



# Bioinformatics pipeline development for integrative transcriptomic analysis across cell lines exposed to NPLs

R. Egea, J. Gutierrez, R. Marcos, A. Hernández

Group of Mutagenesis, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Cerdanyola del Vallès (Barcelona), Spain

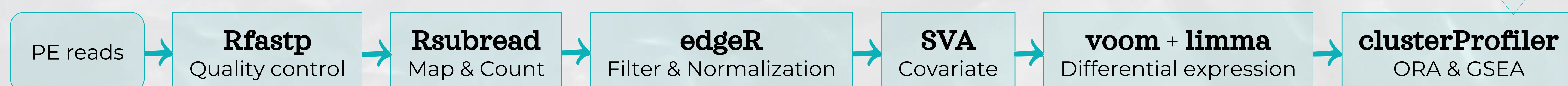
## # Summary

Humans are exposed to Nanoplastics (NPLs), although there is evidence about NPLs internalization, there is limited information on the potential impact on human health. Transcriptomic analyses allow to measure the biological response of NPLs exposure through the comparison of gene expression levels between different samples under different conditions. Several analyses evaluating different types of nanoparticles, exposure times and coexposures have been performed. As NPLs inhalation and ingestion are a mayor way of human exposure, the analyses have been performed on cells lines related to gastrointestinal and respiratory tract barriers.

A bioinformatic pipeline has been developed to analyze the transcriptomic results enabling the identification of Differentially Expressed Genes (DEGs) or Overrepresented pathways for each analysis. The molecular mechanisms altered during the biological response will be unveiled, allowing data-driven selection of the most appropriate functional analysis for their validation. The combination of all these results into an integrative analysis will provide a list of common genes or pathways from the NPLs exposure response. Therefore, it will be possible to select potential biomarkers to facilitate decision-makers to evaluate the risk of NPLs for humans or to biomonitoring.



## # Pipeline schema



## # Results

The pipeline works on R on any operating system. It analyses the raw FASTQ files and in a few hours generates excel files with the results for the differentially expressed genes and overrepresented pathways. Other R libraries are used to generate plots to visualize the results.

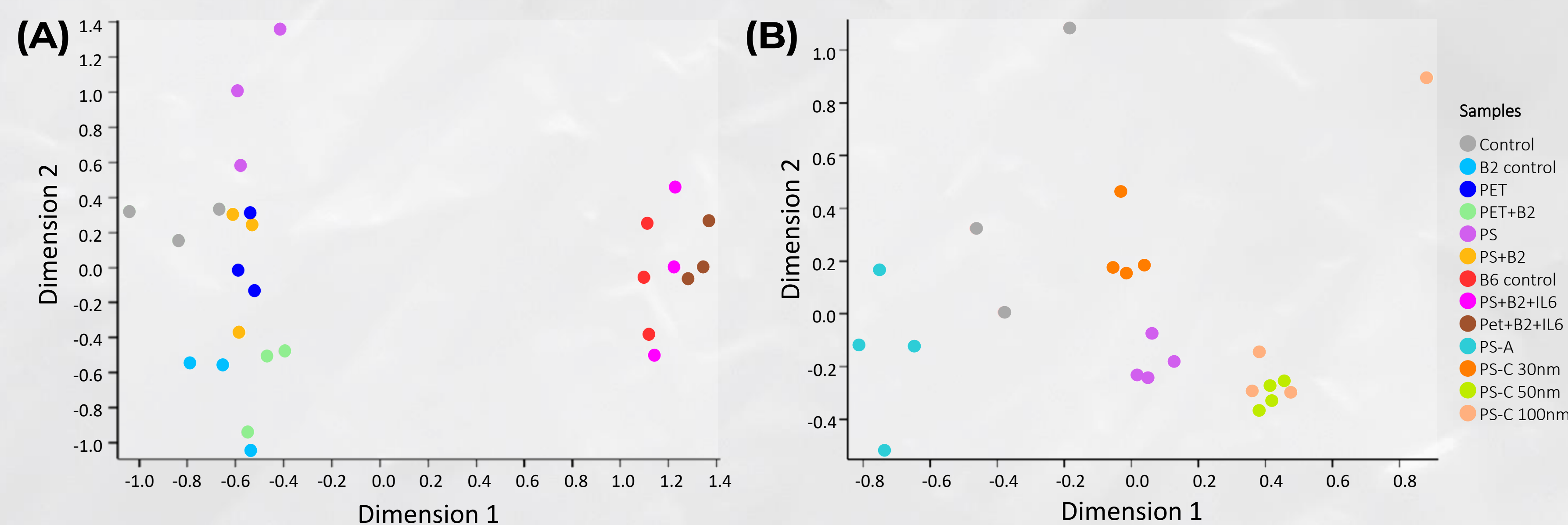


Figure 1. MultiDimensional Scaling plot for (A) MCF10A and (B) HUVEC cell line analyses.

