Informe PAC 1

Guillem Casanovas Sanglas

Contents

1 Resum

S'explora la caquèxia, un estat de desnutrició i atrofia muscular relacionat amb el càncer, mitjançant l'anàlisi de dades de 77 pacients amb càncer, que inclou pacients amb caquèxia i pacients controls. S'han analitzat les concentracions de diferents metabòlits a partir de mostres d'orina utilitzant RStudio. Els resultats indiquen que hi ha determinats metabòlits, com la Leucina i la Creatinina, que presenten diferències significatives en la seva concentració entre ambdós grups. La identificació d'aquests elements pot ser clau per a la creació d'un panell de biomarcadors que permetin diagnosticar pacients caquèxics.

2 Objectius

La caquèxia es un estat d'extrema desnturició, atròfia muscular i anorèxia que pateix una persona la qual no esta tractant de perdre pes. El més normal es perdre massa muscular i en alguns casos es perd teixit adipòs, però no sempre. Pot estar relacionada amb altres patologies, essent un símptoma de moltes patologies com per exemple càncer. Però també pot tenir relació amb altres patologies com tuberculosi, sida i malalties autoimmunes. Apareix en el 70% de pacients amb cancer, essent el responsable del 22% de les morts de pacients amb càncer [1]. Els pacients que la pateixen presenten feblesa física que pot arribar a portar a un estat d'immobilitat degut a la important perdua muscular. És molt díficil diagnosticar-la ja que no hi ha un criteri ben definit, sobretot en els estadis inicials on es podria prevenir [2].

L'objectiu d'aquest informe es el de realitzar una exploració de les dades de 77 pacients amb càncer, alguns dels quals pateixen caquèxia i altres no. D'aquesta manera amb la exploració de les dades que tenim, es pretè intentar trobar diferencies en els valors dels metàbolits analitzats que permetin diferenciar entre la població que pateix caquèxia i la que no. El qual podria ser útil en futurs estudis de pacients amb condicions similars per a poder intentar establir biomarcadors que permetin diagnosticar la patologia.

3 Materials i Metodes

Pel que fa a les dades usades en l'informe, no són dades originals, sinò que son proporcionades pel tutor de l'assignatura. Les quals s'obtenen de la web MetaboAnalyst i contenen informació metabolòmica de mostres d'oringa de pacients amb càncer, tant controls com pacients caquèxics. En particular s'analitza la concentració d'orina d'una gran mostra de metabòlits.

Per a l'analisis i tractament de les dades s'ha usat exclusivament el software *Rstudio* amb el qual s'ha generat tant l'informe com el fitxer que conté el contenidor *SummarizedExperiment*. Al treballar les dades s'ha treballat amb el llenguatge de programació R. Gran part de els analisis s'han realitzat fent servir funcions propies de la versio base del software. A més a més però també s'ha fet servir els paquets *BiocManager*, *SummarizedExperiment* i *ggfortify*, que proporcionen funcions específiques per a la manipulació i visualització de les dades.

El procés seguit per l'analisis ha sigut el que es descriu a continuació. En primer lloc s'estudien les dades per observar si es necessita cap manipulació previa a poder treballar-hi. Seguidament es crea el contenidor *SummarizedExpermient*, el qual conté tota la informació relativa a les dades de l'experiement, aixi com també les metadades associades. Posteriorment es procedeix al analisis exploratori de les dades, ja que es un dels objectius del informe. On es miren les característiques a nivell general, per finalment entrar en analisis més detallats buscant patrons o agrupaments intrínsecs en les dades, que permetin trobar diferències significatives i rellevants entre els grups de pacients.

4 Resultats

4.1 Observació general dades

En primer lloc es realitza una observació general del contingut del contenidor SE. Per veure quines dades estem tractant.

```
## class: SummarizedExperiment
## dim: 6 77
## metadata(3): description dataType measurementType
## assays(1): counts
## rownames(6): Muscle.loss X1.6.Anhydro.beta.D.glucose ...
##   X2.Hydroxyisobutyrate X2.Oxoglutarate
## rowData names(1): feature
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2
## colData names(2): col.names condition
```

```
## DataFrame with 64 rows and 1 column
##                                         feature
##                                     <character>
## Muscle.loss                          Muscle.loss
## X1.6.Anhydro.beta.D.glucose X1.6.Anhydro.beta.D...
## X1.Methylnicotinamide        X1.Methylnicotinamide
## X2.Aminobutyrate                  X2.Aminobutyrate
## X2.Hydroxyisobutyrate        X2.Hydroxyisobutyrate
## ...                                            ...
## cis.Aconitate                        cis.Aconitate
## myo.Inositol                          myo.Inositol
## trans.Aconitate                    trans.Aconitate
## pi.Methylhistidine              pi.Methylhistidine
## tau.Methylhistidine            tau.Methylhistidine
```

```
## DataFrame with 77 rows and 2 columns
##                 col.names condition
##                   <factor>  <factor>
## PIF_178       PIF_178      cachexic
## PIF_087       PIF_087      cachexic
## PIF_090       PIF_090      cachexic
## NETL_005_V1   NETL_005_V1  cachexic
## PIF_115       PIF_115      cachexic
## ...                   ...       ...
## NETCR_019_V2 NETCR_019_V2   control
## NETL_012_V1  NETL_012_V1    control
## NETL_012_V2  NETL_012_V2    control
## NETL_003_V1  NETL_003_V1    control
## NETL_003_V2  NETL_003_V2    control
```

```
## $description
## [1] "Dades metabolòmiques d'orina de pacients amb càncer control i amb caquèxica"
##
## $dataType
## [1] "Metabolomica"
##
```

```
## $measurementType
## [1] "Concentració en unitats arbitraries (U.A)"
```

Veiem com estem tractant amb un conjunt de dades que conté observacions de pacients amb cànceramb caquèxia i pacients controls. Cada pacient té un identificador únic. Per a cada pacient s'han pres mostres d'orina i s'ha analitzat la concentració de diferents metabòlits en unitats arbitraries (U.A).

Seguim amb l'ànalisis del conjunt de dades.

S'evaluen les dimensions de les dades d'estudi, per veure l'extensió de les dades que es tracten.

```
## Les dimensions de la base de dades són 64 files i 77 columnes. Les 64 files corresponen a 63
## metabolits analitzats i la condició del malalt (caquèxic o control),i les 77 columnes
## es el nombre de pacients del estudi.
```

4.2 Proporció de pacients en l'estudi

A continuació visualitzarem quants dels pacients que tenim corresponen a pacients control (pacients amb càncer) i quants d'ells pateixen caquèxia. Es rellevant determinar-ho ja que si tenim massa pocs pacients de qualsevol dels dos grups ens pot ser perjudicial. Una baixa mostra de qualsevol dels dos grups faria que els valors obtinguts poguessin no ser representatius del grup al qual pertanyen, fent que els resultats ens portessin a interpertacions erronies.

```
##          Recompte pacients Proporció pacients
## Caquèxia                47               0.61
## Control                 30               0.39
```

Recompte de pacients amb caquèxia i pacients control en l'estudi.

Tal i com podem veure hi ha un major nombre de pacients amb caquèxia, concretament hi ha 17 pacients més. En proporció tenim un 60% de la mostra que correspon a pacients amb caquèxia i el 40% restant són pacients control. Pel que podem considerar que tenim una mostra bastant similar de cada un dels grups.

Com que ambdós grups es troben bastan ben representats podem procedir amb l'exploració de les dades.

4.3 Valors omessos

Un altre analisis interessant es buscar si hi ha cap valor omès al llarg del conjunt de dades. Ja que si no els idenitifiquem, podrien ser considerats pel software d'analisis com a 0, fet que podria alterar significativament els valors obtinguts en les variables afectades.

```
## El nombre de valors omessos es 0
```

Per sort tenim un conjunt de dades que no conté valors omessos per tant podem procedir amb l'exploració sense necessitat de tractar les dades.

4.4 Analisis de possibles patrons

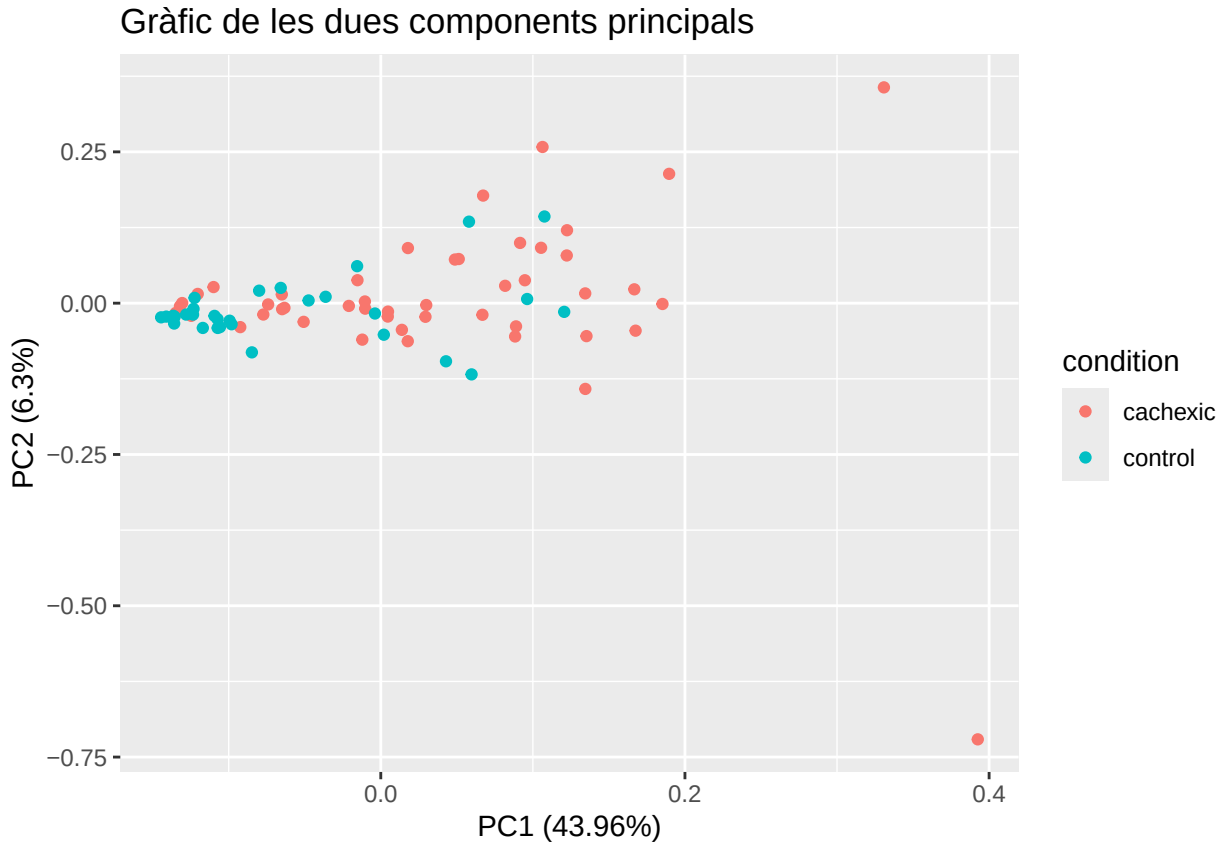
Farem un estudi de components principals (PCA), per a tal d'observar si hi ha cap patró que ens permeti estudiar la major part de la variabilitat del conjunt de dades.

```
##                             PC1       PC2       PC3       PC4       PC5
## Standard deviation     0.9327665 0.3530814 0.3188944 0.3116965 0.294107
## Proportion of Variance 0.4396200 0.0629900 0.0513800 0.0490900 0.043710
## Cumulative Proportion  0.4396200 0.5026200 0.5540000 0.6030900 0.646800
##                             PC6       PC7       PC8       PC9      PC10
## Standard deviation     0.2706839 0.2643263 0.2327817 0.2231789 0.2031929
## Proportion of Variance 0.0370200 0.0353000 0.0273800 0.0251700 0.0208600
## Cumulative Proportion  0.6838200 0.7191200 0.7465000 0.7716700 0.7925300
```

Mostra del resum obtingut del PCA on es veu la contribució de cada component principal (PC).

Veient el resum del PCA podem observar com hi ha una component molt rellevant, ja que per si sola explica el 44% de la variancia total de les dades. Si agafem fins la PC7 arribariem a explicar el 72% de la variancia, pel que podriem reduir les dimensions del model dràsticament i relativament mantenir gran part de la informació.

Es fa un gràfic amb les dues components principals per observar com separen les mostres segons la condició del pacient.

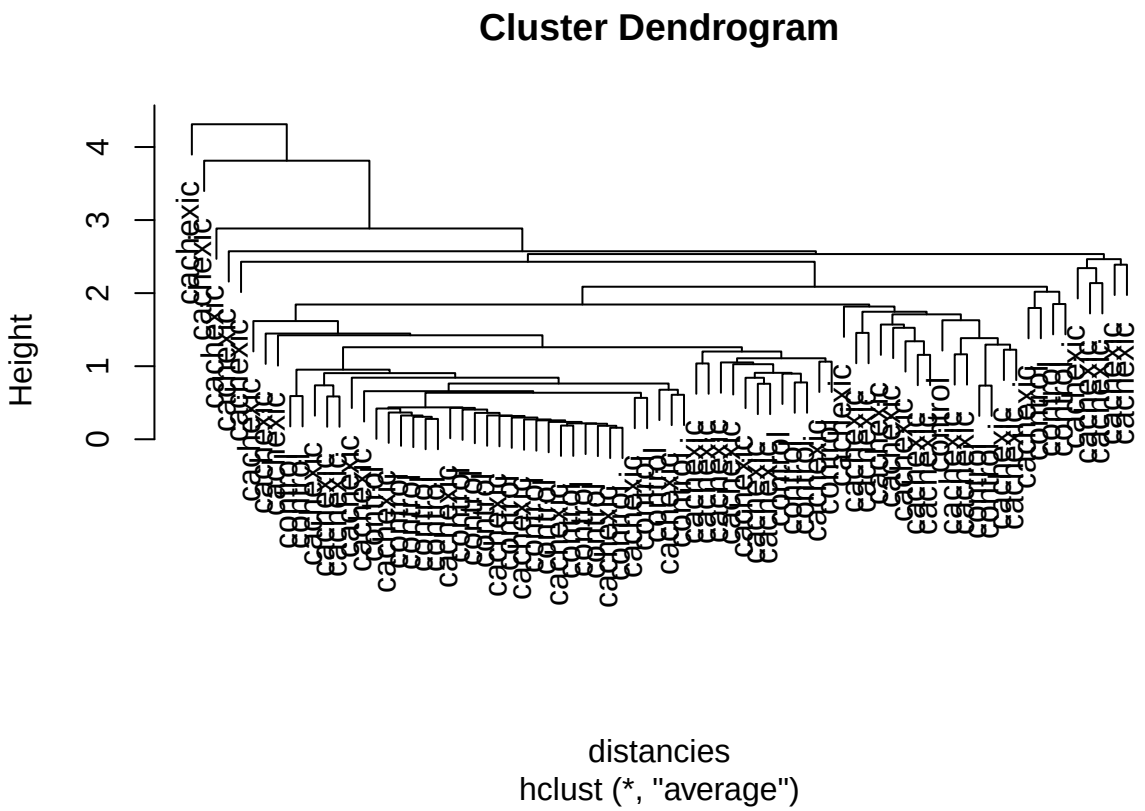


Gràfic que representa com es separen les dades segons si són pacients caquèxics o control gràcies a la variancia explicada per les dues components principals.

