

PAC_1

Laia Ortega i Flores

2025-03-25

Sobre R Markdown

Benvingut/da al Rmd la PAC 1. Abans de començar, establiré working directory i packages que necessitare.

```
# Carrego les llibreries necessaries  
library("SummarizedExperiment")
```

```
## S'està carregant el paquet requerit: MatrixGenerics
```

```
## S'està carregant el paquet requerit: matrixStats
```

```
## Warning: el paquet 'matrixStats' es va construir amb la versió d'R 4.4.3
```

```
##
```

```
## S'està adjuntant el paquet: 'MatrixGenerics'
```

```
## Els següents objectes estan emmascarats des de 'package:matrixStats':
```

```
##
```

```
## colAlls, colAnyNAs, colAnys, colAvgPerRowSet, 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, rowAvgPerColSet,  
## 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
```

```
## S'està carregant el paquet requerit: GenomicRanges
```

```
## Warning: el paquet 'GenomicRanges' es va construir amb la versió d'R 4.4.1
```

```
## S'està carregant el paquet requerit: stats4
```

```

## S'està carregant el paquet requerit: BiocGenerics

##
## S'està adjuntant el paquet: 'BiocGenerics'

## Els següents objectes estan emmascarats des de 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## Els següents objectes estan emmascarats des de '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

## S'està carregant el paquet requerit: S4Vectors

## Warning: el paquet 'S4Vectors' es va construir amb la versió d'R 4.4.1

##
## S'està adjuntant el paquet: 'S4Vectors'

## L'objecte següent està emmascarat per 'package:utils':
##
##     findMatches

## Els següents objectes estan emmascarats des de 'package:base':
##
##     expand.grid, I, unname

## S'està carregant el paquet requerit: IRanges

## Warning: el paquet 'IRanges' es va construir amb la versió d'R 4.4.1

##
## S'està adjuntant el paquet: 'IRanges'

## L'objecte següent està emmascarat per 'package:grDevices':
##
##     windows

## S'està carregant el paquet requerit: GenomeInfoDb

## S'està carregant el paquet requerit: Biobase

```

```

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

##
## S'està adjuntant el paquet: 'Biobase'

## L'objecte següent està emmascarat per 'package:MatrixGenerics':
##
##   rowMedians

## Els següents objectes estan emmascarats des de 'package:matrixStats':
##
##   anyMissing, rowMedians

library("limma")

## Warning: el paquet 'limma' es va construir amb la versió d'R 4.4.1

##
## S'està adjuntant el paquet: 'limma'

## L'objecte següent està emmascarat per 'package:BiocGenerics':
##
##   plotMA

library("ggplot2")

## Warning: el paquet 'ggplot2' es va construir amb la versió d'R 4.4.3

library("dplyr")

## Warning: el paquet 'dplyr' es va construir amb la versió d'R 4.4.3

##
## S'està adjuntant el paquet: 'dplyr'

## L'objecte següent està emmascarat per 'package:Biobase':
##
##   combine

## Els següents objectes estan emmascarats des de 'package:GenomicRanges':
##
##   intersect, setdiff, union

## L'objecte següent està emmascarat per 'package:GenomeInfoDb':
##
##   intersect

```

```
## Els següents objectes estan emmascarats des de 'package:IRanges':
##
## collapse, desc, intersect, setdiff, slice, union

## Els següents objectes estan emmascarats des de 'package:S4Vectors':
##
## first, intersect, rename, setdiff, setequal, union

## Els següents objectes estan emmascarats des de 'package:BiocGenerics':
##
## combine, intersect, setdiff, union

## L'objecte següent està emmascarat per 'package:matrixStats':
##
## count

## Els següents objectes estan emmascarats des de 'package:stats':
##
## filter, lag

## Els següents objectes estan emmascarats des de 'package:base':
##
## intersect, setdiff, setequal, union
```

```
library("pheatmap")
```

```
## Warning: el paquet 'pheatmap' es va construir amb la versió d'R 4.4.3
```

```
library("readxl")
```

```
## Warning: el paquet 'readxl' es va construir amb la versió d'R 4.4.3
```

