UNIVERSITÀ DI TRENTO
Department of Cellular, Computational and Integrative Biology - CIBIO

Computational analysis of RNA expression among survival groups in amyotrophic lateral sclerosis

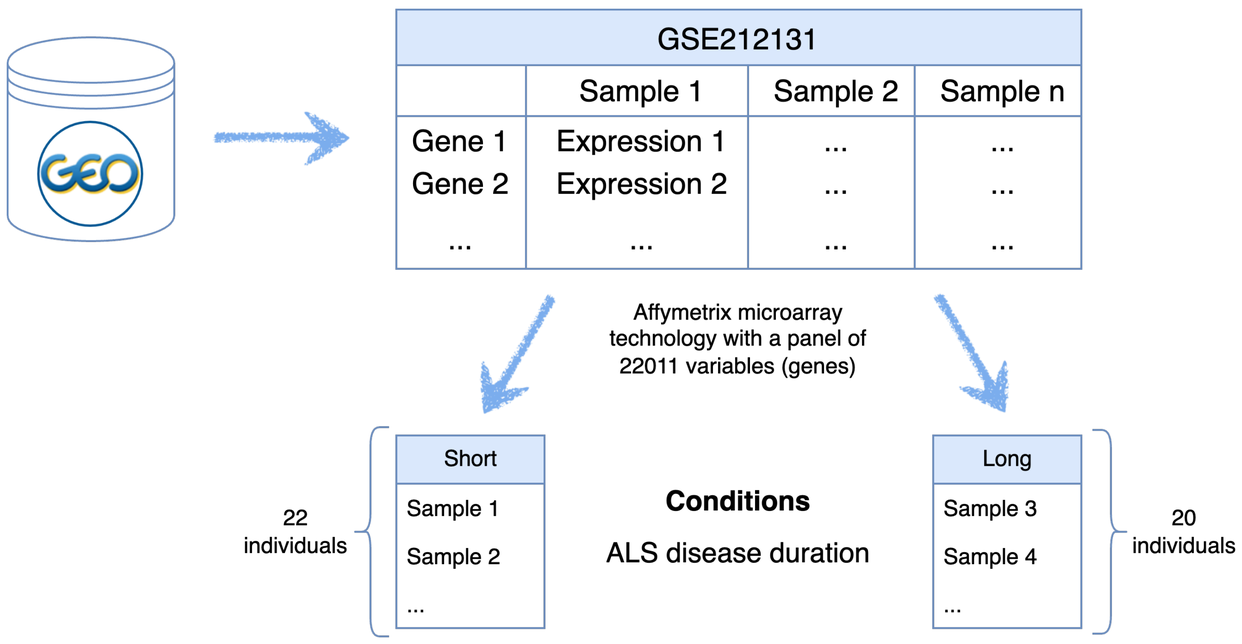
Baldinelli Sara, Lauria Mario

SIBBM

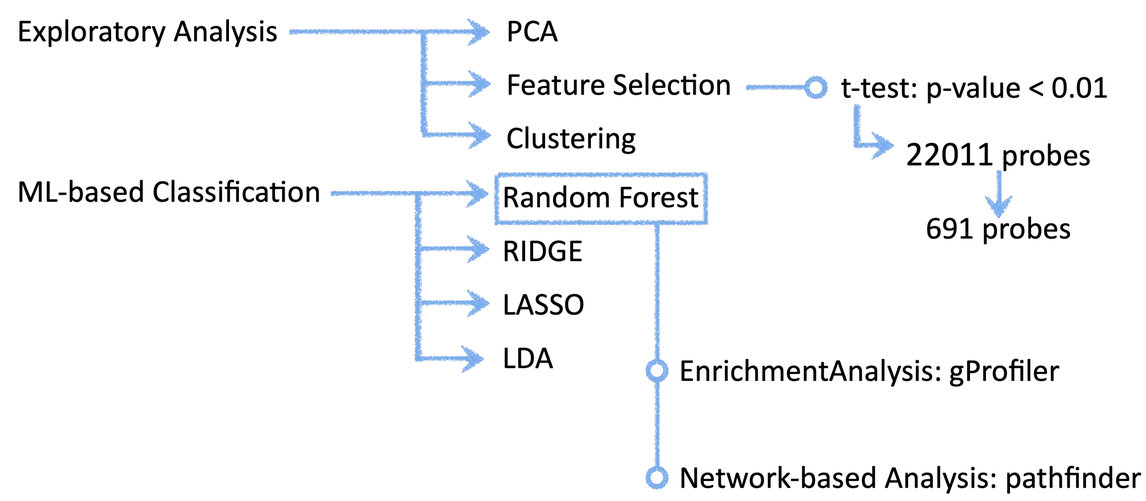
Background

Amyotrophic lateral sclerosis (ALS) is a neural disorder characterized by the degeneration of motor neurons, loss of voluntary muscle control and premature mortality. While ALS is primarily considered a sporadic condition. One of the key aspects of ALS lies in its heterogeneity, evident in the spectrum of survival times observed among affected individuals. This not only makes it harder to find a prognosis but also underscores the need for a deeper understanding of the underlying molecular mechanisms driving disease progression and survival outcomes.