En el gràfic, tot i que veiem que hi ha moltes mostres barrejades, s'observa una certa tendencia a que les mostres de pacients caquèxics es desplaçen cap a valors positius de la PC1. Fet que encaixa amb el fet que aquesta variable explica practicament la meitat de la variancia del conjunt de dades. Tot i això nomes estem representant el 50% de la variancia, per això hi ha moltes mostres que no es troben del tot separades, fent que la separació encara no sigui optima.

4.5 Evaluació de possibles agrupaments

Realitzarem un dendograma per a veure si podem observar cap cluster o patró que ens permeti identificar alguna relació entre determinades mostres que ens permetin diferenciar clarament entre els dos grups estudiats.

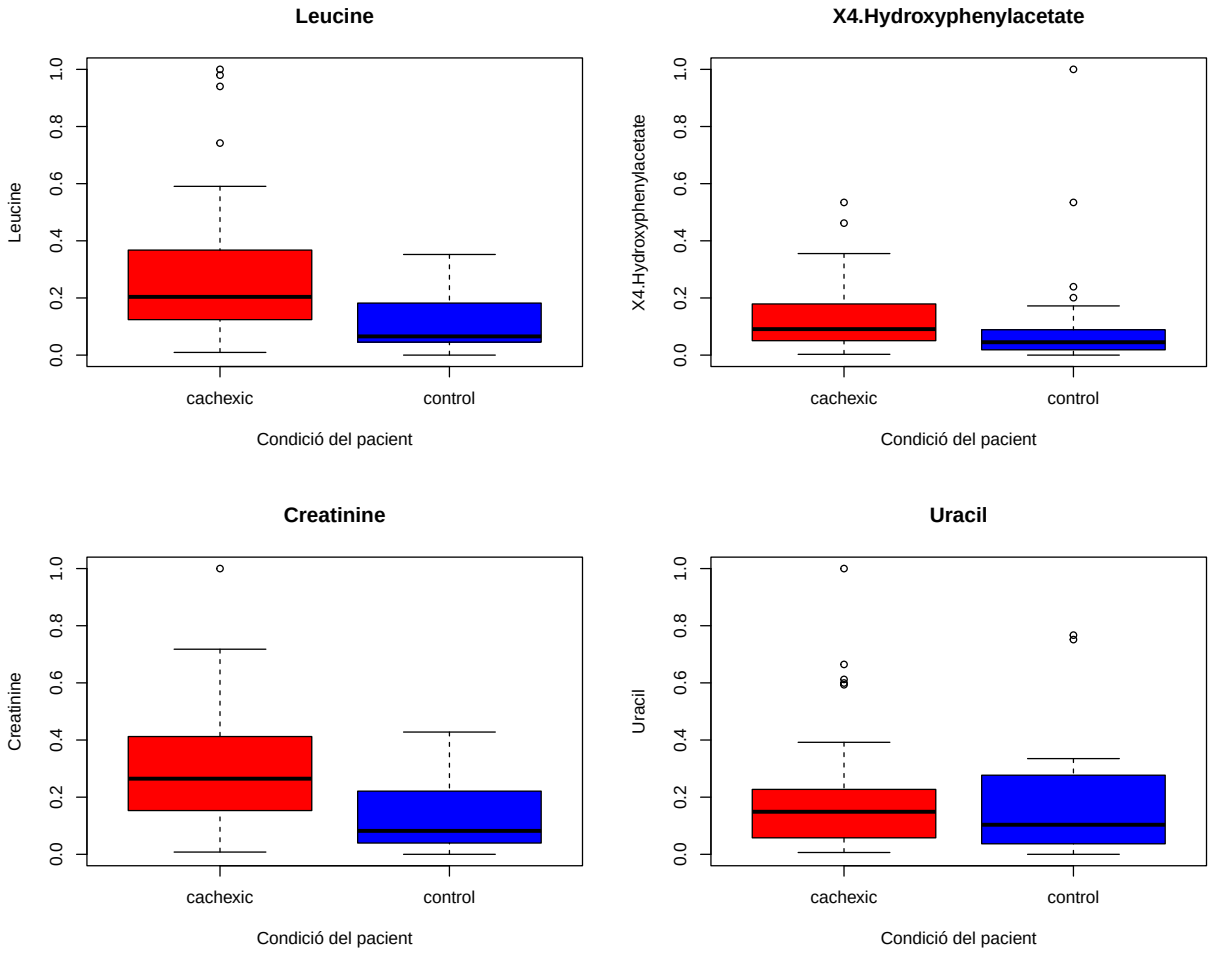


Representació del dendograma on veiem cada observació segons la condició del pacient.

Veient el dendograma veiem com en molts punts tenim tant observacions de pacients amb caquèxia com pacients control, fet que no ens permet observar cap patró clar. Per desgràcia el dendograma no ens es de massa ajuda per a poder observar si es forma cap cluster amb les dades, pel que no podem extreure'n cap relació.

4.6 Comparació concentracions de metabolits segons la condició del pacient

Veient el conjunt de dades que tenim, sembla interessant mirar si entre els dos grups hi ha cap variable que presenti uns valors d'expressió relativament diferencials entre ells. De manera que ens pogués ser d'ajuda per poder usar aquelles variables com a possibles biomarcadors pel diagnostic de la caquèxia. Ara passarem a mirar els valors de cada variable segons la condició del pacient. Es fa l'analisis amb les dades escalades per facilitar la comprensió. Tal i com podem veure en l'Apèndix en la secció del **Gràfic de caixes segons cada condició,** hi ha algunes variables que si semblen presentar diferencies entre cada condició. A continuació es seleccionen alguns dels gràfics més representatius.



Comparació dels nivells de Leucina, Hydroxypenylacetat, Creatinina i Uracil en pacients amb i sense caquèxia.

Hem seleccionat 4 gràfics que mostren diferents situacions. En el cas de Leucina i Creatinina, s'observa una gran diferencia en els valors de concentració dels metabolits en pacients caquexics i control, essent superior en els pacients caquexics en ambdós cassos. Pel que fa a Hydroxypenylacetat, la diferencia no es tant important, pel que no podem afirmar que hi ha diferencies clares entre els dos grups. I finalment pel que fa a Uracil, veiem com els valors són similars.

Es seleccionen aquests gràfics per criteris educatius, seleccionant dues variables que presentin grans diferencies a nivell visual, una altra que la diferencia sigui menor, pero tot i això apreciable, i una que aparentment no hi ha diferencies. Aixi a posteriori, podem comparar les mitjanes i fer un estudi estadístic amb dos cassos diferents.

Es calcula la mitjana sense escalar de la concentració dels metabòlits destacats, per veure fins a quin punt hi ha tanta diferencia.

```
## La concentració mitjana de Leucina en els pacients controls es de 13.55667 U.A.
## I en els pacients caquexics és 31.2617 U.A.
```

```
## La concentració mitjana de Hydroxypenylacetat en els pacients controls es de 99.79867 U.A.
## I en els pacients caquexics és 119.8226 U.A.
```

```
## La concentració mitjana de Creatinina en els pacients controls es de 5619.175 U.A.
## I en els pacients caquexics és 10722.14 U.A.
```

```
## La concentració mitjana d'Uracil en els pacients controls es de 32.49333 U.A.
## I en els pacients caquexics és 37.51362 U.A.
```

Veiem com en el cas de la Leucina i la Creatinina les diferencies són importants, en canvi en Hydroxypenylacetat i Uracil les diferencies són menors.

Mirem-ho en percentatge per tenir una idea més clara de les diferencies entre elles.

```
## La concentració de Leucina es un 130.6 % superior en pacients amb caquèxia
## en comparació al control.
```

```
## La concentració de Hydroxypenylacetat es un 20.06 % superior en
## pacients amb  caquèxia en comparació al control.
```

```
## La concentració de Creatinina es un 90.81 % superior en pacients amb caquèxia
## en comparació al control.
```

```
## La concentració d'Uracil es un 15.45 % superior en pacients amb  caquèxia
## en comparació al control.
```

Veiem com la concentració de Leucina i Creatinina en pacients amb caquèxia es molt més gran que la que presenten els pacients control en orina. Essent el doble pel que fa a la Leucina i gairebè el doble en Creatinina. Pel que fa a Hydroxypenyalcetat i Uracil la diferencia es menor, en els dos casos la diferencia es menor al 20%.

Aquesta valoració s'ha de prendre amb pinces ja que no disposem de dades estadístiques que ho recolzin. Pel que es realitza un breu estudi estadístic per donar més pes al que hem vist.

Es realitza un analisis ANOVA que ens permet comparar si els nivells de concentració dels dos metàbolits estudiats són significativament diferents segons la condició.

```
## El p_value obtingut per la Leucina es 3e-04 .
```

```
## El p_value obtingut per la Hydroxypenyalcetat es 0.4818 .
```

```
## El p_value obtingut per la Creatinina es 5e-04 .
```

```
## El p_value obtingut per la Uracil es 0.5429 .
```

Com podem veure el breu estudi que hem fet ens permet observar com entre les dues condicions si que hi ha diferencies significatives pel que fa a la concentració de Leucina i Creatinina, ja que ambdós p-values son signifiactivament menors a 0.05, pel que considerem que la diferencia de mitjanes es estadísitcament significant al 0.05 de significancia. En canvi pel que fa a Hydroxypenyalcetat i Uracil, no hi ha prou evidencia per a poder acceptar aquesta hipotesis. Per tant no podem considerar que les diferencies vistes siguin estadísticament significatives.

5 Discussió

Tal i com hem vist al llarg del informe hem pogut realitzar una exploració de les dades. Aquest anàlisis ens ha permés identificar dos punts rellevants.

Gràcies al Analisis de possibles patrons hem pogut observar com tot i que el model conté moltes variables i per tant moltes components principals, gairebe la meitat de la variancia de les dades s'explica per una sola component principals. Tot i això la resta de components aporten molt poc en comparació, ja que només hi ha 4 components principals més que expliquin més del 5% de variancia, fet que fa que necessitem moltes components principals per a poder representar un valor de variancia raonable. Per poder explicar el 80% de la variancia necessitem fins a la PC10, i fins a la PC18 no expliquem el 90% de la variancia. Encara que la seva aportació pugui ser infíma, poden contenir informació vital, pel que no s'han de menysprear. Això ens porta a que si seleccionem només les dues primeres components principals, el model que tenim no permeti realitzar una separació massa bona de les dades segons la condició de cada mostra, pel que si es pretengués fer una reducció de la dimensionalitat de les dades s'hauria d'anar amb cautela, ja que una reducció massa grossa ens faria perdre informació rellevant.

Amb els resultats de Comparació concentracions de metabolits segons la condició del pacient, hem pogut identificar dos metabòlits que semblen presentar concentracions estadísticament diferents en pacients caquèxics i control. COncretament hem vist com la Leucina i Creatinina, presenten valors superiors en pacient caquèxics en comparació al control. Això ho hem verificat per un test estadístic. Aquest fet es rellevant, ja que aquesta troballa podria significar que aquests dos metàbolits possiblement podrien formar part d'un panell de biomarcadors útils per a identificar pacients amb caquèxia. Això també va ser observat per [3] i [4] que en ambdós estudis es va trobar que aquests metàbolits presentaven perfils de concentració en orina diferents segons si el pacient era caquèxic o control. Suggerint que poden ajudar a crear models de diganostic per la caquèxia.

Es important remarcar que la mida del estudi potser no es suficientment gran, ja que en el nostre cas i inclós en [3] i [4], cap estudi supera les 100 mostres. Pel que cal prendre les afirmacions que s'han fet amb cura, ja que possiblement necessitem un major volum de mostra per a poder establir biomarcadors útils que siguin estadísticament rellevants. Però tot i això, la troballa d'aquests marcadors diferencials en diferents estudis pot ser indicador que cal seguir analitzant perfils metabolomics en pacients caquèxics, que ens permetin augmentar el volum de mostres i poguem reafirmar aquestes observacions.

6 Bibliografia

[1] Setiawan, T., Sari, I. N., Wijaya, Y. T., Julianto, N. M., Muhammad, J. A., Lee, H., Chae, J. H., & Kwon, H. Y. (2023). Cancer cachexia: molecular mechanisms and treatment strategies. Journal of hematology & oncology, 16(1), 54. https://doi.org/10.1186/s13045-023-01454-0

[2] Peixoto da Silva, S., Santos, J. M. O., Costa E Silva, M. P., Gil da Costa, R. M., & Medeiros, R. (2020). Cancer cachexia and its pathophysiology: links with sarcopenia, anorexia and asthenia. Journal of cachexia, sarcopenia and muscle, 11(3), 619–635. https://doi.org/10.1002/jcsm.12528

[3]Yang, Q. J., Zhao, J. R., Hao, J., Li, B., Huo, Y., Han, Y. L., Wan, L. L., Li, J., Huang, J., Lu, J., Yang, G. J., & Guo, C. (2018). Serum and urine metabolomics study reveals a distinct diagnostic model for cancer cachexia. Journal of cachexia, sarcopenia and muscle, 9(1), 71–85. https://doi.org/10.1002/jcsm.12246

[4]Eisner, R., Stretch, C., Eastman, T.B., Xia, J., Hau, D.D., Damaraju, S., Greiner, R., Wishart, D.S., & Baracos, V.E. (2011). Learning to predict cancer-associated skeletal muscle wasting from 1H-NMR profiles of urinary metabolites. Metabolomics, 7, 25-34. doi:10.1007/s11306-010-0232-9

7 Repositori GitHub

Link al repositori de GitHub: https://github.com/Garroww/Casanovas-Sanglas-Guillem-PEC1.git

8 Apendix

A continuació s'inclouen tots aquells codis que s'han anat executant al llarg del informe però que no han aparegut. Així com també versions completes de resultats que s'han mostrat parcialment. També s'inclou altres fragments de codi usats que han sigut utils. Els fragments de codi que són identics als que s'han usat en l'apartat de Resultats es mostren sense executar per estalviar espai. Cada fragment de codi incloul'explicació dels passos rellevants, així com també algun comentari no rellevant que no s'ha afegit en l'informe.

