

Anàlisi Metabolòmica de la Caquèxia amb SummarizedExperiment

Antoni Mitjana Campos

2024-11-03

```
# Instal·la i carrega les biblioteques necessàries
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("SummarizedExperiment")
library(SummarizedExperiment)
library(readxl)
library(ggplot2)
library(pheatmap)

# Llegeix el dataset de l'arxiu Excel
data <- read_excel("human_cachexia.xlsx")

# Separa les dades de metabolits i les metadades
metabolite_data <- as.data.frame(data[, -(1:2)]) # Exclou les dues primeres columnes
rownames(metabolite_data) <- data$`Patient ID`

# Transposa les dades de metabolits per tenir els pacients com a columnes
metabolite_data <- t(metabolite_data)

# Crea el DataFrame de metadades amb els pacients com a files
metadata <- data.frame(
  Patient_ID = data$`Patient ID`,
  Muscle_loss = data$`Muscle loss`,
  row.names = data$`Patient ID`
)

# Elimina duplicats i valors buits
metadata <- metadata[!duplicated(rownames(metadata)) & !is.na(rownames(metadata)), ]
metabolite_data <- metabolite_data[, !duplicated(colnames(metabolite_data)) & !is.na(colnames(metabolite_data))]

# Filtra els pacients comuns
common_patients <- intersect(rownames(metadata), colnames(metabolite_data))
metabolite_data <- metabolite_data[, common_patients, drop = FALSE]
metadata <- metadata[common_patients, , drop = FALSE]

# Verifica les dimensions després del filtratge
cat("Dimensions després del filtratge:", dim(metabolite_data), dim(metadata), "\n")

## Dimensions després del filtratge: 63 77 77 2
```

```

# Comprova si hi ha pacients que no coincideixen
cat("Pacients en metabolite_data però no en metadata:", setdiff(colnames(
