

# PAC1

Xènia Castellà

2025-03-30

## Contents

<b>Resum</b>	<b>1</b>
<b>Objectius</b>	<b>2</b>
<b>Mètodes</b>	<b>2</b>
<b>Resultats</b>	<b>2</b>
Anàlisi estadístic univariant . . . . .	3
Visualització de dades multivariants . . . . .	7
<b>Discussió</b>	<b>9</b>
<b>Conclusions</b>	<b>9</b>
<b>Referències</b>	<b>10</b>
<b>Annex A</b>	<b>10</b>
<b>Annex B</b>	<b>12</b>

## Resum

En aquesta primera prova d'avaluació contínua es crea a partir de dades de **caquèxia** un objecte de classe **SummarizedExperiment**. A partir del maneig de l'objecte on tenim identificadors dels pacients i pèrdua del múscul com a covariables característiques i les covariables mostres que són les mostres dels metabòlits que s'han pres, fem una anàlisi exploratòria, tant una anàlisi univariant com multivariant. Gràcies a l'anàlisi univariant hem observat que per comparar les mostres entre elles hem hagut de normalitzar prenent logaritmes. Així que finalment, hem obtingut a partir del mètode de les principals components (PCA) i de l'agrupació jeràrquica que hi ha 4 mostres que s'agrupen amb diferenciació a les restants. Aquest fet ens dona informació que hi pot haver efecte **batch**.

## Objectius

Per tal d'afrontar la realització d'aquesta PAC, hem de saber treballar amb algunes de les eines que han estat introduïdes durant el primer repete. Per tant, l'objectiu d'aquesta entrega és estar familiaritzats i tenir destresa amb:

- Les tecnologies òmiques.
- Bioconductor i les classes que s'utilitzen per emmagatzemar dades òmiques, com els `expressionSets` i `SummarizedExperiment`.
- Git com a eina de control de versions, com GitHub.
- Les eines estadístiques d'exploració de dades introduïdes.
- La pràctica d'eines i mètodes treballades.
- Planificar i executar una versió simplificada del procés d'anàlisi de dades òmiques.

## Mètodes

En aquest entrega ens hem centrat amb les dades proporcionades pel repositori de GitHub de l'usuari `alexsanchezpla`. En concret, hem seleccionat i descarregat un dataset de *metabolòmica*, el de *caquèxia*. El dataset de *caquèxia* inclou un fitxer `.csv` anomenat `human_cachexia.csv` on inclou 77 files i 65 columnes. Les 65 columnes proporcionen dades de sobre la identificació del pacient, la pèrdua de múscul i 63 mostres de metabòlits. A més, hi ha un document anomenat `description.md` que ens proporciona informació sobre el dataset. Ens informa que les mostres no estan aparellades, hi ha 2 grups que es detecten a les mostres, tots els valors de les dades són numèriques, que no hi ha cap valor `NULL` i que aquestes dades han passat la comprovació de sanitat. Per poder treballar amb aquestes dades, hem hagut de familiaritzar-nos amb la classe `SummarizedExperiment` per tal de crear un objecte d'aquesta classe. Per això, s'ha requerit les llibreries de `SummarizedExperiment` i `BiocManager`. Per fer l'anàlisi exploratòria del dataset per tal d'obtenir-ne una visió general, hem utilitzat comandaments bàsics per veure'n el seu aspecte. Un cop realitzat, hem continuat fent un anàlisi estadístic univariant. Hem observat cada mostra com es distribueix tant amb un histograma com amb boxplot. Així que boxplot ens ha permès poder tenir totes les mostres graficades conjuntament. Així que hem observat que no era fàcil comparar-les, i hem optat per normalitzar les mostres prenent logaritmes. Seguidament, hem passat a analitzar les dades a nivell multivariant. Hem utilitzat les eines de les components principals (PCA) i de l'agrupació jeràrquica. On en tots dos casos hem detectat que es diferencien dos grups.

## Resultats

Primer de tot, llegim el fitxer que conté les dades i mirem quina estructura té. Veure resultats a Annex A.

```
dades <- read.csv(myfile, header = TRUE, sep = ",", stringsAsFactors = FALSE,
                  check.names = FALSE)
dim(dades)
```

```
## [1] 77 65
```

```
table(dades$`Muscle loss`)
```

```
##
## cachexic  control
##      47      30
```

En el nostre cas és un fitxer de dades de metabolòmica que tracta la caquèxia amb 77 pacients i 65 variables. Observem que tenim dues columnes, `Patient ID` i `Muscle loss`, que són les covariables que ens donen informació de les característiques. La resta de les columnes ens serveixen per construir una matriu de dades, ja que són els metabolits i són covariables de mostres.

Procedim a organitzar aquesta informació del dataset en els conjunts de metadades, característiques i de la matriu per crear l'objecte de classes `SummarizedExperiment`. El conjunt de metadades, que estarà assignat a `col_dades`, contindrà la informació de les mostres dels metabolits. El conjunt de característiques, és a dir, les columnes del `PatientID` i `Muscle loss`, estarà contingut dins del `DataFrame` de `row_dades`. Finalment, la matriu de les dades de les mostres ho contindrà `assay_dades`.

```
row_dades <- DataFrame(
  Patient_ID = dades$`Patient ID`,
  Muscle_loss = dades$`Muscle loss`
)

assay_dades <- as.matrix(dades[, -c(1,2)])
rownames(assay_dades) <- dades$`Patient ID`

col_dades <- DataFrame(
  Sample = colnames(dades[, -c(1,2)])
)
```

Amb aquestes assignacions realitzades, ja podem construir el nostre objecte de la classe `SummarizedExperiment`.

```
se <- SummarizedExperiment(assays=list(counts=assay_dades),
  colData = col_dades,
  rowData = row_dades)
```

Com que la classe `SummarizedExperiment` és una extensió d' `ExpressionSet`, és molt similar. No obstant, la principal diferència és que `SummarizedExperiment` és més flexible amb la informació de les files, per exemple, permet tant `GRanges` based com els `DataFrames`.

Un cop creat l'objecte de la classe `SummarizedExperiment`, continuem fent una anàlisi exploratòria. Primer de tot, observem quin aspecte té el nostre objecte.

