Anàlisi de dades òmiques (M0-157). PAC1.

Marc Anton

2025-03-26

1. Introducció

Aquest arxiu conté les sol·lucions als problemes plantejats en la PAC1 *Anàlisi de dades òmiques (M0-157)*. El projecte total del treball es pot consultar al: meu repositori github, cliqueu **aquí**

2. Carrega de dades

A continuació, procedeixo a resoldre la PAC. Abans que res, carreguem les dades al nostre entorn de treball. Treballarem amb l'exemple de mostra, veure **aquí**. Carreguem els dos fitxers:

```
DataInfo_S013 <- read.csv("DataInfo_S013.csv", sep = ",")
DataValues_S013 <- read.csv("DataValues_S013.csv", sep = ",")
```

Com que ja sabem que aquests fitxers solen portar molts camps, mirem primer l'estructura dels dos.

3. Eploració previa de les dades descarregades

```
dim(DataInfo_S013)
```

```
## [1] 695   4
```

```
dim(DataValues_S013)
```

```
## [1]  39 696
```

Aqui ja veiem que en l'arxiu Values tenim 696 columnes, el que correspon a un index únic de cada cas d'estudi més les 695 variables que es tenen de cada un d'ells. Per veure que hi ha, mirem les 20 primeres files d'Info i podem fer-nos una idea

```
head(DataInfo_S013,20)
```

```
##                   X      VarName    varTpe Description
## 1          SUBJECTS     SUBJECTS   integer    dataDesc
## 2           SURGERY      SURGERY character    dataDesc
## 3               AGE          AGE   integer    dataDesc
## 4            GENDER       GENDER character    dataDesc
## 5             Group        Group   integer    dataDesc
## 6          MEDDM_T0     MEDDM_T0   integer    dataDesc
## 7         MEDCOL_T0    MEDCOL_T0   integer    dataDesc
## 8         MEDINF_T0    MEDINF_T0   integer    dataDesc
## 9         MEDHTA_T0    MEDHTA_T0   integer    dataDesc
## 10          GLU_T0       GLU_T0   integer    dataDesc
## 11          INS_T0       INS_T0   numeric    dataDesc
```

```
## 12          HOMA_T0          HOMA_T0   numeric   dataDesc
## 13          HBA1C_T0          HBA1C_T0   numeric   dataDesc
## 14 HBA1C.mmol.mol_T0 HBA1C.mmol.mol_T0   numeric   dataDesc
## 15           PESO_T0          PESO_T0   integer   dataDesc
## 16            bmi_T0           bmi_T0   numeric   dataDesc
## 17             CC_T0            CC_T0   numeric   dataDesc
## 18           CINT_T0          CINT_T0   integer   dataDesc
## 19            CAD_T0           CAD_T0   integer   dataDesc
## 20            TAD_T0           TAD_T0   integer   dataDesc
```

Aquí ja veiem que hi ha 9 primers camps que semblen ser els que tenen la informació sobre el pacient mostra. Mirem que contenen aquests camps a values:

```
# En mirem 10 en realitat, per comprovar que la primera columna és un index
head(DataValues_S013[,1:10])
```

```
##   X.1 SUBJECTS SURGERY AGE GENDER Group MEDDM_T0 MEDCOL_T0 MEDINF_T0 MEDHTA_T0
## 1   1        1 by pass  27      F     1        0         0         0         1
## 2   2        2 by pass  19      F     2        0         0         0         0
## 3   3        3 by pass  42      F     1        0         0         0         0
## 4   4        4 by pass  37      F     2        0         0         0         0
## 5   5        5 tubular  42      F     1        0         0         0         0
## 6   6        6 by pass  24      F     2        0         0         0         0
```

Com suposavem, aquestes columnes corresponen a: 1. Un codi de pacient (que està repetit a les dues primeres columnes) 2. La cirugia a la que va ser sotmés el pacient 3. L'edat i el gènere 4. Un grup de tractament 5. 4 tractaments que venen amb 0 i 1

Per veure que la resta són metabolits, mirem 5 columnes més

```
head(DataValues_S013[,11:15],10)
```

```
##    GLU_T0 INS_T0 HOMA_T0 HBA1C_T0 HBA1C.mmol.mol_T0
## 1      85  11.40    2.40       NA                NA
## 2      78  12.10    2.32       NA                NA
## 3      75   8.41    1.56      5.4             35.51
## 4      71  12.80    2.25      5.1             32.23
## 5      82   6.01    1.22      5.6             37.69
## 6      71   9.88    1.73      5.1             32.23
## 7      80   9.20    1.82      5.6             37.69
## 8      90   3.40    0.76      5.5             36.60
## 9      92   5.43    1.23      5.7             38.78
## 10     84   6.98    1.45      5.5             36.60
```

Obviament ja ho veiem i també podem veure que hi ha més d'un valor que no es té (NA)

4. Generació de l'obsject SummarizedExperiment

Ara que ja sabem que tenim a les dades, podem procedir a carregar-les en un objecte *SummarizedExperiment*. Per a fer-ho, convertirem però les dades a un data frame i li traurem la primera columna que ja hem vist que és redundant amb la segona. Aleshores ja ho podem carregar. En concret, carregarem:

1. L'arxiu DataVAlues com a assays que és on es guarden les dades de les mostres que tenim
2. L'arxiu DataInfo com a colData que és on es guarda la informació de com és cada columna d'assays.

```
# Primer carreguem la llibreria per poder tractar el SummarizedExperiment
library(SummarizedExperiment)
```

```
## Loading required package: MatrixGenerics

## Loading required package: matrixStats

##
## Attaching package: '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

## Loading required package: GenomicRanges

## Loading required package: stats4

## Loading required package: BiocGenerics

##
## Attaching package: '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, saveRDS, setdiff,
##     table, tapply, union, unique, unsplit, which.max, which.min

## Loading required package: S4Vectors

##
## Attaching package: 'S4Vectors'

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

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

## Loading required package: IRanges
```

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

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

## Loading required package: GenomeInfoDb

## 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")'.

## 
## Attaching package: 'Biobase'

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

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

```r
# Primer eliminem la primera columna de DataValues i ho guardem com un data frame que és el que accepta
DataValues_S013_matrix <- DataValues_S013[,-1]
# Ara ja podem carregar el SummarizedExperiment
seMA <- SummarizedExperiment(assays = list(counts = DataValues_S013_matrix),
                             colData = DataInfo_S013)
seMA
```

```
## class: SummarizedExperiment
## dim: 39 695
## metadata(0):
## assays(1): counts
## rownames(39): 1 2 ... 38 39
## rowData names(0):
## colnames(695): SUBJECTS SURGERY ... SM.C24.0_T5 SM.C24.1_T5
## colData names(4): X VarName varTpe Description
```

