Anàlisi de dades òmiques (M0-157)

Primera prova d'avaluació contínua

Mar Llorca Torres

1. Introducció.

Aquesta PEC està planificada com una introducció a les òmiques mitjançant un exercici de repàs, ampliant l'ús d'algunes eines com Bioconductor i l' exploració multivariant de dades.

1.1. Objectius.

L'objectiu principal d'aquest treball és planificar i executar una versió simplificada del procés d'anàlisi de dades òmiques. Es demana fer un anàlisi exploratori d'unes dades de metabolòmica, utilitzant el programa estadístic R. És requisit crear un objecte de classe SummarizedExperiment per emmagatzemar les dades i, posteriorment, realitzar l'anàlisi de les dades a través d'aquest objecte.

L'objectiu desglossat consisteix en:

- Seleccionar i descarregar un dataset de metabolòmica.
- Crear un objecte de classe SummarizedExperiment que contingui les dades i les metadades del conjunt de dades.
- Realitzar una anàlisi exploratòria de les dades.
- Interpretar els resultats de l'anàlisi.
- Desenvolupar un informe detallat amb els resultats de l'anàlisi.
- Crear un repositori GitHub.

2. Materials i Mètodes.

2.1. Selecció de les dades per a l'estudi.

Les dades de metabolòmica que s'utilitzen en aquesta PAC es descarreguen del repositori GitHub proporcionat a l'enunciat. Es selecciona el dataset 2024-Cachexia.

La caquèxia és una síndrome metabòlica complexa associada a diverses malalties cròniques, com el càncer, la insuficiència cardíaca congestiva, la malaltia renal crònica i altres que es caracteritza per la pèrdua de pes i massa muscular. En aquest context, l'anàlisi d'aquestes dades de metabolòmica pot ser molt útil per identificar biomarcadors potencials per a un diagnòstic precoç i per ampliar i millorar l'enteniment dels mecanismes biològics implicats.

3. Resultats.

3.1. Estructura de les dades i de l'estudi.

Aquestes dades son sobre la caquèxia, un síndrome metabòlic complex associat a una malaltia subjacent i que es caracteritza per la pèrdua de múscul amb o sense pèrdua de massa grassa. El dataset consisteix en un total de 77 mostres d'orina, 47 d'elles pacients amb caquèxia i 30 pacients control.

Les dades contenen la següent informació:

- **Patient.ID**: Identificador del pacient.
- **Muscle.loss**: Indica si el pacient té caquèxia ("cachexic") o és control ("control").
- **Mesures de metabolits**: concentracions en orina de 63 metabolits diferents.

3.2 Creació de l'objecte SummarizedExperiment.

Per a emmagatzemar les dades de manera estructurada i facilitar l'anàlisi, es crea un objecte de classe SummarizedExperiment. Aquest tipus d'objecte permet gestionar tant les dades de mesures (concentracions dels metabolits) com les metadades associades (informació adicional com la ID del pacient i el grup de pèrdua muscular). S'adjunta el codi i la Taula 1 en Annexes .

Per a l'estudi es separen les dades en dues parts:

- Dades: conjunt de dades amb totes les columnes de concentracions de metabolits. Les dades que interessen per a l'anàlisi de metabolòmica (columnes de la 3 a la 65).
- Metadades: conjunt que inclou les metadades ID del pacient i pèrdua muscular, que ajudaràn a interpretar l'anàlisi (columnes 1 i 2).

Tant SummarizedExperiment com ExpressionSet s'utilitzen per gestionar dades experimentals en biologia computacional. Les seves principals diferències son:

- ExpressionSet: dissenyada principalment per a experiments basats en microarrays, on les files representen característiques com gens o sondes, i les columnes representen les mostres. Son matrius amb una gran quantitat de metadades. Conté tres components principals: assayData (matriu de dades), phenoData (metadades de les mostres) i featureData (metadades de les features).

- SummarizedExperiment: dissenyada principalment per a experiments basats en seqüenciació i altres dades òmiques (metabolòmica, transcriptòmica, etc.). Permet múltiples assays (matrius) dins del mateix objecte, cadascun representant diferents tipus de dades experimentals (sempre que tinguin les mateixes dimensions). Les metadades de les files es gestionen amb rowData i les de les columnes es gestionen amb colData. És més flexible que la ExpressionSet.