## Anàlisi estadístic univariant

```
se
```

```
## class: SummarizedExperiment
## dim: 77 63
## metadata(0):
## assays(1): counts
## rownames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## rowData names(2): Patient_ID Muscle_loss
## colnames(63): 1,6-Anhydro-beta-D-glucose 1-Methylnicotinamide ...
## pi-Methylhistidine tau-Methylhistidine
## colData names(1): Sample
```

```
head(assay(se), n=2)
```

```
##      1,6-Anhydro-beta-D-glucose 1-Methylnicotinamide 2-Aminobutyrate
## PIF_178      40.85      65.37      18.73
## PIF_087      62.18      340.36      24.29
##      2-Hydroxyisobutyrate 2-Oxoglutarate 3-Aminoisobutyrate
## PIF_178      26.05      71.52      1480.30
## PIF_087      41.68      67.36      116.75
##      3-Hydroxybutyrate 3-Hydroxyisovalerate 3-Indoxylsulfate
## PIF_178      56.83      10.07      566.80
## PIF_087      43.82      79.84      368.71
##      4-Hydroxyphenylacetate Acetate Acetone Adipate Alanine Asparagine
## PIF_178      120.30 126.47   9.49   38.09 314.19   159.17
## PIF_087      432.68 212.72  11.82 327.01 871.31   157.59
##      Betaine Carnitine Citrate Creatine Creatinine Dimethylamine
## PIF_178 109.95 265.07 3714.50 196.37 16481.60   632.70
## PIF_087 244.69 120.30 2617.57 212.72 15835.35   607.89
##      Ethanolamine Formate Fucose Fumarate Glucose Glutamine Glycine
## PIF_178 645.48 441.42 336.97   7.69 395.44   871.31 2038.56
## PIF_087 487.85 252.14 198.34  18.92 8690.62   601.85 1107.65
##      Glycolate Guanidoacetate Hippurate Histidine Hypoxanthine Isoleucine
## PIF_178 685.40      154.47 4582.50  925.19      97.51      5.58
## PIF_087 651.97      109.95 1737.15  845.56      82.27      8.17
##      Lactate Leucine Lysine Methylamine Methylguanidine N,N-Dimethylglycine
## PIF_178 106.70  42.10 146.94      52.46      9.97      23.34
## PIF_087 368.71  77.48 284.29      23.57      7.69      87.36
##      O-Acetylcarnitine Pantothenate Pyroglutamate Pyruvate Quinolate
## PIF_178      52.98      25.79      437.03   21.12   165.67
## PIF_087      50.40      186.79      437.03   36.97    72.97
##      Serine Succinate Sucrose Tartrate Taurine Threonine Trigonelline
## PIF_178 284.29  154.47  45.15   97.51 1919.85   184.93   943.88
## PIF_087 391.51  244.69 459.44   32.79 1261.43   198.34   208.51
##      Trimethylamine N-oxide Tryptophan Tyrosine Uracil Valine Xylose
## PIF_178      2121.76   259.82  290.03 111.05  86.49  72.24
## PIF_087      639.06    83.10  167.34  46.99 109.95 192.48
##      cis-Aconitate myo-Inositol trans-Aconitate pi-Methylhistidine
## PIF_178      237.46    135.64      51.94      157.59
## PIF_087      333.62    376.15      217.02      307.97
##      tau-Methylhistidine
## PIF_178      160.77
## PIF_087      130.32
```

```
dim(assay(se))
```

```
## [1] 77 63
```

```
colData(se)
```

```
## DataFrame with 63 rows and 1 column
##                               Sample
##                               <character>
## 1,6-Anhydro-beta-D-glucose 1,6-Anhydro-beta-D-g..
```

```
## 1-Methylnicotinamide      1-Methylnicotinamide
## 2-Aminobutyrate           2-Aminobutyrate
## 2-Hydroxyisobutyrate      2-Hydroxyisobutyrate
## 2-Oxoglutarate            2-Oxoglutarate
## ...                       ...
## cis-Aconitate              cis-Aconitate
## myo-Inositol               myo-Inositol
## trans-Aconitate            trans-Aconitate
## pi-Methylhistidine         pi-Methylhistidine
## tau-Methylhistidine        tau-Methylhistidine
```

```
dim(colData(se))
```

```
## [1] 63  1
```

```
head(rowData(se), n=3)
```

```
## DataFrame with 3 rows and 2 columns
##      Patient_ID Muscle_loss
##      <character> <character>
## PIF_178      PIF_178      cachexic
## PIF_087      PIF_087      cachexic
## PIF_090      PIF_090      cachexic
```

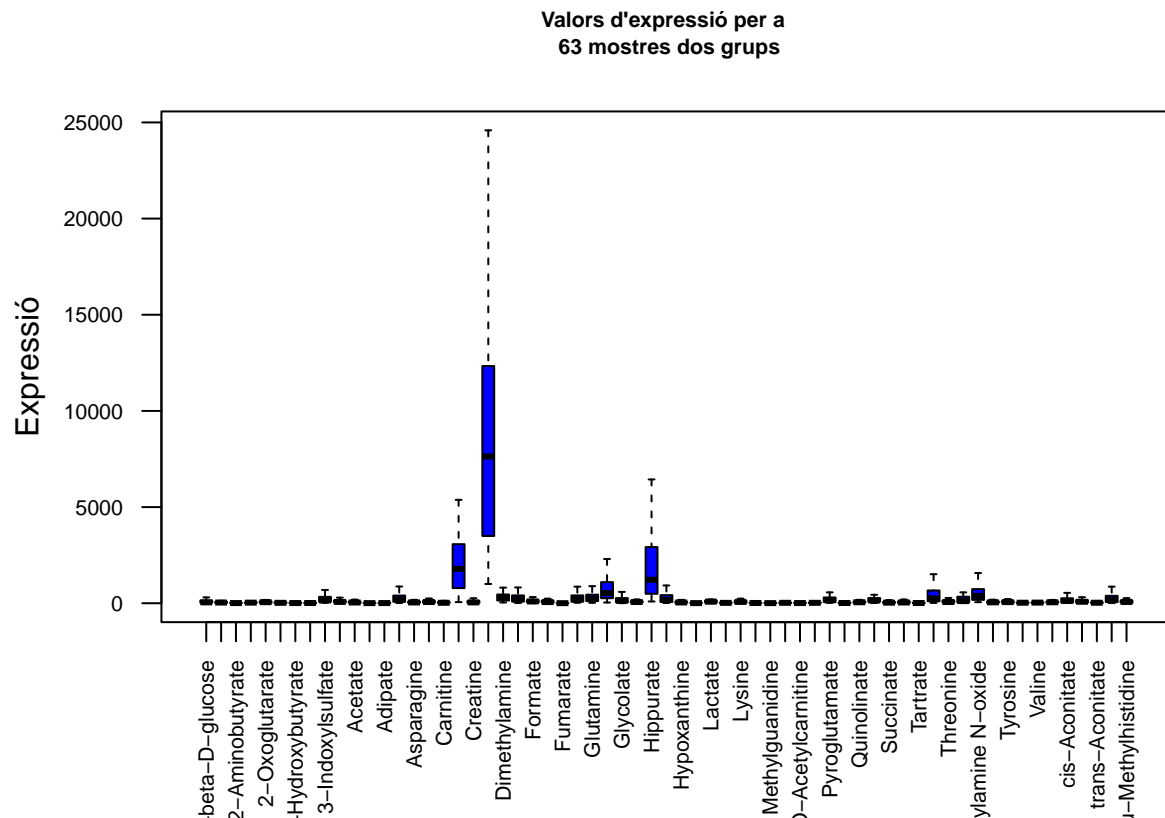
```
dim(rowData(se))
```

```
## [1] 77  2
```

Ara, utilitzem eines estadístiques bàsiques i ho representem gràficament. Primer apliquem la funció `summary()` a les dades i les representem gràficament amb histograma. Veure resultats a l'annex B.

