Pipeline corresponent a nualart_oriol_ADO_PEC1.pdf

Oriol Nualart Mundó

25/4/2020

Codi utilitzat als apartats 3.2 i 4 de l'informe nualart_oriol_ADO_PEC1.pdf.

CONTINGUTS

3.2.1. Preparació de les dades

- Càrrega de l'arxiu targets.csv.

```
targets <- read.csv2("./data/targets.csv", header = TRUE, sep = ";")
```

- Lectura dels arxius .CEL.

```
library(oligo)
```

```
## Loading required package: BiocGenerics
```

```
## Loading required package: parallel
```

```
##
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:parallel':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB
```

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

```
## The following objects are masked from 'package:base':
##
##     anyDuplicated, 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, sort, table, tapply,
##     union, unique, unsplit, which, which.max, which.min
```

```
## Loading required package: oligoClasses
```

```
## Welcome to oligoClasses version 1.48.0
```

```
## Loading required package: Biobase
```

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

```
## Loading required package: Biostrings
```

```
## Loading required package: S4Vectors
```

```
## Loading required package: stats4
```

```
##
## Attaching package: 'S4Vectors'
```

```
## The following object is masked from 'package:base':
##
##     expand.grid
```

```
## Loading required package: IRanges
```

```
##
## Attaching package: 'IRanges'

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

## Loading required package: XVector

##
## Attaching package: 'Biostrings'

## The following object is masked from 'package:base':
##
##     strsplit

## No methods found in package 'RSQLite' for request: 'dbListFields' when loading 'oligo'

## ========================================================================

## Welcome to oligo version 1.50.0

## ========================================================================
```

```
celFiles <- list.celfiles("./data", full.names = TRUE)
library(Biobase)
my.targets <-read.AnnotatedDataFrame(file.path("./data","targets.csv"),
                                     header = TRUE, row.names = 1,
                                     sep=";")
rawData <- read.celfiles(celFiles, phenoData = my.targets)
```

```
## Loading required package: pd.mogene.2.0.st

## Loading required package: RSQLite

## Loading required package: DBI

## Platform design info loaded.

## Reading in : ./data/GSM3561725_NC_mouse_liver-1.CEL
## Reading in : ./data/GSM3561726_NC_mouse_liver-2.CEL
## Reading in : ./data/GSM3561727_NC_mouse_liver-3.CEL
## Reading in : ./data/GSM3561728_NC_mouse_liver-4.CEL
## Reading in : ./data/GSM3561729_NC_mouse_liver-5.CEL
## Reading in : ./data/GSM3561730_HFD_mouse_liver-1.CEL
## Reading in : ./data/GSM3561731_HFD_mouse_liver-2.CEL
## Reading in : ./data/GSM3561732_HFD_mouse_liver-3.CEL
## Reading in : ./data/GSM3561733_HFD_mouse_liver-4.CEL
## Reading in : ./data/GSM3561734_HFD_mouse_liver-5.CEL
## Reading in : ./data/GSM3561735_CPT_mouse_liver-1.CEL
## Reading in : ./data/GSM3561736_CPT_mouse_liver-2.CEL
## Reading in : ./data/GSM3561737_CPT_mouse_liver-3.CEL
## Reading in : ./data/GSM3561738_CPT_mouse_liver-4.CEL
## Reading in : ./data/GSM3561739_CPT_mouse_liver-5.CEL

## Warning in read.celfiles(celFiles, phenoData = my.targets): 'channel'
## automatically added to varMetadata in phenoData.
```

- Canvi del nom llarg de les mostres pel *ShortName*.

```
my.targets@data$ShortName->rownames(pData(rawData))
colnames(rawData) <-rownames(pData(rawData))
```

3.2.2. Control de qualitat de les dades crues

- Controls de qualitat del paquet *arrayQualityMetrics*.

```
library(arrayQualityMetrics)
arrayQualityMetrics(rawData, intgroup = c("Diet", "ShortName"), outdir = "qcRaw",
                    force = TRUE)
```

```
## The report will be written into directory 'qcRaw'.
```

```
## Warning in maximumLevels(fac, n = length(colors)): A factor was provided with 15 levels, but the colo
```

```
## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value
```

```
## (loaded the KernSmooth namespace)

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
```

```
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'
```

- Visualització de la taula resum del control de qualitat i del heatmap. Nota: la imatge de la taula resum és una captura treta de l'informe de resultats.

```
library(knitr)
include_graphics(c("extra_pics/mdRaw.png", "qcRaw/hm.png"))
```

	array	sampleNames	*1	*2	*3	Diet	ShortName
☐	1	NC.1				NC	NC.1
☐	2	NC.2				NC	NC.2
☐	3	NC.3				NC	NC.3
☐	4	NC.4				NC	NC.4
☐	5	NC.5	x		x	NC	NC.5
☐	6	HFD.1				HFD	HFD.1
☐	7	HFD.2				HFD	HFD.2
☐	8	HFD.3				HFD	HFD.3
☐	9	HFD.4				HFD	HFD.4
☐	10	HFD.5				HFD	HFD.5
☐	11	CPT.1	x			CPT	CPT.1
☐	12	CPT.2				CPT	CPT.2
☐	13	CPT.3				CPT	CPT.3
☐	14	CPT.4		x		CPT	CPT.4
☐	15	CPT.5				CPT	CPT.5