metabolite_data), rownames(metadata)))

## Pacients en metabolite_data però no en metadata:

cat("Pacients en metadata però no en metabolite_data:", setdiff(rownames(metadata), colnames(
metabolite_data)))

## Pacients en metadata però no en metabolite_data:

# Crea el SummarizedExperiment si les dimensions coincideixen
if (ncol(metabolite_data) == nrow(metadata)) {
  se <- SummarizedExperiment(
    assays = list(counts = as.matrix(metabolite_data)),
    colData = metadata
  )
  print("SummarizedExperiment creat amb èxit")
  print(se)
} else {
  stop("Les dimensions de metabolite_data i metadata no coincideixen.")
}

## [1] "SummarizedExperiment creat amb èxit"
## class: SummarizedExperiment
## dim: 63 77
## metadata(0):
## assays(1): counts
## rownames(63): 1,6-Anhydro-beta-D-glucose 1-Methylnicotinamide ...
## pi-Methylhistidine tau-Methylhistidine
## rowData names(0):
## colnames(77): NETCR_002_V1 NETCR_003_V1 ... PIF_192 PIF_195
## colData names(2): Patient_ID Muscle_loss

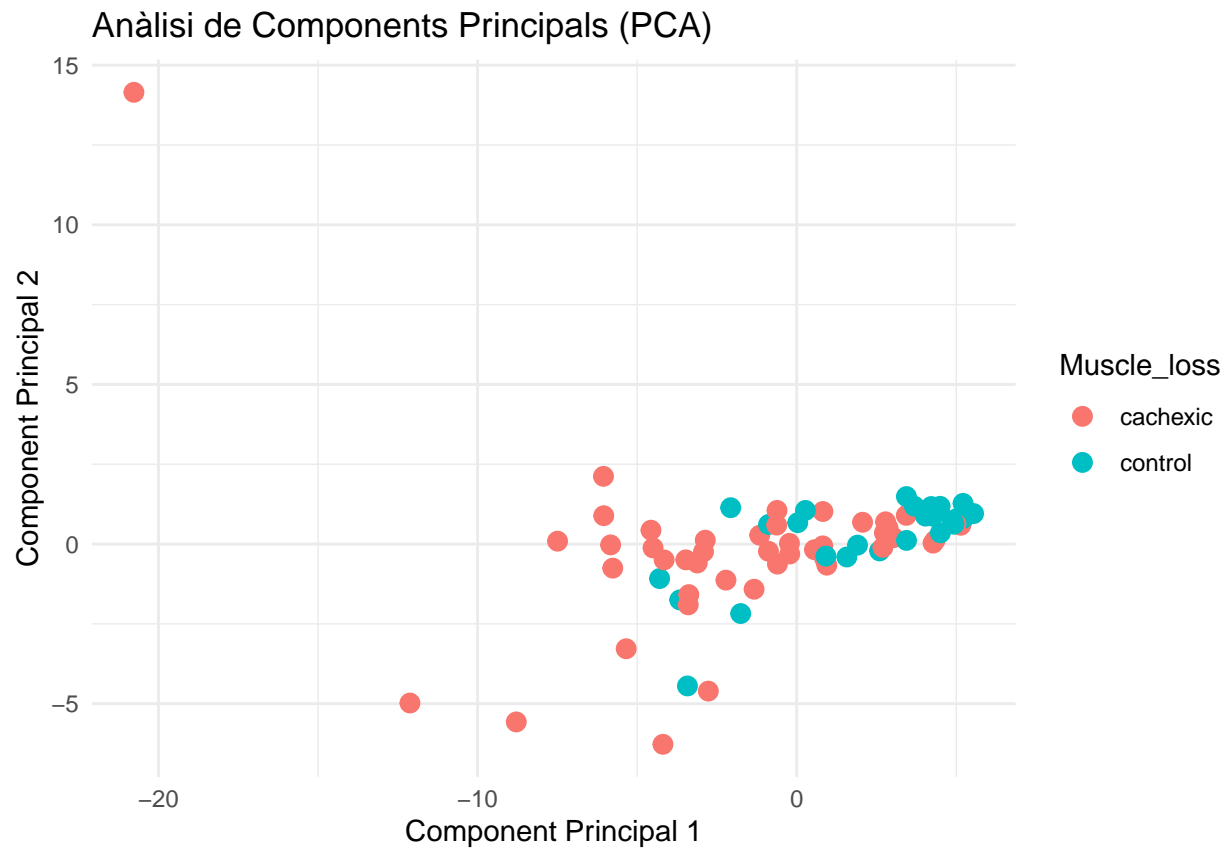
# Guardar l'objecte
save(se, file = "human_cachexia.Rda")

# Anàlisi PCA
pca <- prcomp(t(assay(se)), scale. = TRUE)

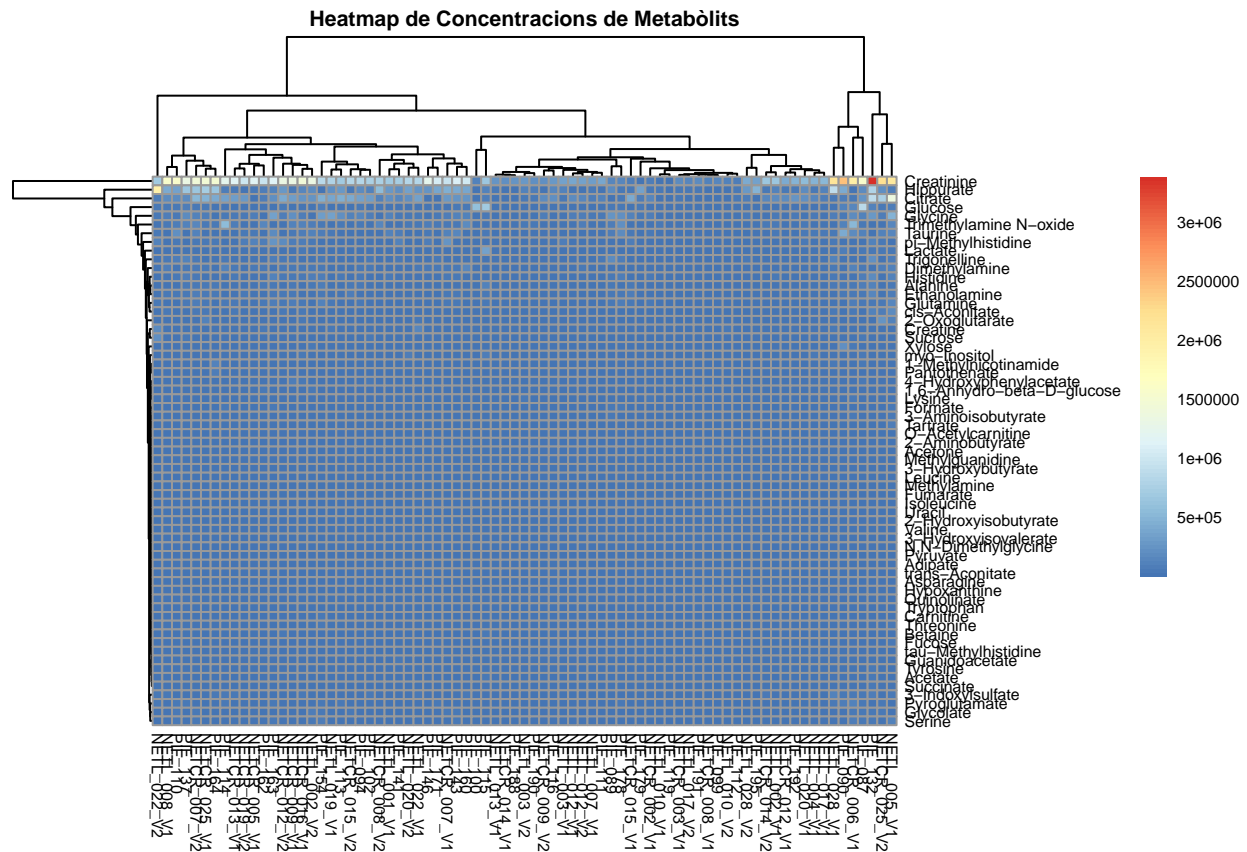
# Convertir el resultat del PCA en un data frame per a ggplot2
pca_data <- data.frame(PC1 = pca$x[, 1], PC2 = pca$x[, 2], Muscle_loss = colData(se)$Muscle_loss)

# Crear el gràfic PCA amb ggplot2
ggplot(pca_data, aes(x = PC1, y = PC2, color = Muscle_loss)) +
  geom_point(size = 3) +
  ggtitle("Anàlisi de Components Principals (PCA)") +
  xlab("Component Principal 1") +
  ylab("Component Principal 2") +
  theme_minimal()