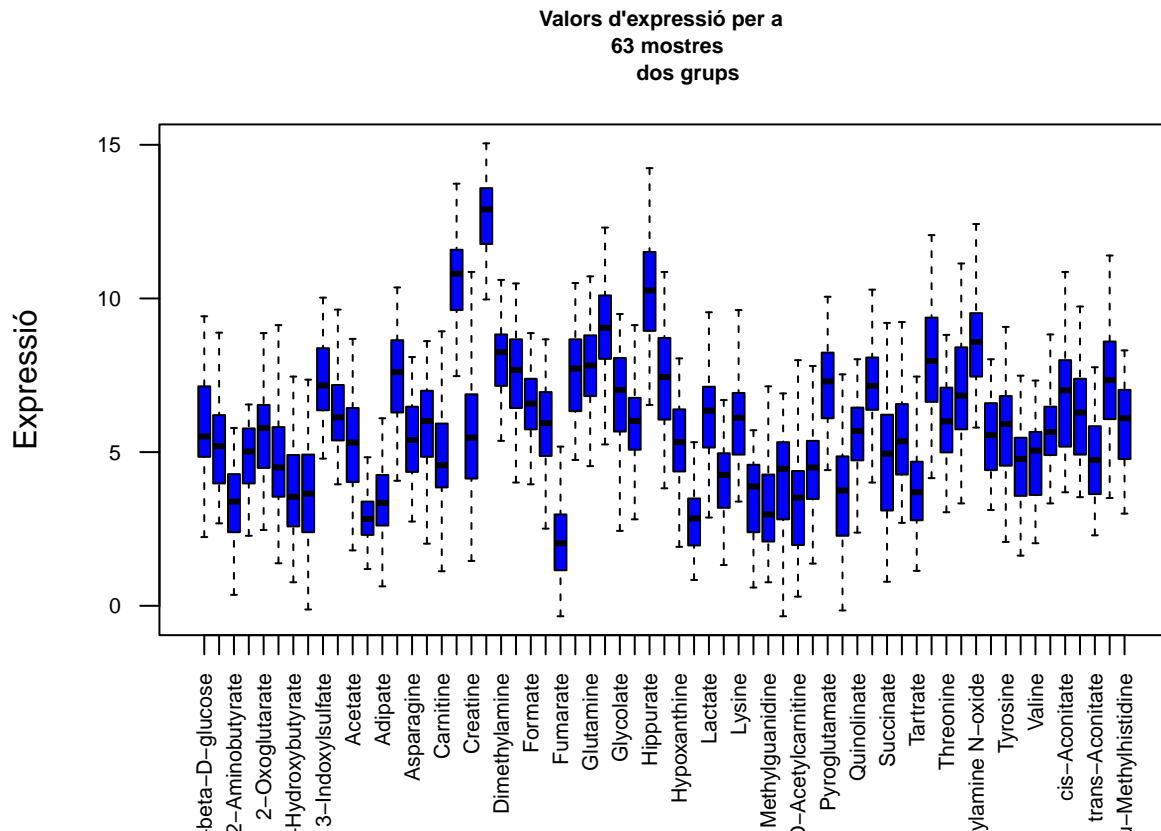
També ho representem amb boxplot per veure totes les mostres juntes.

```
boxplot(assay(se), col = "blue", main = "Valors d'expressió per a \n 63 mostres dos grups",
        ylab="Expressió",
        las=2, cex.axis=0.7, cex.main=0.7, outline = FALSE)
```



Acabem de veure que si ho grafiquem amb boxplot no podem comparar els metabolits, ja que tenen valors molt diferents. Per tant, normalitzem les dades fent ús de logaritmes.

```
boxplot(log2(assay(se)), col = "blue", main = "Valors d'expressió per a \n 63 mostres dos grups",
        ylab="Expressió", las=2, cex.axis=0.7, cex.main=0.7, outline = FALSE)
```



Veiem que és millor i per tant, d'ara endavant, prenem logaritmes a les dades del nostre objecte.

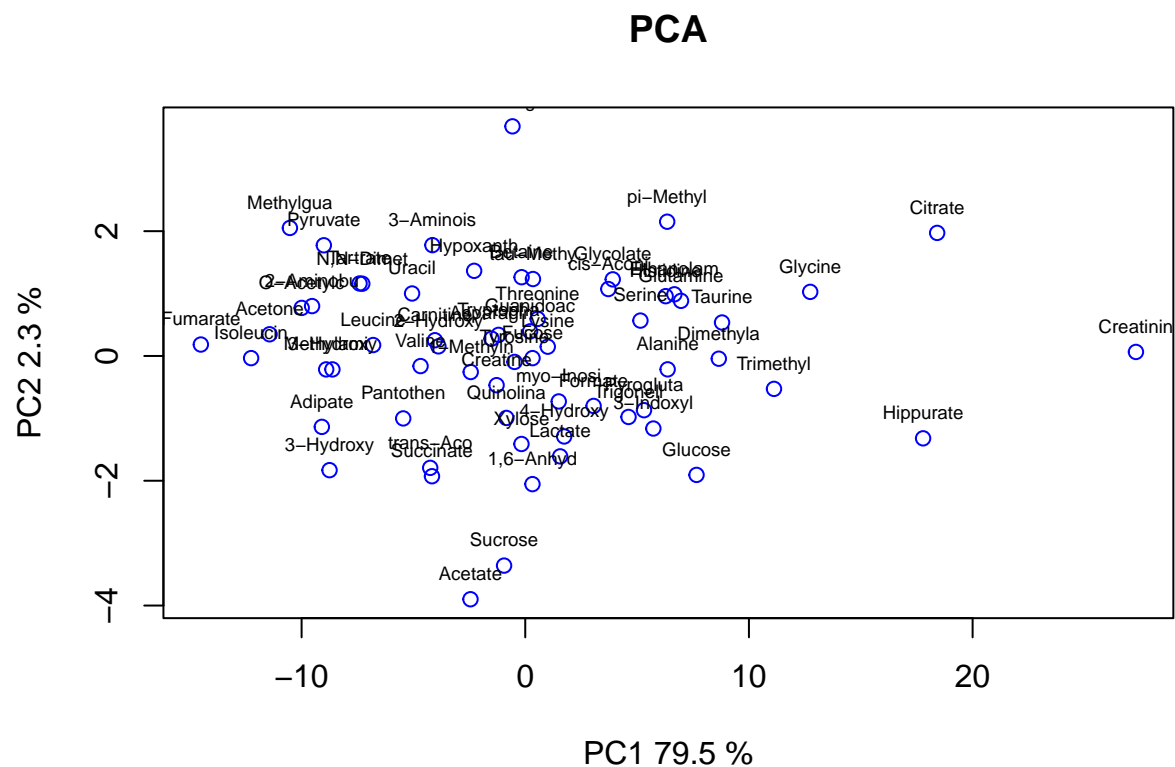
## Visualització de dades multivariants

Calcular les components principals per mostres ens ajudarà per detectar mostres inusuals o efectes `batch`. Així que primer de tot, calculem les components principals (PCA).

```
pcX <- prcomp(t(log2(assay(se))), scale = TRUE)
loads <- round(pcX$sdev^2/sum(pcX$sdev^2)*100,1)
```

Grafiquem les dues primeres components.