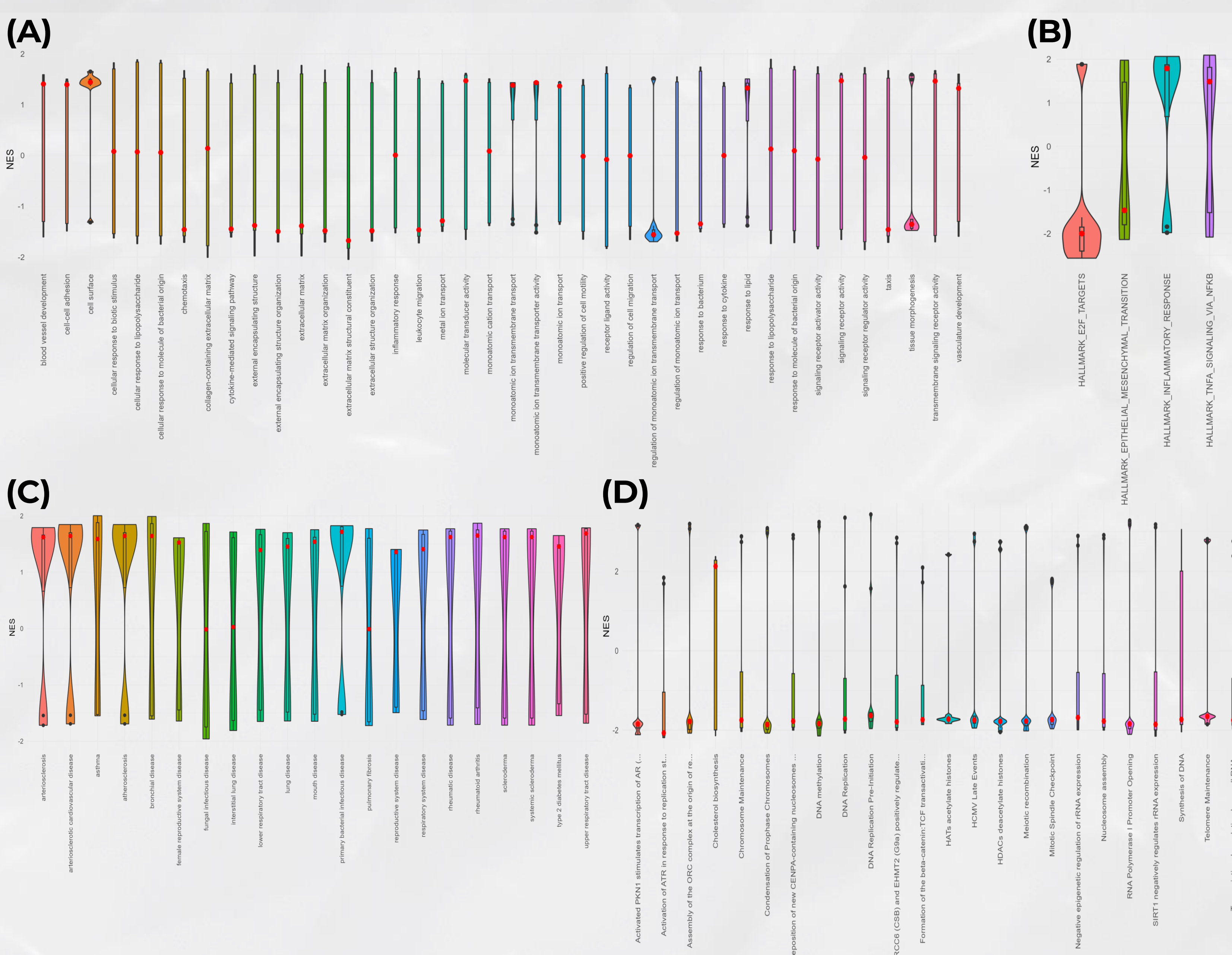


Figure 3. Violin plot for most common enriched pathways among all analyses from (A) Gene Ontology (B) Hallmarks (C) Disease Ontology and (D) Reactome database