Anàlisi de dades

Preparació de dades

Es carreguen les dades i fem parsing

```
# Llegirem l'excel i fem que les dades siguin en format dataframe
raw_data <- read_xlsx("TI02+PTYR-human-MSS+MSIvsPD.XLSX", sheet = 1) %>% as.data.frame()
metadata_pagina <- read_excel("TI02+PTYR-human-MSS+MSIvsPD.XLSX", sheet = 2)
```

```
## New names:
## * `Sample` -> `Sample...1`
## * `Sample` -> `Sample...2`
```

```
# Agafo les columnes "grup" de la pagina de metadades
metadata <- metadata_pagina[, c("Sample...1", "Phenotype")]
```

```
# Em quedare les dades que ens interessin
dades <- raw_data[, c("SequenceModifications",
                     "M1_1_MSS",
                     "M1_2_MSS",
                     "M5_1_MSS",
                     "M5_2_MSS",
                     "T49_1_MSS",
                     "T49_2_MSS",
                     "M42_1_PD",
                     "M42_2_PD",
                     "M43_1_PD",
                     "M43_2_PD",
                     "M64_1_PD",
                     "M64_2_PD")]
```

```
# Tenim dades duplicades, fare la mitjana
dades <- dades %>%
  group_by(SequenceModifications) %>%
  summarise(across(everything(), mean, na.rm = TRUE))
```

```
## Warning: There was 1 warning in `summarise()`.
## i In argument: `across(everything(), mean, na.rm = TRUE)`.
## i In group 1: `SequenceModifications =
##   "AAAAAPASEDEDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37]
##   Phospho"`.
## Caused by warning:
## ! The `...` argument of `across()` is deprecated as of dplyr 1.1.0.
## Supply arguments directly to `.fns` through an anonymous function instead.
##
##   # Previously
##   across(a:b, mean, na.rm = TRUE)
##
##   # Now
##   across(a:b, \(x) mean(x, na.rm = TRUE))
```

```
# converteixo la primera columna en rownames
dades <- as.data.frame(dades)
rownames(dades) <- dades$SequenceModification
dades <- dades[, -1]
```

```
# Comprovem que les dades estiguin be
head(dades, 3)
```

```
##
## AAAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho M1_1_MSS
## AADSDDGAVSAPAASDGGVSK[4] Phospho 22156.32
## AAESSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 409568.35
##
## AAAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho M1_2_MSS
## AAAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 11826.017
```

```

## AADSDDGAVSAPAASDGGVSK[4] Phospho 9682.942
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 351462.833
## M5_1_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 145141.0
## AADSDDGAVSAPAASDGGVSK[4] Phospho 18736.1
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 1781241.6
## M5_2_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 82939.65
## AADSDDGAVSAPAASDGGVSK[4] Phospho 31641.42
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 1323561.47
## T49_1_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 124955.57
## AADSDDGAVSAPAASDGGVSK[4] Phospho 66936.35
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 33974.62
## T49_2_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 195168.08
## AADSDDGAVSAPAASDGGVSK[4] Phospho 49893.63
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 31569.15
## M42_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 117271.94
## AADSDDGAVSAPAASDGGVSK[4] Phospho 36887.28
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 239551.23
## M42_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 29053.85
## AADSDDGAVSAPAASDGGVSK[4] Phospho 12865.21
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 84362.86
## M43_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 26620.14
## AADSDDGAVSAPAASDGGVSK[4] Phospho 24920.65
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 90274.96
## M43_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 5462.714
## AADSDDGAVSAPAASDGGVSK[4] Phospho 8349.087
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 12687.618
## M64_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 159310.72
## AADSDDGAVSAPAASDGGVSK[4] Phospho 78210.24
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 76375.54
## M64_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 212134.64
## AADSDDGAVSAPAASDGGVSK[4] Phospho 64258.48
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 81279.75

```

Un cop tenim les dades carregades i arreglades creem l'objecte summarizedExperiment

```

# Necessitare dos valors, nom i grups
# Noms contindrà els id de cada modificació
noms <- colnames(dades)

# Grups servirà per dividir les dades entre MSS i PD
grups <- sub(".*_(.*)$", "\\1", noms)

# Ara faig un dataframe on es miren les dades amb la classificació de grups
dataframe_metadata <- data.frame(

```

```

mostra_ID = noms,
grup = factor(grups)
)

# Ara podem veure si les dades s'han classificat correctament
head(dataframe_metadata, 3)

```

```

##  mostra_ID grup
## 1  M1_1_MSS  MSS
## 2  M1_2_MSS  MSS
## 3  M5_1_MSS  MSS