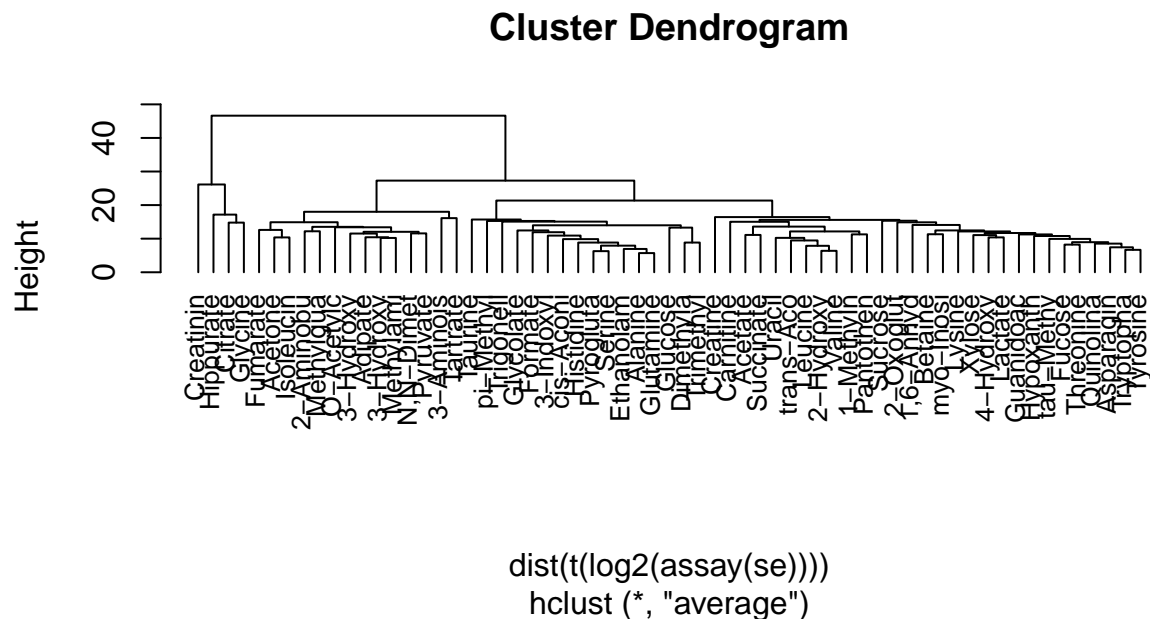
```
xlab <- c(paste("PC1", loads[1], "%"))
ylab <- c(paste("PC2", loads[2], "%"))
plot(pcX$x[,1:2], xlab=xlab, ylab = ylab, col = "blue", main = "PCA")
names2plot <- substr(colnames(assay(se)), 1, 9)
text(pcX$x[,1], pcX$x[,2], names2plot, pos = 3, cex = .6)
```



La següent graficació és una agrupació jeràrquica per visualitzar qualsevol mena d'agrupació entre les mostres.

```
colnames(se) <- names2plot
clust.euclid.average <- hclust(dist(t(log2(assay(se))))), method = "average")
par(mar = c(10,4,4,2))
plot(clust.euclid.average, hang = -1, cex = 0.8)
```





Tant a la PCA com a l'agrupació jeràrquica podem observar que hi ha dues agrupacions clares. Per una banda hi ha Creatinine, Hippurate, Citrate i Glycine que s'agrupen separades de la resta, fet que ens indica que hi ha efecte batch o mostres inusuals.

## Discussió

Un cop feta l'anàlisi exploratòria del dataset *caquèxia* i hem pogut observar la seva informació a través d'un objecte de la classe `SummarizedExperiment`, podem discutir sobre la informació que hem proporcionat i també les seves limitacions. Per una banda, tenim 77 pacients diferents dels quals se'ls hi ha pres mostres de 63 metabòlits. A més, també tenim una informació prou interessant que és la pèrdua de múscul que pot ser de dos tipus, o bé *controlada* o bé amb *caquèxia*. Al final d'aquesta anàlisi obtenim que hi ha efecte batch. Així que una limitació és que no proporcionem cap correcció per revertir els efectes de batch. Amb les dades proporcionades tampoc sabem quin és l'error primari que comporta que hi hagi aquests valors inusuals. Una altra limitació seria que no hi ha hagut una conclusió clara entre els tipus de pèrdua de múscul amb les diferents mostres dels metabòlits. En conclusió, tot i que s'ha pogut explorar les dades amb cert detall gràcies a l'objecte creat, tots els càlculs i gràfics ens han ajudat a entendre més les dades i a corroborar entre les diferents eines allà on hi ha limitacions.

## Conclusions

En conclusió, la creació de l'objecte de classe `SummarizedExperiment` permet que les dades estiguin en constant sincronització, cosa que facilita la consistència de les dades. Hem detectat que de tots els pacients, hi ha dos grups diferenciats en la 'pèrdua de múscul', ja sigui controlat o per *caquèxia*. Hem estudiat 63

mostres, tant una per una com conjuntament fent una anàlisi univariant i multivariant. Hem observat que per poder-les comparar les havíem de normalitzar amb logaritme. Això ens ha permès poder observar tant a través del gràfic de PCA com de l'agrupació jeràrquica que hi ha 4 mostres amb valors inusuals, amb efecte batch.

## Referències

Podeu trobar el codi del repositori a GitHub o sinó <https://github.com/xcasteca/Castella-Camps-Xenia-PAC1>.