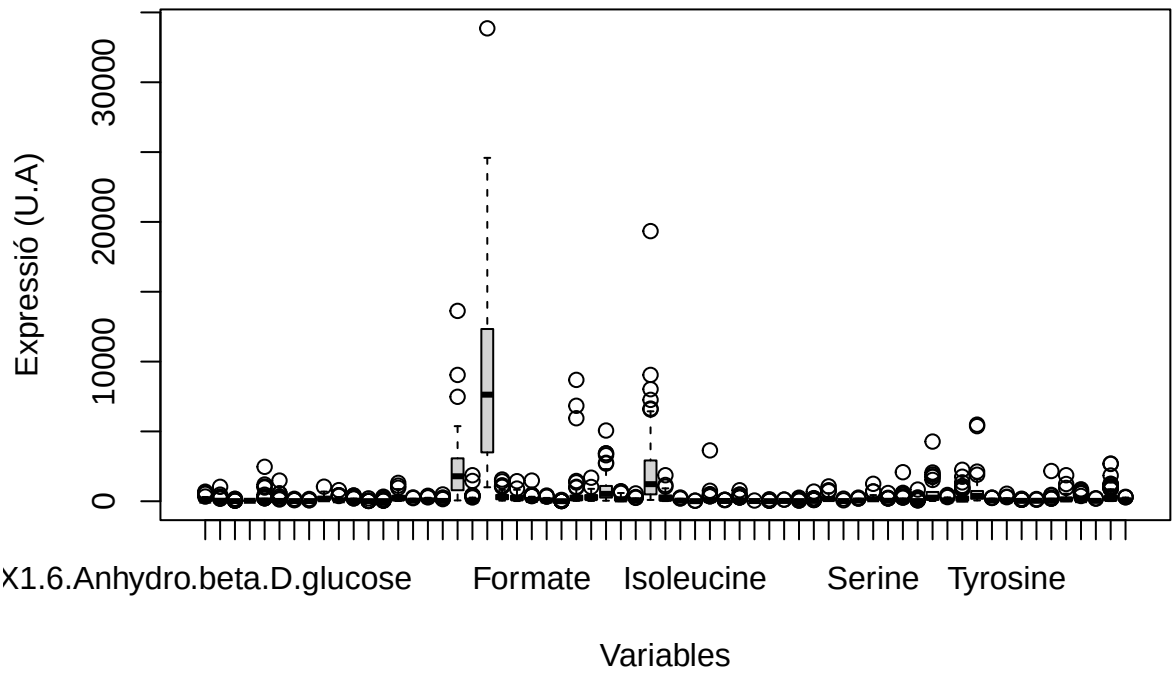
8.1 Carregar les dades de treball

```
#Carreguem el fitxer .csv
dades<-read.csv("human_cachexia.csv",row.names = 1)
```

8.2 Comprovació rang de valors de les variables

Observem com efectivament hi ha molta diferencia en el rang en determinades variables, tot i que en general la inmensa majoria de variables prenen valors similars. Pel que necessitarem escalar les dades, per fer-ho mes entendible i interpretable.

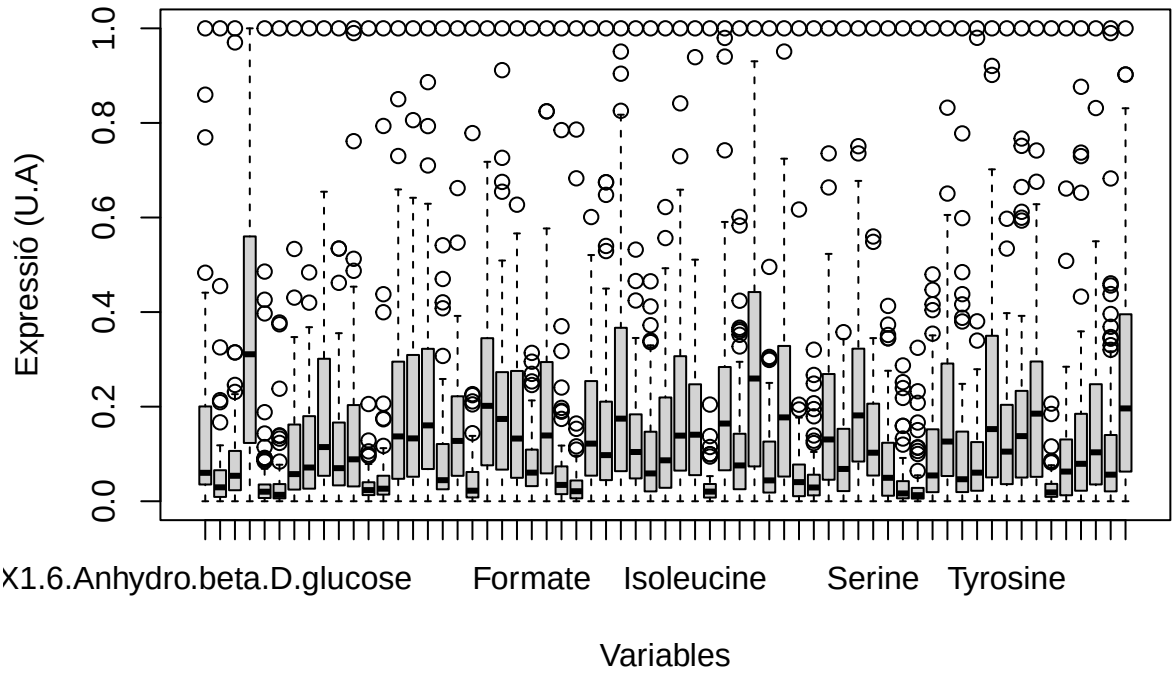
```
#Separem les dades numeriques de la resta
dades_numeriques<- dades[,2:ncol(dades)]
#FEm un grafic de caixes per veure els valors que pren cada variable, aixi podem
#observar si tenim cap variable que té un rang de valors molt diferent al de la resta de dades.
boxplot(dades_numeriques, xlab="Variables", ylab="Expressió (U.A)")
```



8.3 Escalatge de dades

Vist que el rang de algunes variables es molt variat, l'escalarem per a poder treballar més comodament. Aquest pas es realitza previament a la creació del contenidor SummarizedExperiment.

```
#Escalarem les dades per tal que siguin més comparables i entendibles. Les
#escalarem en un rang de valors de 0 a 1.
escala <- function(x) {
    (x - min(x)) / (max(x) - min(x))
}
dades_numeriques_escalades <- apply(dades_numeriques, 2, escala)
#Abans però guardem en un objecte nou la columna de la condició
dades_no_numeriques <- dades[, 1, drop = FALSE]
#Ara ho guardem en un nou objecte juntament amb la condició de cada pacient
dades_escalades<- cbind(dades_no_numeriques, dades_numeriques_escalades)
dades_escalades<-as.data.frame(dades_escalades)

#Repetim de nou el boxplot per verificar
boxplot(dades_numeriques_escalades,xlab="Variables", ylab="Expressió (U.A)")
```



8.4 Creació contenidor Summarized Experiment

Es troba en el document específic "Summarized Experiment.rmd". El contenidor creat ja conté les dades escalades, que s'han obtingut en l'apartat Escalatge de dades

```
#Carreguem el contenidor per a poder-hi treballar
load(file = "SummarizedExperiment.rda")
```

8.5 Observació general dades

En primer lloc es realitza una observació general del contingut del contenidor SE. Per veure quines dades estem tractant

```
#Seleccionem els components principals del objecte SE per fer-nos una idea del contingut d'aquest.
head(SE)
rowData(SE)
colData(SE)
metadata(SE)
```

```
cat("Les dimensions de la base de dades són", dim(assay(SE))[1],"files i",
    dim(assay(SE))[2], "columnes. Les 64 files corresponen a 63
    metabolits analitzats i la condició del malalt (caquèxic o control),
    i les 77 columnes es el nombre de pacients del estudi.\n")
```

8.6 Proporció de pacients en l'estudi

```
#En primer lloc calculem el nombre de pacients segons la condició
Numero_pacients<-table(colData(SE)$condition)
#Fem el mateix en proporció
Percentatge_pacients<-prop.table(Numero_pacients)
#Preparem les columnes i files de la taula que farem
Columna_taula<-c("Recompte pacients", "Proporció pacients")
Fila_taula<-c("Caquèxia", "Control")
#Creem la taula amb les dades, ajustant-les segons la seva naturalesa
Taula_recompte<-data.frame(Numero=as.integer(Numero_pacients),
                           Percentatge= round(as.numeric(Percentatge_pacients),2))
#Afegim els noms de les files i les columnes
colnames(Taula_recompte)<-Columna_taula
rownames(Taula_recompte)<-Fila_taula
print(Taula_recompte)
```

8.7 Valors omessos

```
#Farem un sumatori del nombre de valors omesos (NA), en cas que en trobem
#procedirem a identificar-los.
cat("El nombre de valors omessos es",sum(is.na.data.frame(dades_expressio)),
    "\n")
```

8.8 Analisis de components principals (PCA)

```
#Ara podem fer el PCA, abans pero cal transposar les dades ja que el PCA vol
#les variables en columnes i observacions en files
PCA<-prcomp(t(dades_expressio))
#Representarem nomes una part de les dades del resum,ja que sino es massa llarg
resum_pca<-summary(PCA)
#Mirem les components
resum_pca
```

```
## Importance of components:
##                           PC1     PC2     PC3     PC4     PC5     PC6    PC7
## Standard deviation     0.9328 0.35308 0.31889 0.31170 0.29411 0.27068 0.2643
## Proportion of Variance 0.4396 0.06299 0.05138 0.04909 0.04371 0.03702 0.0353
## Cumulative Proportion  0.4396 0.50262 0.55400 0.60309 0.64680 0.68382 0.7191
##                           PC8     PC9    PC10    PC11    PC12    PC13    PC14
## Standard deviation     0.23278 0.22318 0.20319 0.19417 0.18137 0.17948 0.17384
## Proportion of Variance 0.02738 0.02517 0.02086 0.01905 0.01662 0.01628 0.01527
## Cumulative Proportion  0.74650 0.77167 0.79253 0.81158 0.82820 0.84448 0.85975
##                           PC15    PC16    PC17    PC18    PC19    PC20   PC21
## Standard deviation     0.16298 0.16193 0.15304 0.14877 0.14008 0.13336 0.1274
## Proportion of Variance 0.01342 0.01325 0.01183 0.01118 0.00991 0.00899 0.0082
## Cumulative Proportion  0.87317 0.88642 0.89825 0.90944 0.91935 0.92834 0.9365
##                           PC22    PC23    PC24    PC25    PC26    PC27    PC28
```

```
## Standard deviation     0.12448 0.11534 0.10878 0.10157 0.09515 0.09143 0.08534
## Proportion of Variance 0.00783 0.00672 0.00598 0.00521 0.00457 0.00422 0.00368
## Cumulative Proportion  0.94437 0.95109 0.95707 0.96228 0.96685 0.97108 0.97476
##                            PC29    PC30    PC31    PC32    PC33    PC34    PC35
## Standard deviation     0.07638 0.07502 0.06760 0.06621 0.05943 0.05773 0.05520
## Proportion of Variance 0.00295 0.00284 0.00231 0.00222 0.00178 0.00168 0.00154
## Cumulative Proportion  0.97770 0.98055 0.98286 0.98507 0.98686 0.98854 0.99008
##                            PC36    PC37    PC38    PC39    PC40    PC41    PC42
## Standard deviation     0.04982 0.04752 0.04647 0.04231 0.04057 0.03620 0.03377
## Proportion of Variance 0.00125 0.00114 0.00109 0.00090 0.00083 0.00066 0.00058
## Cumulative Proportion  0.99134 0.99248 0.99357 0.99447 0.99530 0.99597 0.99654
##                            PC43    PC44    PC45    PC46    PC47    PC48    PC49
## Standard deviation     0.03019 0.02951 0.02711 0.02696 0.02534 0.02374 0.02142
## Proportion of Variance 0.00046 0.00044 0.00037 0.00037 0.00032 0.00028 0.00023
## Cumulative Proportion  0.99700 0.99744 0.99781 0.99818 0.99851 0.99879 0.99902
##                            PC50    PC51    PC52    PC53    PC54    PC55    PC56
## Standard deviation     0.01887 0.01715 0.01598 0.01475 0.01424 0.01256 0.01129
## Proportion of Variance 0.00018 0.00015 0.00013 0.00011 0.00010 0.00008 0.00006
## Cumulative Proportion  0.99920 0.99935 0.99948 0.99959 0.99969 0.99977 0.99984
##                            PC57     PC58     PC59     PC60     PC61     PC62
## Standard deviation     0.009606 0.008874 0.008178 0.006287 0.004606 0.004287
## Proportion of Variance 0.000050 0.000040 0.000030 0.000020 0.000010 0.000010
## Cumulative Proportion  0.999880 0.999920 0.999960 0.999980 0.999990 1.000000
##                            PC63
## Standard deviation     0.002788
## Proportion of Variance 0.000000
## Cumulative Proportion  1.000000
```

```
#Seleccionem nomes les 10 primeres components
resum_pca$importance[,1:10]
```

```
##                             PC1       PC2       PC3       PC4       PC5
## Standard deviation     0.9327665 0.3530814 0.3188944 0.3116965 0.294107
## Proportion of Variance 0.4396200 0.0629900 0.0513800 0.0490900 0.043710
## Cumulative Proportion  0.4396200 0.5026200 0.5540000 0.6030900 0.646800
##                             PC6       PC7       PC8       PC9      PC10
## Standard deviation     0.2706839 0.2643263 0.2327817 0.2231789 0.2031929
## Proportion of Variance 0.0370200 0.0353000 0.0273800 0.0251700 0.0208600
## Cumulative Proportion  0.6838200 0.7191200 0.7465000 0.7716700 0.7925300
```

```
#Carreguem el paquet ggfortify que fa gràfics de PCA d'una manera molt #simple i efectiva
library(ggfortify)
#Fem el gràfic representant les dades de les dues principals components i #que es diferencin cada condi
autoplot(PCA, data=colData(SE), colour = 'condition',
         main="Gràfic de les dues components principals")
```