```

```

# Amb el dataframe anterior, podem crear summarized experiment deixant els noms de assay i colData com al
se <- SummarizedExperiment(
  assays = list(counts = as.matrix(dades)),
  colData = dataframe_metadata
)

# Primera observació del se
se

```

```

## class: SummarizedExperiment
## dim: 1437 12
## metadata(0):
## assays(1): counts
## rownames(1437): AAAAAPASEDEDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8]
##   Phospho|[35] Oxidation|[37] Phospho AADSDDGAVSAPAASDGGVSK[4] Phospho
##   ... YYRPTEVDFLQGDCTK[2] Phospho|[14] Carbamidomethyl
##   YYSPCEEHPAETNQNESESGETIR[3] Phospho|[5] Carbamidomethyl
## rowData names(0):
## colnames(12): M1_1_MSS M1_2_MSS ... M64_1_PD M64_2_PD
## colData names(2): mostra_ID grup

```

Summary

Ara ja podem començar a fer una mica d'anàlisi

```

# Veiem les 3 primeres rows de Se, l'assay
head(assay(se), 3)

```

```

##                                                                 M1_1_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 17313.81
## AADSDDGAVSAPAASDGGVSK[4] Phospho 22156.32
## AAESSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 409568.35
##                                                                 M1_2_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 11826.017
## AADSDDGAVSAPAASDGGVSK[4] Phospho 9682.942
## AAESSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 351462.833
##                                                                 M5_1_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDDEEDDSEEEAMETTPAK[8] Phospho|[35] Oxidation|[37] Phospho 145141.0
## AADSDDGAVSAPAASDGGVSK[4] Phospho 18736.1

```

```
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 1781241.6
## M5_2_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 82939.65
## AADSDDGAVSAPAASDGGVSK[4] Phospho 31641.42
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 1323561.47
## T49_1_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 124955.57
## AADSDDGAVSAPAASDGGVSK[4] Phospho 66936.35
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 33974.62
## T49_2_MSS
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 195168.08
## AADSDDGAVSAPAASDGGVSK[4] Phospho 49893.63
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 31569.15
## M42_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 117271.94
## AADSDDGAVSAPAASDGGVSK[4] Phospho 36887.28
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 239551.23
## M42_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 29053.85
## AADSDDGAVSAPAASDGGVSK[4] Phospho 12865.21
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 84362.86
## M43_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 26620.14
## AADSDDGAVSAPAASDGGVSK[4] Phospho 24920.65
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 90274.96
## M43_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 5462.714
## AADSDDGAVSAPAASDGGVSK[4] Phospho 8349.087
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 12687.618
## M64_1_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 159310.72
## AADSDDGAVSAPAASDGGVSK[4] Phospho 78210.24
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 76375.54
## M64_2_PD
## AAAAAPASEDEDEDDEDDEDDEDDEDDEDDSEEEAMETTPAK[8] Phospho[35] Oxidation|[37] Phospho 212134.64
## AADSDDGAVSAPAASDGGVSK[4] Phospho 64258.48
## AAESSSDSSDSDSSEDDEAPSKPAGTTK[5] Phospho 81279.75
```

```
# Veiem les 3 primeres rows de Se, la colData
head(colData(se), 3)
```

```
## DataFrame with 3 rows and 2 columns
##      mostra_ID      grup
##      <character> <factor>
## M1_1_MSS      M1_1_MSS      MSS
## M1_2_MSS      M1_2_MSS      MSS
## M5_1_MSS      M5_1_MSS      MSS
```

```
summary(assay(se))
```

```
##      M1_1_MSS      M1_2_MSS      M5_1_MSS      M5_2_MSS
## Min.   :      0  Min.   :      0  Min.   :      0  Min.   :      0
## 1st Qu.:  5651  1st Qu.:  5488  1st Qu.:  2567  1st Qu.:  3261
```



```
## Median : 30871 Median : 27001 Median : 20853 Median : 26414
## Mean : 225900 Mean : 250298 Mean : 231342 Mean : 259579
## 3rd Qu.: 117475 3rd Qu.: 113194 3rd Qu.: 114138 3rd Qu.: 130208
## Max. :16719906 Max. :43928481 Max. :15135169 Max. :19631820
## T49_1_MSS T49_2_MSS M42_1_PD M42_2_PD
## Min. : 0 Min. : 0 Min. : 0 Min. : 0
## 1st Qu.: 9346 1st Qu.: 8622 1st Qu.: 5323 1st Qu.: 4221
## Median : 55654 Median : 46397 Median : 36887 Median : 30597
## Mean : 525698 Mean : 452818 Mean : 371929 Mean : 319011
## 3rd Qu.: 223267 3rd Qu.: 189197 3rd Qu.: 180508 3rd Qu.: 152696
## Max. :34886738 Max. :29240206 Max. :24091688 Max. :21279848
## M43_1_PD M43_2_PD M64_1_PD M64_2_PD
## Min. : 0 Min. : 0 Min. : 0 Min. : 0
## 1st Qu.: 19633 1st Qu.: 17226 1st Qu.: 11041 1st Qu.: 8677
## Median : 68153 Median : 59616 Median : 52310 Median : 47454
## Mean : 337059 Mean : 349469 Mean : 446015 Mean : 465180
## 3rd Qu.: 205615 3rd Qu.: 201931 3rd Qu.: 210268 3rd Qu.: 206426
## Max. :34431283 Max. :63082982 Max. :49485762 Max. :88912734
```