## Annex A

Estructura de les dades un cop llegides del fitxer `human_cachexia.csv`.

```
head(dades, n=3)
```

```
## Patient ID Muscle loss 1,6-Anhydro-beta-D-glucose 1-Methylnicotinamide
## 1 PIF_178 cachexic 40.85 65.37
## 2 PIF_087 cachexic 62.18 340.36
## 3 PIF_090 cachexic 270.43 64.72
## 2-Aminobutyrate 2-Hydroxyisobutyrate 2-Oxoglutarate 3-Aminoisobutyrate
## 1 18.73 26.05 71.52 1480.30
## 2 24.29 41.68 67.36 116.75
## 3 12.18 65.37 23.81 14.30
## 3-Hydroxybutyrate 3-Hydroxyisovalerate 3-Indoxylsulfate
## 1 56.83 10.07 566.80
## 2 43.82 79.84 368.71
## 3 5.64 23.34 665.14
## 4-Hydroxyphenylacetate Acetate Acetone Adipate Alanine Asparagine Betaine
## 1 120.30 126.47 9.49 38.09 314.19 159.17 109.95
## 2 432.68 212.72 11.82 327.01 871.31 157.59 244.69
## 3 292.95 314.19 4.44 131.63 464.05 89.12 116.75
## Carnitine Citrate Creatine Creatinine Dimethylamine Ethanolamine Formate
## 1 265.07 3714.50 196.37 16481.60 632.70 645.48 441.42
## 2 120.30 2617.57 212.72 15835.35 607.89 487.85 252.14
## 3 25.03 862.64 221.41 24587.66 735.10 407.48 249.64
## Fucose Fumarate Glucose Glutamine Glycine Glycolate Guanidoacetate Hippurate
## 1 336.97 7.69 395.44 871.31 2038.56 685.40 154.47 4582.50
## 2 198.34 18.92 8690.62 601.85 1107.65 651.97 109.95 1737.15
## 3 186.79 7.10 1352.89 301.87 620.17 141.17 183.09 4315.64
## Histidine Hypoxanthine Isoleucine Lactate Leucine Lysine Methylamine
## 1 925.19 97.51 5.58 106.70 42.10 146.94 52.46
## 2 845.56 82.27 8.17 368.71 77.48 284.29 23.57
## 3 284.29 114.43 9.30 749.95 31.50 97.51 18.73
## Methylguanidine N,N-Dimethylglycine O-Acetylcarnitine Pantothenate
## 1 9.97 23.34 52.98 25.79
## 2 7.69 87.36 50.40 186.79
## 3 4.66 24.53 5.58 145.47
## Pyroglutamate Pyruvate Quinolinolate Serine Succinate Sucrose Tartrate Taurine
## 1 437.03 21.12 165.67 284.29 154.47 45.15 97.51 1919.85
```

```
## 2      437.03      36.97      72.97 391.51      244.69 459.44      32.79 1261.43
## 3      713.37      29.37      192.48 295.89      142.59 160.77      16.28 4272.69
##   Threonine Trigonelline Trimethylamine N-oxide Tryptophan Tyrosine Uracil
## 1      184.93      943.88      2121.76      259.82      290.03 111.05
## 2      198.34      208.51      639.06      83.10      167.34 46.99
## 3      109.95      192.48      1152.86      82.27      60.34 31.50
##   Valine Xylose cis-Aconitate myo-Inositol trans-Aconitate pi-Methylhistidine
## 1  86.49  72.24      237.46      135.64      51.94      157.59
## 2 109.95 192.48      333.62      376.15      217.02      307.97
## 3  59.15 2164.62      330.30      86.49      58.56      145.47
##   tau-Methylhistidine
## 1      160.77
## 2      130.32
## 3      83.93
```

```
str(dades)
```

```
## 'data.frame': 77 obs. of 65 variables:
## $ Patient ID : chr "PIF_178" "PIF_087" "PIF_090" "NETL_005_V1" ...
## $ Muscle loss : chr "cachexic" "cachexic" "cachexic" "cachexic" ...
## $ 1,6-Anhydro-beta-D-glucose: num 40.9 62.2 270.4 154.5 22.2 ...
## $ 1-Methylnicotinamide : num 65.4 340.4 64.7 53 73.7 ...
## $ 2-Aminobutyrate : num 18.7 24.3 12.2 172.4 15.6 ...
## $ 2-Hydroxyisobutyrate : num 26.1 41.7 65.4 74.4 83.9 ...
## $ 2-Oxoglutarate : num 71.5 67.4 23.8 1199.9 33.1 ...
## $ 3-Aminoisobutyrate : num 1480.3 116.8 14.3 555.6 29.7 ...
## $ 3-Hydroxybutyrate : num 56.83 43.82 5.64 175.91 76.71 ...
## $ 3-Hydroxyisovalerate : num 10.1 79.8 23.3 25 69.4 ...
## $ 3-Indoxylsulfate : num 567 369 665 412 166 ...
## $ 4-Hydroxyphenylacetate : num 120.3 432.7 292.9 214.9 97.5 ...
## $ Acetate : num 126.5 212.7 314.2 37.3 407.5 ...
## $ Acetone : num 9.49 11.82 4.44 206.44 44.26 ...
## $ Adipate : num 38.1 327 131.6 144 15 ...
## $ Alanine : num 314 871 464 590 1119 ...
## $ Asparagine : num 159.2 157.6 89.1 273.1 42.5 ...
## $ Betaine : num 110 245 117 279 392 ...
## $ Carnitine : num 265.1 120.3 25 200.3 84.8 ...
## $ Citrate : num 3714 2618 863 13630 854 ...
## $ Creatine : num 196.4 212.7 221.4 85.6 105.6 ...
## $ Creatinine : num 16482 15835 24588 20952 6768 ...
## $ Dimethylamine : num 633 608 735 1064 242 ...
## $ Ethanolamine : num 645 488 407 821 365 ...
## $ Formate : num 441 252 250 469 114 ...
## $ Fucose : num 337 198.3 186.8 407.5 26.1 ...
## $ Fumarate : num 7.69 18.92 7.1 96.54 19.69 ...
## $ Glucose : num 395 8691 1353 863 6836 ...
## $ Glutamine : num 871 602 302 1686 433 ...
## $ Glycine : num 2039 1108 620 5064 395 ...
## $ Glycolate : num 685.4 652 141.2 70.8 26.6 ...
## $ Guanidoacetate : num 154 110 183 103 53 ...
## $ Hippurate : num 4582 1737 4316 757 1153 ...
## $ Histidine : num 925 846 284 1043 327 ...
## $ Hypoxanthine : num 97.5 82.3 114.4 223.6 66.7 ...
## $ Isoleucine : num 5.58 8.17 9.3 37.71 40.04 ...
```

```
## $ Lactate : num 107 369 750 369 3641 ...
## $ Leucine : num 42.1 77.5 31.5 103.5 101.5 ...
## $ Lysine : num 146.9 284.3 97.5 290 122.7 ...
## $ Methylamine : num 52.5 23.6 18.7 48.9 27.9 ...
## $ Methylguanidine : num 9.97 7.69 4.66 141.17 5.31 ...
## $ N,N-Dimethylglycine : num 23.3 87.4 24.5 40 46.1 ...
## $ O-Acetylcarnitine : num 52.98 50.4 5.58 254.68 45.6 ...
## $ Pantothenate : num 25.8 186.8 145.5 42.5 74.4 ...
## $ Pyroglutamate : num 437 437 713 567 185 ...
## $ Pyruvate : num 21.1 37 29.4 64.1 12.3 ...
## $ Quinolate : num 165.7 73 192.5 86.5 38.1 ...
## $ Serine : num 284 392 296 1249 206 ...
## $ Succinate : num 154.5 244.7 142.6 144 68.7 ...
## $ Sucrose : num 45.1 459.4 160.8 111 75.2 ...
## $ Tartrate : num 97.51 32.79 16.28 837.15 4.53 ...
## $ Taurine : num 1920 1261 4273 1525 469 ...
## $ Threonine : num 184.9 198.3 110 376.1 64.1 ...
## $ Trigonelline : num 943.9 208.5 192.5 992.3 86.5 ...
## $ Trimethylamine N-oxide : num 2122 639 1153 1451 172 ...
## $ Tryptophan : num 259.8 83.1 82.3 235.1 103.5 ...
## $ Tyrosine : num 290 167.3 60.3 323.8 142.6 ...
## $ Uracil : num 111 47 31.5 30.6 44.3 ...
## $ Valine : num 86.5 110 59.1 102.5 160.8 ...
## $ Xylose : num 72.2 192.5 2164.6 125.2 186.8 ...
## $ cis-Aconitate : num 237 334 330 1863 101 ...
## $ myo-Inositol : num 135.6 376.1 86.5 247.2 750 ...
## $ trans-Aconitate : num 51.9 217 58.6 75.9 98.5 ...
## $ pi-Methylhistidine : num 157.6 308 145.5 249.6 84.8 ...
## $ tau-Methylhistidine : num 160.8 130.3 83.9 254.7 79.8 ...
```

## Annex B

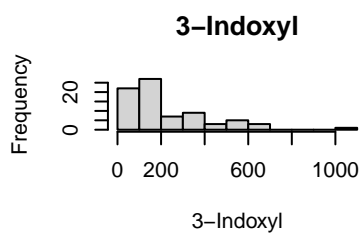
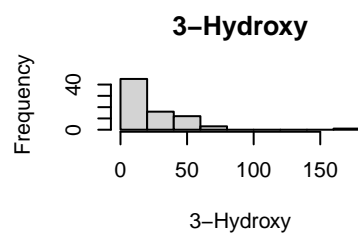
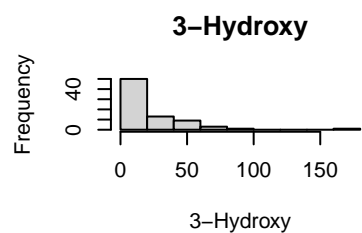
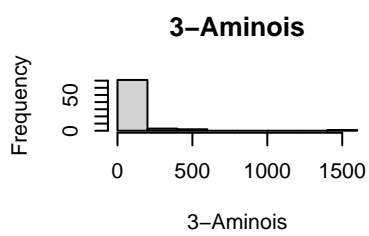
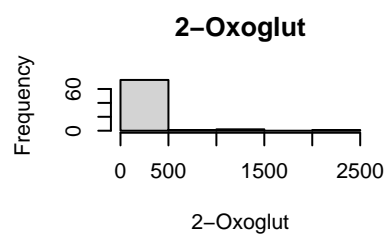
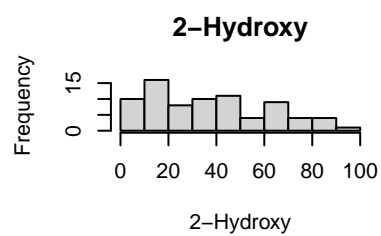
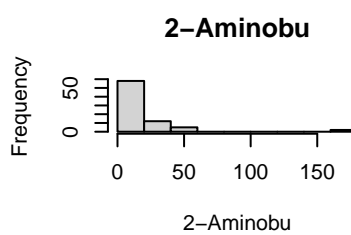
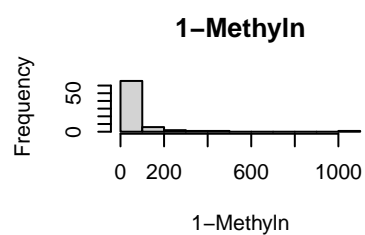
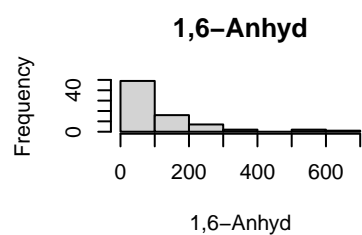
```
round(apply(assay(se), 2, summary))
```