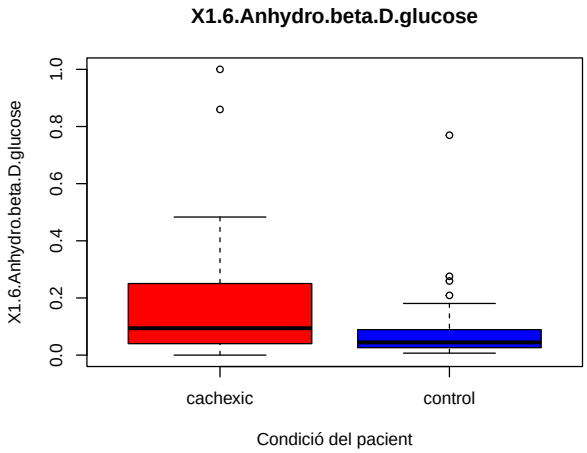
8.9 Dendograma

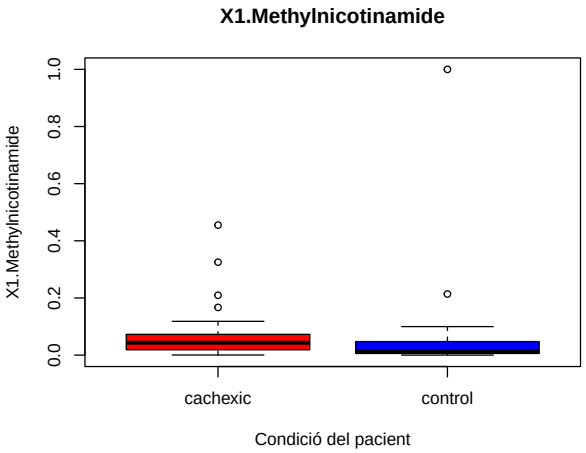
Crearem un dendograma per a veure com es distribuixen les diferents mostres segons la condició del pacient.

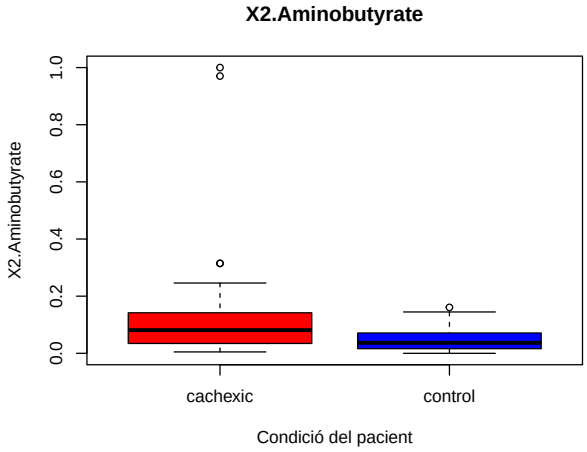
```r
# Calculem la matriu de distancies
distancies <- dist(t(dades_expressio), method = "euclidean")

#Creem l'objecte que crea l'arbre de distancies
cluster <- hclust(distancies, method = "average")
plot(cluster, labels=colData(SE)$condition)
```

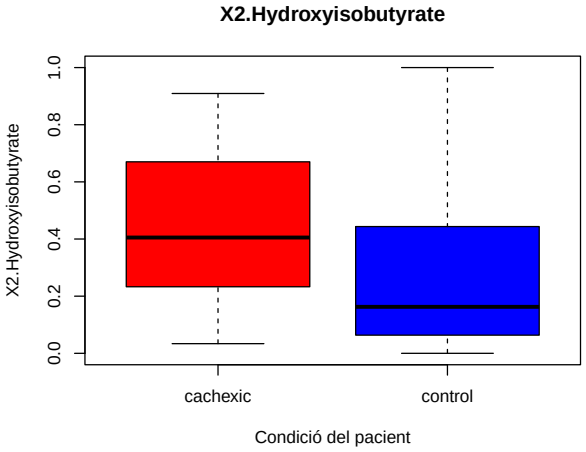
8.10 Gràfic de caixes de cada metàbolit segons la condició del pacient

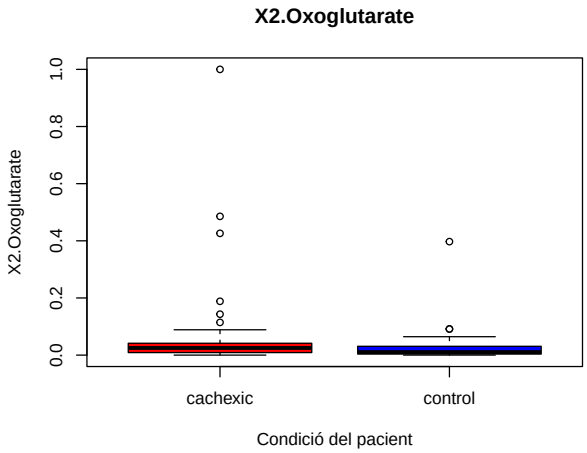
```r
#En primer lloc es guarden els noms de cada fila
noms<-names(SE)

#Eliminem la fila de la condició, ja que no volem fer-ne un gràfic per ella,
#ja que cada gràfic contindrà aquesta variable
noms<-noms[noms!="Muscle.loss"]

#Creem un gràfic de caixa, on compararem els valors que pren cada variable
#segons la condició dels pacients
par(mfrow=c(2,2))
for (i in seq_along(noms)) {
  #Per a fer-ho necessitem que les dades es tractin com a números, ja que sino dona errors. I que es va
  dades_numeriques_2 <- as.numeric(assay(SE, "counts")[noms[i], ])
  #Fem el grafic comparant els valors de cada variable  segons la condicio
  boxplot(dades_numeriques_2 ~ colData(SE)$condition,
          main = noms[i],
          xlab = "Condició del pacient",
          ylab = noms[i],
          col = c("red", "blue"))
}
```

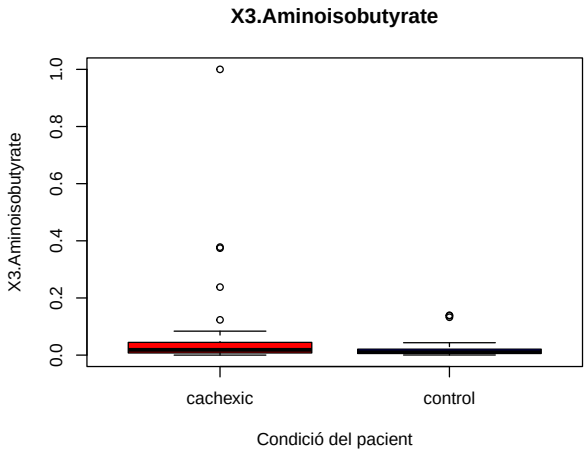

X1.6.Anhydro.beta.D.glucose
X1.6.Anhydro.beta.D.glucose
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

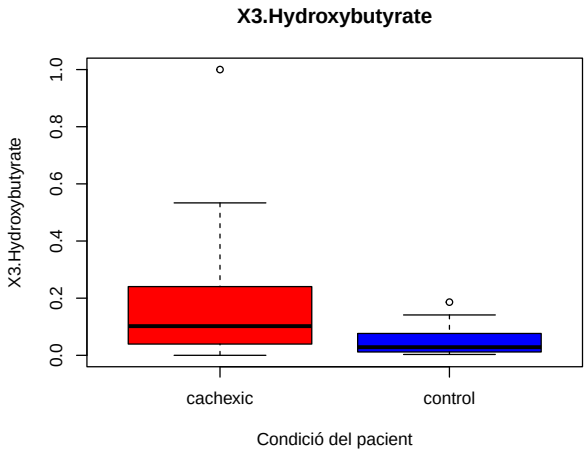

X1.Methylnicotinamide
X1.Methylnicotinamide
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

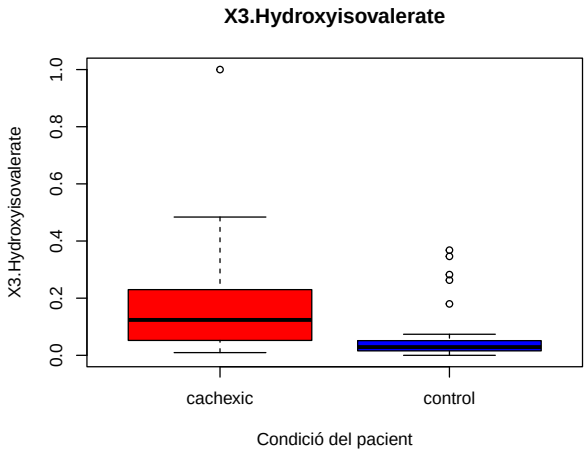

X2.Aminobutyrate
X2.Aminobutyrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

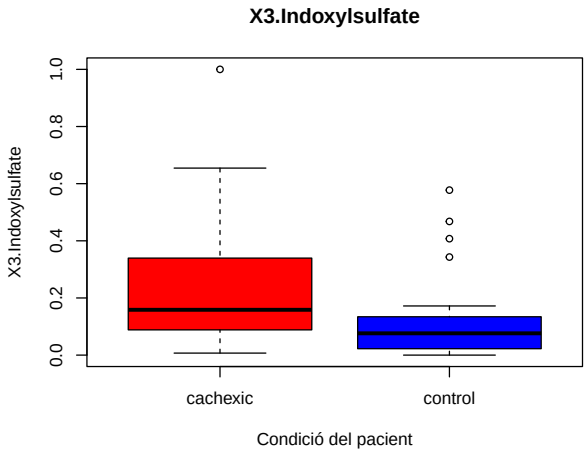

X2.Hydroxyisobutyrate
X2.Hydroxyisobutyrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

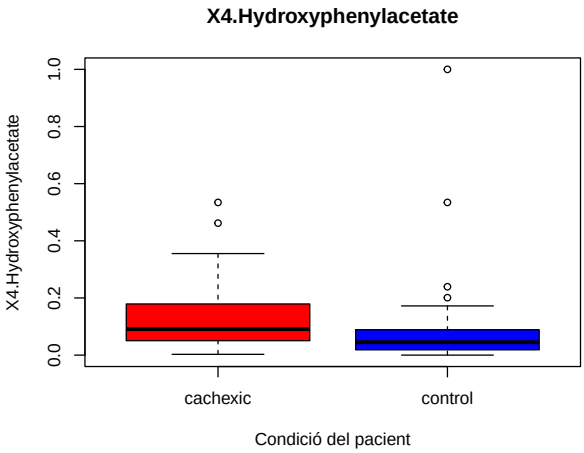

X2.Oxoglutarate
X2.Oxoglutarate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

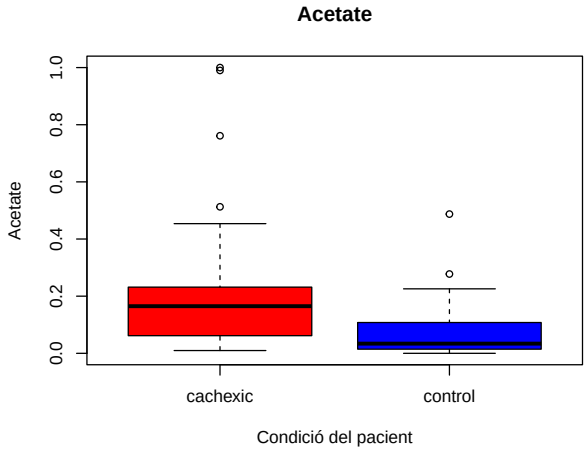

X3.Aminoisobutyrate
X3.Aminoisobutyrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

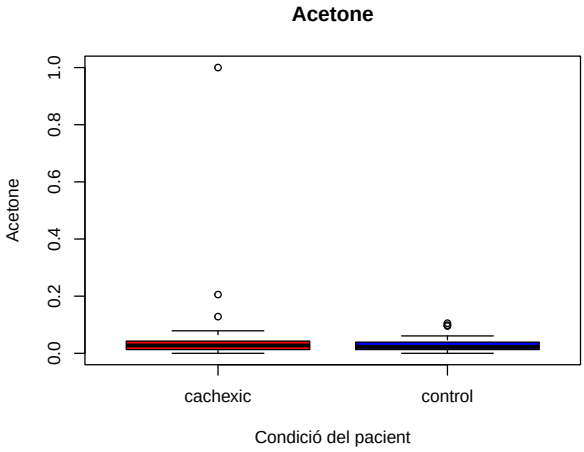

X3.Hydroxybutyrate
X3.Hydroxybutyrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

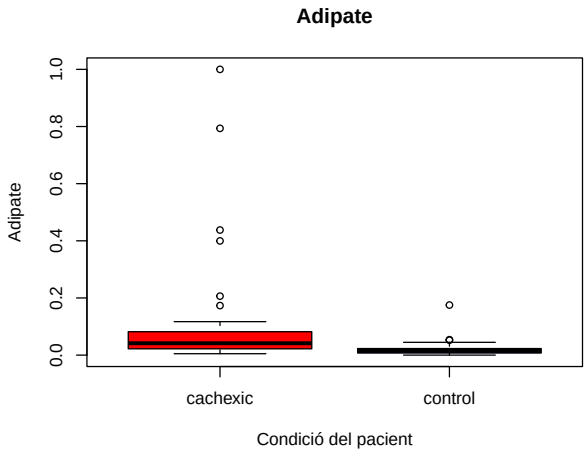

X3.Hydroxyisovalerate
X3.Hydroxyisovalerate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