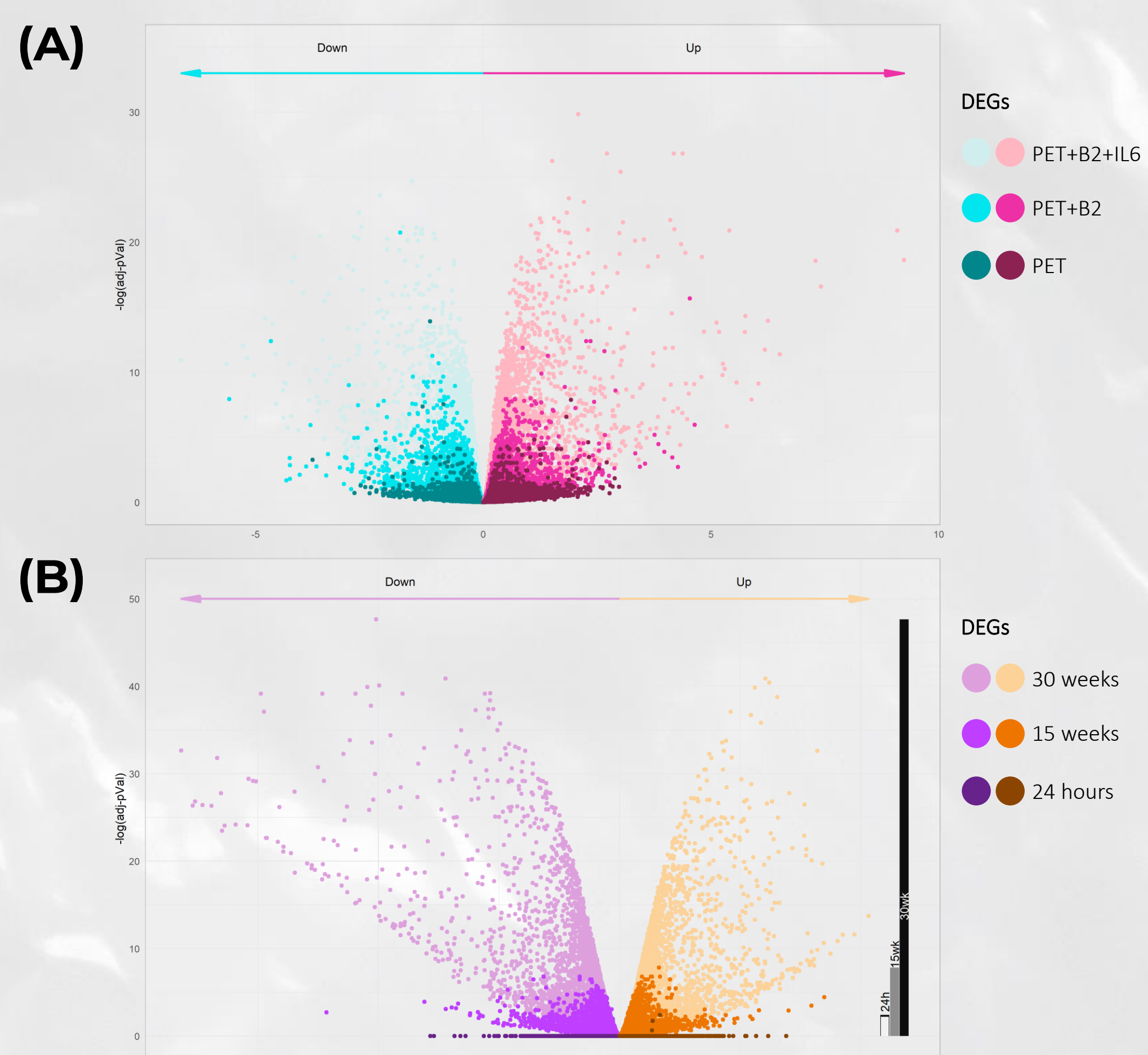


Figure 2. Volcano plot representing differentially expressed genes from (A) MCF10A and (B) BEAS-2B cell line analyses.

## # Conclusions

The pipeline smoothly analyses the gene expression levels from any condition and facilitates result comparison among different analyses.

The integrative module provides an overview of the results allowing to filter by NPLs treatment, cell line or any other condition of interest.

The integrative module will serve as a basis for data integration with other omics results

## # Acknowledgments

Funding: This work was partially supported by the EU Horizon 2020 programme (965196, PlasticHeal), the Spanish Ministry of Science and Innovation (PID2020-116789, RB-C43), the Generalitat de Catalunya (2021-SGR-00731), and the ICREA-Academia programme to Alba Hernandez.