Sembla que l'objecte s'ha generat correctament

Mirem un parell de funcions que es poden fer.

```r
# Comprovem com es diuen les columnes que tenim en cada una de les classes de dins de l'objecte
dimnames(seMA)
```

```
## [[1]]
##  [1] "1"  "2"  "3"  "4"  "5"  "6"  "7"  "8"  "9"  "10" "11" "12" "13" "14" "15"
## [16] "16" "17" "18" "19" "20" "21" "22" "23" "24" "25" "26" "27" "28" "29" "30"
## [31] "31" "32" "33" "34" "35" "36" "37" "38" "39"
## 
## [[2]]
##   [1] "SUBJECTS"          "SURGERY"          "AGE"
##   [4] "GENDER"            "Group"            "MEDDM_T0"
##   [7] "MEDCOL_T0"         "MEDINF_T0"        "MEDHTA_T0"
```

```
##  [10] "GLU_T0"            "INS_T0"              "HOMA_T0"
##  [13] "HBA1C_T0"          "HBA1C.mmol.mol_T0"   "PESO_T0"
##  [16] "bmi_T0"            "CC_T0"               "CINT_T0"
##  [19] "CAD_T0"            "TAD_T0"              "TAS_T0"
##  [22] "TG_T0"             "COL_T0"              "LDL_T0"
##  [25] "HDL_T0"            "VLDL_T0"             "PCR_T0"
##  [28] "LEP_T0"            "ADIPO_T0"            "GOT_T0"
##  [31] "GPT_T0"            "GGT_T0"              "URICO_T0"
##  [34] "CREAT_T0"          "UREA_T0"             "HIERRO_T0"
##  [37] "TRANSF_T0"         "FERR_T0"             "Ile_T0"
##  [40] "Leu_T0"            "Val_T0"              "Ala_T0"
##  [43] "Pro_T0"            "Gly_T0"              "Ser_T0"
##  [46] "Trp_T0"            "Phe_T0"              "Met_T0"
##  [49] "Orn_T0"            "Arg_T0"              "His_T0"
##  [52] "Asn_T0"            "Asp_T0"              "Glu_T0"
##  [55] "Gln_T0"            "Cit_T0"              "Tyr_T0"
##  [58] "Thr_T0"            "Lys_T0"              "Creatinine_T0"
##  [61] "Kynurenine_T0"     "Putrescine_T0"       "Sarcosine_T0"
##  [64] "Serotonin_T0"      "Taurine_T0"          "SDMA_T0"
##  [67] "C0_T0"             "C2_T0"               "C3.OH_T0"
##  [70] "C6..C4.1.DC._T0"   "C5.DC..C6.OH._T0"    "C7.DC_T0"
##  [73] "C8_T0"             "C10_T0"              "C10.1_T0"
##  [76] "C10.2_T0"          "C14.1_T0"            "C14.2_T0"
##  [79] "C16.1_T0"          "C16.2_T0"            "C16.2.OH_T0"
##  [82] "C18.1_T0"          "C18.1.OH_T0"         "C18.2_T0"
##  [85] "lysoPC.a.C16.0_T0" "lysoPC.a.C16.1_T0"  "lysoPC.a.C17.0_T0"
##  [88] "lysoPC.a.C18.0_T0" "lysoPC.a.C18.1_T0"  "lysoPC.a.C18.2_T0"
##  [91] "lysoPC.a.C20.3_T0" "lysoPC.a.C20.4_T0"  "lysoPC.a.C24.0_T0"
##  [94] "lysoPC.a.C26.0_T0" "lysoPC.a.C26.1_T0"  "lysoPC.a.C28.0_T0"
##  [97] "lysoPC.a.C28.1_T0" "PC.aa.C24.0_T0"     "PC.aa.C28.1_T0"
## [100] "PC.aa.C30.0_T0"    "PC.aa.C32.0_T0"     "PC.aa.C32.1_T0"
## [103] "PC.aa.C32.3_T0"    "PC.aa.C34.1_T0"     "PC.aa.C34.2_T0"
## [106] "PC.aa.C34.3_T0"    "PC.aa.C34.4_T0"     "PC.aa.C36.0_T0"
## [109] "PC.aa.C36.1_T0"    "PC.aa.C36.2_T0"     "PC.aa.C36.3_T0"
## [112] "PC.aa.C36.4_T0"    "PC.aa.C36.5_T0"     "PC.aa.C38.0_T0"
## [115] "PC.aa.C38.1_T0"    "PC.aa.C38.3_T0"     "PC.aa.C38.4_T0"
## [118] "PC.aa.C38.5_T0"    "PC.aa.C38.6_T0"     "PC.aa.C40.1_T0"
## [121] "PC.aa.C40.2_T0"    "PC.aa.C40.3_T0"     "PC.aa.C40.4_T0"
## [124] "PC.aa.C40.5_T0"    "PC.aa.C40.6_T0"     "PC.aa.C42.0_T0"
## [127] "PC.aa.C42.1_T0"    "PC.aa.C42.2_T0"     "PC.aa.C42.4_T0"
## [130] "PC.aa.C42.5_T0"    "PC.aa.C42.6_T0"     "PC.ae.C30.0_T0"
## [133] "PC.ae.C32.1_T0"    "PC.ae.C32.2_T0"     "PC.ae.C34.0_T0"
## [136] "PC.ae.C34.1_T0"    "PC.ae.C34.2_T0"     "PC.ae.C34.3_T0"
## [139] "PC.ae.C36.0_T0"    "PC.ae.C36.1_T0"     "PC.ae.C36.2_T0"
## [142] "PC.ae.C36.3_T0"    "PC.ae.C36.4_T0"     "PC.ae.C36.5_T0"
## [145] "PC.ae.C38.0_T0"    "PC.ae.C38.2_T0"     "PC.ae.C38.3_T0"
## [148] "PC.ae.C38.4_T0"    "PC.ae.C38.5_T0"     "PC.ae.C38.6_T0"
## [151] "PC.ae.C40.1_T0"    "PC.ae.C40.2_T0"     "PC.ae.C40.3_T0"
## [154] "PC.ae.C40.4_T0"    "PC.ae.C40.5_T0"     "PC.ae.C40.6_T0"
## [157] "PC.ae.C42.1_T0"    "PC.ae.C42.2_T0"     "PC.ae.C42.3_T0"
## [160] "PC.ae.C42.4_T0"    "PC.ae.C42.5_T0"     "PC.ae.C44.3_T0"
## [163] "PC.ae.C44.4_T0"    "PC.ae.C44.5_T0"     "PC.ae.C44.6_T0"
## [166] "SM..OH..C14.1_T0"  "SM..OH..C16.1_T0"   "SM..OH..C22.1_T0"
## [169] "SM..OH..C22.2_T0"  "SM..OH..C24.1_T0"   "SM.C16.0_T0"
```

```
## [172] "SM.C16.1_T0"        "SM.C18.0_T0"        "SM.C18.1_T0"
## [175] "SM.C20.2_T0"        "SM.C24.0_T0"        "SM.C24.1_T0"
## [178] "MEDDM_T2"           "MEDCOL_T2"          "MEDINF_T2"
## [181] "MEDHTA_T2"          "GLU_T2"             "INS_T2"
## [184] "HOMA_T2"            "HBA1C_T2"           "HBA1C.mmol.mol_T2"
## [187] "PESO_T2"            "bmi_T2"             "CC_T2"
## [190] "CINT_T2"            "CAD_T2"             "TAD_T2"
## [193] "TAS_T2"             "TG_T2"              "COL_T2"
## [196] "LDL_T2"             "HDL_T2"             "VLDL_T2"
## [199] "PCR_T2"             "LEP_T2"             "ADIPO_T2"
## [202] "GOT_T2"             "GPT_T2"             "GGT_T2"
## [205] "URICO_T2"           "CREAT_T2"           "UREA_T2"
## [208] "HIERRO_T2"          "TRANSF_T2"          "FERR_T2"
## [211] "Ile_T2"             "Leu_T2"             "Val_T2"
## [214] "Ala_T2"             "Pro_T2"             "Gly_T2"
## [217] "Ser_T2"             "Trp_T2"             "Phe_T2"
## [220] "Met_T2"             "Orn_T2"             "Arg_T2"
## [223] "His_T2"             "Asn_T2"             "Asp_T2"
## [226] "Glu_T2"             "Gln_T2"             "Cit_T2"
## [229] "Tyr_T2"             "Thr_T2"             "Lys_T2"
## [232] "Creatinine_T2"      "Kynurenine_T2"      "Putrescine_T2"
## [235] "Sarcosine_T2"       "Serotonin_T2"       "Taurine_T2"
## [238] "SDMA_T2"            "C0_T2"              "C2_T2"
## [241] "C3.OH_T2"           "C6..C4.1.DC._T2"    "C5.DC..C6.OH._T2"
## [244] "C7.DC_T2"           "C8_T2"              "C10_T2"
## [247] "C10.1_T2"           "C10.2_T2"           "C14.1_T2"
## [250] "C14.2_T2"           "C16.1_T2"           "C16.2_T2"
## [253] "C16.2.OH_T2"        "C18.1_T2"           "C18.1.OH_T2"
## [256] "C18.2_T2"           "X"                  "lysoPC.a.C14.0_T2"
## [259] "lysoPC.a.C16.0_T2"  "lysoPC.a.C16.1_T2"  "lysoPC.a.C17.0_T2"
## [262] "lysoPC.a.C18.0_T2"  "lysoPC.a.C18.1_T2"  "lysoPC.a.C18.2_T2"
## [265] "lysoPC.a.C20.3_T2"  "lysoPC.a.C20.4_T2"  "lysoPC.a.C24.0_T2"
## [268] "lysoPC.a.C26.0_T2"  "lysoPC.a.C26.1_T2"  "lysoPC.a.C28.0_T2"
## [271] "lysoPC.a.C28.1_T2"  "PC.aa.C24.0_T2"     "PC.aa.C28.1_T2"
## [274] "PC.aa.C30.0_T2"     "PC.aa.C32.0_T2"     "PC.aa.C32.1_T2"
## [277] "PC.aa.C32.3_T2"     "PC.aa.C34.1_T2"     "PC.aa.C34.2_T2"
## [280] "PC.aa.C34.3_T2"     "PC.aa.C34.4_T2"     "PC.aa.C36.0_T2"
## [283] "PC.aa.C36.1_T2"     "PC.aa.C36.2_T2"     "PC.aa.C36.3_T2"
## [286] "PC.aa.C36.4_T2"     "PC.aa.C36.5_T2"     "PC.aa.C38.0_T2"
## [289] "PC.aa.C38.1_T2"     "PC.aa.C38.3_T2"     "PC.aa.C38.4_T2"
## [292] "PC.aa.C38.5_T2"     "PC.aa.C38.6_T2"     "PC.aa.C40.1_T2"
## [295] "PC.aa.C40.2_T2"     "PC.aa.C40.3_T2"     "PC.aa.C40.4_T2"
## [298] "PC.aa.C40.5_T2"     "PC.aa.C40.6_T2"     "PC.aa.C42.0_T2"
## [301] "PC.aa.C42.1_T2"     "PC.aa.C42.2_T2"     "PC.aa.C42.4_T2"
## [304] "PC.aa.C42.5_T2"     "PC.aa.C42.6_T2"     "PC.ae.C30.0_T2"
## [307] "PC.ae.C32.1_T2"     "PC.ae.C32.2_T2"     "PC.ae.C34.0_T2"
## [310] "PC.ae.C34.1_T2"     "PC.ae.C34.2_T2"     "PC.ae.C34.3_T2"
## [313] "PC.ae.C36.0_T2"     "PC.ae.C36.1_T2"     "PC.ae.C36.2_T2"
## [316] "PC.ae.C36.3_T2"     "PC.ae.C36.4_T2"     "PC.ae.C36.5_T2"
## [319] "PC.ae.C38.0_T2"     "PC.ae.C38.2_T2"     "PC.ae.C38.3_T2"
## [322] "PC.ae.C38.4_T2"     "PC.ae.C38.5_T2"     "PC.ae.C38.6_T2"
## [325] "PC.ae.C40.1_T2"     "PC.ae.C40.2_T2"     "PC.ae.C40.3_T2"
## [328] "PC.ae.C40.4_T2"     "PC.ae.C40.5_T2"     "PC.ae.C40.6_T2"
## [331] "PC.ae.C42.1_T2"     "PC.ae.C42.2_T2"     "PC.ae.C42.3_T2"
```

```
## [334] "PC.ae.C42.4_T2"     "PC.ae.C42.5_T2"     "PC.ae.C44.3_T2"
## [337] "PC.ae.C44.4_T2"     "PC.ae.C44.5_T2"     "PC.ae.C44.6_T2"
## [340] "SM..OH..C14.1_T2"   "SM..OH..C16.1_T2"   "SM..OH..C22.1_T2"
## [343] "SM..OH..C22.2_T2"   "SM..OH..C24.1_T2"   "SM.C16.0_T2"
## [346] "SM.C16.1_T2"        "SM.C18.0_T2"        "SM.C18.1_T2"
## [349] "SM.C20.2_T2"        "SM.C24.0_T2"        "SM.C24.1_T2"
## [352] "MEDDM_T4"           "MEDCOL_T4"          "MEDINF_T4"
## [355] "MEDHTA_T4"          "GLU_T4"             "INS_T4"
## [358] "HOMA_T4"            "HBA1C_T4"           "HBA1C.mmol.mol_T4"
## [361] "PESO_T4"            "bmi_T4"             "CC_T4"
## [364] "CINT_T4"            "CAD_T4"             "TAD_T4"
## [367] "TAS_T4"             "TG_T4"              "COL_T4"
## [370] "LDL_T4"             "HDL_T4"             "VLDL_T4"
## [373] "PCR_T4"             "LEP_T4"             "ADIPO_T4"
## [376] "GOT_T4"             "GPT_T4"             "GGT_T4"
## [379] "URICO_T4"           "CREAT_T4"           "UREA_T4"
## [382] "HIERRO_T4"          "TRANSF_T4"          "FERR_T4"
## [385] "Ile_T4"             "Leu_T4"             "Val_T4"
## [388] "Ala_T4"             "Pro_T4"             "Gly_T4"
## [391] "Ser_T4"             "Trp_T4"             "Phe_T4"
## [394] "Met_T4"             "Orn_T4"             "Arg_T4"
## [397] "His_T4"             "Asn_T4"             "Asp_T4"
## [400] "Glu_T4"             "Gln_T4"             "Cit_T4"
## [403] "Tyr_T4"             "Thr_T4"             "Lys_T4"
## [406] "Creatinine_T4"      "Kynurenine_T4"      "Putrescine_T4"
## [409] "Sarcosine_T4"       "Serotonin_T4"       "Taurine_T4"
## [412] "SDMA_T4"            "C0_T4"              "C2_T4"
## [415] "C3.OH_T4"           "C6..C4.1.DC._T4"    "C5.DC..C6.OH._T4"
## [418] "C7.DC_T4"           "C8_T4"              "C10_T4"
## [421] "C10.1_T4"           "C10.2_T4"           "C14.1_T4"
## [424] "C14.2_T4"           "C16.1_T4"           "C16.2_T4"
## [427] "C16.2.OH_T4"        "C18.1_T4"           "C18.1.OH_T4"
## [430] "C18.2_T4"           "lysoPC.a.C16.0_T4"  "lysoPC.a.C16.1_T4"
## [433] "lysoPC.a.C17.0_T4"  "lysoPC.a.C18.0_T4"  "lysoPC.a.C18.1_T4"
## [436] "lysoPC.a.C18.2_T4"  "lysoPC.a.C20.3_T4"  "lysoPC.a.C20.4_T4"
## [439] "lysoPC.a.C24.0_T4"  "lysoPC.a.C26.0_T4"  "lysoPC.a.C26.1_T4"
## [442] "lysoPC.a.C28.0_T4"  "lysoPC.a.C28.1_T4"  "PC.aa.C24.0_T4"
## [445] "PC.aa.C28.1_T4"     "PC.aa.C30.0_T4"     "PC.aa.C32.0_T4"
## [448] "PC.aa.C32.1_T4"     "PC.aa.C32.3_T4"     "PC.aa.C34.1_T4"
## [451] "PC.aa.C34.2_T4"     "PC.aa.C34.3_T4"     "PC.aa.C34.4_T4"
## [454] "PC.aa.C36.0_T4"     "PC.aa.C36.1_T4"     "PC.aa.C36.2_T4"
## [457] "PC.aa.C36.3_T4"     "PC.aa.C36.4_T4"     "PC.aa.C36.5_T4"
## [460] "PC.aa.C38.0_T4"     "PC.aa.C38.1_T4"     "PC.aa.C38.3_T4"
## [463] "PC.aa.C38.4_T4"     "PC.aa.C38.5_T4"     "PC.aa.C38.6_T4"
## [466] "PC.aa.C40.1_T4"     "PC.aa.C40.2_T4"     "PC.aa.C40.3_T4"
## [469] "PC.aa.C40.4_T4"     "PC.aa.C40.5_T4"     "PC.aa.C40.6_T4"
## [472] "PC.aa.C42.0_T4"     "PC.aa.C42.1_T4"     "PC.aa.C42.2_T4"
## [475] "PC.aa.C42.4_T4"     "PC.aa.C42.5_T4"     "PC.aa.C42.6_T4"
## [478] "PC.ae.C30.0_T4"     "PC.ae.C32.1_T4"     "PC.ae.C32.2_T4"
## [481] "PC.ae.C34.0_T4"     "PC.ae.C34.1_T4"     "PC.ae.C34.2_T4"
## [484] "PC.ae.C34.3_T4"     "PC.ae.C36.0_T4"     "PC.ae.C36.1_T4"
## [487] "PC.ae.C36.2_T4"     "PC.ae.C36.3_T4"     "PC.ae.C36.4_T4"
## [490] "PC.ae.C36.5_T4"     "PC.ae.C38.0_T4"     "PC.ae.C38.2_T4"
## [493] "PC.ae.C38.3_T4"     "PC.ae.C38.4_T4"     "PC.ae.C38.5_T4"
```

```