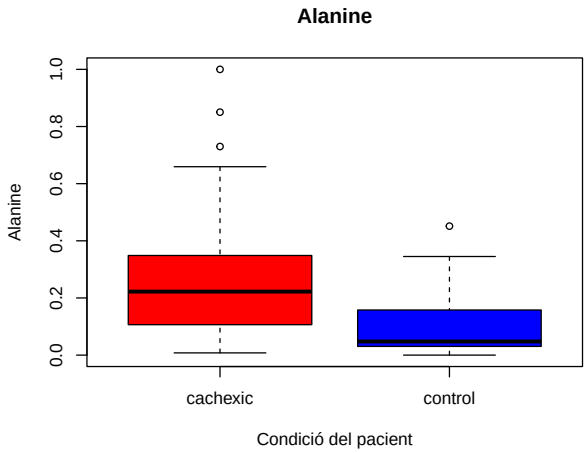

X3.Indoxylsulfate
X3.Indoxylsulfate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

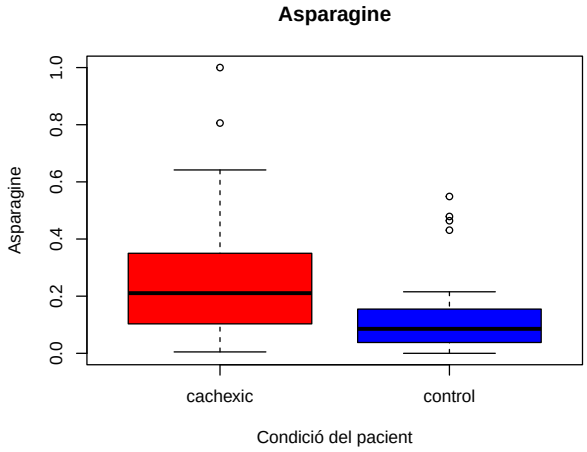

X4.Hydroxyphenylacetate
X4.Hydroxyphenylacetate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