3.3. Anàlisi exploratòria de les dades.

Primer, s'utilitzen funcions bàsiques per tenir una visió general com:

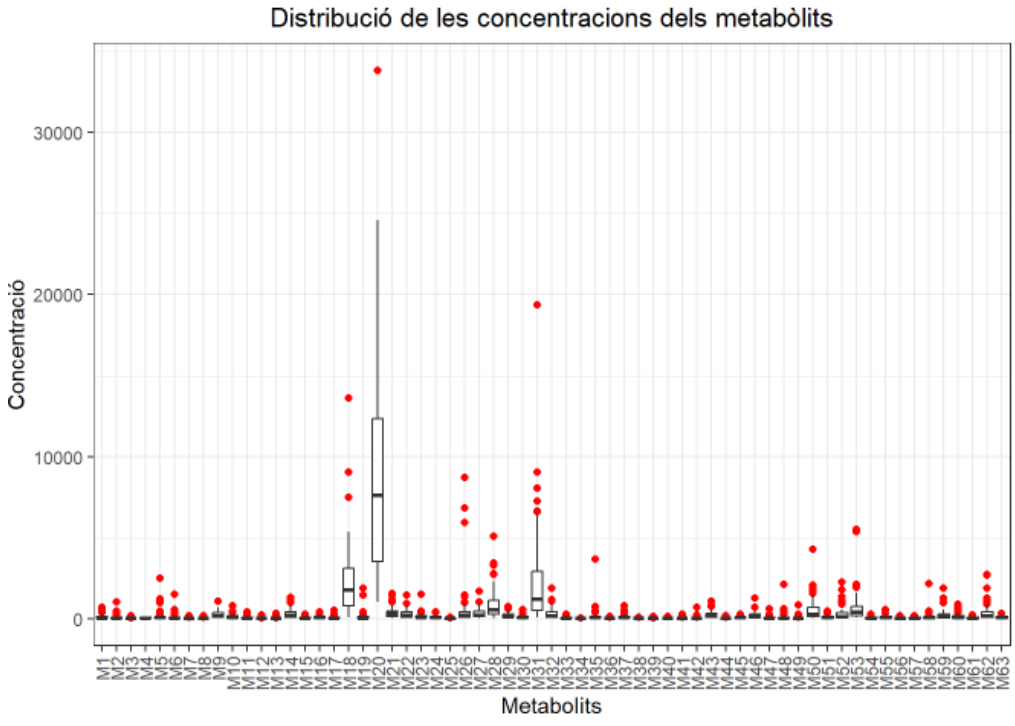
- dim(assays(se_cachexia)$raw): Per accedir a les dades d'expressió (assays)
- head(rowData(se_cachexia)): Per accedir a les files (metabolits)
- head(colData(se_cachexia)): Per accedir a les metadades de les columnes (ID pacient i pèrdua muscular)

Després d'analitzar-les es pot concluir que aquest SummaizedExperiment conté una matriu de dimensions 63x77, on les files son els diferents metabolits i les columnes les diferents mostres.

A més, les diferents mostres estàn vinculades amb les metadades corresponents: el ID del pacient i el tipus de pèrdua muscular (cachexic vs control). I les files de metabolits estàn vinculades amb el seu nom complet.

3.3.1. Exploració univariant.

En primer lloc, res realitza un anàlisi univariant mitjançant gràfics senzills per explorar la distribució general de les variables individuals (concentració de metabolits). S'ha escollit una representació de tipus boxplot per visualitzar els valors de concentracions per als diferents metabolits:



Donat que l'estudi està enfocat a una comparació entre el grup de pacients amb cachexia i el grup de pacients de control, es realitza una taula amb algunes dades estadístiques importants a tenir en compte per a l'anàlisi (Taula 2). La taula conté les següents columnes:

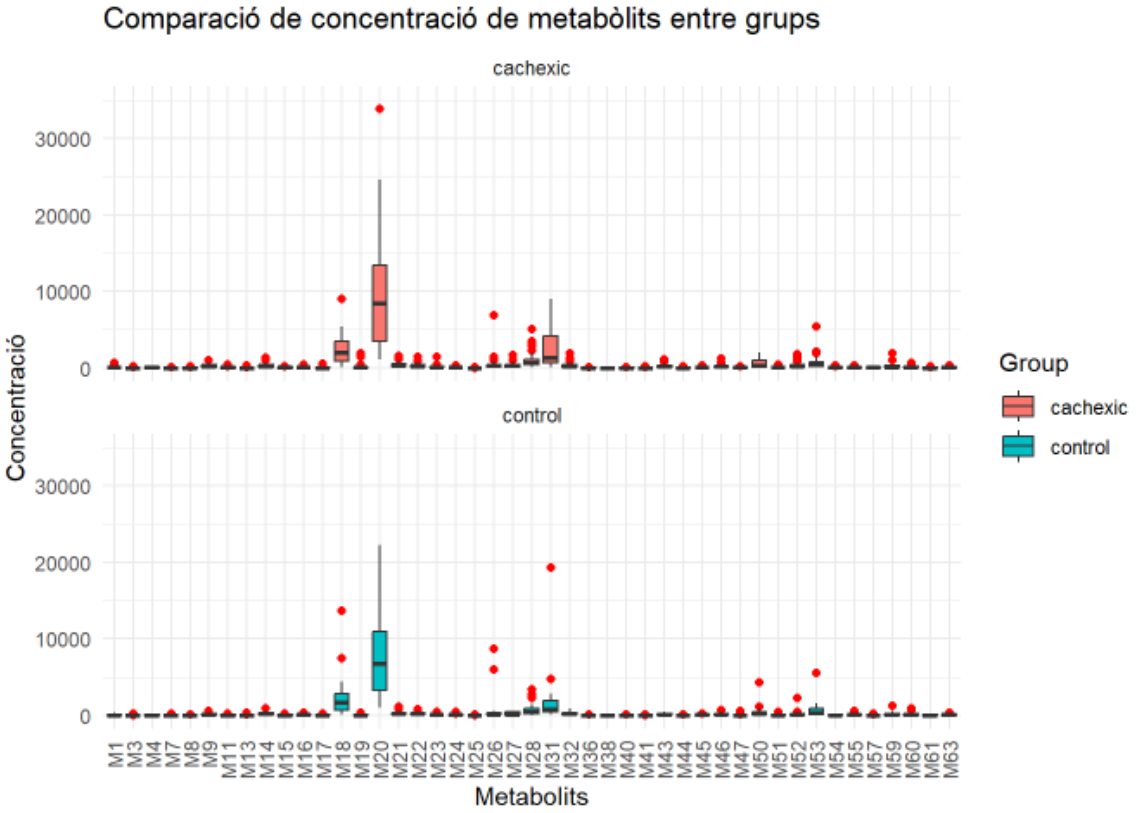
- Metabolite: Nom del metabolit.
- TotalMissing: Nombre total de valors perduts. Si un metabolit té molts valors perduts (TotalMissing), pot ser menys fiable per a conclusions estadístiques.
- TTestPvalue: P-value del test t. Un valor de TTestPvalue<0.05 indica que la diferència entre els dos grups no és deguda a l'atzar (comprova si la diferència observada entre Cachexic_Mean i Control_Mean és estadísticament significativa).
- Cachexic_Mean: Mitjana del grup cachexic.
- Control_Mean: Mitjana del grup control.

Amb la comparació entre Cachexic_Mean i Control_Mean es poden identificar potencials biomarcadors quan mitjana d'un metabolit és molt diferent entre els dos grups. Els 10 metabolits amb major diferencia de mitjanes son: M20, M31, M18, M26, M28, M53, M50, M21, M52, M27.

Arrel de la taula anterior (Taula 2), s'utilitza el Test estadístic per seleccionar els metabolits que tenen diferencies significatives entre grups.

Dels 63 metabolits presents a les mostres, només 45 d'ells tenen diferencies estadisticament significatives entre el grup de pacients amb la malaltia i el grup de pacients de control.

Es repeteix el boxplot de la distribució de les concentracions dels metabolits, seleccionant només els metabolits significatius i separant el grup cachexia del de control.