## [496] "PC.ae.C38.6_T4"       "PC.ae.C40.1_T4"       "PC.ae.C40.2_T4"
## [499] "PC.ae.C40.3_T4"       "PC.ae.C40.4_T4"       "PC.ae.C40.5_T4"
## [502] "PC.ae.C40.6_T4"       "PC.ae.C42.1_T4"       "PC.ae.C42.2_T4"
## [505] "PC.ae.C42.3_T4"       "PC.ae.C42.4_T4"       "PC.ae.C42.5_T4"
## [508] "PC.ae.C44.3_T4"       "PC.ae.C44.4_T4"       "PC.ae.C44.5_T4"
## [511] "PC.ae.C44.6_T4"       "SM..OH..C14.1_T4"     "SM..OH..C16.1_T4"
## [514] "SM..OH..C22.1_T4"     "SM..OH..C22.2_T4"     "SM..OH..C24.1_T4"
## [517] "SM.C16.0_T4"          "SM.C16.1_T4"          "SM.C18.0_T4"
## [520] "SM.C18.1_T4"          "SM.C20.2_T4"          "SM.C24.0_T4"
## [523] "SM.C24.1_T4"          "MEDDM_T5"             "MEDCOL_T5"
## [526] "MEDINF_T5"            "MEDHTA_T5"            "GLU_T5"
## [529] "INS_T5"               "HOMA_T5"              "HBA1C_T5"
## [532] "HBA1C.mmol.mol_T5"    "PESO_T5"              "bmi_T5"
## [535] "CC_T5"                "CINT_T5"              "CAD_T5"
## [538] "TAD_T5"               "TAS_T5"               "TG_T5"
## [541] "COL_T5"               "LDL_T5"               "HDL_T5"
## [544] "VLDL_T5"              "PCR_T5"               "LEP_T5"
## [547] "ADIPO_T5"             "GOT_T5"               "GPT_T5"
## [550] "GGT_T5"               "URICO_T5"             "CREAT_T5"
## [553] "UREA_T5"              "HIERRO_T5"            "TRANSF_T5"
## [556] "FERR_T5"              "Ile_T5"               "Leu_T5"
## [559] "Val_T5"               "Ala_T5"               "Pro_T5"
## [562] "Gly_T5"               "Ser_T5"               "Trp_T5"
## [565] "Phe_T5"               "Met_T5"               "Orn_T5"
## [568] "Arg_T5"               "His_T5"               "Asn_T5"
## [571] "Asp_T5"               "Glu_T5"               "Gln_T5"
## [574] "Cit_T5"               "Tyr_T5"               "Thr_T5"
## [577] "Lys_T5"               "Creatinine_T5"        "Kynurenine_T5"
## [580] "Putrescine_T5"        "Sarcosine_T5"         "Serotonin_T5"
## [583] "Taurine_T5"           "SDMA_T5"              "C0_T5"
## [586] "C2_T5"                "C3.OH_T5"             "C6..C4.1.DC._T5"
## [589] "C5.DC..C6.OH._T5"     "C7.DC_T5"             "C8_T5"
## [592] "C10_T5"               "C10.1_T5"             "C10.2_T5"
## [595] "C14.1_T5"             "C14.2_T5"             "C16.1_T5"
## [598] "C16.2_T5"             "C16.2.OH_T5"          "C18.1_T5"
## [601] "C18.1.OH_T5"          "C18.2_T5"             "lysoPC.a.C16.0_T5"
## [604] "lysoPC.a.C16.1_T5"    "lysoPC.a.C17.0_T5"    "lysoPC.a.C18.0_T5"
## [607] "lysoPC.a.C18.1_T5"    "lysoPC.a.C18.2_T5"    "lysoPC.a.C20.3_T5"
## [610] "lysoPC.a.C20.4_T5"    "lysoPC.a.C24.0_T5"    "lysoPC.a.C26.0_T5"
## [613] "lysoPC.a.C26.1_T5"    "lysoPC.a.C28.0_T5"    "lysoPC.a.C28.1_T5"
## [616] "PC.aa.C24.0_T5"       "PC.aa.C28.1_T5"       "PC.aa.C30.0_T5"
## [619] "PC.aa.C32.0_T5"       "PC.aa.C32.1_T5"       "PC.aa.C32.3_T5"
## [622] "PC.aa.C34.1_T5"       "PC.aa.C34.2_T5"       "PC.aa.C34.3_T5"
## [625] "PC.aa.C34.4_T5"       "PC.aa.C36.0_T5"       "PC.aa.C36.1_T5"
## [628] "PC.aa.C36.2_T5"       "PC.aa.C36.3_T5"       "PC.aa.C36.4_T5"
## [631] "PC.aa.C36.5_T5"       "PC.aa.C38.0_T5"       "PC.aa.C38.1_T5"
## [634] "PC.aa.C38.3_T5"       "PC.aa.C38.4_T5"       "PC.aa.C38.5_T5"
## [637] "PC.aa.C38.6_T5"       "PC.aa.C40.1_T5"       "PC.aa.C40.2_T5"
## [640] "PC.aa.C40.3_T5"       "PC.aa.C40.4_T5"       "PC.aa.C40.5_T5"
## [643] "PC.aa.C40.6_T5"       "PC.aa.C42.0_T5"       "PC.aa.C42.1_T5"
## [646] "PC.aa.C42.2_T5"       "PC.aa.C42.4_T5"       "PC.aa.C42.5_T5"
## [649] "PC.aa.C42.6_T5"       "PC.ae.C30.0_T5"       "PC.ae.C32.1_T5"
## [652] "PC.ae.C32.2_T5"       "PC.ae.C34.0_T5"       "PC.ae.C34.1_T5"
## [655] "PC.ae.C34.2_T5"       "PC.ae.C34.3_T5"       "PC.ae.C36.0_T5"
```

```
## [658] "PC.ae.C36.1_T5"    "PC.ae.C36.2_T5"    "PC.ae.C36.3_T5"
## [661] "PC.ae.C36.4_T5"    "PC.ae.C36.5_T5"    "PC.ae.C38.0_T5"
## [664] "PC.ae.C38.2_T5"    "PC.ae.C38.3_T5"    "PC.ae.C38.4_T5"
## [667] "PC.ae.C38.5_T5"    "PC.ae.C38.6_T5"    "PC.ae.C40.1_T5"
## [670] "PC.ae.C40.2_T5"    "PC.ae.C40.3_T5"    "PC.ae.C40.4_T5"
## [673] "PC.ae.C40.5_T5"    "PC.ae.C40.6_T5"    "PC.ae.C42.1_T5"
## [676] "PC.ae.C42.2_T5"    "PC.ae.C42.3_T5"    "PC.ae.C42.4_T5"
## [679] "PC.ae.C42.5_T5"    "PC.ae.C44.3_T5"    "PC.ae.C44.4_T5"
## [682] "PC.ae.C44.5_T5"    "PC.ae.C44.6_T5"    "SM..OH..C14.1_T5"
## [685] "SM..OH..C16.1_T5"  "SM..OH..C22.1_T5"  "SM..OH..C22.2_T5"
## [688] "SM..OH..C24.1_T5"  "SM.C16.0_T5"       "SM.C16.1_T5"
## [691] "SM.C18.0_T5"       "SM.C18.1_T5"       "SM.C20.2_T5"
## [694] "SM.C24.0_T5"       "SM.C24.1_T5"
```

```
# I ara mirem les 10 primeres columnes i les 5 primeres files d'assays
assay(seMA)[1:5, 1:10]
```

```
##   SUBJECTS SURGERY AGE GENDER Group MEDDM_T0 MEDCOL_T0 MEDINF_T0 MEDHTA_T0
## 1        1 by pass  27      F     1        0         0         0         1
## 2        2 by pass  19      F     2        0         0         0         0
## 3        3 by pass  42      F     1        0         0         0         0
## 4        4 by pass  37      F     2        0         0         0         0
## 5        5 tubular  42      F     1        0         0         0         0
##   GLU_T0
## 1     85
## 2     78
## 3     75
## 4     71
## 5     82
```

```
# I també mirem com ha quedat colData
colData(seMA)
```

```
## DataFrame with 695 rows and 4 columns
##                       X     VarName     varTpe Description
##             <character> <character> <character> <character>
## SUBJECTS       SUBJECTS    SUBJECTS     integer    dataDesc
## SURGERY         SURGERY     SURGERY   character    dataDesc
## AGE                 AGE         AGE     integer    dataDesc
## GENDER           GENDER      GENDER   character    dataDesc
## Group             Group       Group     integer    dataDesc
## ...                 ...         ...         ...         ...
## SM.C18.0_T5 SM.C18.0_T5 SM.C18.0_T5     numeric    dataDesc
## SM.C18.1_T5 SM.C18.1_T5 SM.C18.1_T5     numeric    dataDesc
## SM.C20.2_T5 SM.C20.2_T5 SM.C20.2_T5     numeric    dataDesc
## SM.C24.0_T5 SM.C24.0_T5 SM.C24.0_T5     numeric    dataDesc
## SM.C24.1_T5 SM.C24.1_T5 SM.C24.1_T5     numeric    dataDesc
```