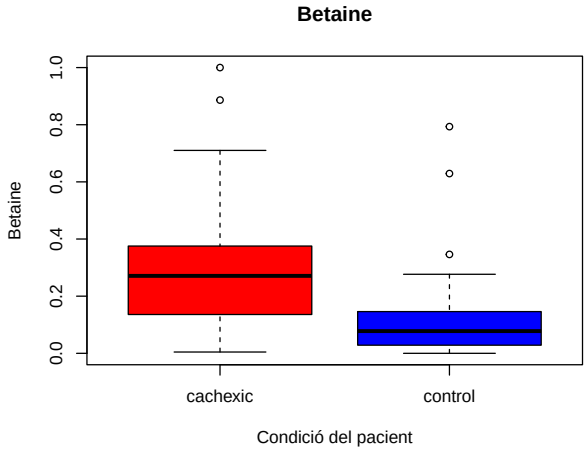

Acetate
Acetate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

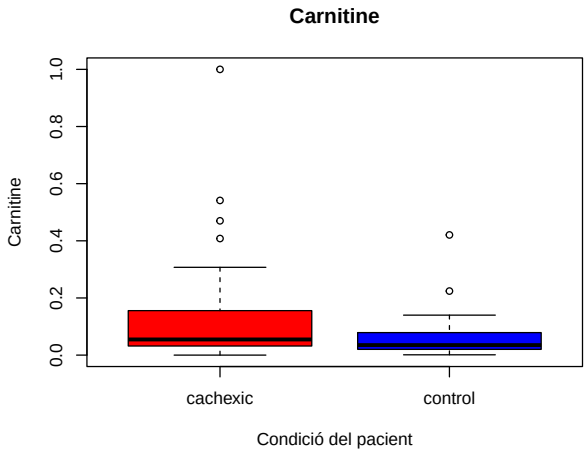

Acetone
Acetone
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

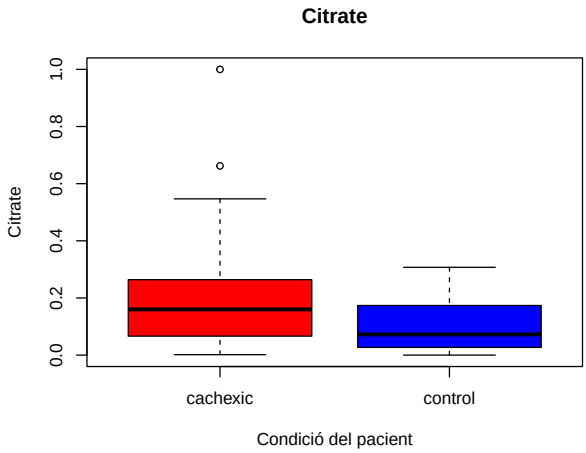

Adipate
Adipate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

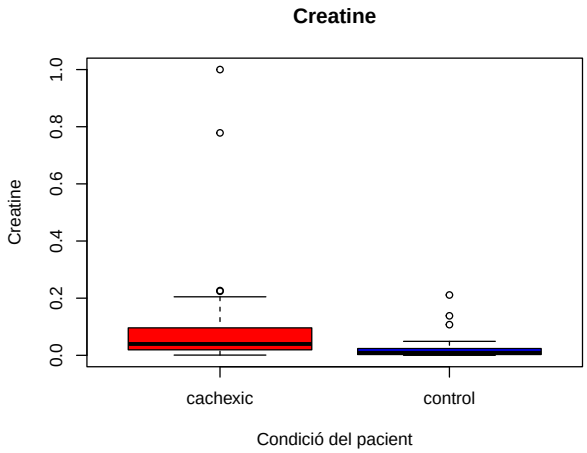

Alanine
Alanine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

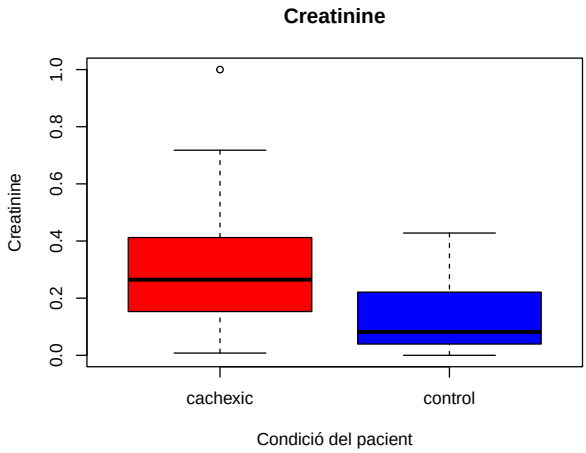

Asparagine
Asparagine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

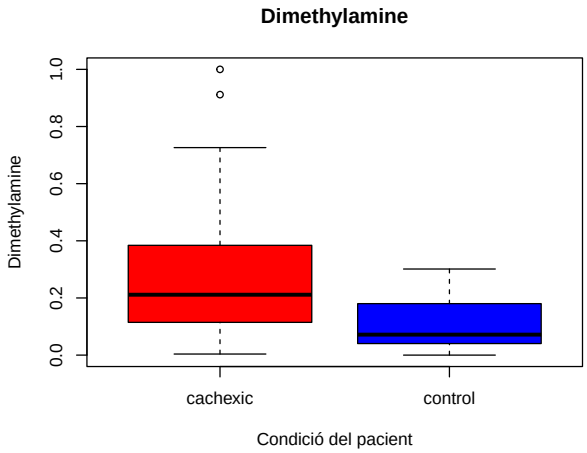

Betaine
Betaine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

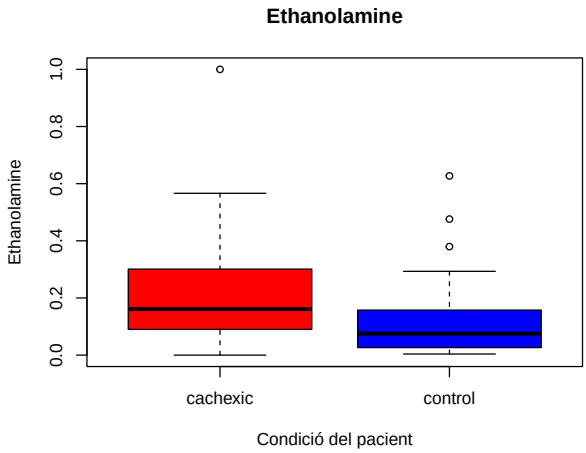

Carnitine
Carnitine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

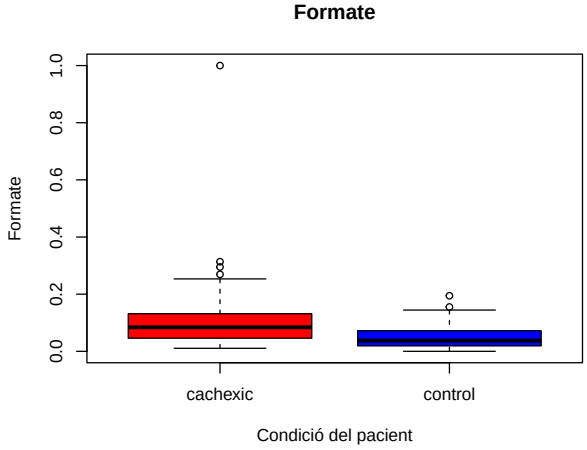

Citrate
Citrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

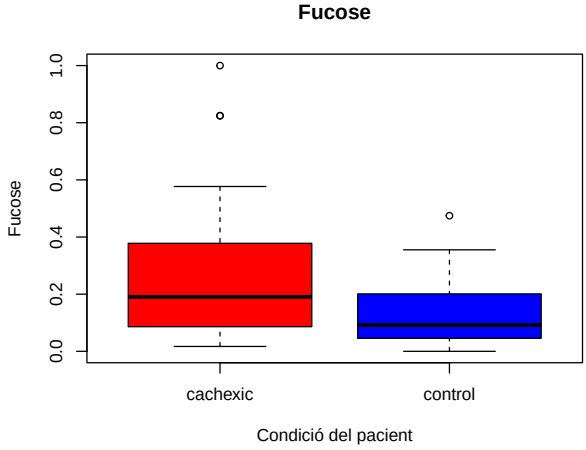

Creatine
Creatine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

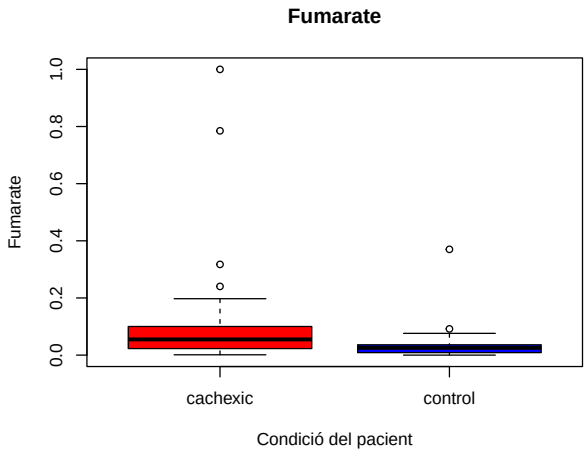

Creatinine
Creatinine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

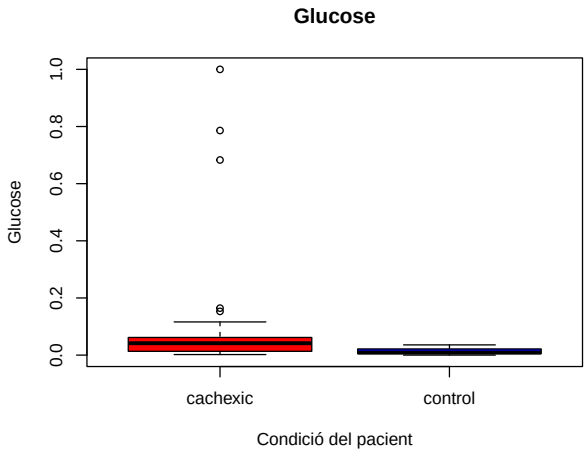

Dimethylamine
Dimethylamine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

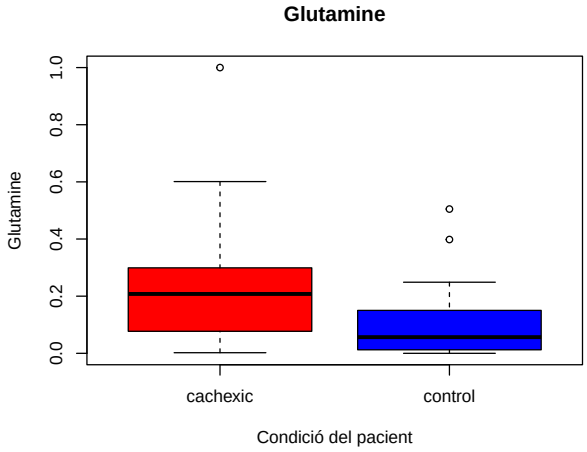

Ethanolamine
Ethanolamine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

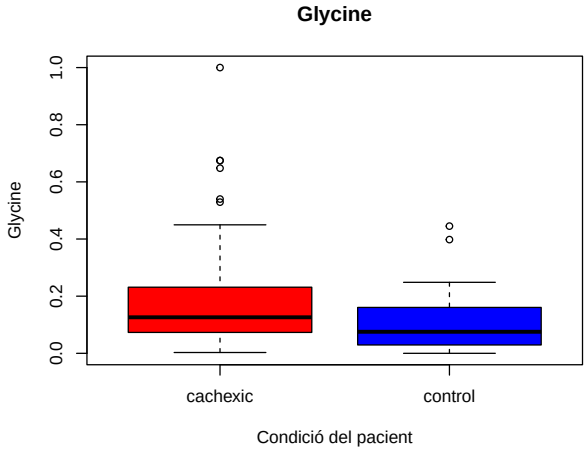

Formate
Formate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

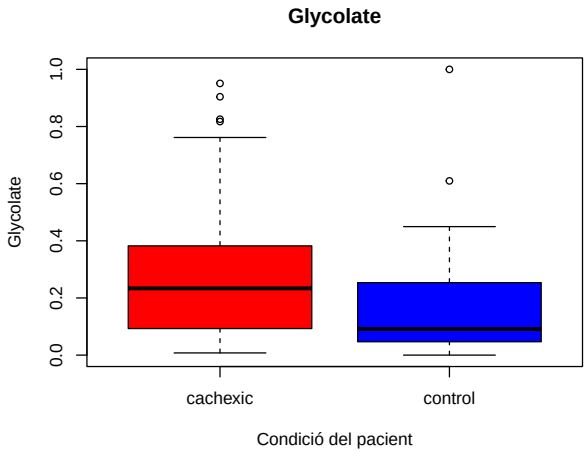

Fucose
Fucose
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

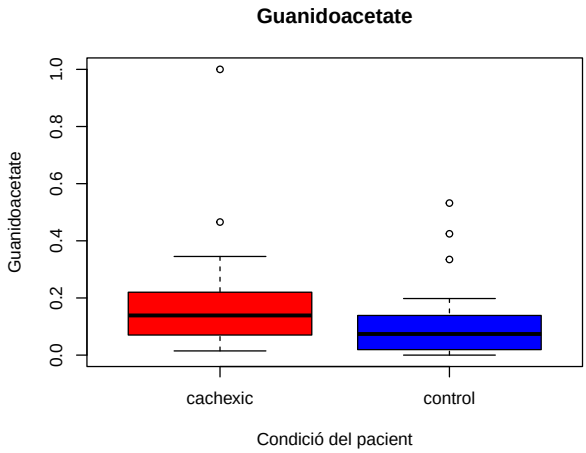

Fumarate
Fumarate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

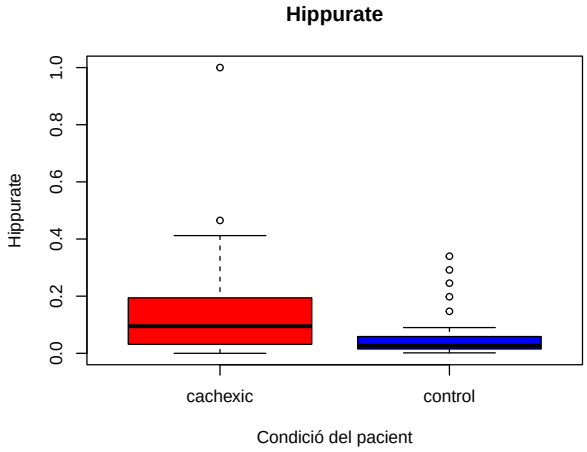

Glucose
Glucose
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

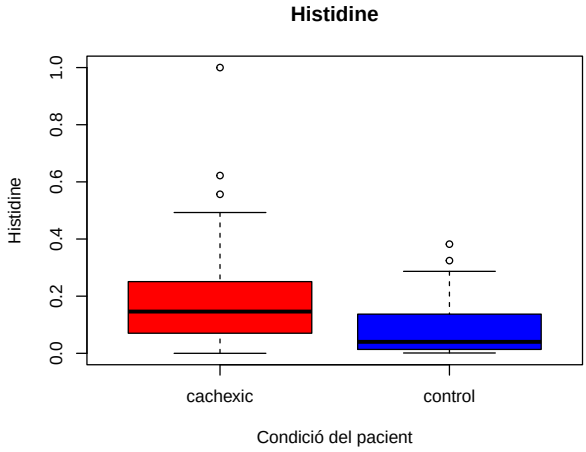

Glutamine
Glutamine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

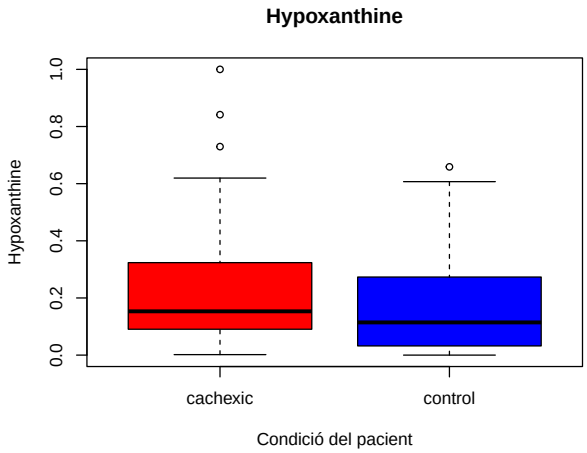

Glycine
Glycine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

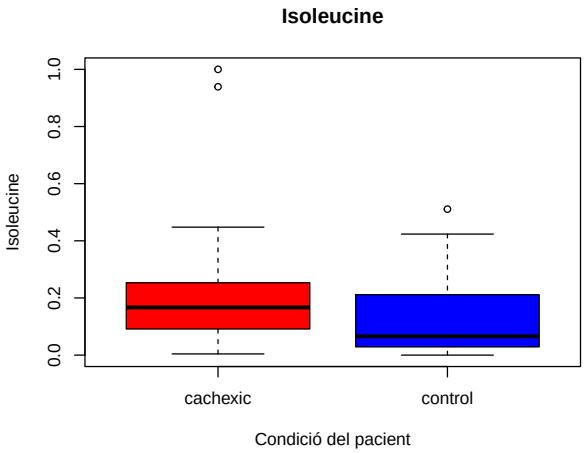

Glycolate
Glycolate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

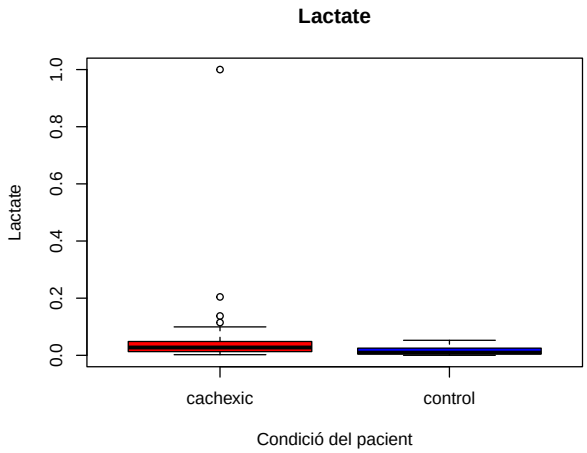

Guanidoacetate
Guanidoacetate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

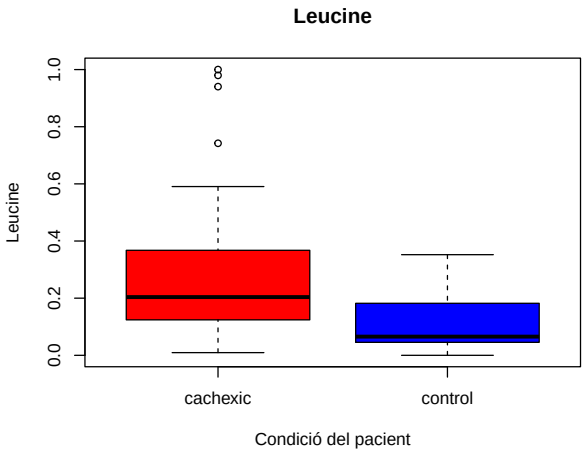

Hippurate
Hippurate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

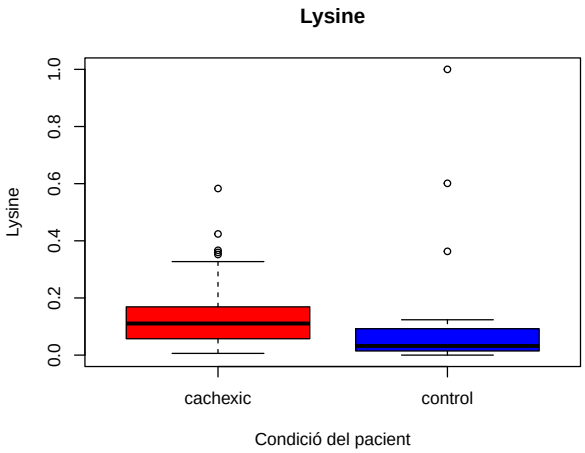

Histidine
Histidine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

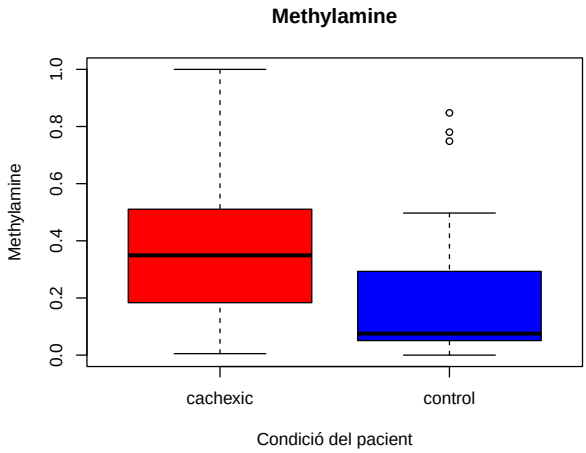

Hypoxanthine
Hypoxanthine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

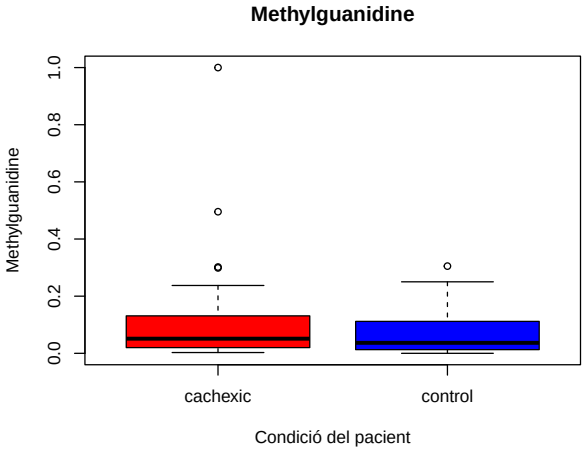

Isoleucine
Isoleucine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

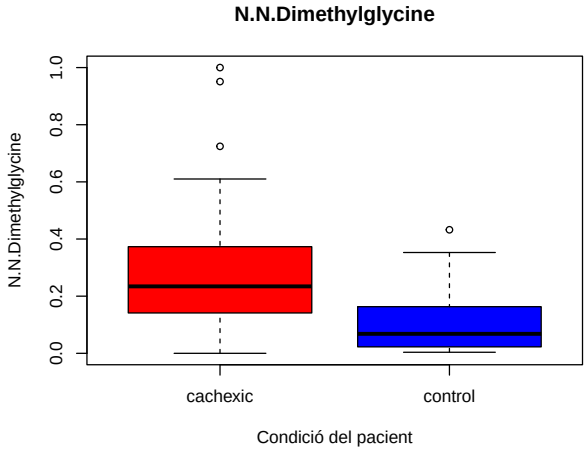

Lactate
Lactate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

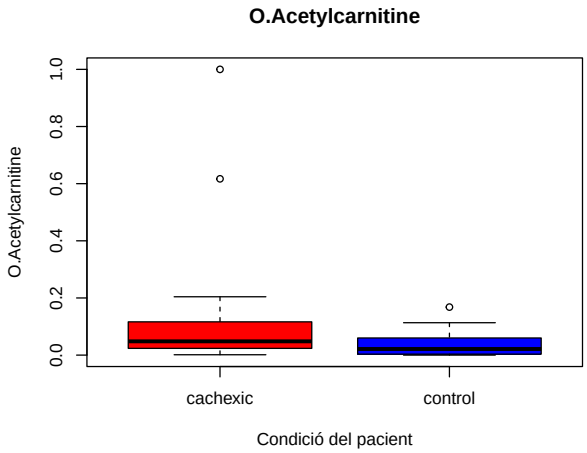

Leucine
Leucine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

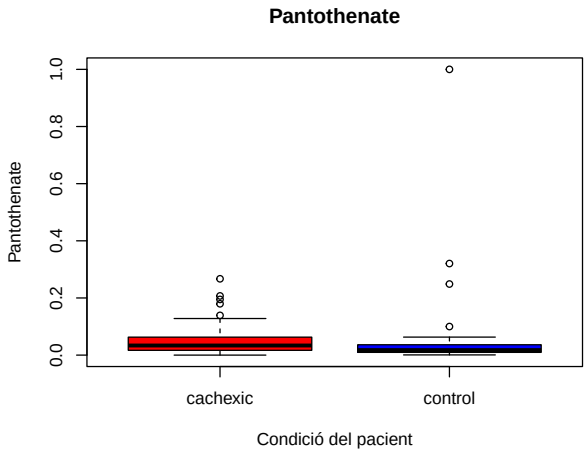

Lysine
Lysine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

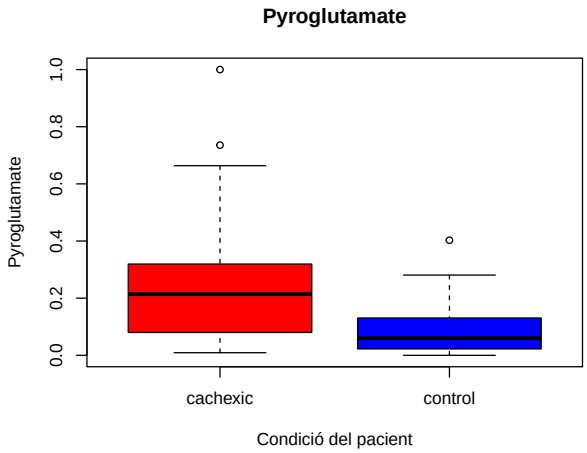

Methylamine
Methylamine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

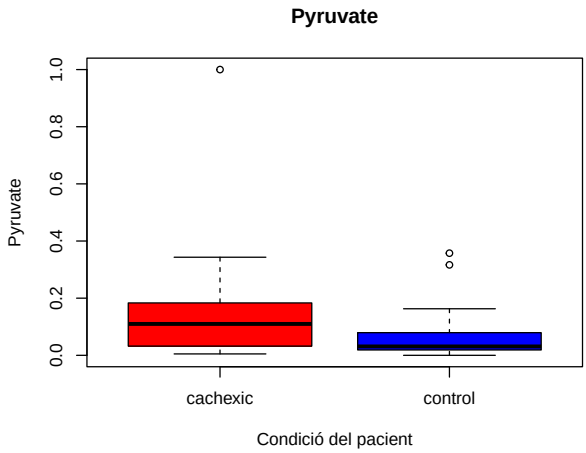

Methylguanidine
Methylguanidine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

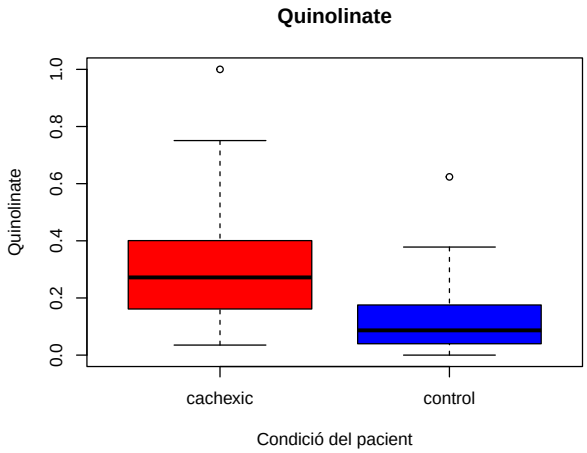

N.N.Dimethylglycine
N.N.Dimethylglycine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

