

RFLOMICS: R package and Shiny interface for Integrative analysis of omics data

Nadia Bessoltane, Christine Paysant-Le-Roux, Gwendal Cueff, Audrey Hulot,

Delphine Charif

▶ To cite this version:

Nadia Bessoltane, Christine Paysant-Le-Roux, Gwendal Cueff, Audrey Hulot, Delphine Charif. RFLOMICS: R package and Shiny interface for Integrative analysis of omics data. Journées ouvertes en biologie, informatique et mathématiques (JOBIM), Jul 2022, Rennes, France. hal-04546139

HAL Id: hal-04546139 https://hal.science/hal-04546139

Submitted on 15 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

	RFLOMICS: R package and Shiny interface for Integrative analysis of omics data	ting
Nadia BESSOLTANE ¹ , Christine PAYSANT-LE-ROUX ² , Gwendal CUEFF ¹ , Audrey HULOT ¹ , Delphine CHARIF ¹ 1 Université Paris-Saclay, INRAE, AgroParisTech, Institut Jean-Pierre Bourgin, 78000, Versailles, France 2 Université Paris-Saclay, CNRS, INRAE, Univ Evry, Institute of Plant Sciences Paris-Saclay (IPS2), 91405, Orsay, France. <u>nadia.bessoltane@inrae.fr</u> & <u>delphine.charif@inrae.fr</u>		

The acquisition of multi-omics data in the context of complex experimental design is a widely used practice to identify entities and characterise related biological processes. The investigation of each omic layer is a good way to explore and extract relevant biological variability, and restrict statistical integration to only pertinent omics levels and features. The analysis of heterogeneous data remains a technical challenge with the needs of relevant methods and parameters to take into account data specificity. Moreover, applying different statistical methods from several tools could also implies important data management issues. In this context, we developed RFLOMICS, an R package with a shiny interface: it offers a guided, comprehensive and reproducible analysis in a framework which can manage several omics-data. The interface offers an interesting flexibility between the visualization of the results and the data manipulation (filtering, setting).

RFLOMICS is available at : <u>https://forgemia.inra.fr/flomics/rflomics</u>

The list and order of conditions

01. Load data

02. Set statistical model

elect a model formula

Valid model choice

Models are written from the simplest one to the most complete one. Only the two orders interaction terms between the biological factors will appear. Batch factor will never appear in interaction terms



- computes the list of possible model formulas from experimental design setup - all possible contrasts are automatically generated from the chosen model formula



Contrast type : simp

(temperature

(imbibition) I

(imbibition) I

Contrast type : average

ibitionD

ow - temperatureElevated) in imbibitionDS	(temperatureLow - temperatureElevated) in mean		
edium - temperatureElevated) in imbibitionDS	(temperatureMedium - temperatureElevated) in me		
edium - temperatureLow) in imbibitionDS	(temperatureMedium - temperatureLow) in mean		
ow - temperatureElevated) in imbibitionEl	(imbibitionEl - imbibitionDS) in mean		
edium - temperatureElevated) in imbibitionEl	(imbibitionLI - imbibitionDS) in mean		
edium - temperatureLow) in imbibitionEl	(imbibitionLI - imbibitionEI) in mean		
ow - temperatureElevated) in imbibitionLI			
edium - temperatureElevated) in imbibitionLI			
edium - temperatureLow) in imbibitionLl	Contrast type : interaction		
mbibitionDS) in temperatureElevated	 (temperatureLow - temperatureElevated) in imbibit (temperatureLow - temperatureElevated) in imbibit (temperatureMedium - temperatureElevated) in imbibit 		
mbibitionDS) in temperatureLow			
mbibitionDS) in temperatureMedium			
mbibitionDS) in temperatureElevated	(temperatureMedium - temperatureElevated) in int		
mbibitionDS) in temperatureLow	 (temperatureMedium - temperatureLow) in imbibiti (temperatureMedium - temperatureLow) in imbibiti 		
	(F		







Technical choices

- multiAssayExperiment harmonizes data management of different omics levels (matrix, results, parameters...) - **Clustermq** allows the remote computing for resource consuming tasks



6

Acknowledgements

RNAseq transfo

limma (voom

Scale views

This project is part of CATI SySmics [2]. We thank ML. Martin-Magniette, V. Brunaud, G. Rigaill (IPS2) for helpful discussion and feedback. We also thank L. Rajou (IJPB) for given access to Ecoseed project data.

- **rmarkdown** is used to generate a dynamic html report
- **Docker** technology is user to facilitate deployment of RFLOMICS and ensure the reproducibility of analyses

Future developments

- code and interface improvement
- include **clusterProfiler** for functional enrichment
- include missing value imputation strategies for proteomics/metabolomics data
- include supervised integration methods (mixOmics)



References

1. Lambert, I, et al. (2020). "DiCoExpress: a tool to process multifactorial RNAseq experiments from quality controls to co-expression analysis through differential analysis based on contrasts inside GLM models". Plant Methods.

2. https://sysmics.cati.inrae.fr/

- 3. Efstathiou G, et al. (2017). ProteoSign: an end-user online differential proteomics statistical analysis platform. Nucleic Acids Res.
- 4. Argelaguet, Ricard, et al. "MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data." Genome biology 21.1 (2020): 1-17.



Institut Jean-Pierre Bourgin Centre Ile-de-France - Versailles - Grignon Route de Saint-Cyr 78000 Versailles