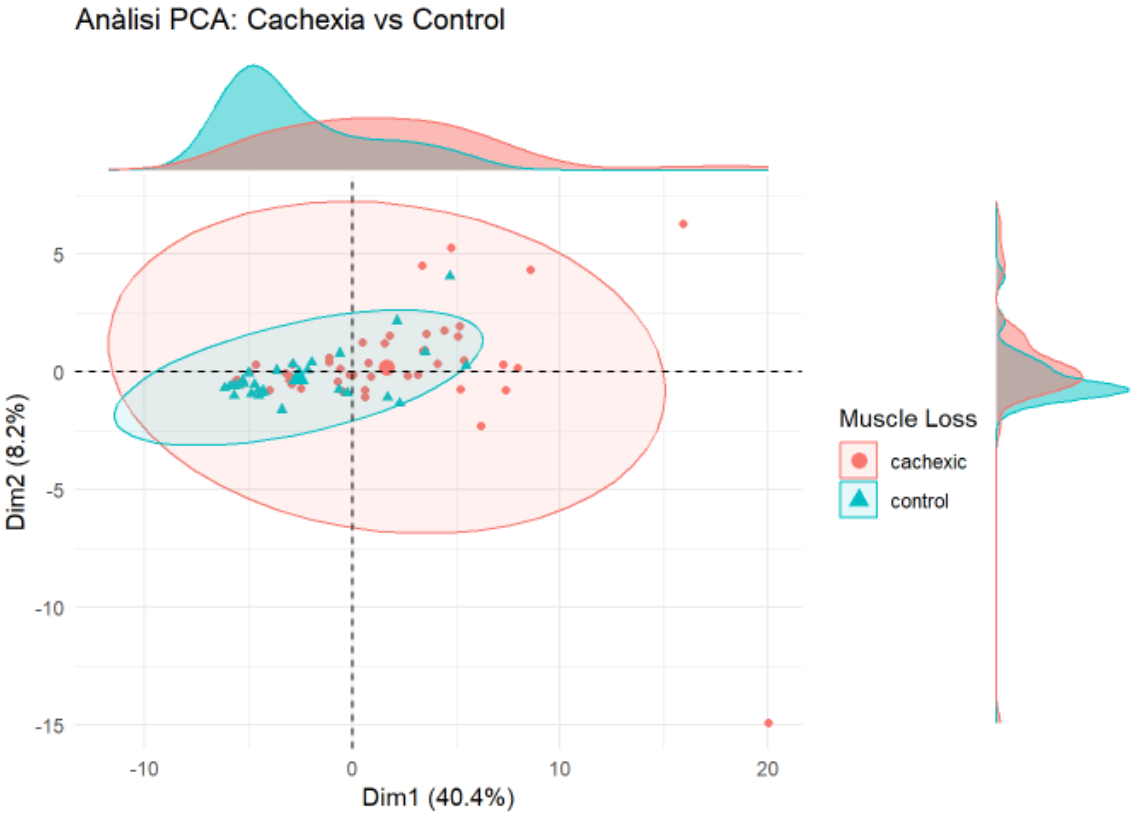


3.3.2. Exploració multivariant.

Posteriorment, es procedeix a fer un l'anàlisi multivariant. Es realitza l'Anàlisi de Components Principals (PCA) per reduir la dimensió de les dades i identificar possibles patrons associats amb la caquèxia. Això ajuda a visualitzar com es distribueixen les mostres en funció de les components principals i si existeixen diferències clares entre els grups. També es realitza una matriu de correlació visualitzada en forma de HeatMap.

- Anàlisi de PCA.

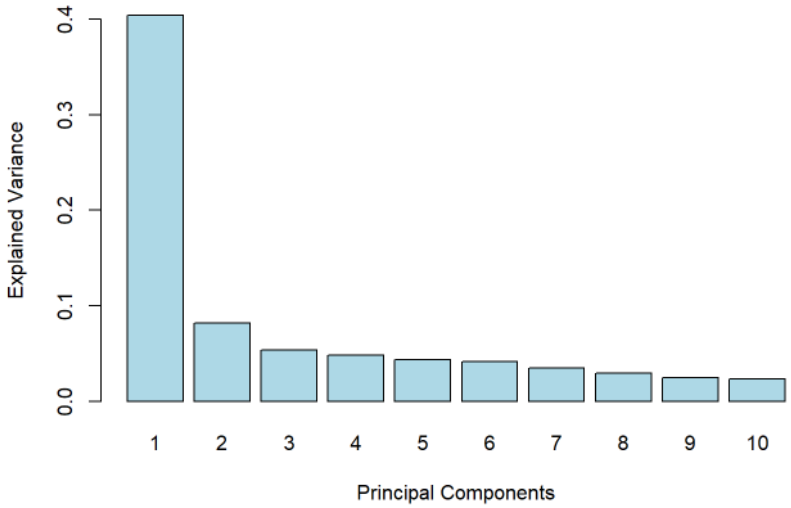
Es realitza un anàlisi de components principals (PCA) diferenciant el grup cachexia amb el control. A més, s'afegeixen els eixos marginals amb les densitats per visualitzar la distribució de les dades per a cada grup.