```



```
# Generació de Heatmap amb font més petita
pheatmap(assay(se), main = "Heatmap de Concentracions de Metabòlits",
          cluster_rows = TRUE, cluster_cols = TRUE, fontsize = 6)
```



```
# Aplicació de proves estadístiques amb correcció FDR
resultats <- apply(assay(se), 1, function(x) {
  test <- t.test(x ~ se$Muscle_loss)
  test$p.value
})

# Correcció dels p-valors amb FDR
p_vals_adjusted <- p.adjust(resultats, method = "fdr")

# Seleccionar els metabòlits amb p-valor ajustat < 0.05
significatius <- rownames(assay(se))[p_vals_adjusted < 0.05]

# Comprovar si hi ha resultats significatius
if (length(significatius) > 0) {
  print(paste("Metabòlits significatius:", significatius))
} else {
  print("No s'han trobat metabòlits significatius.")
}

## [1] "Metabòlits significatius: 2-Aminobutyrate"
## [2] "Metabòlits significatius: 2-Hydroxyisobutyrate"
## [3] "Metabòlits significatius: 3-Hydroxybutyrate"
## [4] "Metabòlits significatius: 3-Hydroxyisovalerate"
## [5] "Metabòlits significatius: 3-Indoxylsulfate"
## [6] "Metabòlits significatius: Acetate"
## [7] "Metabòlits significatius: Adipate"
```

```

## [8] "Metabòlits significatius: Alanine"
## [9] "Metabòlits significatius: Asparagine"
## [10] "Metabòlits significatius: Betaine"
## [11] "Metabòlits significatius: Citrate"
## [12] "Metabòlits significatius: Creatine"
## [13] "Metabòlits significatius: Creatinine"
## [14] "Metabòlits significatius: Dimethylamine"
## [15] "Metabòlits significatius: Formate"
## [16] "Metabòlits significatius: Fucose"
## [17] "Metabòlits significatius: Fumarate"
## [18] "Metabòlits significatius: Glucose"
## [19] "Metabòlits significatius: Glutamine"
## [20] "Metabòlits significatius: Glycine"
## [21] "Metabòlits significatius: Hippurate"
## [22] "Metabòlits significatius: Histidine"
## [23] "Metabòlits significatius: Leucine"
## [24] "Metabòlits significatius: Methylamine"
## [25] "Metabòlits significatius: N,N-Dimethylglycine"
## [26] "Metabòlits significatius: Pyroglutamate"
## [27] "Metabòlits significatius: Pyruvate"
## [28] "Metabòlits significatius: Quinolinate"
## [29] "Metabòlits significatius: Serine"
## [30] "Metabòlits significatius: Succinate"
## [31] "Metabòlits significatius: Taurine"
## [32] "Metabòlits significatius: Threonine"
## [33] "Metabòlits significatius: Trigonelline"
## [34] "Metabòlits significatius: Trimethylamine N-oxide"
## [35] "Metabòlits significatius: Tryptophan"
## [36] "Metabòlits significatius: Valine"
## [37] "Metabòlits significatius: cis-Aconitate"
## [38] "Metabòlits significatius: myo-Inositol"
## [39] "Metabòlits significatius: tau-Methylhistidine"

```