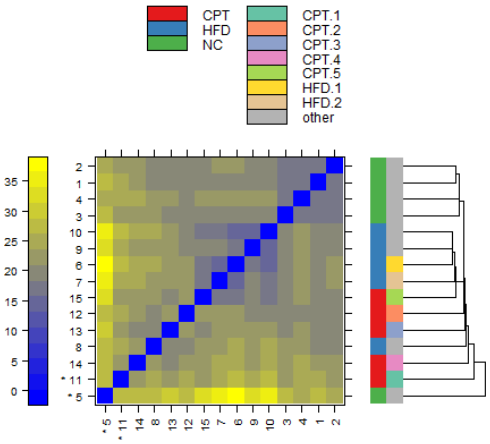


Figure 1: Taula resum del control de qualitat de les dades crues i heatmap

- Càrrega de les dades modificades.

```
targets <- read.csv2("./moddata/targets.csv", header = TRUE, sep = ";")

library(oligo)
celFiles <- list.celfiles("./moddata", full.names = TRUE)
library(Biobase)
my.targets <-read.AnnotatedDataFrame(file.path("./moddata","targets.csv"),
                                     header = TRUE, row.names = 1,
                                     sep=";")
rawData <- read.celfiles(celFiles, phenoData = my.targets)
```

```
## Platform design info loaded.
```

```
## Reading in : ./moddata/GSM3561725_NC_mouse_liver-1.CEL
## Reading in : ./moddata/GSM3561726_NC_mouse_liver-2.CEL
## Reading in : ./moddata/GSM3561727_NC_mouse_liver-3.CEL
## Reading in : ./moddata/GSM3561728_NC_mouse_liver-4.CEL
## Reading in : ./moddata/GSM3561730_HFD_mouse_liver-1.CEL
## Reading in : ./moddata/GSM3561731_HFD_mouse_liver-2.CEL
```

```
## Reading in : ./moddata/GSM3561732_HFD_mouse_liver-3.CEL
## Reading in : ./moddata/GSM3561733_HFD_mouse_liver-4.CEL
## Reading in : ./moddata/GSM3561734_HFD_mouse_liver-5.CEL
## Reading in : ./moddata/GSM3561735_CPT_mouse_liver-1.CEL
## Reading in : ./moddata/GSM3561736_CPT_mouse_liver-2.CEL
## Reading in : ./moddata/GSM3561737_CPT_mouse_liver-3.CEL
## Reading in : ./moddata/GSM3561738_CPT_mouse_liver-4.CEL
## Reading in : ./moddata/GSM3561739_CPT_mouse_liver-5.CEL

## Warning in read.celfiles(celFiles, phenoData = my.targets): 'channel'
## automatically added to varMetadata in phenoData.
```

```
my.targets@data$ShortName->rownames(pData(rawData))
colnames(rawData) <-rownames(pData(rawData))
```

- Repetició del control de qualitat.

```
arrayQualityMetrics(rawData, intgroup = c("Diet", "ShortName"), outdir = "qcMod",
                    force = TRUE)
```

```
## The report will be written into directory 'qcMod'.

## Warning in maximumLevels(fac, n = length(colors)): A factor was provided with 14 levels, but the colo

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value
```

```
## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value
## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
```

```
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'

## Warning in KernSmooth::bkde2D(x, gridsize = nbin, bandwidth = bandwidth):
## Binning grid too coarse for current (small) bandwidth: consider increasing
## 'gridsize'
```

- Visualització de la taula resum i el heatmap.

```
include_graphics(c("extra_pics/mdMod.png", "qcMod/hm.png"))
```

	array	sampleNames	*1	*2	*3	Diet	ShortName
☐	1	NC.1				NC	NC.1
☐	2	NC.2				NC	NC.2
☐	3	NC.3				NC	NC.3
☐	4	NC.4				NC	NC.4
☐	5	HFD.1				HFD	HFD.1
☐	6	HFD.2				HFD	HFD.2
☐	7	HFD.3				HFD	HFD.3
☐	8	HFD.4				HFD	HFD.4
☐	9	HFD.5				HFD	HFD.5
☐	10	CPT.1	x			CPT	CPT.1
☐	11	CPT.2				CPT	CPT.2
☐	12	CPT.3				CPT	CPT.3
☐	13	CPT.4		x		CPT	CPT.4
☐	14	CPT.5				CPT	CPT.5

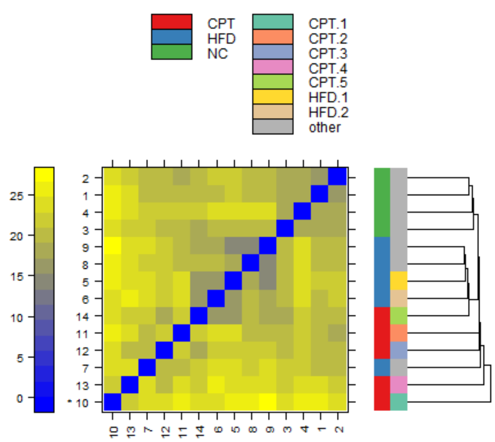


Figure 2: Taula resum del control de qualitat de les dades modificades i heatmap

3.2.3. Normalització

- Correcció del soroll de fons, normalització i sumarització de les dades.

```
library(oligo)
normData <- rma(rawData)
```

```
## Background correcting
## Normalizing
## Calculating Expression
```

3.2.4. Control de qualitat de les dades normalitzades

```
arrayQualityMetrics(normData, intgroup = c("Diet", "ShortName"), outdir = "qcNorm",
                    force = TRUE)
```

```
## The report will be written into directory 'qcNorm'.
```

```
## Warning in maximumLevels(fac, n = length(colors)): A factor was provided with 14 levels, but the colo
```

```
## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value

## Warning in svgStyleAttributes(style, svgdev): Removing non-SVG style attribute
## name(s): subscripts, group.number, group.value
```

- Visualització de la taula resum i el heatmap.

```
include_graphics(c("extra_pics/mdNorm.png", "qcNorm/hm.png"))
```

	array	sampleNames	*1	*2	*3	Diet	ShortName
☐	1	NC.1				NC	NC.1
☐	2	NC.2				NC	NC.2
☐	3	NC.3				NC	NC.3
☐	4	NC.4				NC	NC.4
☐	5	HFD.1				HFD	HFD.1
☐	6	HFD.2				HFD	HFD.2
☐	7	HFD.3				HFD	HFD.3
☐	8	HFD.4				HFD	HFD.4
☐	9	HFD.5				HFD	HFD.5
☐	10	CPT.1	x			CPT	CPT.1
☐	11	CPT.2				CPT	CPT.2
☐	12	CPT.3				CPT	CPT.3
☐	13	CPT.4	x			CPT	CPT.4
☐	14	CPT.5				CPT	CPT.5

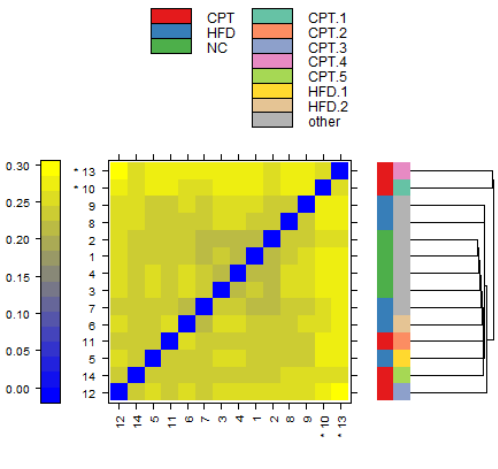


Figure 3: Taula resum del control de qualitat de les dades normalitzades i heatmap

3.2.5. Filtratge no específic

- Filtratge dels gens menys variables.

```
library(genefilter)
library(mogene20sttranscriptcluster.db)
```

```
## Loading required package: AnnotationDbi
```

```
## Loading required package: org.Mm.eg.db
```

```
##
```

```
##
```

```
annotation(normData) <- "mogene20sttranscriptcluster.db"
filtered <- nsFilter(normData,
                     require.entrez = TRUE, remove.dupEntrez = TRUE,
                     var.filter=TRUE, var.func=IQR, var.cutoff=0.5,
                     filterByQuantile=TRUE, feature.exclude = "^AFFX")
filtData <- filtered$eset
```

- Visualització del número de gens descartats.

```
print(filtered$filter.log)
```

```
## $numDupsRemoved
## [1] 671
##
## $numLowVar
## [1] 11982
##
## $numRemoved.ENTREZID
## [1] 16710
```