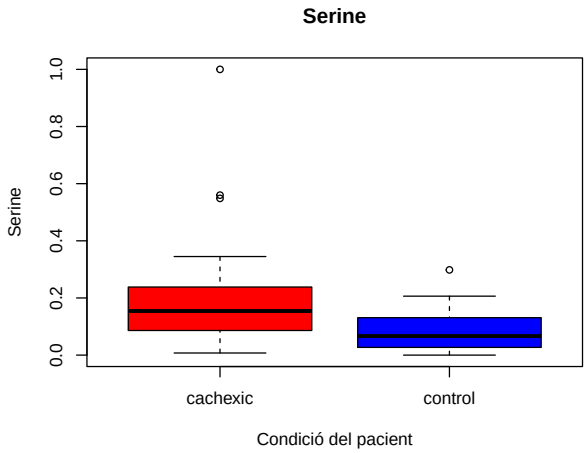

O.Acetylcarnitine
O.Acetylcarnitine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

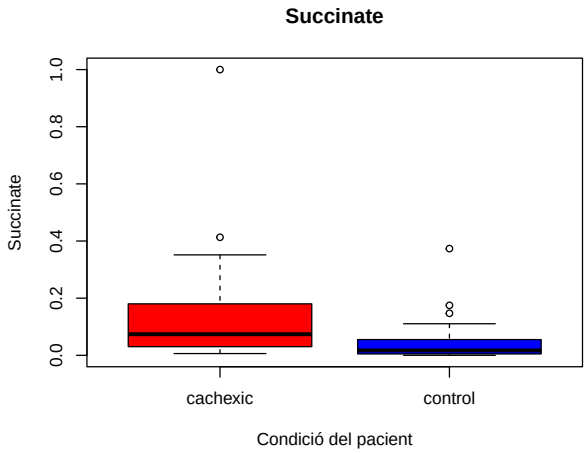

Pantothenate
Pantothenate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

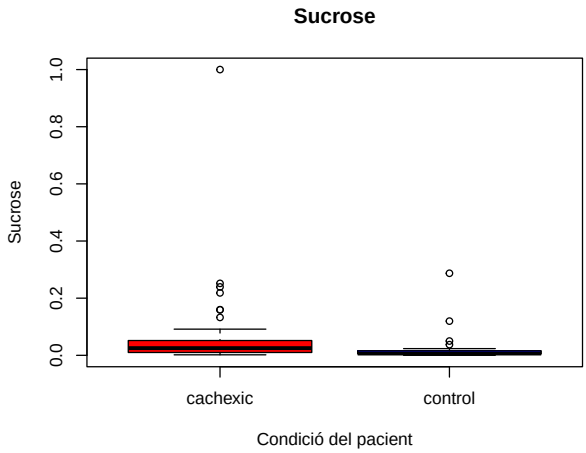

Pyroglutamate
Pyroglutamate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

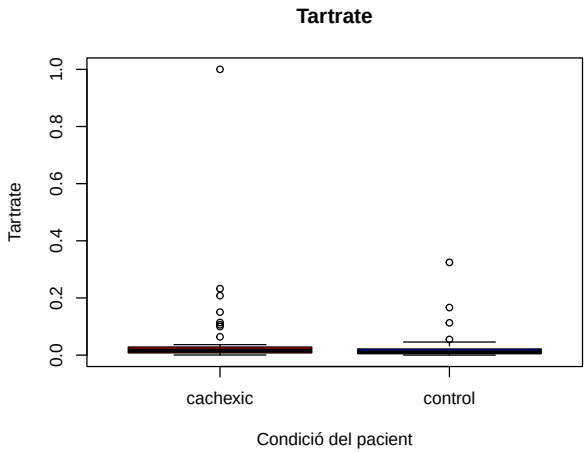

Pyruvate
Pyruvate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

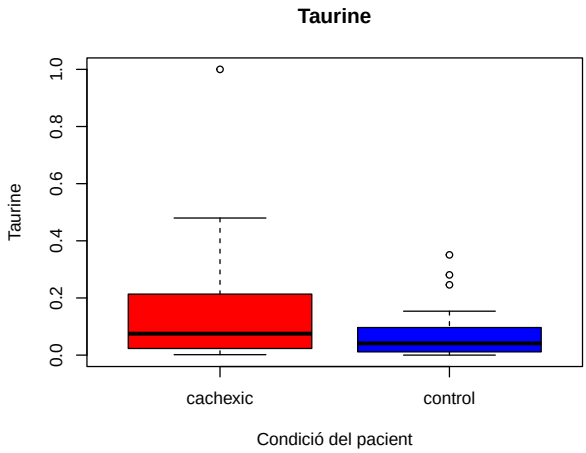

Quinolinate
Quinolinate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

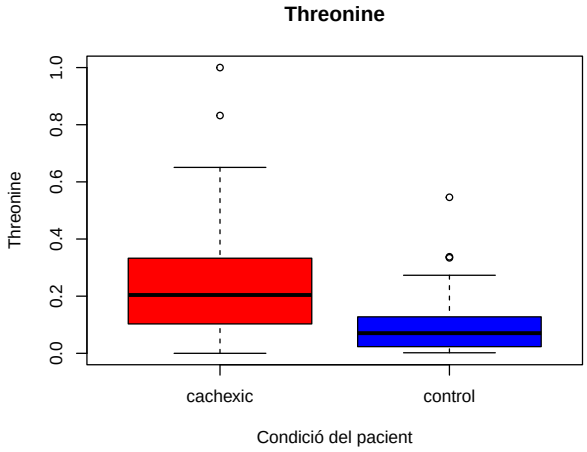

Serine
Serine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

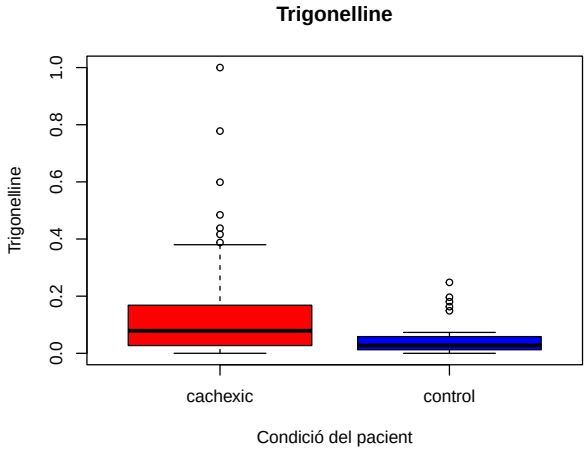

Succinate
Succinate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

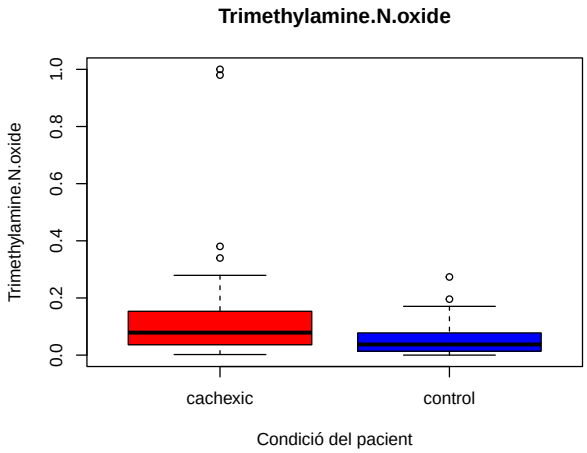

Sucrose
Sucrose
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

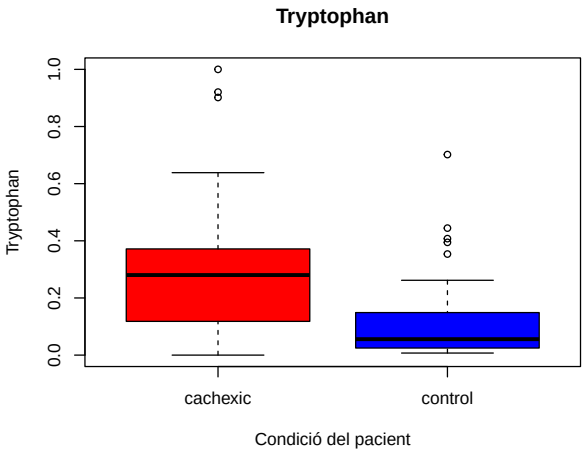

Tartrate
Tartrate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

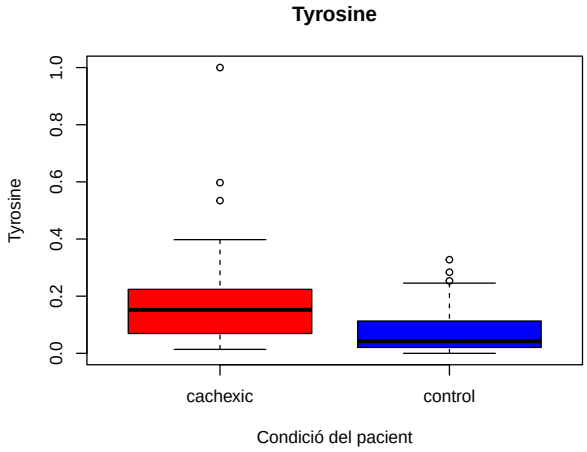

Taurine
Taurine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

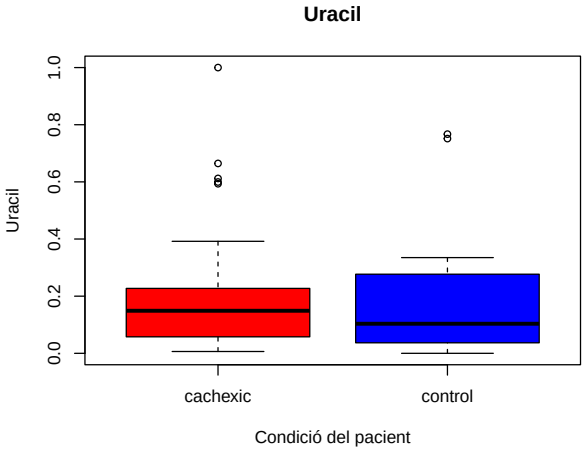

Threonine
Threonine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

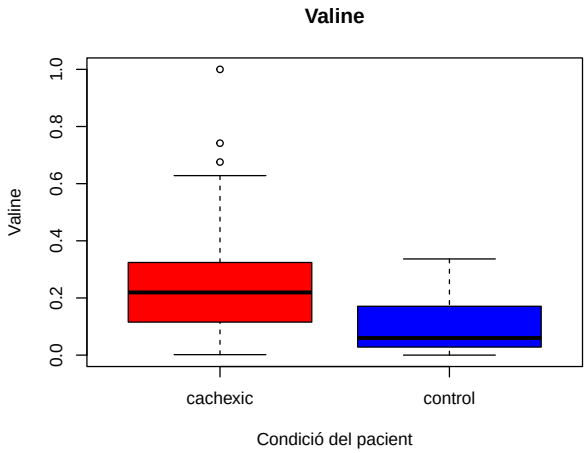

Trigonelline
Trigonelline
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

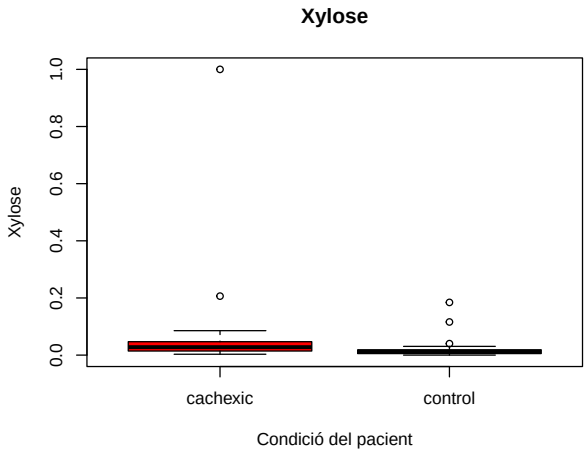

Trimethylamine.N.oxide
Trimethylamine.N.oxide
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