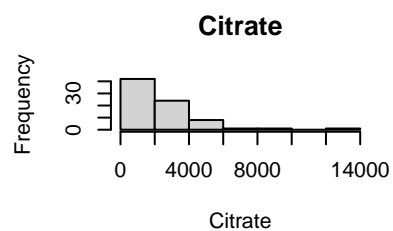
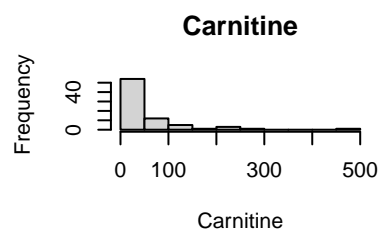
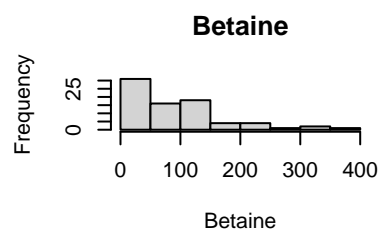
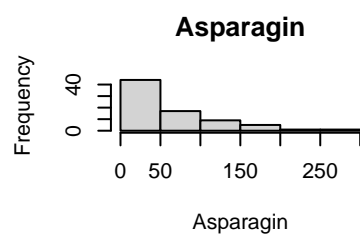
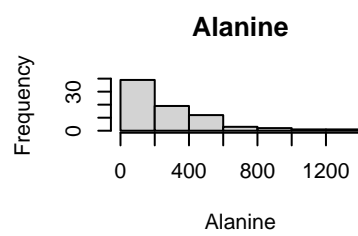
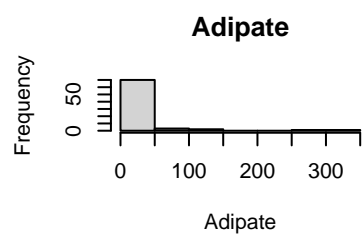
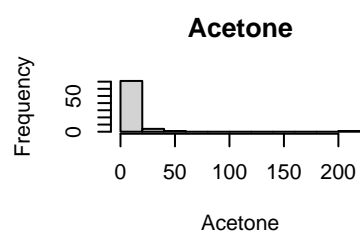
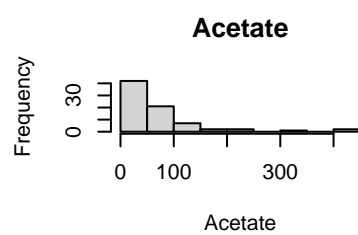
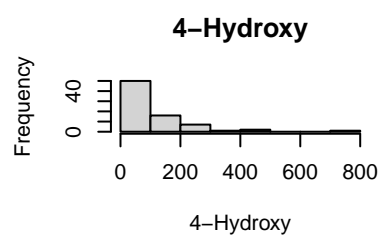
```
##      1,6-Anhyd 1-Methyln 2-Aminobu 2-Hydroxy 2-Oxoglut 3-Aminois 3-Hydroxy
## Min.      5         6         1         5         6         3         2
## 1st Qu.    29        16         5        16        22        12        6
## Median     46        37        10        32        55        23       12
## Mean      106        72        18        37       145        77       22
## 3rd Qu.    141        74        19        55        93        56       30
## Max.      685       1033       172        94      2465       1480      176
##      3-Hydroxy 3-Indoxyl 4-Hydroxy Acetate Acetone Adipate Alanine Asparagin
## Min.      1         28        15         3         2         2        17        7
## 1st Qu.     5         82        42        16         5         6       78       20
## Median     13        144        70        40         7        10      194       42
## Mean      22        219       112        66        11       25      274       62
## 3rd Qu.     30        334       145        86        10       19      399       89
## Max.      164       1043       796       412       206      327     1313     273
##      Betaine Carnitine Citrate Creatine Creatinin Dimethyla Ethanolam
## Min.      2         2        60         3      1002        41       16
## 1st Qu.    29        14       788        18     3498       143       86
```

