ARRANZ-García-Alba-PEC1

Alba Arranz

2024-11-04

# Tabla de contenidos

1. Abstract
2. Objetivos del estudio
3. Materiales y métodos
4. Resultados

# Abstract

Este estudio presenta un proceso de análisis de datos hecho con R. Los datos han sido descargados de la base de datos MetabolomicsWorkbechR en formato Summary Experiment. Los datos han sido cargados en el paquete POMA para hacer un análi- sis exploratorio que incluye un análisis univariante y multivariante. Los resul- tados sugieren un posible cambio en los metabolitos aunque no son concluyentes.

# Objetivos del estudio

El objetivo de este estudio es la obtención de unos datos metabolómicos de un repositorio y llevar a cabo un análisis exploratorio de estos.

# Materiales y métodos

Los datos corresponden a diferentes metabolitos detectados por espectrometria de masas (GCMS) en muestras intestinales pre y post transplante. Los datos han sido obtenidos del paquete ‘metabolomicsWorkbenchR’ en un contenedor SummarizedExperiment del paquete con el mismo nombre. Para el análisis explora- torio se ha utilizado el paquete POMA.Primero se han buscado valores perdidos, y después los datos han sido normalizados. Los datos normalizados se han grafi- cado en boxplots y también se han graficado los posibles outliers.

Para el análisis exploratorio primero se ha hecho un análisis univariante con el modelo t-student para hacer una comparación de la presencia de los diferen- tes metabolitos en los dos grupos antes y después del transplante.Después se ha hecho un análisis multivariante mediante el análisis de componentes principales (PCA).

# Resultados

## Datos

Los datos del estudio contienen 12 muestras, de las cuáles 6 han sido obtenidas antes de un transplante y las otras 6 restantes después. Para estas 12 muestras se ha analizado por espectrometría de masas la presencia de 142 metabolitos di- ferentes.

Con los datos normalizados, mediante un boxplot se observa que las muestras an- teriores al transplante presentan una ligera disminución en la presencia de me- tabólitos excepto para una muestra de post transplante que es la que tiene la media más elevada de metabólitos. A pesar de haber cierta dispersión dentro de muestras del mismo grupo no se detecta ningún outlayer en ellas.

## Welcome to POMA!  
## Version 1.14.0  
## POMAShiny app: https://github.com/pcastellanoescuder/POMAShiny

## Cargando paquete requerido: MatrixGenerics

## Cargando paquete requerido: matrixStats

##   
## Adjuntando el paquete: 'MatrixGenerics'

## The following objects are masked from 'package:matrixStats':  
##   
## colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,  
## colCounts, colCummaxs, colCummins, colCumprods, colCumsums,  
## colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,  
## colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,  
## colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,  
## colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,  
## colWeightedMeans, colWeightedMedians, colWeightedSds,  
## colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgsPerColSet,  
## rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,  
## rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,  
## rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,  
## rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,  
## rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,  
## rowWeightedMads, rowWeightedMeans, rowWeightedMedians,  
## rowWeightedSds, rowWeightedVars

## Cargando paquete requerido: GenomicRanges

## Cargando paquete requerido: stats4

## Cargando paquete requerido: BiocGenerics

##   
## Adjuntando el paquete: 'BiocGenerics'

## The following objects are masked from 'package:stats':  
##   
## IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':  
##   
## anyDuplicated, aperm, append, as.data.frame, basename, cbind,  
## colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,  
## get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,  
## match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,  
## Position, rank, rbind, Reduce, rownames, sapply, setdiff, table,  
## tapply, union, unique, unsplit, which.max, which.min

## Cargando paquete requerido: S4Vectors

##   
## Adjuntando el paquete: 'S4Vectors'

## The following object is masked from 'package:utils':  
##   
## findMatches

## The following objects are masked from 'package:base':  
##   
## expand.grid, I, unname

## Cargando paquete requerido: IRanges

##   
## Adjuntando el paquete: 'IRanges'

## The following object is masked from 'package:grDevices':  
##   
## windows

## Cargando paquete requerido: GenomeInfoDb

## Cargando paquete requerido: Biobase

## Welcome to Bioconductor  
##   
## Vignettes contain introductory material; view with  
## 'browseVignettes()'. To cite Bioconductor, see  
## 'citation("Biobase")', and for packages 'citation("pkgname")'.

##   
## Adjuntando el paquete: 'Biobase'

## The following object is masked from 'package:MatrixGenerics':  
##   
## rowMedians

## The following objects are masked from 'package:matrixStats':  
##   
## anyMissing, rowMedians

##   
## Adjuntando el paquete: 'magrittr'

## The following object is masked from 'package:GenomicRanges':  
##   
## subtract

##   
## Adjuntando el paquete: 'dplyr'

## The following object is masked from 'package:Biobase':  
##   
## combine

## The following objects are masked from 'package:GenomicRanges':  
##   
## intersect, setdiff, union

## The following object is masked from 'package:GenomeInfoDb':  
##   
## intersect

## The following objects are masked from 'package:IRanges':  
##   
## collapse, desc, intersect, setdiff, slice, union

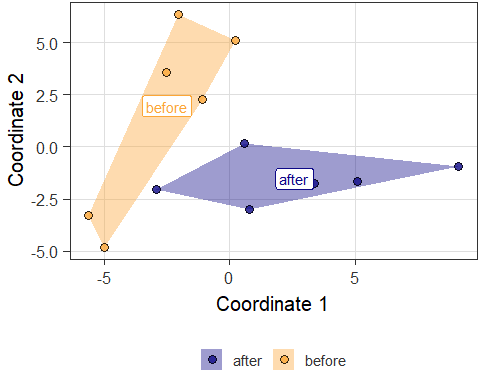
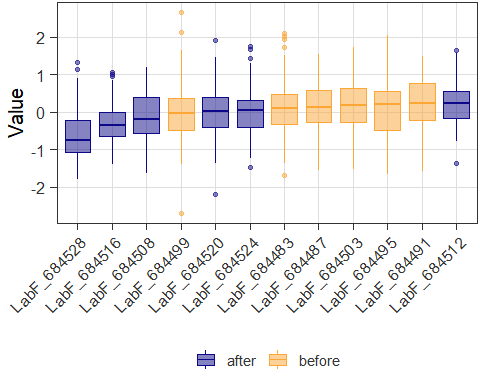
## The following objects are masked from 'package:S4Vectors':  
##   
## first, intersect, rename, setdiff, setequal, union