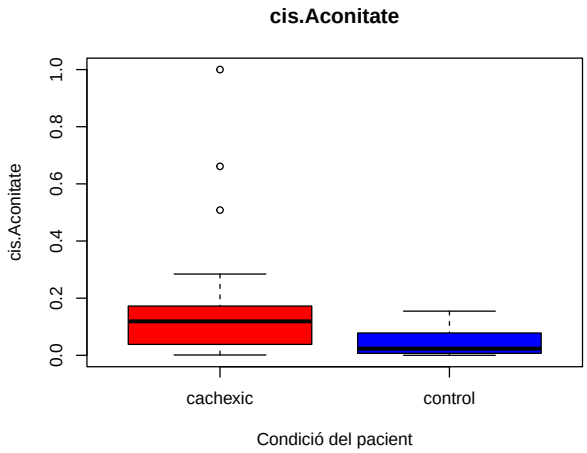

Tryptophan
Tryptophan
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

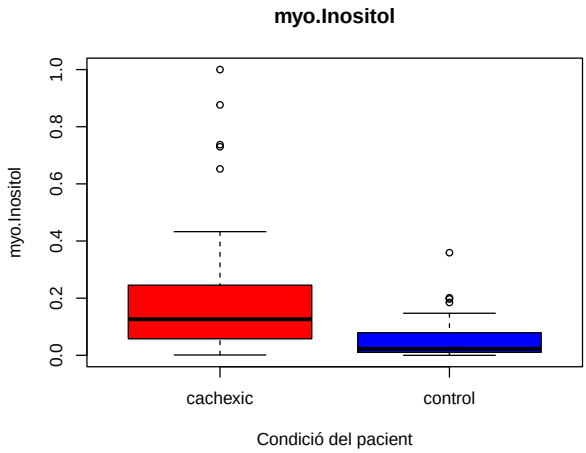

Tyrosine
Tyrosine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

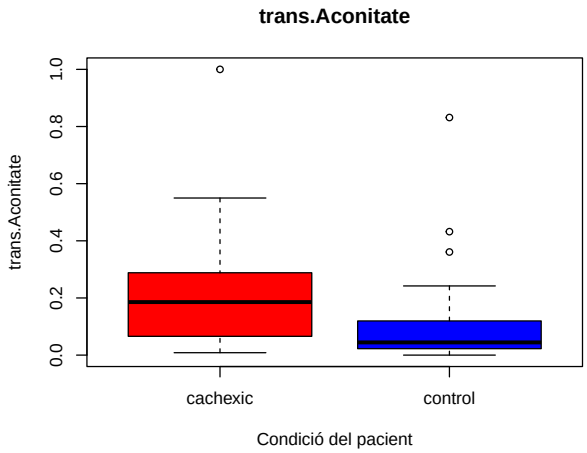

Uracil
Uracil
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

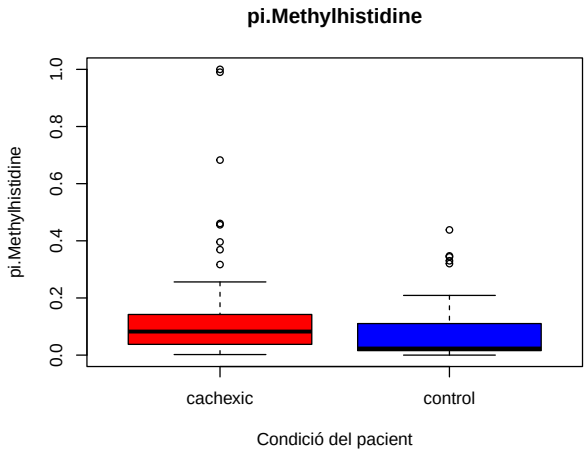

Valine
Valine
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient

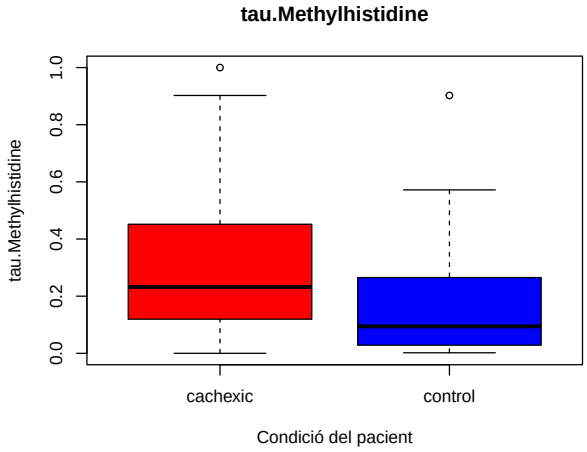

Xylose
Xylose
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient


cis.Aconitate
cis.Aconitate
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient


myo.Inositol
myo.Inositol
0.0
0.2
0.4
0.6
0.8
1.0
cachexic
control
Condició del pacient








```r
#Representem els gràfics de varies varaibles rellevants
#Seleccionem quants gràfics volem represnetar per pagina i fila
par(mfrow = c(2, 2))
#Fem el  grafic de caixes, on es representi els valors de concentració de
#la Leucina segons la condició del pacient
boxplot(as.numeric(assay(SE, "counts")["Leucine", ]) ~ colData(SE)$condition,
        #Afegim elements
        main = "Leucine",
        xlab = "Condició del pacient",
        ylab = "Leucine" ,
        col = c("red", "blue"))
boxplot(as.numeric(assay(SE, "counts")["X4.Hydroxyphenylacetate", ]) ~ colData(SE)$condition,
        main = "X4.Hydroxyphenylacetate",
        xlab = "Condició del pacient",
        ylab = "X4.Hydroxyphenylacetate" ,
        col = c("red", "blue"))
boxplot(as.numeric(assay(SE, "counts")["Creatinine", ]) ~ colData(SE)$condition,
        main = "Creatinine",
        xlab = "Condició del pacient",
        ylab = "Creatinine" ,
        col = c("red", "blue"))
```

```
boxplot(as.numeric(assay(SE, "counts")["Uracil", ]) ~ colData(SE)$condition,
        main = "Uracil",
        xlab = "Condició del pacient",
        ylab = "Uracil" ,
        col = c("red", "blue"))
```

8.11 Càlcul mitjanes

```
#Calculem la mitjana de cada variable segons la condicio
mitjana_leucina_control<-mean(dades$Leucine[dades$Muscle.loss=="control"])
mitjana_leucina_caquexia<-mean(dades$Leucine[dades$Muscle.loss!="control"])
mitjana_X4_control<-mean(dades$X4.Hydroxyphenylacetate
                         [dades$Muscle.loss=="control"])
mitjana_X4_caquexia<-mean(dades$X4.Hydroxyphenylacetate
                          [dades$Muscle.loss!="control"])
mitjana_creatinina_control<-mean(dades$Creatinine[dades$Muscle.loss=="control"])
mitjana_creatinina_caquexia<-mean(dades$Creatinine[dades$Muscle.loss!="control"])
mitjana_Uracil_control<-mean(dades$Uracil
                         [dades$Muscle.loss=="control"])
mitjana_Uracil_caquexia<-mean(dades$Uracil
                          [dades$Muscle.loss!="control"])

cat("La concentració mitjana de Leucina en els pacients controls es de",mitjana_leucina_control,"U.A.
I en els pacients caquexics és", mitjana_leucina_caquexia,"U.A.\n")

cat("La concentració mitjana de Hydroxypenylacetat en els pacients controls es de",mitjana_X4_control,"U
I en els pacients caquexics és", mitjana_X4_caquexia,"U.A.\n")

cat("La concentració mitjana de Creatinina en els pacients controls es de",mitjana_creatinina_control,"U
I en els pacients caquexics és", mitjana_creatinina_caquexia,"U.A.\n")

cat("La concentració mitjana d'Uracil en els pacients controls es de",mitjana_Uracil_control,"U.A.
I en els pacients caquexics és", mitjana_Uracil_caquexia,"U.A.\n")
```

```
#Calculem el % de diferencia entre una mitjana i l'altre i ho arrodonim.
leucina_prop<-round(((mitjana_leucina_caquexia-mitjana_leucina_control)/
                  mitjana_leucina_control)*100,2)
X4_prop <- round(((mitjana_X4_caquexia - mitjana_X4_control) /
                    mitjana_X4_control) * 100, 2)
creatinina_prop<-round(((mitjana_creatinina_caquexia-mitjana_creatinina_control)                    /mit
Uracil_prop <- round(((mitjana_Uracil_caquexia - mitjana_Uracil_control) /
                    mitjana_Uracil_control) * 100, 2)
cat("La concentració de Leucina es un", leucina_prop, "% superior en pacients amb caquèxia
en comparació al control.\n")
cat("La concentració de Hydroxypenylacetat es un", X4_prop, "% superior en
pacients amb  caquèxia en comparació al control.\n")
cat("La concentració de Creatinina es un", creatinina_prop, "% superior en pacients amb caquèxia
en comparació al control.\n")
cat("La concentració d'Uracil es un", Uracil_prop, "% superior en
pacients amb  caquèxia
en comparació al control.\n")
```

8.12 ANOVA

Es mostra l'analisis ANOVA realitzat amb les dades no escalades. Del qual s'han seleccionat nomes dues variables, una de significativa i una que no ho es per proposits educatius.

```
#Generem un objecte on guardarem els p-valors
p_values<-numeric(length(noms))
#Fem un loop per fer l'ANOVA per totes les variables
for (i in seq_along(noms)){
  #Amb la funció "as.formula" podem fer una formual dinamica on la variable que anirà canviant es "noms
  anova<-aov(as.formula(paste(noms[i],"~ Muscle.loss")),data=dades)
  #Guardem cada resum que fem
  resum_model<-summary(anova)
  #De cada resum en guardem el p_valor corresponent
  p_values[i]<-resum_model[[1]]["Muscle.loss", "Pr(>F)"]
}
#Finalment guardem els noms de les variables amb el p_valor calculat
resultats_pvalue<-data.frame(Variable=noms, P_value=p_values)
print(resultats_pvalue)
```

```
##                       Variable      P_value
## 1  X1.6.Anhydro.beta.D.glucose 0.0507913658
## 2         X1.Methylnicotinamide 0.9343185339
## 3              X2.Aminobutyrate 0.0274454761
## 4         X2.Hydroxyisobutyrate 0.0052880207
## 5               X2.Oxoglutarate 0.2250725245
## 6           X3.Aminoisobutyrate 0.1779554046
## 7            X3.Hydroxybutyrate 0.0011853671
## 8         X3.Hydroxyisovalerate 0.0078274621
## 9             X3.Indoxylsulfate 0.0089161341
## 10     X4.Hydroxyphenylacetate 0.4818107935
## 11                     Acetate 0.0060731953
## 12                     Acetone 0.3723624590
## 13                     Adipate 0.0274331625
## 14                     Alanine 0.0011788609
## 15                  Asparagine 0.0067852184
## 16                     Betaine 0.0029929659
## 17                   Carnitine 0.0621334639
## 18                     Citrate 0.0128569793
## 19                    Creatine 0.0526324364
## 20                  Creatinine 0.0005129808
## 21                Dimethylamine 0.0004460069
## 22                Ethanolamine 0.0265705687
## 23                     Formate 0.0174219763
## 24                      Fucose 0.0058534857
## 25                    Fumarate 0.0590430806
## 26                     Glucose 0.0333063584
## 27                    Glutamine 0.0010607678
## 28                     Glycine 0.0281395160
## 29                    Glycolate 0.0563341550
## 30               Guanidoacetate 0.1378998638
## 31                    Hippurate 0.0232194018
## 32                    Histidine 0.0109704361
## 33                Hypoxanthine 0.2555158507
```

```
## 34                  Isoleucine 0.1326513615
## 35                     Lactate 0.1228434803
## 36                     Leucine 0.0002695046
## 37                      Lysine 0.2817634974
## 38                 Methylamine 0.0019466201
## 39              Methylguanidine 0.2616448016
## 40          N.N.Dimethylglycine 0.0001554346
## 41            O.Acetylcarnitine 0.0616844278
## 42                 Pantothenate 0.5353413018
## 43                Pyroglutamate 0.0004845363
## 44                     Pyruvate 0.0174471086
## 45                  Quinolinate 0.0001185108
## 46                      Serine 0.0037584213
## 47                   Succinate 0.0113771704
## 48                     Sucrose 0.1194324540
## 49                     Tartrate 0.4464672671
## 50                      Taurine 0.0323497353
## 51                    Threonine 0.0032881849
## 52                 Trigonelline 0.0128629342
## 53      Trimethylamine.N.oxide 0.0415941735
## 54                  Tryptophan 0.0018983944
## 55                     Tyrosine 0.0112887118
## 56                      Uracil 0.5429079298
## 57                      Valine 0.0001394238
## 58                      Xylose 0.2154519963
## 59                cis.Aconitate 0.0038978573
## 60                 myo.Inositol 0.0022150347
## 61              trans.Aconitate 0.0220856691
## 62           pi.Methylhistidine 0.1413081041
## 63          tau.Methylhistidine 0.0220399755
```

```
#Ordenem els valos per veure quines son les variables amb diferencies significatives
valors_significatius_ordenats<-resultats_pvalue[order(resultats_pvalue$P_value),]
valors_significatius_ordenats
```

```
##                       Variable      P_value
## 45                  Quinolinate 0.0001185108
## 57                       Valine 0.0001394238
## 40          N.N.Dimethylglycine 0.0001554346
## 36                       Leucine 0.0002695046
## 21                  Dimethylamine 0.0004460069
## 43                  Pyroglutamate 0.0004845363
## 20                     Creatinine 0.0005129808
## 27                     Glutamine 0.0010607678
## 14                       Alanine 0.0011788609
## 7             X3.Hydroxybutyrate 0.0011853671
## 54                    Tryptophan 0.0018983944
## 38                   Methylamine 0.0019466201
## 60                  myo.Inositol 0.0022150347
## 16                       Betaine 0.0029929659
## 51                    Threonine 0.0032881849
## 46                       Serine 0.0037584213
## 59                cis.Aconitate 0.0038978573
## 4          X2.Hydroxyisobutyrate 0.0052880207
```

```
## 24                     Fucose 0.0058534857
## 11                    Acetate 0.0060731953
## 15                 Asparagine 0.0067852184
## 8        X3.Hydroxyisovalerate 0.0078274621
## 9            X3.Indoxylsulfate 0.0089161341
## 32                   Histidine 0.0109704361
## 55                   Tyrosine 0.0112887118
## 47                   Succinate 0.0113771704
## 18                     Citrate 0.0128569793
## 52               Trigonelline 0.0128629342
## 23                     Formate 0.0174219763
## 44                    Pyruvate 0.0174471086
## 63         tau.Methylhistidine 0.0220399755
## 61             trans.Aconitate 0.0220856691
## 31                   Hippurate 0.0232194018
## 22                Ethanolamine 0.0265705687
## 13                     Adipate 0.0274331625
## 3             X2.Aminobutyrate 0.0274454761
## 28                     Glycine 0.0281395160
## 50                     Taurine 0.0323497353
## 26                     Glucose 0.0333063584
## 53      Trimethylamine.N.oxide 0.0415941735
## 1  X1.6.Anhydro.beta.D.glucose 0.0507913658
## 19                    Creatine 0.0526324364
## 29                   Glycolate 0.0563341550
## 25                    Fumarate 0.0590430806
## 41           O.Acetylcarnitine 0.0616844278
## 17                   Carnitine 0.0621334639
## 48                     Sucrose 0.1194324540
## 35                     Lactate 0.1228434803
## 34                  Isoleucine 0.1326513615
## 30              Guanidoacetate 0.1378998638
## 62          pi.Methylhistidine 0.1413081041
## 6          X3.Aminoisobutyrate 0.1779554046
## 58                      Xylose 0.2154519963
## 5              X2.Oxoglutarate 0.2250725245
## 33                Hypoxanthine 0.2555158507
## 39             Methylguanidine 0.2616448016
## 37                     Lysine 0.2817634974
## 12                    Acetone 0.3723624590
## 49                    Tartrate 0.4464672671
## 10     X4.Hydroxyphenylacetate 0.4818107935
## 42                Pantothenate 0.5353413018
## 56                     Uracil 0.5429079298
## 2        X1.Methylnicotinamide 0.9343185339
```

```
#Fem recompte de quantes variables presenten diferencies significatives
count(valors_significatius_ordenats$P_value<0.05)
```

```
## [1] 40
```