- Guardem les dades, un cop normalitzades i un cop filtrades, en arxius .csv.

```
write.csv(exprs(normData), file="./results/normData.csv")
write.csv(exprs(filtData), file="./results/filtData.csv")
```

3.2.6. Identificació de gens diferencialment expressats

- Matriu de disseny.

```
library(limma)
```

```
##
## Attaching package: 'limma'
```

```
## The following object is masked from 'package:oligo':
##
##     backgroundCorrect
```

```
## The following object is masked from 'package:BiocGenerics':
##
##     plotMA
```

```
designMat<- model.matrix(~0+Diet, pData(filtData))
colnames(designMat) <- c("CPT", "HFD", "NC")
print(designMat)
```

```
##       CPT HFD NC
## NC.1    0   0  1
## NC.2    0   0  1
## NC.3    0   0  1
## NC.4    0   0  1
## HFD.1   0   1  0
## HFD.2   0   1  0
## HFD.3   0   1  0
## HFD.4   0   1  0
## HFD.5   0   1  0
## CPT.1   1   0  0
## CPT.2   1   0  0
## CPT.3   1   0  0
## CPT.4   1   0  0
## CPT.5   1   0  0
## attr(,"assign")
## [1] 1 1 1
## attr(,"contrasts")
## attr(,"contrasts")$Diet
## [1] "contr.treatment"
```

- Matriu de contrastos.

```
contMat <- makeContrasts (HFDvsCPT = CPT-HFD,
                          NCvsHFD = HFD-NC,
                          NCvsCPT = CPT-NC,
                          levels=designMat)
print(contMat)
```

```
##       Contrasts
## Levels HFDvsCPT NCvsHFD NCvsCPT
##    CPT        1       0       1
##    HFD       -1       1       0
##    NC         0      -1      -1
```

- Estimació del model.

```
library(limma)
fit<-lmFit(filtData, designMat)
```

- Estimació dels contrastos.

```
fit.main<-contrasts.fit(fit, contMat)
fit.main<-eBayes(fit.main)
```

- Guardem els resultats en un arxiu.

```
save(fit.main, file="./results/liver.fit.main.Rda")
```

- Exploració dels gens més diferencialment expressats en cada comparació.

  HFD vs CPT:

```
top_HFDvsCPT <- topTable (fit.main, number=nrow(fit.main), coef="HFDvsCPT",
                          adjust="fdr")
head(top_HFDvsCPT)
```

```
##               logFC  AveExpr         t      P.Value  adj.P.Val        B
## 17335540  0.9994833 5.082067  7.319599 1.569727e-06 0.01880847 4.793782
## 17364642  0.8786221 7.263510  6.269726 1.041219e-05 0.04481886 3.285387
## 17411955 -1.5712617 2.671452 -6.229835 1.122155e-05 0.04481886 3.224270
## 17246967  1.2477222 7.255852  5.909740 2.062156e-05 0.05603154 2.723798
## 17260644 -0.8431216 5.115747 -5.732672 2.904490e-05 0.05603154 2.439352
## 17481725  0.8290971 4.589389  5.690607 3.152622e-05 0.05603154 2.370993
```

NC vs HFD:

```
top_NCvsHFD <- topTable (fit.main, number=nrow(fit.main), coef="NCvsHFD", adjust="fdr")
head(top_NCvsHFD)
```

```
##              logFC  AveExpr         t      P.Value    adj.P.Val        B
## 17468759 -0.972245 2.589143 -9.402244 5.620902e-08 0.0006217612 8.051459
## 17254395 -1.409242 5.008661 -8.992759 1.037825e-07 0.0006217612 7.550835
## 17269521 -1.835453 6.236747 -7.838127 6.512347e-07 0.0015196324 6.007204
## 17273348 -2.156673 6.020434 -7.831958 6.579512e-07 0.0015196324 5.998408
## 17362717 -1.577277 7.702808 -7.811196 6.810921e-07 0.0015196324 5.968761
## 17342642 -1.996605 6.308200 -7.661105 8.759012e-07 0.0015196324 5.752406
```

NC vs CPT:

```
top_NCvsCPT <- topTable (fit.main, number=nrow(fit.main), coef="NCvsCPT",
                          adjust="fdr")
head(top_NCvsCPT)
```

```
##              logFC  AveExpr          t      P.Value    adj.P.Val        B
## 17362717 -2.128860 7.702808 -10.542817 1.120887e-08 0.0001132889 9.530148
## 17468759 -1.050836 2.589143 -10.162270 1.890984e-08 0.0001132889 9.107213
## 17254395 -1.525307 5.008661  -9.733403 3.470289e-08 0.0001386034 8.608845
## 17435934 -1.150220 9.633207  -9.181977 7.798893e-08 0.0001979705 7.932493
## 17482897 -1.932815 4.681236  -9.143625 8.261161e-08 0.0001979705 7.883909
## 17399266 -2.061990 7.776123  -9.018480 9.980473e-08 0.0001993100 7.723951
```

3.2.7. Anotació dels resultats

- Funció *annotatedTopTable*.

```
annotatedTopTable <- function(topTab, anotPackage)
{
 topTab <- cbind(PROBEID=rownames(topTab), topTab)
 myProbes <- rownames(topTab)
 thePackage <- eval(parse(text = anotPackage))
 geneAnots <- select(thePackage, myProbes, c("SYMBOL", "ENTREZID", "GENENAME"))
 annotatedTopTab<- merge(x=geneAnots, y=topTab, by.x="PROBEID", by.y="PROBEID")
return(annotatedTopTab)
}
```

- Generació de les anotacions.

```
library(mogene20sttranscriptcluster.db)

annot_HFDvsCPT <- annotatedTopTable(top_HFDvsCPT,
                                    anotPackage="mogene20sttranscriptcluster.db")
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
annot_NCvsHFD <- annotatedTopTable(top_NCvsHFD,
                                   anotPackage="mogene20sttranscriptcluster.db")
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
annot_NCvsCPT <- annotatedTopTable(top_NCvsCPT,
                                   anotPackage="mogene20sttranscriptcluster.db")
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

- Guardem els *data frames* obtinguts en arxius .csv.