## The following objects are masked from 'package:BiocGenerics':  
##   
## combine, intersect, setdiff, union

## The following object is masked from 'package:matrixStats':  
##   
## count

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union



## Análisis univariante

Del análisis univariante podemos observar 5 metabolitos con un p valor inferior a 0.1 y por tanto con un diferencia significativa. Estos 5 metabolitos son: áci- do fosfórico, ácido treónico, piruvato, 2-monooctadecanoato, levanbiosa. Única- mente el ácido fosfórico aumenta tras el transplante, mientras que el resto disminuyen.

## method argument is empty. T-test will be used

## [1] "data.frame"

## result.fold\_change result.diff\_means result.pvalue result.adj\_pvalue  
## 1 -1 -1.158 0.002151823 0.08904407  
## 2 -1 1.001 0.002246306 0.08904407  
## 3 -1 1.214 0.002516894 0.08904407  
## 4 -1 1.218 0.002773876 0.08904407  
## 5 -1 1.388 0.003135355 0.08904407  
## result.mean\_after result.mean\_before result.sd\_after result.sd\_before  
## 1 0.5789953 -0.5789953 0.5515047 0.2912252  
## 2 -0.5005952 0.5005952 0.3983429 0.4483731  
## 3 -0.6070583 0.6070583 0.5942935 0.3135637  
## 4 -0.6089716 0.6089716 0.2971105 0.6047527  
## 5 -0.6940017 0.6940017 0.4160816 0.7062611

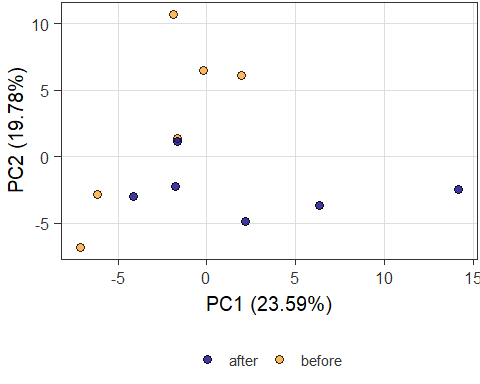
## metabolitos result.fold\_change result.diff\_means result.pvalue  
## 1 Ácido fosfórico -1 -1.158 0.002151823  
## 2 Ácido treónico -1 1.001 0.002246306  
## 3 Piruvato -1 1.214 0.002516894  
## 4 2-monooctadecanoato -1 1.218 0.002773876  
## 5 Levanbiosa -1 1.388 0.003135355  
## result.adj\_pvalue result.mean\_after result.mean\_before result.sd\_after  
## 1 0.08904407 0.5789953 -0.5789953 0.5515047  
## 2 0.08904407 -0.5005952 0.5005952 0.3983429  
## 3 0.08904407 -0.6070583 0.6070583 0.5942935  
## 4 0.08904407 -0.6089716 0.6089716 0.2971105  
## 5 0.08904407 -0.6940017 0.6940017 0.4160816  
## result.sd\_before  
## 1 0.2912252  
## 2 0.4483731  
## 3 0.3135637  
## 4 0.6047527  
## 5 0.7062611

Variación significativa de metabólitos

| metabolitos | result.fold\_change | result.diff\_means | result.pvalue | result.adj\_pvalue | result.mean\_after | result.mean\_before | result.sd\_after | result.sd\_before |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Ácido fosfórico | -1 | -1.158 | 0.0021518 | 0.0890441 | 0.5789953 | -0.5789953 | 0.5515047 | 0.2912252 |
| Ácido treónico | -1 | 1.001 | 0.0022463 | 0.0890441 | -0.5005952 | 0.5005952 | 0.3983429 | 0.4483731 |
| Piruvato | -1 | 1.214 | 0.0025169 | 0.0890441 | -0.6070583 | 0.6070583 | 0.5942935 | 0.3135637 |
| 2-monooctadecanoato | -1 | 1.218 | 0.0027739 | 0.0890441 | -0.6089716 | 0.6089716 | 0.2971105 | 0.6047527 |
| Levanbiosa | -1 | 1.388 | 0.0031354 | 0.0890441 | -0.6940017 | 0.6940017 | 0.4160816 | 0.7062611 |

## Análisis multivariante

El análisis de las componentes principales muestra un ligero agrupamiento de las muestras de antes del transplante más abajo mientras que las de después se agru- pan más a la derecha. Este agrupamiento no es muy obvio y además las dos pri- meras componentes explican un 43.29% de la variabilidad.



# Conclusiones y limitaciones.

Los resultados obtenidos del análisis realizado sugieren que podría haber cierto cambio en la presencia de metabolitos en muestras intestinales antes y después del transplante tendiendo a disminuir. Pero es cierto que en el análisis de compo- nentes principales podría mostrar cierto efecto batch en los resultados. También el bajo número de muestras analizadas impide sacar conclusiones más conclusivas. Se sugiere en un futuro repetir el experimento aumentando el número de muestras a analizar.