5. Anàlisi exploratoria de les dades

Com que el joc de dades que hé triat conté 395 camps i es fa molt dificil fer un anàlisi complet, provem a
fer una cosa una mica més ajustada i agafaré només les dades de descripció del casos analitzats i alguns
metabolits. En concret em pararé en el ferro. és a dir que agafaré les primeres 40 columnes. En un context
real s'hauria de fer amb tot i l'anàlisi seria molt llarg i tediós.

Per tant, comencem per fer un subset del nostre SummarizedExperiment i el guardem com un data frame per

poder treballar amb ell.

```
analisi <- seMA[,1:40]
analisi
```

```
## class: SummarizedExperiment
## dim: 39 40
## metadata(0):
## assays(1): counts
## rownames(39): 1 2 ... 38 39
## rowData names(0):
## colnames(40): SUBJECTS SURGERY ... Ile_T0 Leu_T0
## colData names(4): X VarName varTpe Description
```

Ara, per tant, tenim un segon SummarizedExperiment que és un subset del primer. Conté nomnés les dades d'identificació de tots els pacients (columnes 1-9) i els primers 31 metabolits que es van recollir

5.1. Estadistics bàsics

Amb això, podem mirar una mica el que tenim

```
str(assay(analisi))
```

```
## 'data.frame':    39 obs. of  40 variables:
##  $ SUBJECTS        : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ SURGERY         : chr  "by pass" "by pass" "by pass" "by pass" ...
##  $ AGE             : int  27 19 42 37 42 24 33 55 40 47 ...
##  $ GENDER          : chr  "F" "F" "F" "F" ...
##  $ Group           : int  1 2 1 2 1 2 1 1 1 1 ...
##  $ MEDDM_T0        : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ MEDCOL_T0       : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ MEDINF_T0       : int  0 0 0 0 0 0 0 1 0 0 ...
##  $ MEDHTA_T0       : int  1 0 0 0 0 0 0 0 0 0 ...
##  $ GLU_T0          : int  85 78 75 71 82 71 80 90 92 84 ...
##  $ INS_T0          : num  11.4 12.1 8.41 12.8 6.01 9.88 9.2 3.4 5.43 6.98 ...
##  $ HOMA_T0         : num  2.4 2.32 1.56 2.25 1.22 1.73 1.82 0.76 1.23 1.45 ...
##  $ HBA1C_T0        : num  NA NA 5.4 5.1 5.6 5.1 5.6 5.5 5.7 5.5 ...
##  $ HBA1C.mmol.mol_T0: num  NA NA 35.5 32.2 37.7 ...
##  $ PESO_T0         : int  151 139 84 136 121 148 109 109 114 120 ...
##  $ bmi_T0          : num  62.9 47 29.8 53.1 46.6 48.8 43.7 41.8 44 40.6 ...
##  $ CC_T0           : num  0.7 NA 0.7 1 0.9 0.7 0.9 0.9 0.9 NA ...
##  $ CINT_T0         : int  116 NA 90 157 123 110 122 124 136 NA ...
##  $ CAD_T0          : int  167 NA 126 162 132 148 141 136 148 NA ...
##  $ TAD_T0          : int  125 NA 79 73 84 74 65 75 91 NA ...
##  $ TAS_T0          : int  174 NA 111 127 122 131 121 128 128 NA ...
##  $ TG_T0           : int  147 150 45 109 30 61 75 53 164 145 ...
##  $ COL_T0          : int  256 180 211 205 102 121 192 181 154 220 ...
##  $ LDL_T0          : num  167 94 114 146 24 60.8 96 99.4 67.2 160 ...
##  $ HDL_T0          : int  60 56 88 37 72 48 81 71 54 31 ...
##  $ VLDL_T0         : num  29.4 30 9 21.8 6 12.2 15 10.6 32.8 29 ...
##  $ PCR_T0          : num  10.2 9 3.05 8.89 NA 1.6 4.52 NA NA NA ...
##  $ LEP_T0          : num  155 84 27 46 NA 38 61 NA NA NA ...
##  $ ADIPO_T0        : num  8.15 7.94 16.7 4.68 NA 11.8 6.51 NA NA NA ...
##  $ GOT_T0          : int  21 16 24 21 42 11 14 50 21 15 ...
##  $ GPT_T0          : int  33 40 39 37 68 26 25 24 11 30 ...
##  $ GGT_T0          : int  22 25 20 20 18 12 16 36 16 52 ...
```

```
##  $ URICO_T0         : num  5.7 NA 2.7 4.9 5.2 3.7 3.8 5.6 5.1 5 ...
##  $ CREAT_T0         : num  0.8 0.8 0.8 0.7 0.8 0.7 0.6 0.7 0.7 1 ...
##  $ UREA_T0          : int  33 29 28 22 29 14 26 49 24 24 ...
##  $ HIERRO_T0        : int  77 113 142 64 94 102 72 13 36 113 ...
##  $ TRANSF_T0        : int  NA 290 251 289 252 200 NA 300 288 235 ...
##  $ FERR_T0          : num  53 25 15 18 39.4 25 24 23 48 101 ...
##  $ Ile_T0           : num  53.9 95.7 57.8 84.5 72 109 73.3 91.7 83.8 113 ...
##  $ Leu_T0           : num  105 188 119 180 140 205 147 191 183 225 ...
```

```r
summary(assay(analisi))
```

```
