

1. IRCCS Azienda Ospedaliera Metropolitana, LISCOMP Lab, Genoa, Italy 2 University of Genoa, DINOGMI, Genoa, Italy 3 University of Genoa, DIMA (LISCOMP Lab), Genoa, Italy



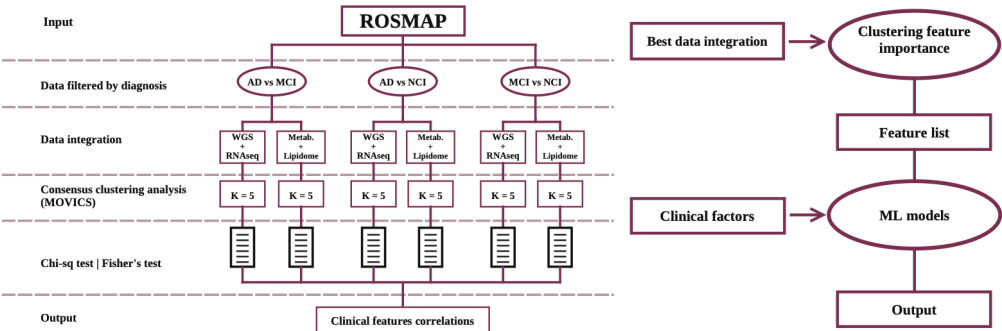
<https://mida.unige.it/liscomp>

INTRODUCTION

Alzheimer’s disease (AD) is the most common **neurodegenerative disorder** in the elderly. While genetic predisposition plays a major role, it is not sufficient to fully explain disease onset and progression [1][2]. The increasing availability of large **multimodal** or **multiomics** datasets offers an opportunity to investigate AD **molecular heterogeneity** through integrative approaches [3][4][5].

We analyzed subjects from the **ROSMAP** [6][7] study, including AD, mild cognitive impairment (MCI), and no cognitive impairment (NCI) individuals. Whole-genome sequencing (WGS), RNA-seq, proteomics, lipidomics, and metabolomics data were integrated using **MOVICS**, an **R-based** multi-omics clustering framework [8]. An explorative analysis was conducted through clustering performed across all pairwise and combined group comparisons and correlation of the identified clusters with the following clinical features: sex, diagnosis, education level, APOE genotype, and **CERAD** score.

METHODS

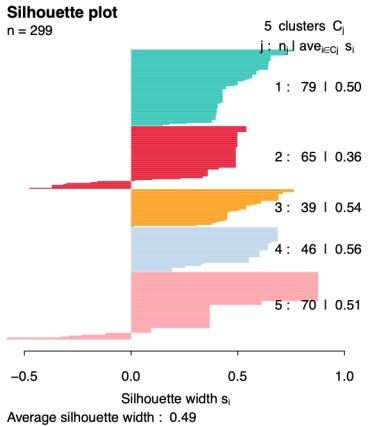
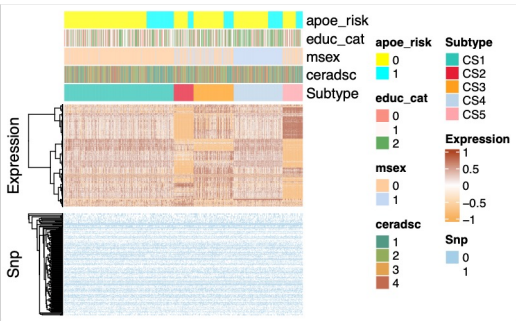


CONCLUSIONS

Our findings demonstrate that comparisons among AD, MCI, and NCI subjects uncover **distinct molecular subtypes** whose clinical relevance depends on the type of data integration performed, with each integration strategy capturing complementary aspects of clinical variability. In the near future we plan to add a supervised analysis, that for each data integration type **will assess the classification** performance of the list of features of the most **statistically robust clusters** (having the highest average silhouette score). This classification power will evaluate the ability to discriminate between 1) the clusters of subjects found by MOVICS and 2) the classes of the clinical features found significant by statistical tests.

Furthermore, we plan to validate the robustness of the clusters and of the features lists in independent studies (e.g., ADNI [9], MSBB) of multi omics data, in order to assess the reproducibility of the identified molecular signatures towards the precision medicine objective in AD.

RESULTS



Comparison	ARI	NMI
All vs SNP+rnaSeq	0.93	0.88
All vs Met-Lip	-0.03	0.03
SNP+rnaSeq vs Met-Lip	-0.008	0.02

A key finding of this study is that across all **group comparisons** and integration strategies, clustering solutions showed consistently high silhouette scores, highlighting the statistical robustness of these clusters. Another key finding is that WGS–RNA-seq **integration generated molecular clusters** distinct from those obtained via lipidomic and metabolomic; moreover these clusters showed **significant associations** with different and often complementary, sets of clinical features, with respect to the clusters identified in the lipidomic and metabolomic integrations.

REFERENCES

- [1] Van Cauwenbergh, C. et al. "The genetic landscape of Alzheimer disease: Clinical implications and perspectives". Genet. Med., 2016.
- [2] Naj, A. C. et al. "Genomic variants, genes, and pathways of Alzheimer's disease: An overview". Am. J. Med. Genet. B. Neuropsychiatr. Genet., 2017.
- [3] R. Sims et al. "The multiplex model of the genetics of Alzheimer's disease." Nat. Neurosci., 2020.
- [4] Tamsi R, Squillario M, et al. "An overview of data integration in neuroscience with focus on Alzheimer's Disease". IEEE J Biomed Health Inform., 2023.
- [5] Squillario M et al. "A telescope GWAS analysis strategy, based on SNPs-genes-pathways ensemble and on multivariate algorithms, to characterize late onset Alzheimer's disease." Sci Rep. 2020.
- [6] ROSMAP: <https://adknowledgeportal.synapse.org/Explore/Studies/DetailsPage/StudyDetails?StudySyn3219045>
- [7] Pérez-González AP et al. The ROSMAP project: aging and neurodegenerative diseases through omic sciences. Front Neuroinform. 2024 Sep 16;18:1443865.
- [8] MOVICS: Lu X et al. "MOVICS: an R package for multi-omics integration and visualization in cancer subtyping". Bioinformatics. 2021 Apr 1.
- [9] ADNI: <https://adni.loni.usc.edu/>
- [10] SKAT: <https://github.com/leelabs/skat>

CONTACTS



maurizio.fedriga97@gmail.com

margherita.squillario@aomliguria.it