Creació del docuemnt summarizedExperiment en format binari

```
# Com ja tenim el se definit, nomes cal guardarlo
save(se, file = "summarized_experiment.Rda")
```

Anàlisi abundància entre MSS i PD

Ho faig mitjançant un limma anàlisi i ho represento amb un volcanoplot

```
# Fem la matriu amb les dades de assay
matriu <- assay(se)
# Prepraració de les dades de la metadata
metadata_prep <- colData(se)

# Ara comprovem que els noms de les dades de la matriu son tots diferents
rownames(matriu) <- rownames(se)

# Creem la matriu de disseny de limma
matriu_disseny_limma <- model.matrix(~ 0 + metadata_prep$grup)
colnames(matriu_disseny_limma) <- c("MSS", "PD")

#Ara miro la matriu
print(matriu_disseny_limma)
```

```
##      MSS PD
## 1      1  0
## 2      1  0
## 3      1  0
## 4      1  0
## 5      1  0
## 6      1  0
## 7      0  1
```

```
## 8    0  1
## 9    0  1
## 10   0  1
## 11   0  1
## 12   0  1
## attr("assign")
## [1] 1 1
## attr("contrasts")
## attr("contrasts")$`metadata_prep$grup`
## [1] "contr.treatment"
```

El següent pas, es aplicar el model linial

```
model_linial <- lmFit(matriu, matriu_disseny_limma)

# Faig la antriu de contrast MSS vs PD
matriu_contrast <- makeContrasts(MSS_vs_PD = MSS - PD, levels = matriu_disseny_limma)

# Ara aplico el contrast
model_contrastat <- contrasts.fit(model_linial, matriu_contrast)

# Apliquem el model estadístic de Bayes
model_contrastat <- eBayes(model_contrastat)
```

```
## Warning: Zero sample variances detected, have been offset away from zero
```

```
# Ara mirem els fosfopeptids que estan diferenciats per abundància
resultats_model <- topTable(model_contrastat, coef = "MSS_vs_PD", adjust = "fdr", number = Inf)

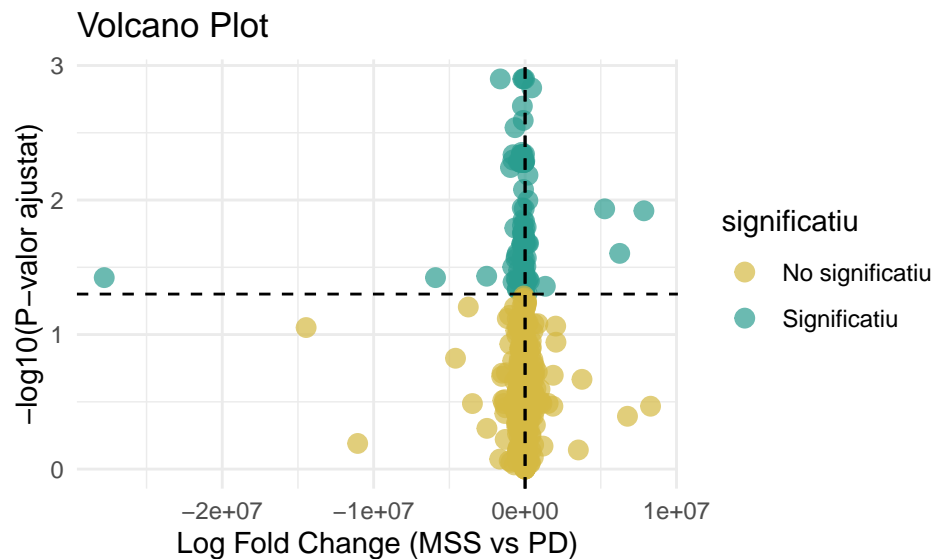
# Veiem els resultats
head(resultats_model)
```