##     SUBJECTS      SURGERY               AGE           GENDER
##  Min.   : 1.0   Length:39          Min.   :19.00   Length:39
##  1st Qu.:10.5   Class :character   1st Qu.:35.00   Class :character
##  Median :20.0   Mode  :character   Median :41.00   Mode  :character
##  Mean   :20.0                      Mean   :40.79
##  3rd Qu.:29.5                      3rd Qu.:46.00
##  Max.   :39.0                      Max.   :59.00
##
##      Group          MEDDM_T0   MEDCOL_T0          MEDINF_T0
##  Min.   :1.000   Min.   :0   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:1.000   1st Qu.:0   1st Qu.:0.00000   1st Qu.:0.0000
##  Median :1.000   Median :0   Median :0.00000   Median :0.0000
##  Mean   :1.385   Mean   :0   Mean   :0.02632   Mean   :0.1316
##  3rd Qu.:2.000   3rd Qu.:0   3rd Qu.:0.00000   3rd Qu.:0.0000
##  Max.   :2.000   Max.   :0   Max.   :1.00000   Max.   :1.0000
##                  NA's   :1   NA's   :1          NA's   :1
##     MEDHTA_T0         GLU_T0          INS_T0          HOMA_T0
##  Min.   :0.0000   Min.   : 71.0   Min.   : 3.40   Min.   : 0.760
##  1st Qu.:0.0000   1st Qu.: 91.0   1st Qu.:11.85   1st Qu.: 2.435
##  Median :0.0000   Median :103.0   Median :16.00   Median : 4.210
##  Mean   :0.2368   Mean   :106.5   Mean   :17.60   Mean   : 4.890
##  3rd Qu.:0.0000   3rd Qu.:109.5   3rd Qu.:21.15   3rd Qu.: 5.720
##  Max.   :1.0000   Max.   :263.0   Max.   :43.00   Max.   :13.600
##  NA's   :1
##     HBA1C_T0      HBA1C.mmol.mol_T0    PESO_T0          bmi_T0
##  Min.   :5.100   Min.   :32.23      Min.   : 84.0   Min.   :29.80
##  1st Qu.:5.400   1st Qu.:35.51      1st Qu.:119.5   1st Qu.:44.40
##  Median :5.600   Median :37.69      Median :135.0   Median :48.80
##  Mean   :5.592   Mean   :37.60      Mean   :140.0   Mean   :50.52
##  3rd Qu.:5.800   3rd Qu.:39.88      3rd Qu.:155.0   3rd Qu.:55.35
##  Max.   :6.400   Max.   :46.44      Max.   :200.0   Max.   :68.60
##  NA's   :15      NA's   :15
##     CC_T0           CINT_T0         CAD_T0          TAD_T0
##  Min.   :0.7000   Min.   : 90.0   Min.   : 85.0   Min.   : 60.00
##  1st Qu.:0.9000   1st Qu.:122.5   1st Qu.:136.0   1st Qu.: 74.50
##  Median :0.9000   Median :133.0   Median :148.0   Median : 83.00
##  Mean   :0.9343   Mean   :135.7   Mean   :147.1   Mean   : 83.58
##  3rd Qu.:1.0000   3rd Qu.:147.5   3rd Qu.:158.5   3rd Qu.: 90.00
##  Max.   :1.7000   Max.   :186.0   Max.   :182.0   Max.   :125.00
##  NA's   :4        NA's   :4       NA's   :4       NA's   :8
##     TAS_T0           TG_T0          COL_T0          LDL_T0
##  Min.   :111.0   Min.   : 30.0   Min.   :102.0   Min.   : 24.0
##  1st Qu.:121.0   1st Qu.: 84.0   1st Qu.:180.5   1st Qu.:100.3
##  Median :128.0   Median :118.0   Median :207.0   Median :130.0
```

```
## Mean   :133.5   Mean   :129.1   Mean   :206.6   Mean   :128.2
## 3rd Qu.:144.0   3rd Qu.:151.0   3rd Qu.:236.0   3rd Qu.:158.8
## Max.   :180.0   Max.   :268.0   Max.   :295.0   Max.   :191.0
## NA's   :8                                       NA's   :1
##     HDL_T0          VLDL_T0          PCR_T0          LEP_T0
## Min.   : 11.00   Min.   : 6.00   Min.   : 0.00   Min.   : 27.00
## 1st Qu.: 39.25   1st Qu.:16.80   1st Qu.: 3.51   1st Qu.: 49.25
## Median : 47.00   Median :23.60   Median : 8.25   Median : 67.20
## Mean   : 49.97   Mean   :25.81   Mean   :10.13   Mean   : 78.33
## 3rd Qu.: 59.50   3rd Qu.:30.20   3rd Qu.:13.90   3rd Qu.: 98.25
## Max.   :107.00   Max.   :53.60   Max.   :29.80   Max.   :155.00
## NA's   :1                        NA's   :14      NA's   :13
##    ADIPO_T0          GOT_T0          GPT_T0          GGT_T0
## Min.   : 1.720   Min.   :11.0    Min.   : 11.00  Min.   :12.00
## 1st Qu.: 4.700   1st Qu.:15.0    1st Qu.: 31.00  1st Qu.:20.00
## Median : 7.460   Median :20.0    Median : 40.00  Median :26.00
## Mean   : 7.702   Mean   :23.9    Mean   : 44.54  Mean   :31.41
## 3rd Qu.: 8.590   3rd Qu.:25.0    3rd Qu.: 48.50  3rd Qu.:37.00
## Max.   :17.100   Max.   :83.0    Max.   :157.00  Max.   :86.00
## NA's   :14                                       NA's   :2
##    URICO_T0         CREAT_T0         UREA_T0         HIERRO_T0
## Min.   :2.700   Min.   :0.5000   Min.   :14.00   Min.   : 13.00
## 1st Qu.:4.700   1st Qu.:0.7000   1st Qu.:24.00   1st Qu.: 50.00
## Median :5.500   Median :0.8000   Median :29.00   Median : 69.00
## Mean   :5.501   Mean   :0.7923   Mean   :31.38   Mean   : 75.61
## 3rd Qu.:6.200   3rd Qu.:0.9000   3rd Qu.:38.00   3rd Qu.: 92.25
## Max.   :8.600   Max.   :1.2000   Max.   :56.00   Max.   :166.00
## NA's   :2                                        NA's   :1
##    TRANSF_T0        FERR_T0          Ile_T0          Leu_T0
## Min.   :159.0   Min.   : 11.10  Min.   : 53.50  Min.   : 90.1
## 1st Qu.:236.5   1st Qu.: 25.00  1st Qu.: 83.55  1st Qu.:150.5
## Median :268.5   Median : 49.00  Median : 91.70  Median :183.0
## Mean   :262.9   Mean   : 77.06  Mean   : 98.98  Mean   :185.2
## 3rd Qu.:289.8   3rd Qu.: 97.00  3rd Qu.:116.00  3rd Qu.:217.0
## Max.   :374.0   Max.   :372.00  Max.   :150.00  Max.   :281.0
## NA's   :9       NA's   :2
```

Ja només amb aquesta primera mirada, podem comprovar que les dades tenen una enorme variabilitat de rangs el que pot portar algun problema a l'hora de fer anàlisi més adelantats i que calgui escalar les variables en alguns casos