Dataset



Methods

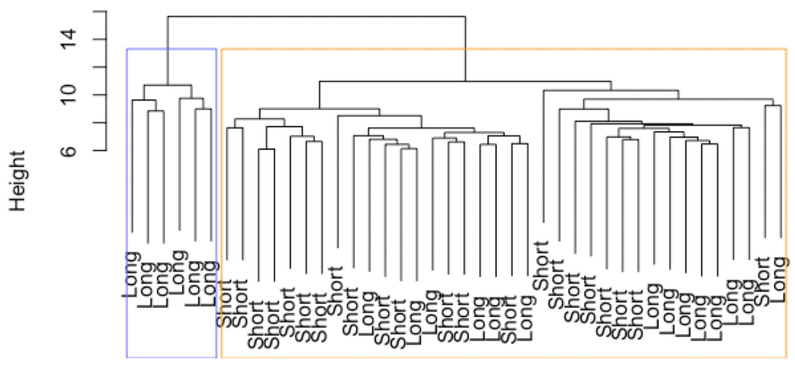


In this study, we employed different methodologies to get meaningful insights from the transcriptomic data, taking advantadge of both supervised and unsupervised approaches.

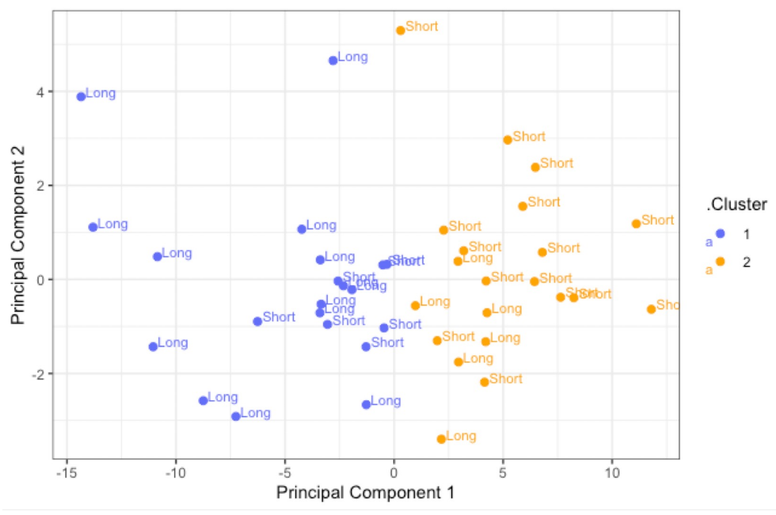
Results

Unsupervised Learning – Clustering

Hierarchical clustering

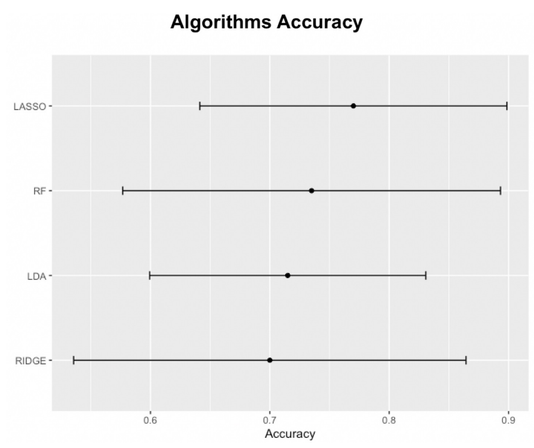


k-means clustering

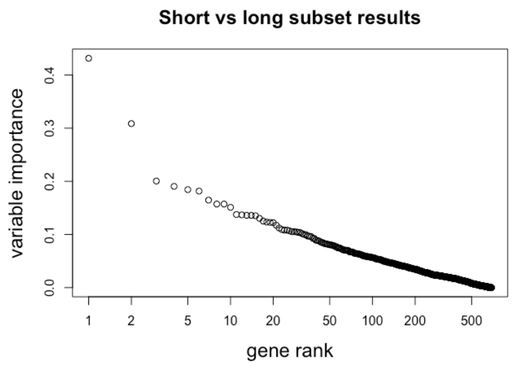


Two clustering methods, k-means and hierarchical clustering, were utilized to identify natural groupings within the dataset. The unsupervised analysis did not manage in a clear distinction between the two groups of the study.
That is the reason why we decided to proceed with supervised method for the classification of the saples.

Supervised Learning – Random Forest

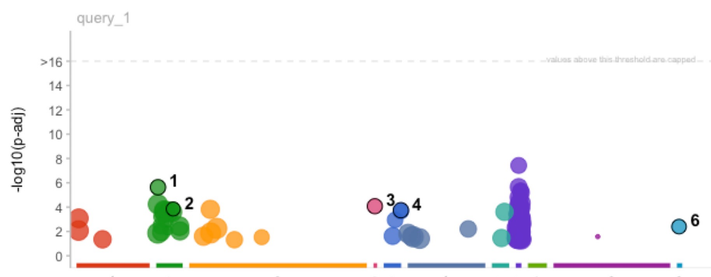


Accuracy differences for classification methods using cross validation.