Tot i que sí s'observen diferències entre ambdós grups, no es veu una separació clara entre ells. Tot i això, el gràfic suggereix que la major font de variació, explicada per els dos primers components, es pot atribuir a les mostres de cachexia.

S'amplia l'anàlisi PCA utilitzant el paquet mixOmics. Es tracta d'un paquet d'anàlisi multivariant que es fa servir en l'anàlisi de dades òmiques que ofereix funcions avançades per a l'anàlisi de components principals (PCA).

S'utilitza la funció tune.pca() per calcular la proporció acumulada de variància explicada per a un gran nombre de components principals (en aquest cas, per a 10 components) i es genera el gràfic de desplaçament de la proporció de variança explicada en relació a la variança total de les dades per a cada component principal.

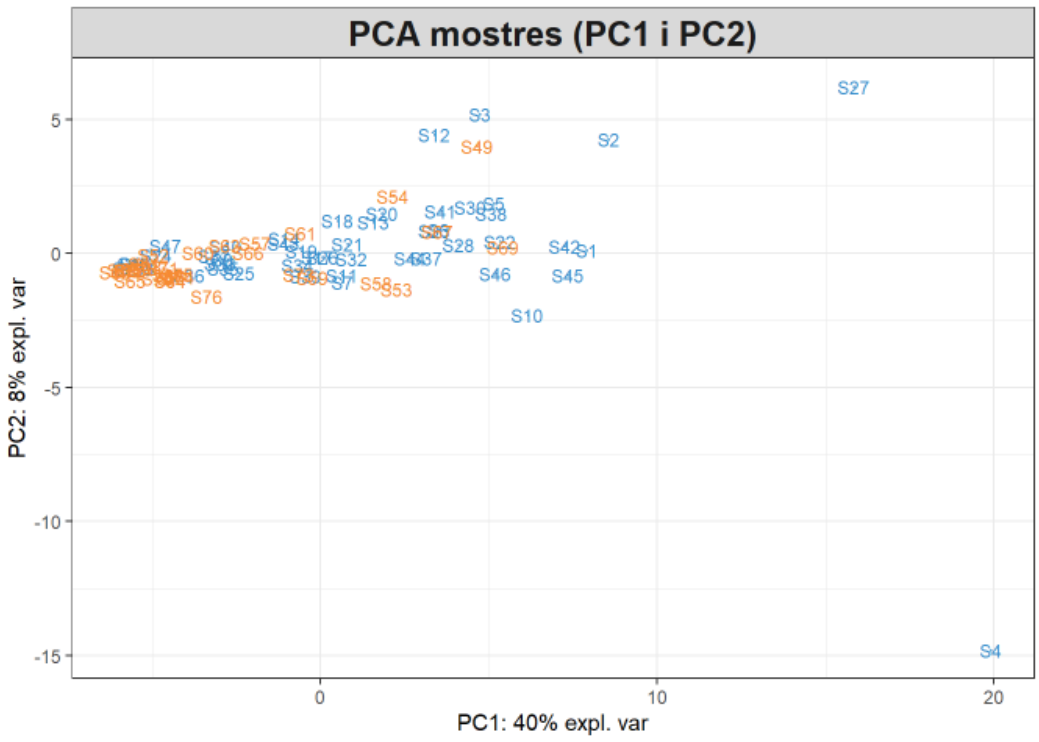


Proporció acumulada de variança;

- PC1: Explica el 40,43% de la variança total.
- PC3: Junts, PC1, PC2 i PC3 expliquen el 53,94%.
- PC6: Els primers 6 components expliquen el 67.29%.
- PC10: Els primers 10 components expliquen el 78,45%.