```
write.csv(annot_HFDvsCPT, file="./results/annot_HFDvsCPT.csv")
write.csv(annot_NCvsHFD, file="./results/annot_NCvsHFD.csv")
write.csv(annot_NCvsCPT, file="./results/annot_NCvsCPT.csv")
```

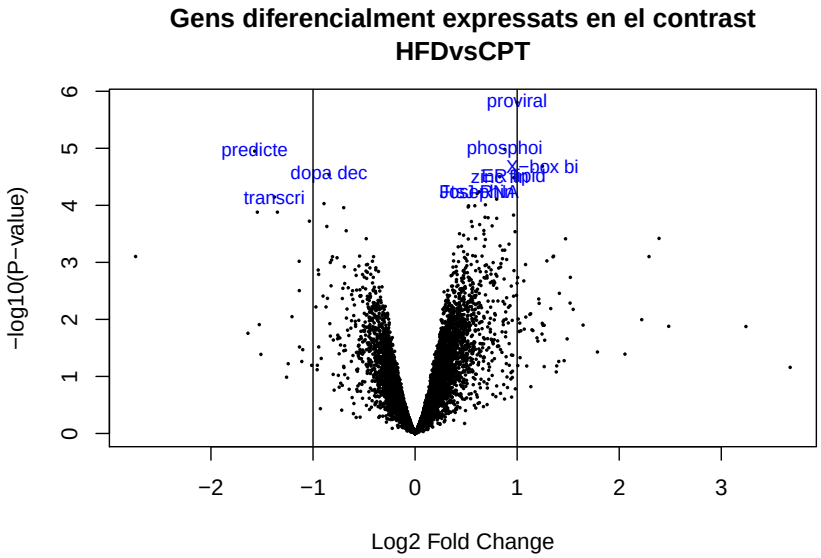
- Volcano plots.

```
geneNames <- select(mogene20sttranscriptcluster.db, rownames(fit.main), c("GENENAME"))
```

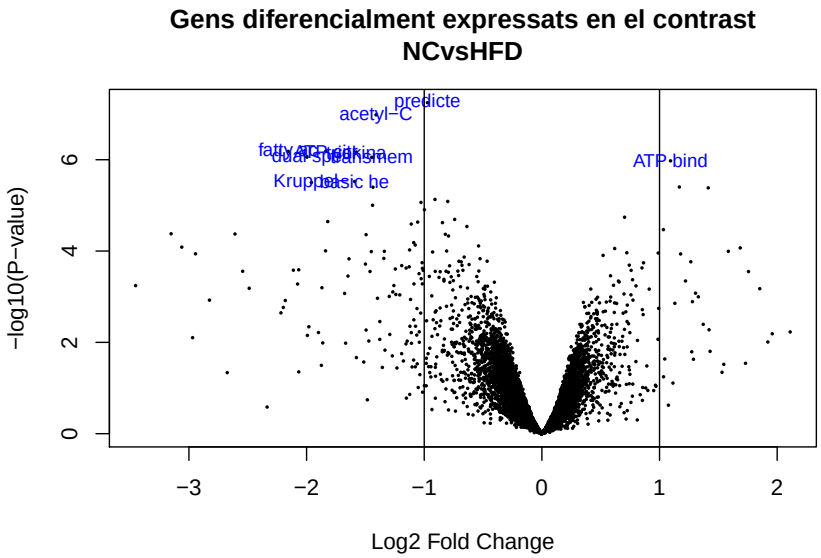
```
## 'select()' returned 1:1 mapping between keys and columns
```

```
geneNames <- geneNames$GENENAME
opt <- par(cex.lab = 0.7)
```

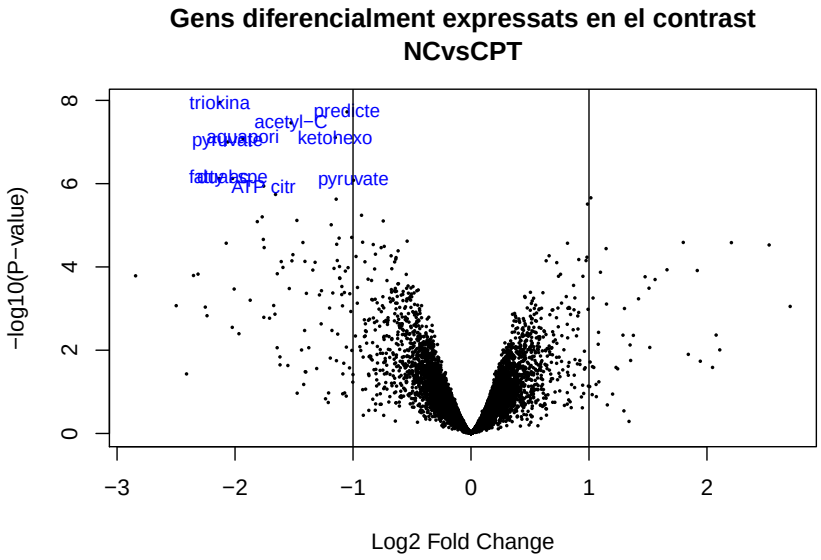
```
 HFD vs CPT:
```



NC vs HFD:



NC vs CPT:



- Selecció dels gens diferencialment expressats.

```
library(limma)
res <- decideTests(fit.main, method="separate", adjust.method="fdr", p.value=0.1, lfc=0.75)
sum.res.rows <- apply(abs(res), 1, sum)
res.selected <- res[sum.res.rows!=0,]
```

- *Heatmap*

```
library(gplots)
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:IRanges':
##
##     space
```

```
## The following object is masked from 'package:S4Vectors':
##
##     space
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

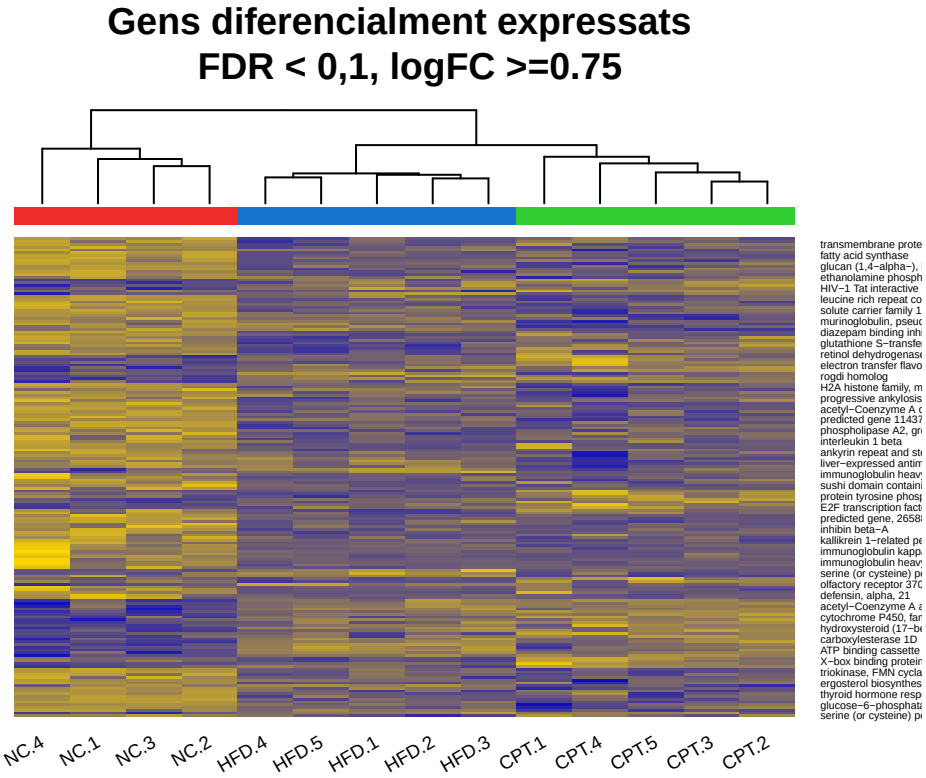
```
genesHeatmap <- rownames(res.selected)
heatData <- exprs(filtData)[rownames(exprs(filtData)) %in% genesHeatmap,]
geneNames <- select(mogene20sttranscriptcluster.db, rownames(heatData), c("GENENAME"))
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
GENENAMES<- geneNames$GENENAME
rownames(heatData) <- GENENAMES
my_palette <- colorRampPalette(c("blue3", "gold"))(n = 299)
```

```
heatmap.2(heatData,
          Rowv = TRUE,
          Colv = TRUE,
```

```
          dendrogram = "column",
          main = "Gens diferencialment expressats \n FDR < 0,1, logFC >=0.75",
          scale = "row",
          col = my_palette,
          sepcolor = "white",
          sepwidth = c(0.05,0.05),
          cexRow = 0.5,
          cexCol = 0.9,
          key = FALSE,
          density.info = "histogram",
          ColSideColors = c(rep("firebrick2",4), rep("dodgerblue3",5),
                            rep("limegreen",5)),
          tracecol = NULL,
          srtCol = 30)
```