5.2. Analisi univariant

Ho fem amb histogrames. Farem 9 de les variables per veure una mica com es distribueixen

```
opt <- par(mfrow=c(3,3))
for (i in 12:20)
  hist(assay(analisi)[,i], xlab = colData(analisi)[i,2], main = colData(analisi)[i,2])
```
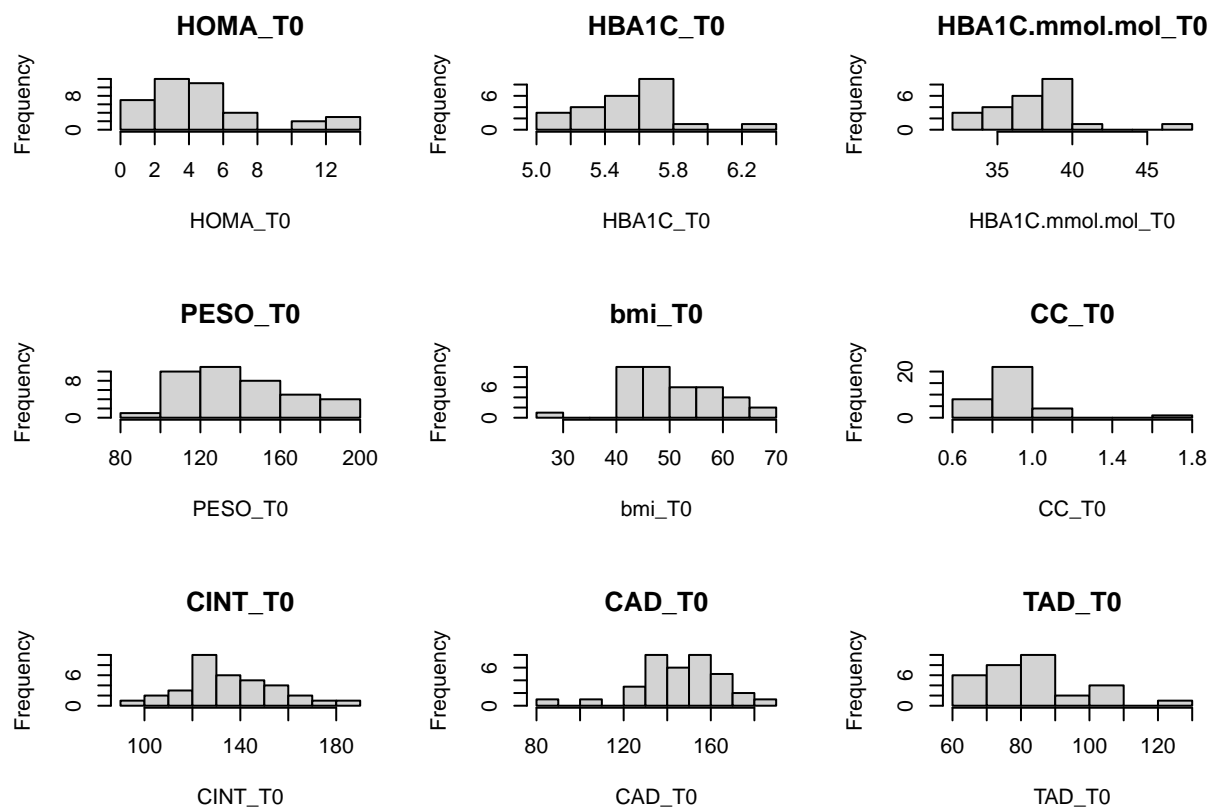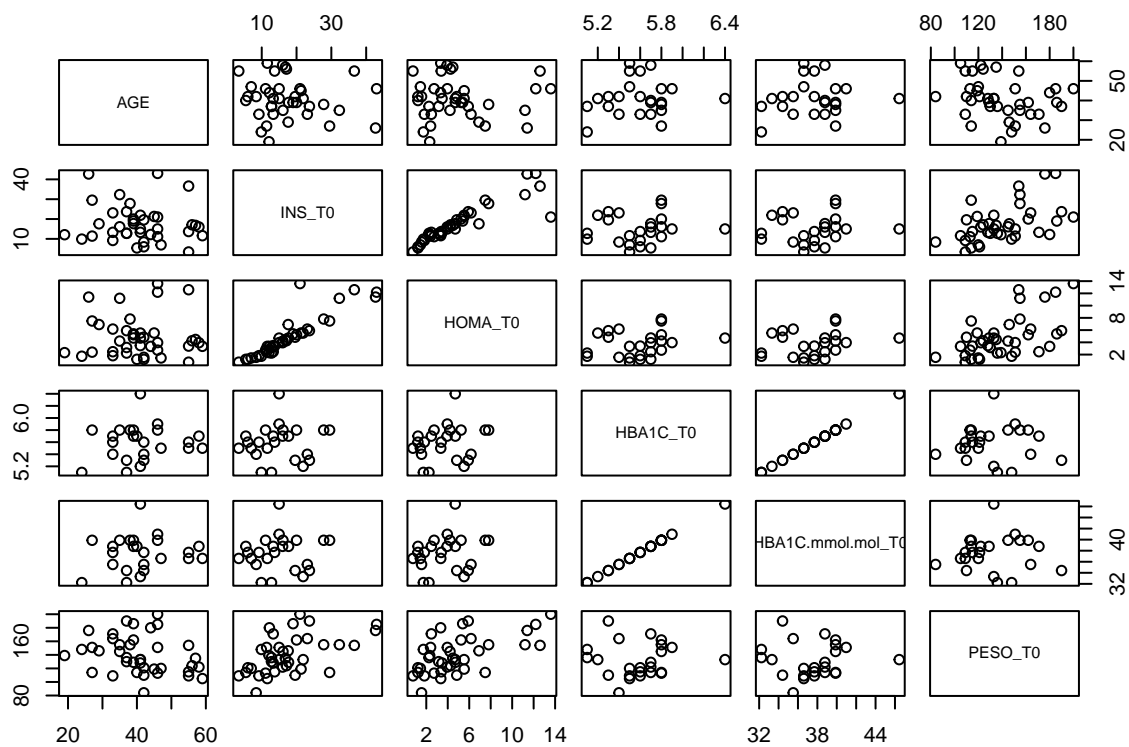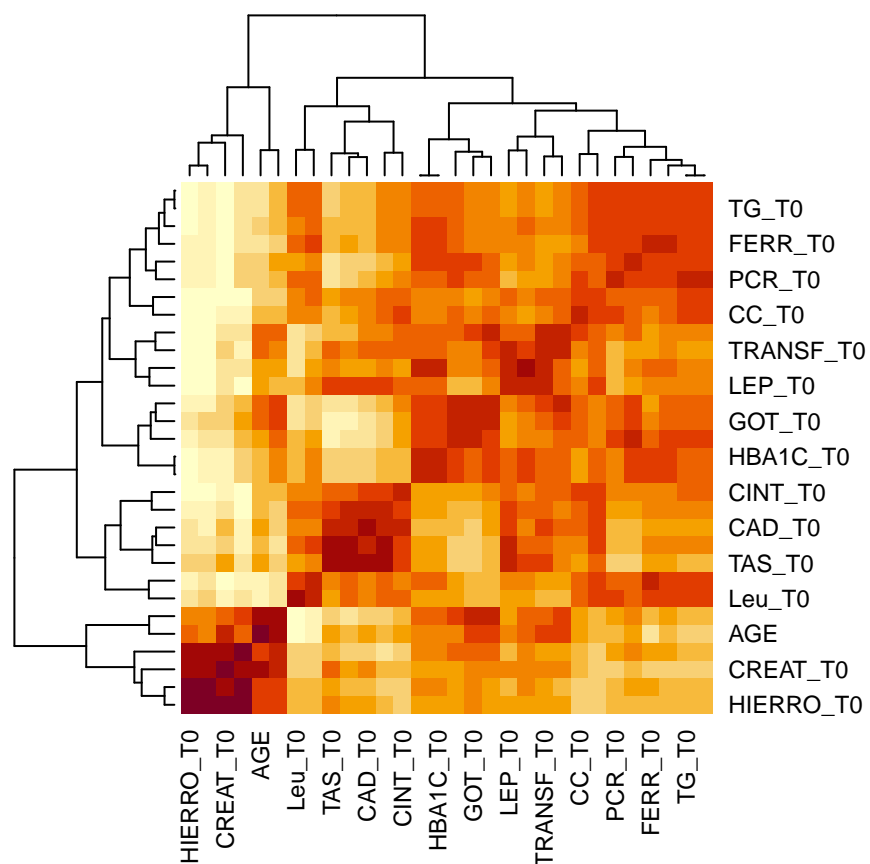
```
par(opt)
```