Plot showing in a decreasing order the gene rank for importance of random forest classification.

Enrichment Analysis – gProfiler

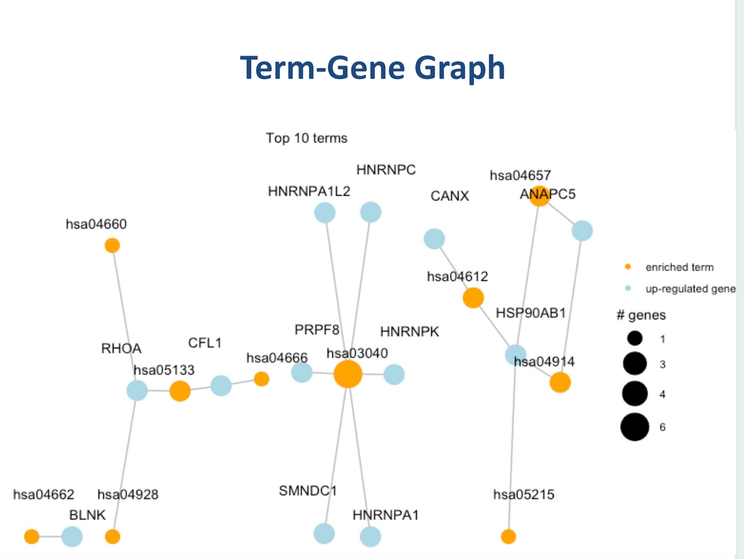


id	source	term_id	term_name	term_size	p_value
1	GO:CC	GO:0005681	spliceosomal complex	206	2.3e-06
2	GO:CC	GO:0071013	catalytic step 2 spliceosome	89	1.4e-04
3	KEGG	KEGG:03040	Spliceosome	174	8.4e-05
4	REAC	REAC:R-HSA-72163	mRNA Splicing - Major Pathway	202	1.8e-04
6	WP	WP:WP411	mRNA processing	126	4.0e-03

g:Profiler (biit.cs.ut.ee/gprofiler)

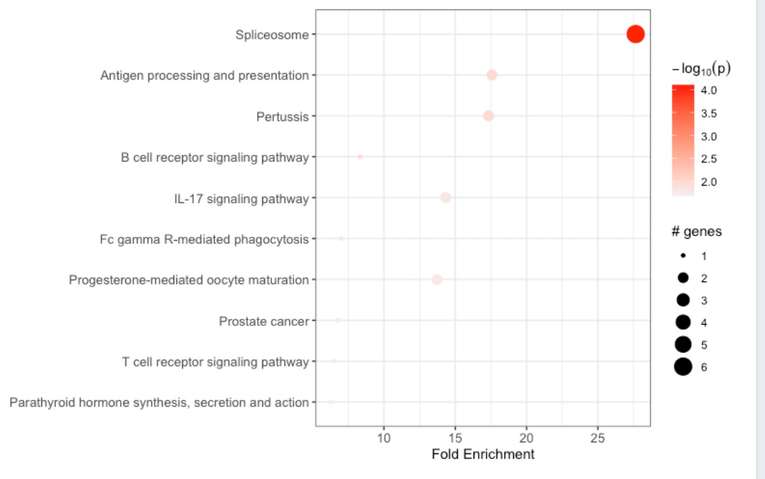
Network Analysis – pathfindR

Term-Gene Graph



pathfindR network of interactions, achieved giving as input the list of the 200 most important genes for the RF classification.

Enrichment Chart



We focused on the heterogeneous nuclear ribonucleoproteins (hnRNPs), that contribute to multiple aspects of nucleic acid metabolism including alternative splicing, and transcriptional and translational regulation. At this point, through both gprofiler2 enrichment and pathfindR network, we have two hints suggesting the potential involvement of splicing in ALS disease duration; but we can also confirm this result through the enrichment chart: the most enriched term is the Spliceosome.

Our findings highlighted the critical role of RNA splicing in ALS, as evidenced by the consistent enrichment of mRNA splicing-related terms in both gprofiler2 and pathfindR analyses. This suggests an involvement of splicing deregulation in ALS pathogenesis, corresponding also with existing literature on the subject. Moreover, the identification of hnRNPs as significant players in this process may offer potential therapeutic targets. Based on our findings, future research could explore the development of targeted therapies aimed at modulating RNA splicing and immune response pathways. Implementing personalized medicine approaches by stratifying ALS patients based on their molecular profiles could enable tailored therapeutic interventions, potentially improving outcomes and quality of life for individuals with ALS.

Contact
Email: sara.baldinelli@unitn.it

SIBBM 2024 - Frontiers in Molecular Biology
17 June 2024 – 19 June 2023