3.2.8. Comparació entre comparacions

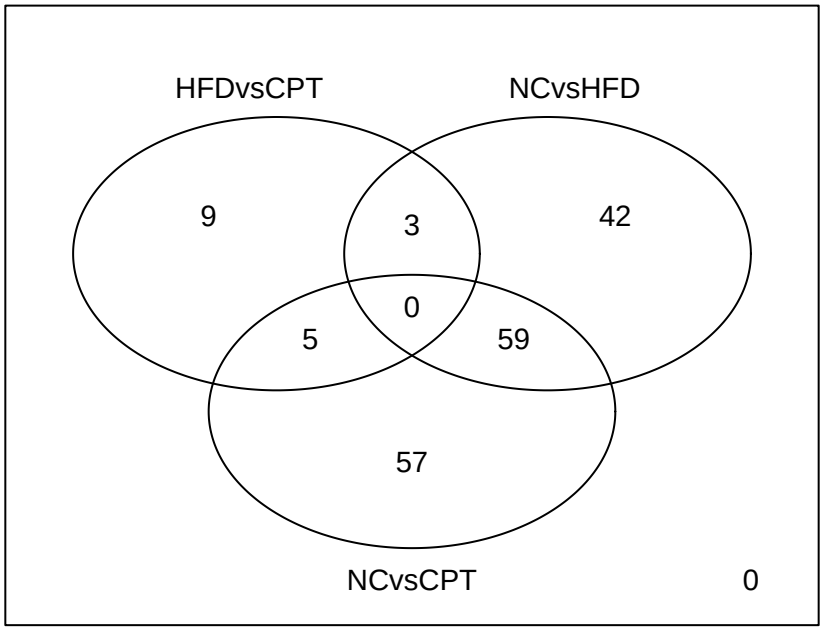
- Número de gens diferencialment expressats en cada comparació.

```
print(summary(res))
```

```
##        HFDvsCPT NCvsHFD NCvsCPT
## Down          8      81      86
## NotSig    11965   11878   11861
## Up            9      23      35
```

- Diagrama de Venn.

```
vennDiagram (res.selected[,1:3], cex=0.9)
title("Gens diferencialment expressats segons cada comparació.")
```



3.2.9. Anàlisi de significació biològica

- Generació dels llistats d'identificadors *Entrez*.

```
genesHFDvsCPT <- top_HFDvsCPT["adj.P.Val"]<0.25
IDsHFDvsCPT <- rownames(top_HFDvsCPT)[genesHFDvsCPT]
EntrezHFDvsCPT <- select(mogene20sttranscriptcluster.db, IDsHFDvsCPT, c("ENTREZID"))
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
EntrezHFDvsCPT <- EntrezHFDvsCPT$ENTREZID

genesNCvsHFD <- top_NCvsHFD["adj.P.Val"]<0.25
IDsNCvsHFD <- rownames(top_NCvsHFD)[genesNCvsHFD]
EntrezNCvsHFD <- select(mogene20sttranscriptcluster.db, IDsNCvsHFD, c("ENTREZID"))
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
EntrezNCvsHFD <- EntrezNCvsHFD$ENTREZID

genesNCvsCPT <- top_NCvsCPT["adj.P.Val"]<0.25
IDsNCvsCPT <- rownames(top_NCvsCPT)[genesNCvsCPT]
EntrezNCvsCPT <- select(mogene20sttranscriptcluster.db, IDsNCvsCPT, c("ENTREZID"))
```

```
## 'select()' returned 1:1 mapping between keys and columns
```

```
EntrezNCvsCPT <- EntrezNCvsCPT$ENTREZID

selectedIDs <- list(EntrezHFDvsCPT, EntrezNCvsHFD, EntrezNCvsCPT)
names(selectedIDs) <- c("HFDvsCPT", "NCvsHFD", "NCvsCPT")
```

- Número de gens inclosos en cada llistat.

```
sapply(selectedIDs, length)
```

```
## HFDvsCPT  NCvsHFD  NCvsCPT
##      506      456      647
```

- Genergeneració del llistat dels gens de ratolí amb anotacions a GO i a KEGG.

```
library(org.Mm.eg.db)
mapped_genes2GO <- mappedkeys(org.Mm.egGO)
mapped_genes2KEGG <- mappedkeys(org.Mm.egPATH)
mapped_genes <- union(mapped_genes2GO , mapped_genes2KEGG)
```

- Anàlisi d'enriquiment de termes GO.

```
library(clusterProfiler)
```

```
##
```

```
## Registered S3 method overwritten by 'enrichplot':
##   method               from
##   fortify.enrichResult DOSE
```

```
## clusterProfiler v3.14.3  For help: https://guangchuangyu.github.io/software/clusterProfiler
##
## If you use clusterProfiler in published research, please cite:
## Guangchuang Yu, Li-Gen Wang, Yanyan Han, Qing-Yu He. clusterProfiler: an R package for comparing biol
```

```
listOfData <- selectedIDs[1:3]
comparisonsNames <- names(listOfData)

for (i in 1:length(listOfData)){
        genesIn <- listOfData[[i]]
        comparison <- comparisonsNames[i]

        enrich.GO <- enrichGO(gene = genesIn,
                              OrgDb = org.Mm.eg.db,
                              pvalueCutoff = 0.05,
                              pAdjustMethod = "BH",
                              universe = mapped_genes)

        if (length(rownames(enrich.GO@result)) != 0) {
                write.csv(as.data.frame(enrich.GO),
                          file = paste0("./results/", "enrichGO.Results.", comparison,
                                        ".csv"),
                          row.names = FALSE)

                png(file=paste0("./results/","enrichGOBarplot.",comparison,".png"),
                    width = 800)
```

```
                print(barplot(enrich.GO, showCategory = 15, font.size = 8,
                              title = paste0("Anàlisi d'enriquiment de termes GO per ",
                                             comparison)))
                dev.off()

                png(file = paste0("./results/","enrichGOcnetplot.",comparison,".png"))
                print(cnetplot(enrich.GO, categorySize = "geneNum",
                               schowCategory = 15, vertex.label.cex = 0.75))
                dev.off()
        }
}
```