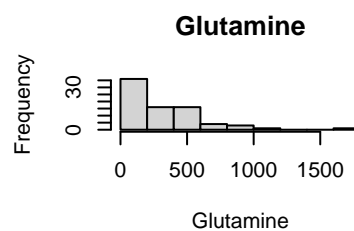
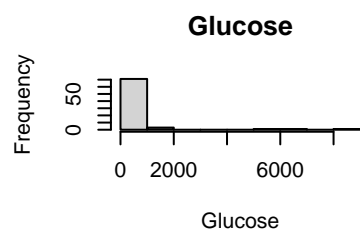
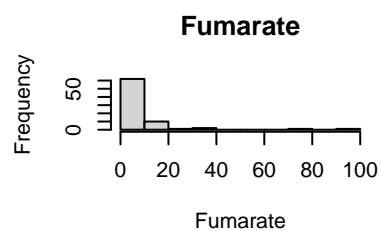
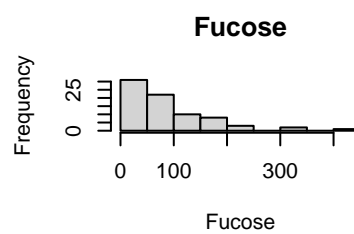
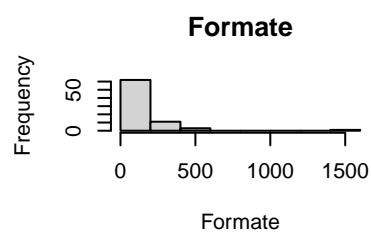
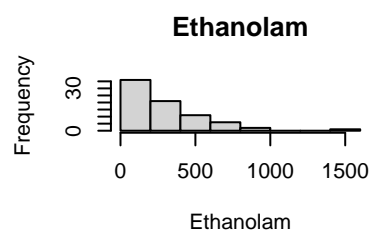
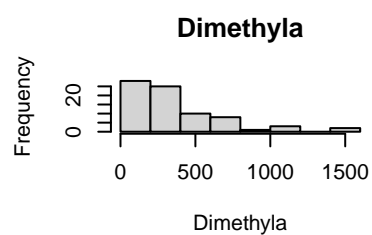
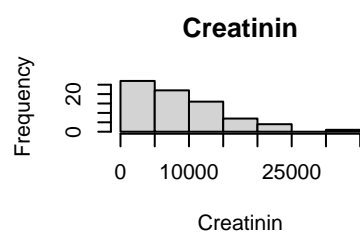
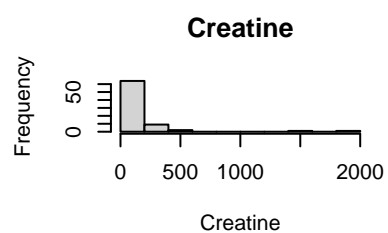
## Median	65	24	1790	44	7631	305	204	
## Mean	90	52	2235	127	8734	358	276	
## 3rd Qu.	128	61	3072	118	12333	455	407	
## Max.	392	488	13630	1863	33860	1556	1437	
##	Formate	Fucose	Fumarate	Glucose	Glutamine	Glycine	Glycolate	Guanidoac
## Min.	6	6	1	27	23	38	5	7
## 1st Qu.	54	29	2	81	113	262	51	34
## Median	96	62	4	211	226	528	130	65
## Mean	147	89	8	560	307	881	188	86
## 3rd Qu.	167	124	8	407	446	1097	268	109
## Max.	1480	407	97	8691	1686	5064	721	561
##	Hippurate	Histidine	Hypoxanth	Isoleucin	Lactate	Leucine	Lysine	
## Min.	93	14	4	2	7	3	10	
## 1st Qu.	493	67	21	4	36	9	30	
## Median	1224	174	40	7	81	19	69	
## Mean	2287	293	61	9	158	24	109	
## 3rd Qu.	2922	420	84	11	140	31	122	
## Max.	19341	1863	265	40	3641	104	788	
##	Methylami	Methylgua	N,N-Dimet	O-Acetyl	Pantothen	Pyrogluta	Pyruvate	
## Min.	2	2	1	1	3	21	1	
## 1st Qu.	5	4	7	4	11	69	5	
## Median	15	8	22	11	23	158	13	
## Mean	17	15	26	20	45	211	21	
## 3rd Qu.	24	19	40	21	41	302	29	
## Max.	52	141	120	255	692	1064	185	
##	Quinolina	Serine	Succinate	Sucrose	Tartrate	Taurine	Threonine	Trigonell
## Min.	5	16	2	6	2	18	8	10
## 1st Qu.	27	83	9	19	7	99	32	54
## Median	51	143	31	41	13	250	64	114
## Mean	66	198	60	113	40	525	95	270
## 3rd Qu.	87	270	74	95	26	665	137	340
## Max.	260	1249	590	2080	837	4273	450	2253
##	Trimethyl	Tryptopha	Tyrosine	Uracil	Valine	Xylose	cis-Aconi	myo-Inosi
## Min.	56	9	4	3	4	10	13	12
## 1st Qu.	176	21	24	12	12	30	36	30
## Median	384	47	60	27	33	50	129	78
## Mean	652	66	82	36	36	101	204	135
## 3rd Qu.	735	97	113	44	50	89	255	167
## Max.	5486	260	539	179	161	2165	1863	854
##	trans-Aco	pi-Methyl	tau-Methy					
## Min.	5	11	8					
## 1st Qu.	12	67	27					
## Median	27	162	69					
## Mean	41	370	90					
## 3rd Qu.	57	388	130					
## Max.	217	2697	317					

Un histograma per cada mostra.

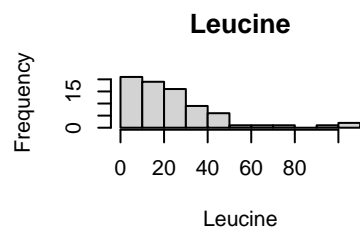
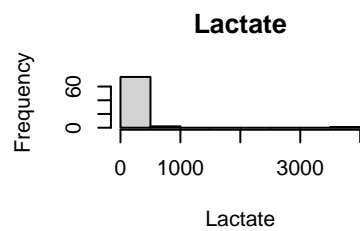
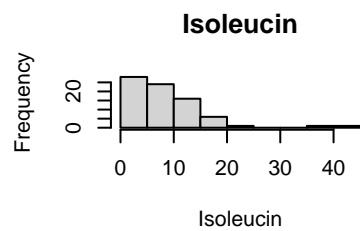
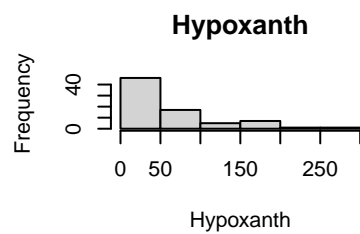
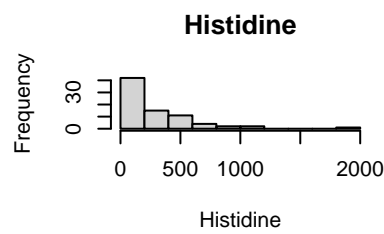
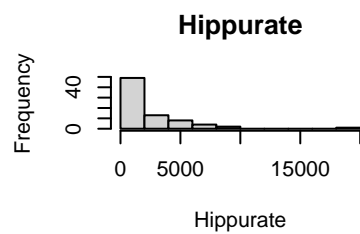
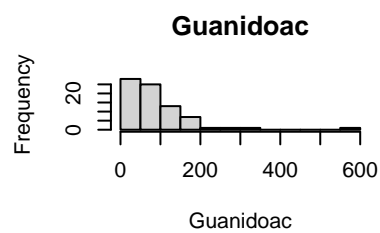
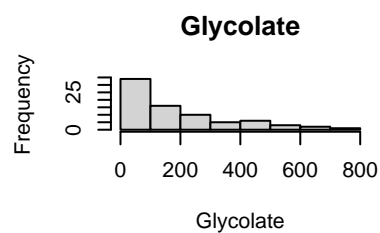
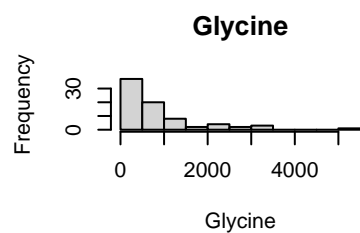
```
opt <- par(mfrow = c(3,3))
for (i in 1:ncol(assay(se)))
  hist(assay(se)[,i], main = colnames(assay(se))[i], xlab = colnames(assay(se))[i])
```