5.3. Analisi de correlacions

Ho fariem amb plot. Un exemple

```
# Mirem per  exemple, si l'edat es correlaciona be amb alguns
plot(assay(analisi)[,c(3,11:15)])
```
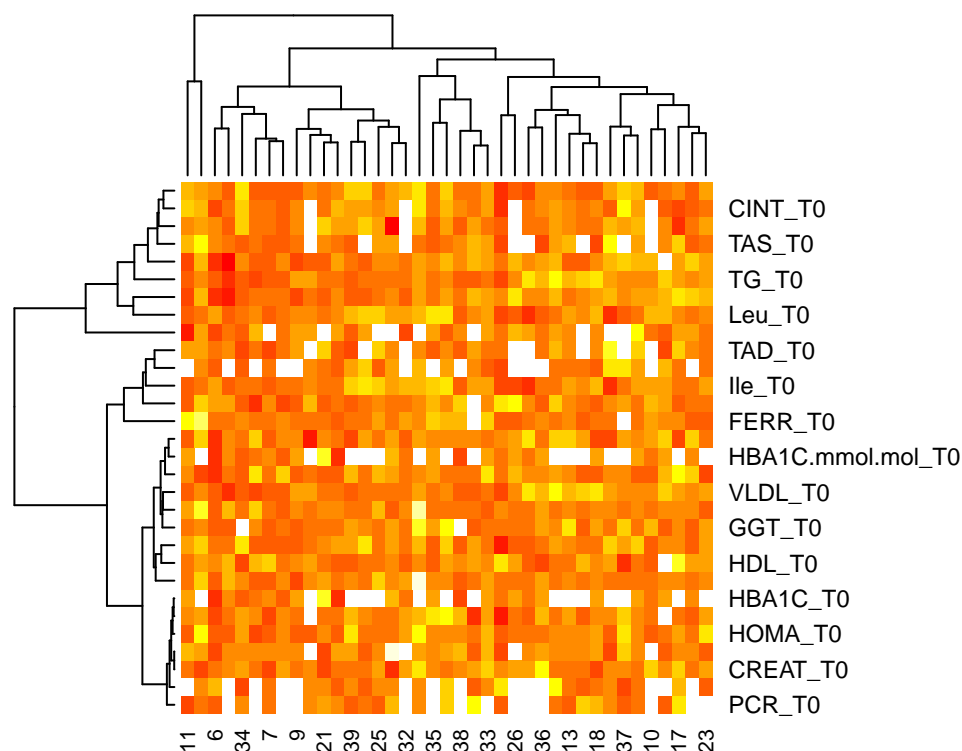
Una altra opció pot ser un heatmap de la matriu de correlacions que permetria veure totes les relacions

```
cor_matrix <- cor(na.omit(assay(analisi)[,c(3,12:40)]))
heatmap(cor_matrix)
```

Un altre heatmap que calcula distancies

```
dist_matrix <- t(assay(analisi)[,c(3,12:40)])
heatmap(dist_matrix, col=heat.colors(16))
```

5.4. Components principals

Primer hem d'eliminar els NA i despres podem. Cal fer-ho amb les columnes de metabolits (11-40)

```r
PCAMA <- prcomp(na.omit(assay(analisi)[,c(3,12:40)]), scale. = TRUE)
PCAMA
```

```
## Standard deviations (1, .., p=5):
## [1] 3.869264e+00 2.708545e+00 2.318885e+00 1.521628e+00 1.443830e-15
##
## Rotation (n x k) = (30 x 5):
##                            PC1          PC2           PC3          PC4
## AGE               -3.879932e-02  0.22167989 -0.3296551604 -0.118639705
## HOMA_T0            2.361425e-01  0.07777135  0.1125121813  0.150887864
## HBA1C_T0           1.848513e-01  0.18356771 -0.0107162236  0.322363548
## HBA1C.mmol.mol_T0  1.848247e-01  0.18361534 -0.0107213019  0.322375631
## PESO_T0            1.030985e-01 -0.29665719 -0.1728637006  0.122169797
## bmi_T0             1.771316e-01 -0.22673533 -0.1680474258  0.023419549
## CC_T0              2.207149e-01  0.00758126 -0.0077235123 -0.341444926
## CINT_T0            1.957452e-01 -0.12561526 -0.1496859137 -0.286552989
## CAD_T0             1.112784e-01 -0.23777875 -0.2598588414 -0.125945898
## TAD_T0             1.607530e-01  0.08101446 -0.2027032128  0.385468271
## TAS_T0             6.116069e-02 -0.23731455 -0.2914386581  0.178911570
## TG_T0              2.382890e-01  0.05363323  0.1546591555 -0.008992496
## COL_T0             1.555571e-01  0.26556346 -0.1365157389 -0.093253755
## LDL_T0             1.731153e-01  0.14942926 -0.2412912905 -0.179315548
```

```
## HDL_T0               -1.731337e-01  0.25966592  0.0425202085  0.142259672
## VLDL_T0               2.382890e-01  0.05363323  0.1546591555 -0.008992496
## PCR_T0                2.012443e-01  0.09770185  0.2063896433 -0.202128531
## LEP_T0                1.695036e-01 -0.06206151 -0.2769408895  0.236209979
## ADIPO_T0             -2.276517e-01  0.11811067  0.0532546355  0.214491905
## GOT_T0                1.227601e-01  0.31588754 -0.0201212506 -0.131704788
## GPT_T0                1.688682e-01  0.26844697  0.0683883809 -0.091181103
## GGT_T0                2.091045e-01  0.18057083  0.1401439169  0.015756329
## URICO_T0              2.537605e-01 -0.06332388 -0.0248458392 -0.037185033
## CREAT_T0             -2.097037e-01  0.10713755 -0.2180727821  0.027140380
## UREA_T0               2.128422e-05  0.32536065 -0.1981896680 -0.072525585
## HIERRO_T0            -2.444839e-01  0.10250032 -0.0003968416  0.110083675
## TRANSF_T0             1.625339e-01  0.07166917 -0.3235396869  0.041264281
## FERR_T0               2.149934e-01  0.01261204  0.1882238050  0.224145251
## Ile_T0                1.941765e-01 -0.15392289  0.1743106342  0.206072925
## Leu_T0                1.372623e-01 -0.20988123  0.2701868494 -0.030995686
##                              PC5
## AGE                  0.225440592
## HOMA_T0             -0.144025054
## HBA1C_T0             0.462598015
## HBA1C.mmol.mol_T0    0.008170488
## PESO_T0             -0.219322952
## bmi_T0               0.164633909
## CC_T0                0.098363609
## CINT_T0              0.057801368
## CAD_T0              -0.112706509
## TAD_T0               0.212400370
## TAS_T0               0.041880828
## TG_T0               -0.058572235
## COL_T0               0.126291305
## LDL_T0              -0.132613519
## HDL_T0              -0.296539028
## VLDL_T0             -0.032781106
## PCR_T0               0.043376648
## LEP_T0              -0.203457228
## ADIPO_T0            -0.146210377
## GOT_T0              -0.165346454
## GPT_T0               0.232846351
## GGT_T0              -0.186318562
## URICO_T0            -0.163492538
## CREAT_T0             0.228432155
## UREA_T0             -0.183839340
## HIERRO_T0           -0.090895613
## TRANSF_T0           -0.281021101
## FERR_T0             -0.196905482
## Ile_T0               0.135385332
## Leu_T0               0.051283737
```

```r
summary(PCAMA)
```

```
## Importance of components:
##                          PC1     PC2    PC3     PC4       PC5
## Standard deviation     3.869  2.7085 2.3189 1.52163 1.444e-15
## Proportion of Variance 0.499  0.2445 0.1792 0.07718 0.000e+00
## Cumulative Proportion  0.499  0.7436 0.9228 1.00000 1.000e+00
```

Les dues primeres components expliquen el 74% de la variancia i les 3 primres, el 93%. Grafiquem les dues primeres

```
loads <- round(PCAMA$sdev^2/sum(PCAMA$sdev^2)*100,1)
groupColors <- c(rep("red", 4), rep("blue", 4))
xlab <- c(paste("PC1",loads[1],"%"))
ylab <- c(paste("PC2",loads[2],"%"))
plot(PCAMA$x[,1:2],
     xlab=xlab,
     ylab=ylab,
     col=groupColors,
     main ="Principal components (PCA)")
filesSenseNA <- complete.cases(assay(analisi)[,c(3,12:40)])
names2plot <- which(filesSenseNA)

text(PCAMA$x[,1],PCAMA$x[,2],names2plot, pos=3, cex=.6)
```

Principal components (PCA)