- Anàlisi d'enriquiment de *pathways* KEGG.

```
for (i in 1:length(listOfData)){
        genesIn <- listOfData[[i]]
        comparison <- comparisonsNames[i]

        enrich.KEGG <- enrichKEGG(gene = genesIn,
                                  organism     = 'mmu',
                                  pvalueCutoff = 0.1,
                                  pAdjustMethod = "BH",
                                  universe = mapped_genes)


        if (length(rownames(enrich.KEGG@result)) != 0) {
                write.csv(as.data.frame(enrich.KEGG),
                          file = paste0("./results/", "enrichKEGG.Results.", comparison,
                                        ".csv"),
                          row.names = FALSE)

                png(file=paste0("./results/","enrichKEGGBarplot.",comparison,".png"),
                    width = 800)
                print(barplot(enrich.KEGG, showCategory = 15, font.size = 8,
                              title = paste0("Anàlisi d'enriquiment de pathways KEGG per ", comparison)))
                dev.off()

                png(file = paste0("./results/","enrichKEGGcnetplot.",comparison,".png"))
                print(cnetplot(enrich.KEGG, categorySize = "geneNum",
                               schowCategory = 15, vertex.label.cex = 0.75))
                dev.off()
        }
}
```

4. Resultats

- Imatges dels gràfics de resultats.

```
library(knitr)
include_graphics("results/enrichGOBarplot.HFDvsCPT.png")
```



```
include_graphics("results/enrichKEGGBarplot.HFDvsCPT.png")
```

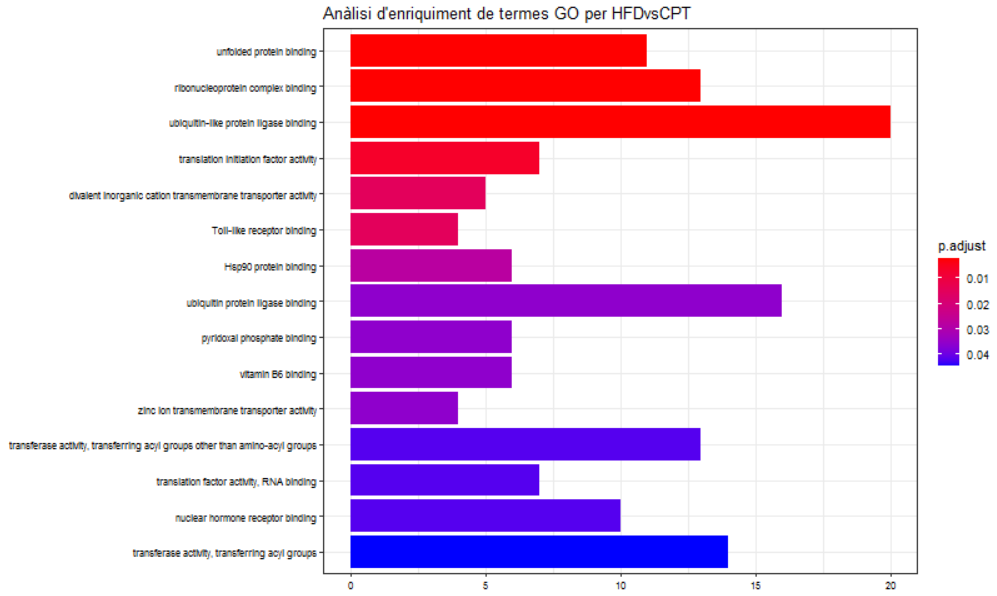


Figure 4: Gràfic de barres de termes GO per HFD vs CPT

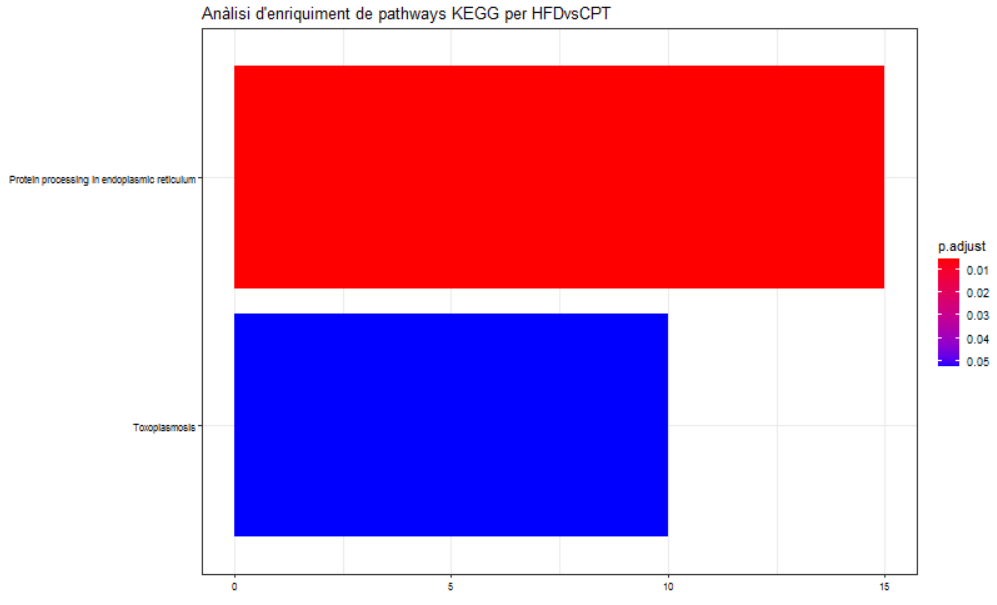


Figure 5: Gràfic de barres de pathways per HFD vs CPT

```
include_graphics(c("results/enrichGOcnetplot.HFDvsCPT.png",
                   "results/enrichKEGGcnetplot.HFDvsCPT.png"))
```