	logFC	AveExpr
## GVGYETILK[4] Phospho	-1637283.85	1193376.08
## KASPEPPDSAEGALK[3] Phospho	-149310.34	147435.55
## AFGYYGPLR[4] Phospho	-27775.59	18777.87
## KDPEDTGAEKSPTTSADLK[14] Phospho	-100676.71	66399.53
## IHDLEDDLEMSSDASDASGEEGGRVPK[10] Oxidation [18] Phospho	435826.87	412752.57
## LDQPVSAPPSPR[10] Phospho	-178774.73	158477.85
##	t	P.Value
## GVGYETILK[4] Phospho	-9.980564	9.705772e-07
## KASPEPPDSAEGALK[3] Phospho	-8.897877	2.925002e-06
## AFGYYGPLR[4] Phospho	-8.888124	2.955616e-06
## KDPEDTGAEKSPTTSADLK[14] Phospho	-8.729772	3.504437e-06
## IHDLEDDLEMSSDASDASGEEGGRVPK[10] Oxidation [18] Phospho	8.387926	5.103973e-06
## LDQPVSAPPSPR[10] Phospho	-7.953677	8.370923e-06
##	adj.P.Val	B
## GVGYETILK[4] Phospho	0.001258969	-4.595119
## KASPEPPDSAEGALK[3] Phospho	0.001258969	-4.595119
## AFGYYGPLR[4] Phospho	0.001258969	-4.595119
## KDPEDTGAEKSPTTSADLK[14] Phospho	0.001258969	-4.595119
## IHDLEDDLEMSSDASDASGEEGGRVPK[10] Oxidation [18] Phospho	0.001466882	-4.595119
## LDQPVSAPPSPR[10] Phospho	0.002004836	-4.595119

Generem un Volcanoplot on es vegi el que s'ha obtingut

```
# Filtrem els valors amb un p < 0.05
resultats_model$significatiu <- ifelse(resultats_model$adj.P.Val < 0.05, "Significatiu", "No significatiu")

# fem un volcanoplot amb les dades Significants
ggplot(resultats_model, aes(x = logFC, y = -log10(adj.P.Val), color = significatiu)) +
  geom_point(alpha = 0.7, size = 3) +
  scale_color_manual(values = c("#D5B942", "#2A9D8F")) +
  geom_vline(xintercept = c(-1, 1), linetype = "dashed") +
  geom_hline(yintercept = -log10(0.05), linetype = "dashed") +
  theme_minimal() +
  labs(title = "Volcano Plot",
       x = "Log Fold Change (MSS vs PD)",
       y = "-log10(P-valor ajustat)")
```



PCA

Seguint els apunts, farem un PCA de les nostres dades

```
# Observo algunes dades que no tenen variància, així que comprovem i arreglem
comprovar_varianca <- apply(assay(se), 1, var, na.rm = TRUE)
sum(comprovar_varianca == 0)
```

```
## [1] 2
```

```
# Trobem 2 fosfolípids que no han tingut canvi, els treiem per fer el PCA
dades_no_var <- assay(se)[comprovar_varianca > 0, ]
dim(dades_no_var)
```

```
## [1] 1435 12
```

```
#Ara podem fer el PCA per veure les agrupacions de mostres dels fosfolipids
pca <- prcomp(t(dades_no_var), scale = TRUE)
```

```
#Ara fem un dataframe epr poder representar-ho en un ggplot
```

```
pca_dades <- data.frame(
  Mostra = colnames(se),
  PC1 = pca$x[, 1],
  PC2 = pca$x[, 2],
  Grup = colData(se)$grup
)
```

```
#Representació en un plot
```

```
ggplot(pca_dades, aes(x = PC1, y = PC2, color = Grup, label = Mostra)) +
  geom_point(size = 2) +
  geom_text(vjust = -1) +
  theme_minimal() +
  labs(title = "PCA d'abundància de fosfopèptids", x = "PC1", y = "PC2")
```