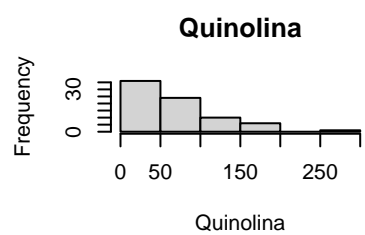
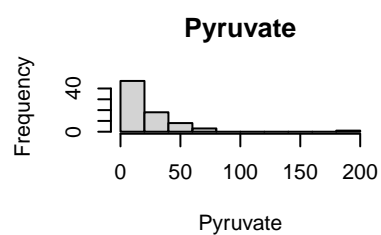
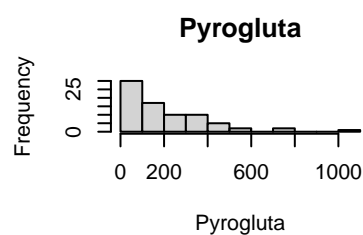
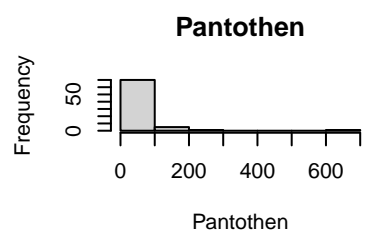
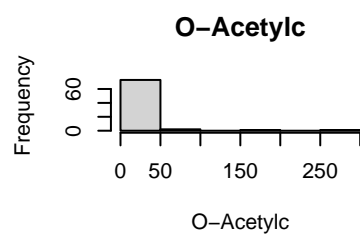
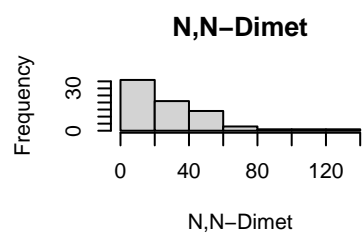
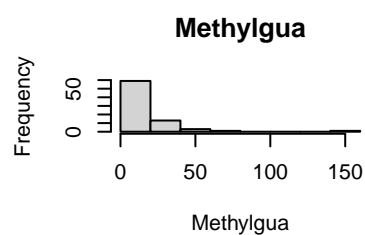
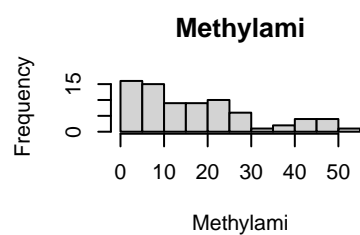
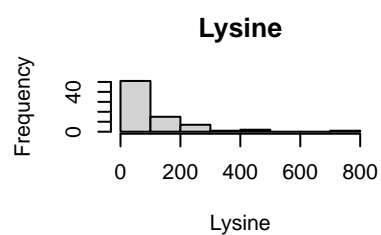


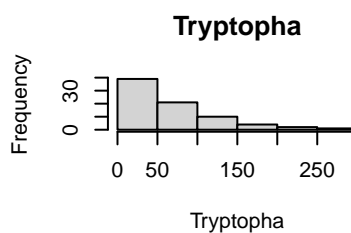
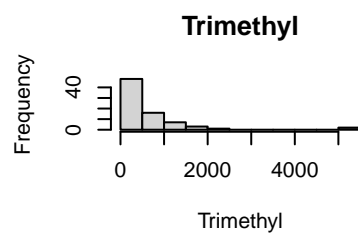
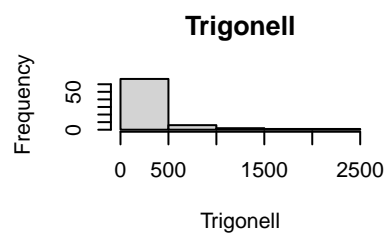
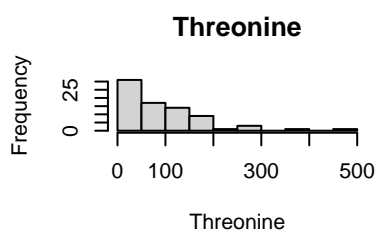
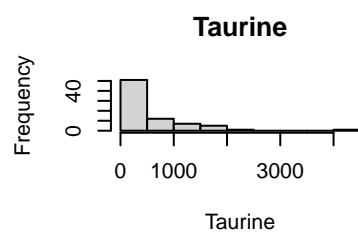
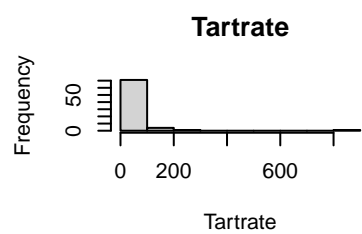
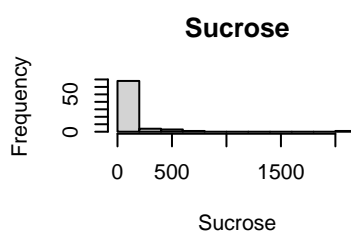
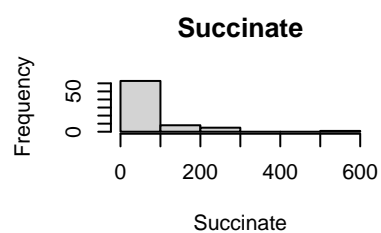
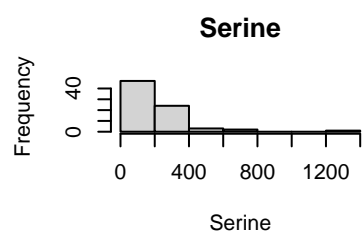


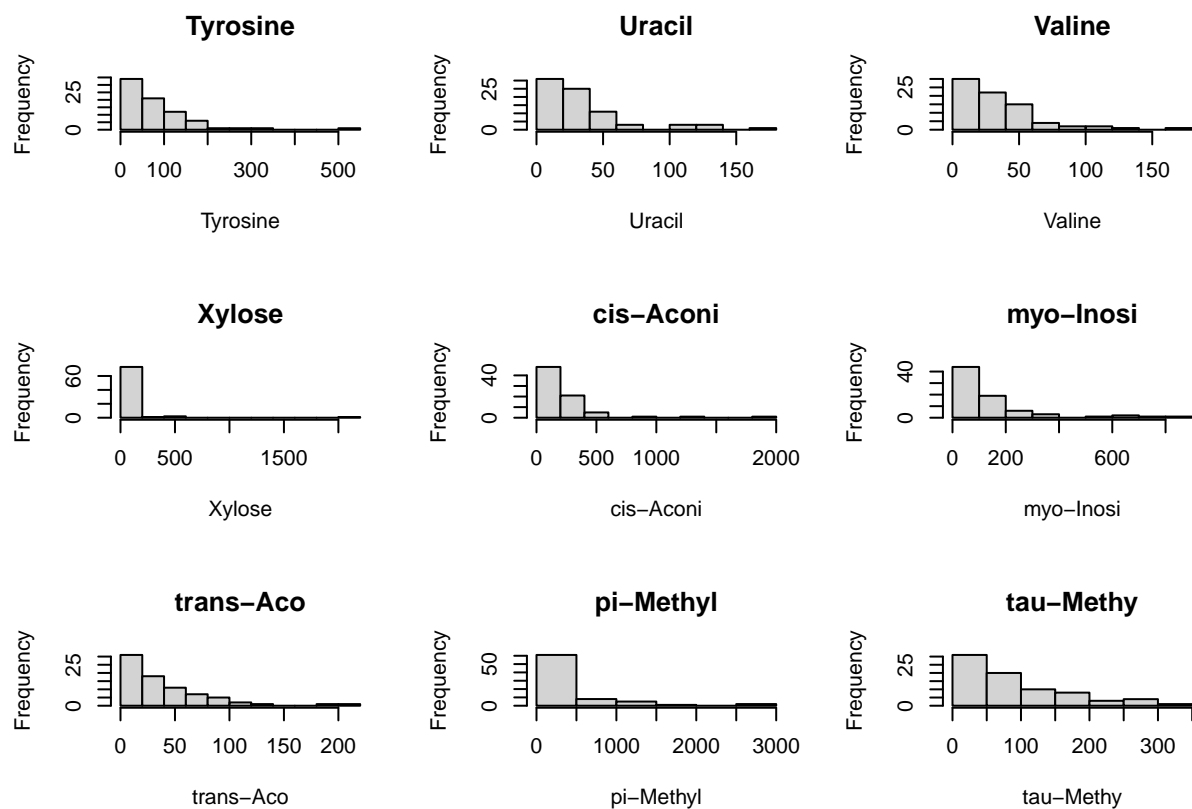












`par(opt)`