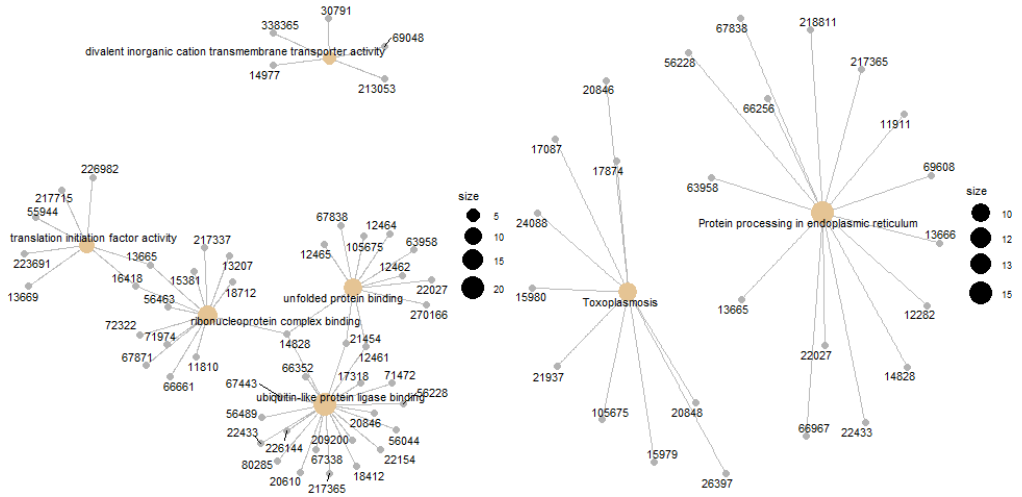


Figure 6: Gràfics de ret de termes GO i de pathways per HFD vs CPT

```
include_graphics("results/enrichGOBarplot.NCvsHFD.png")
```

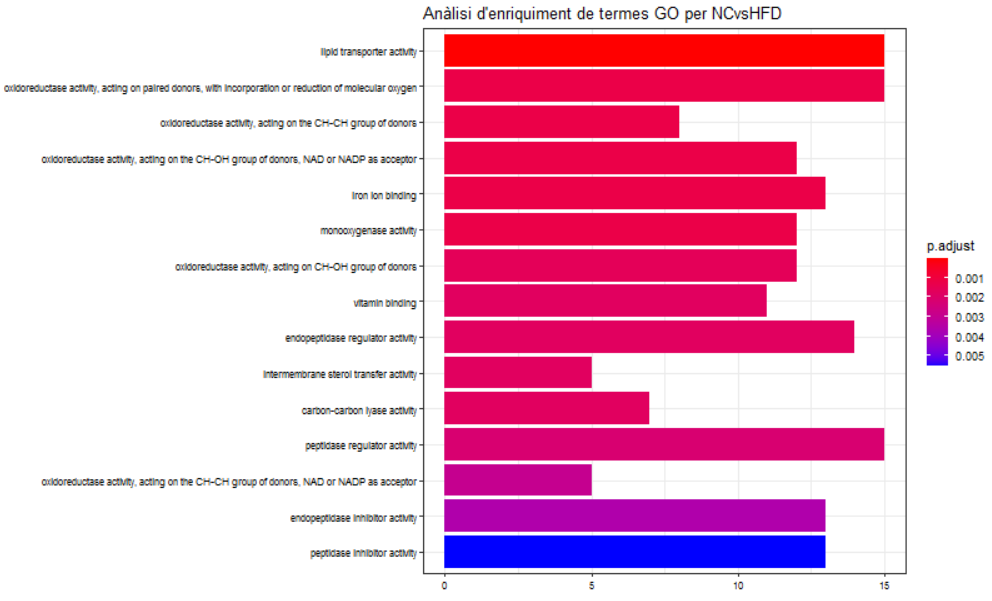


Figure 7: Gràfic de barres de termes GO per NC vs HFD

```
include_graphics("results/enrichKEGGBarplot.NCvsHFD.png")
```

```
include_graphics(c("results/enrichGOcnetplot.NCvsHFD.png",
                   "results/enrichKEGGcnetplot.NCvsHFD.png"))
```

```
include_graphics("results/enrichGOBarplot.NCvsCPT.png")
```

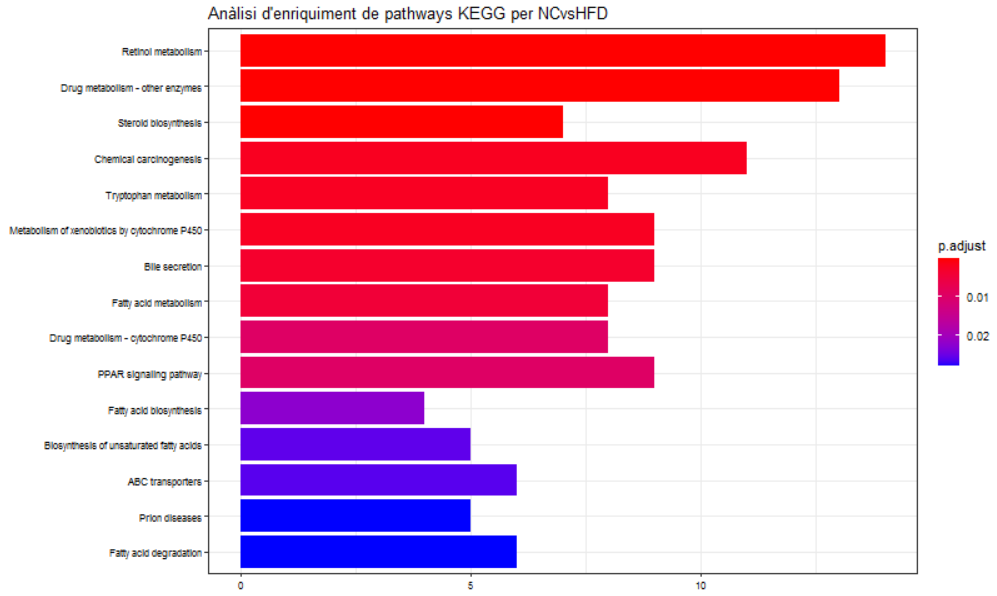


Figure 8: Gràfic de barres de pathways per NC vs HFD

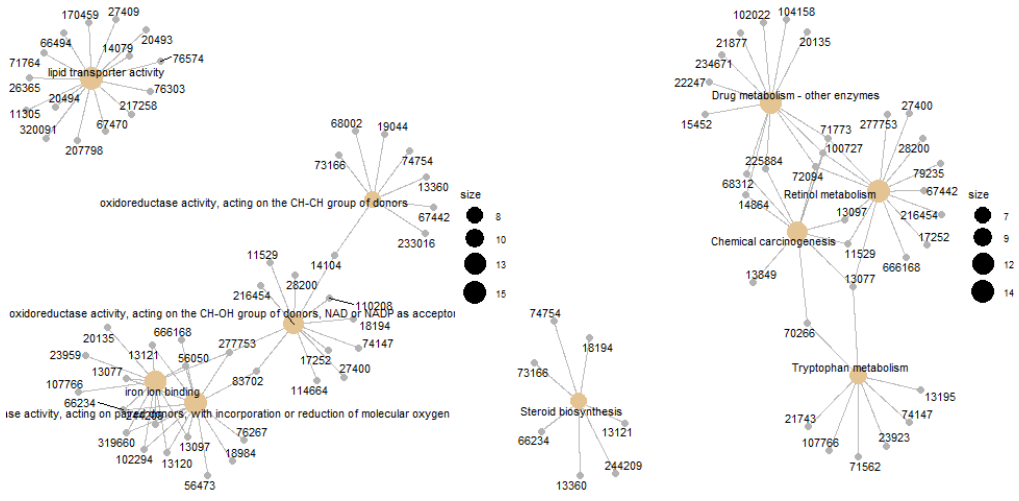


Figure 9: Gràfics de ret de termes GO i de pathways per NC vs HFD

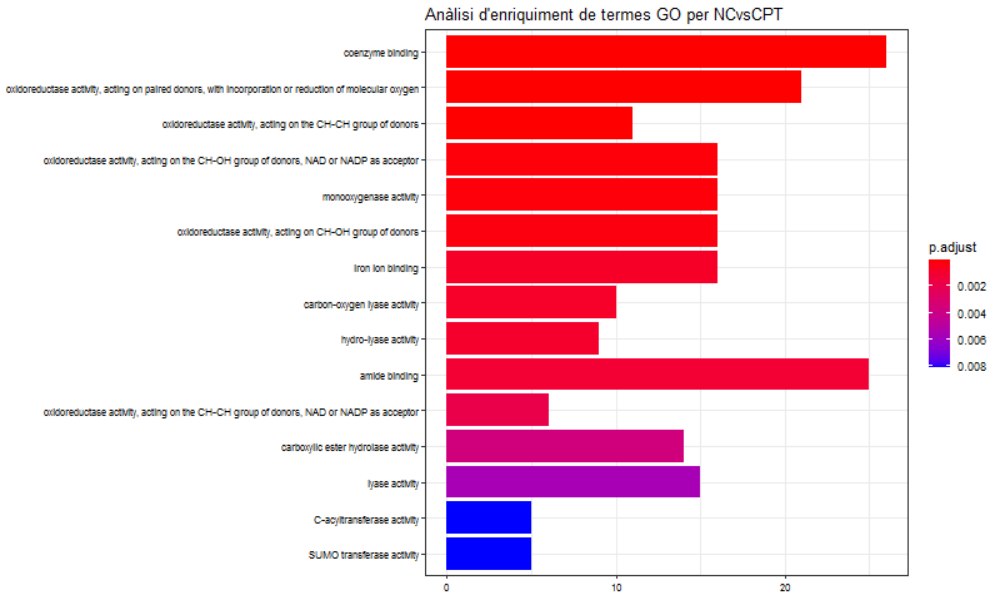


Figure 10: Gràfic de barres de termes GO per NC vs CPT

```
include_graphics("results/enrichKEGGBarplot.NCvsCPT.png")
```

```
include_graphics(c("results/enrichGOcnetplot.NCvsCPT.png",
                   "results/enrichKEGGcnetplot.NCvsCPT.png"))
```

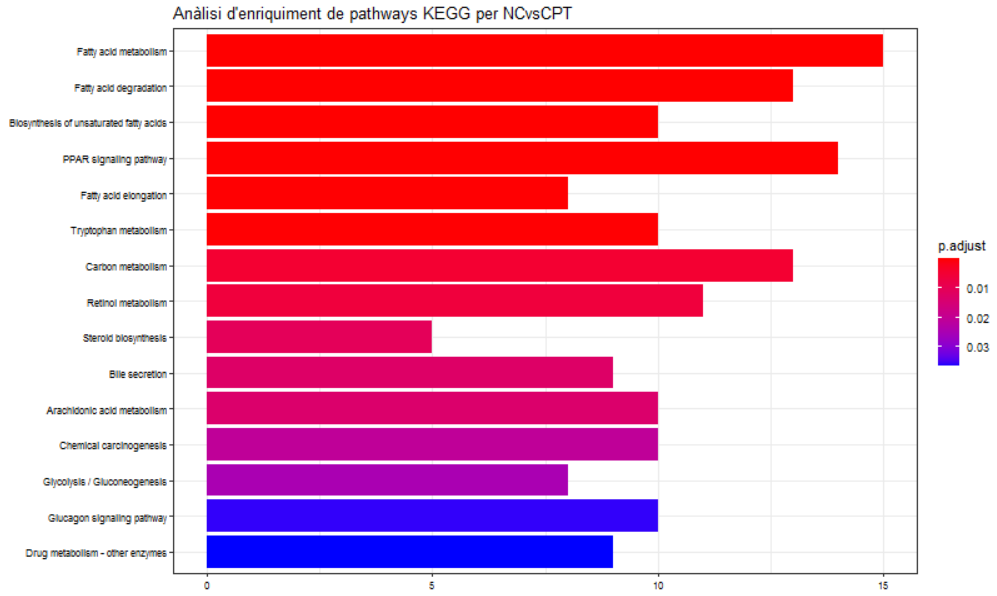


Figure 11: Gràfic de barres de pathways per NC vs CPT

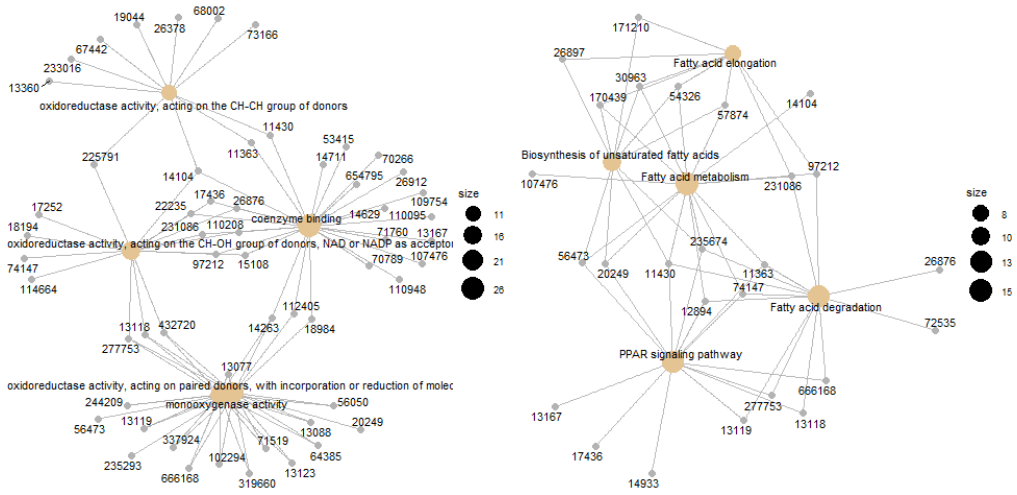


Figure 12: Gràfics de ret de termes GO i de pathways per NC vs CPT