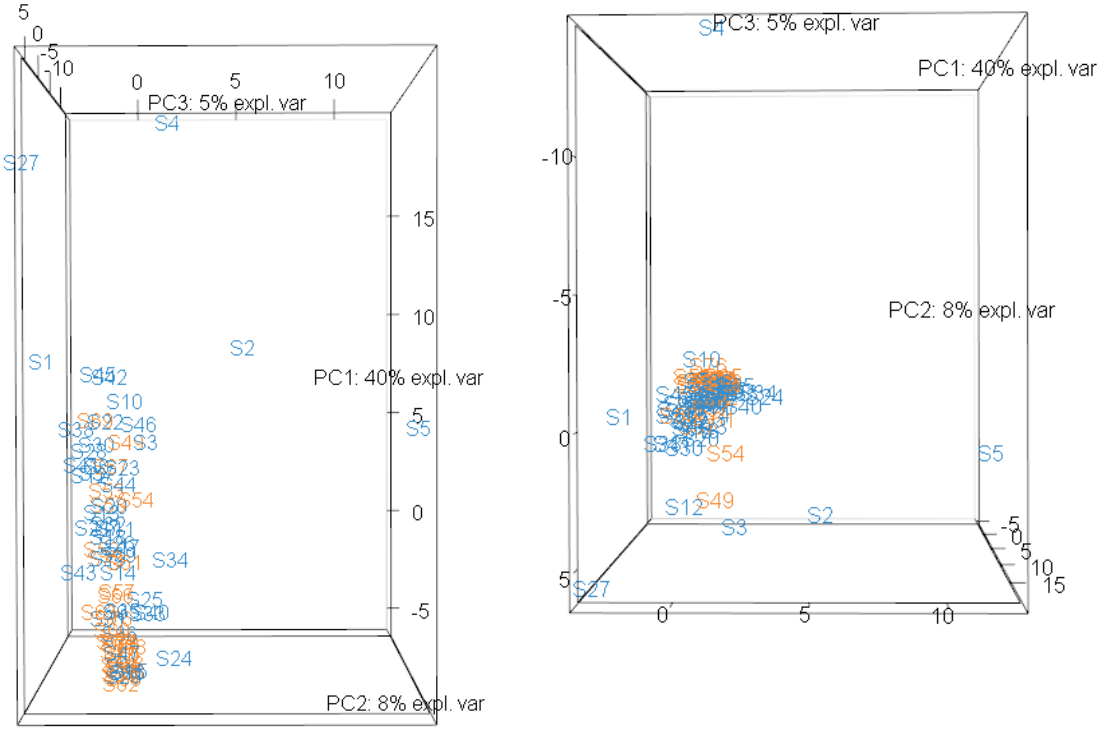
Mitjançant el gràfic s'observa que el nombre òptim de components és 2-3.

Es decideix que el nombre òptim de components es de 3 i es fa el PCA corresponent:

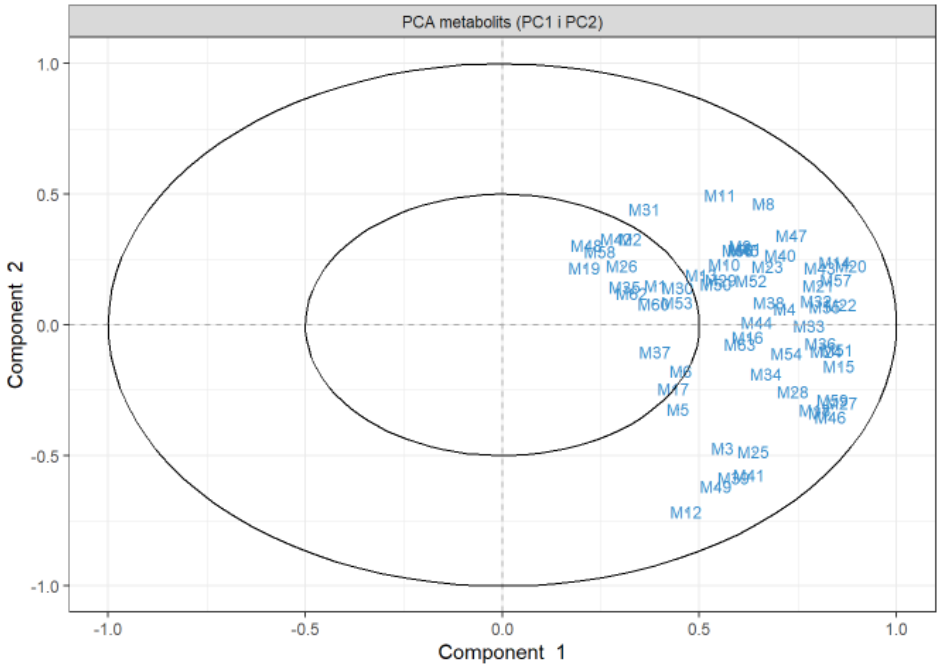


Tal com s'havia vist anteriorment, el gràfic suggereix que la major font de variació, explicada per els dos primers components, es pot atribuir a les mostres de cachexia.

Donat que s'ha executat el PCA per 3 components, es pot examinar aquest tercer component utilitzant un gràfic interactiu 3D. L'adició d'aquest tercer component ressalta el fet que el grup cachexia és la major font de variació, quedant principalment 4 mostres com a possibles valors atípics (S2, S4, S5 i S27).



Es visualitza el gràfic de variables (gràfic de cercle de correlació):



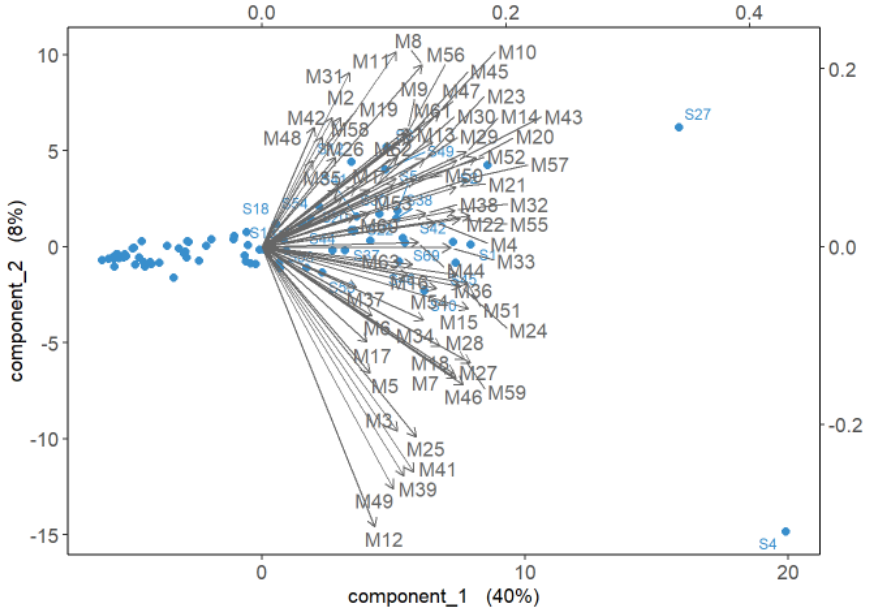
A banda del gràfic, es poden visualitzar les variables més importants per a cada component principal. Primer s'identifiquen amb la funció selectVar (que els ordena per ordre d'importancia) i després es crea una taula per visualitzar els 10 més importants de cada compost (Taula 3).

De la mateixa manera, es poden poden ordenar de forma inversa per veure quins metabolits son menys importants. Es selecciona un llindar amb la influencia mínima desitjada (0.2) i, tots aquells que estiguin per sota d'aquest llindar, seran variables simplificables. Es realitza aquesta operació per a cada component, i finalment es busquen quines variables son simplificables per als tres components.

Per finalitzar, es busquen els metabolits que son poc significatius i, a més, son menys importants respecte als 3 components principals: M1, M4, M7, M9, M13, M14, M15, M16, M17, M18, M19, M20, M21, M22, M23, M24, M27, M28, M31, M32, M36, M38, M40, M43, M44, M45, M46, M47, M50, M51, M52, M53, M54, M55, M57, M59, M61, M63.

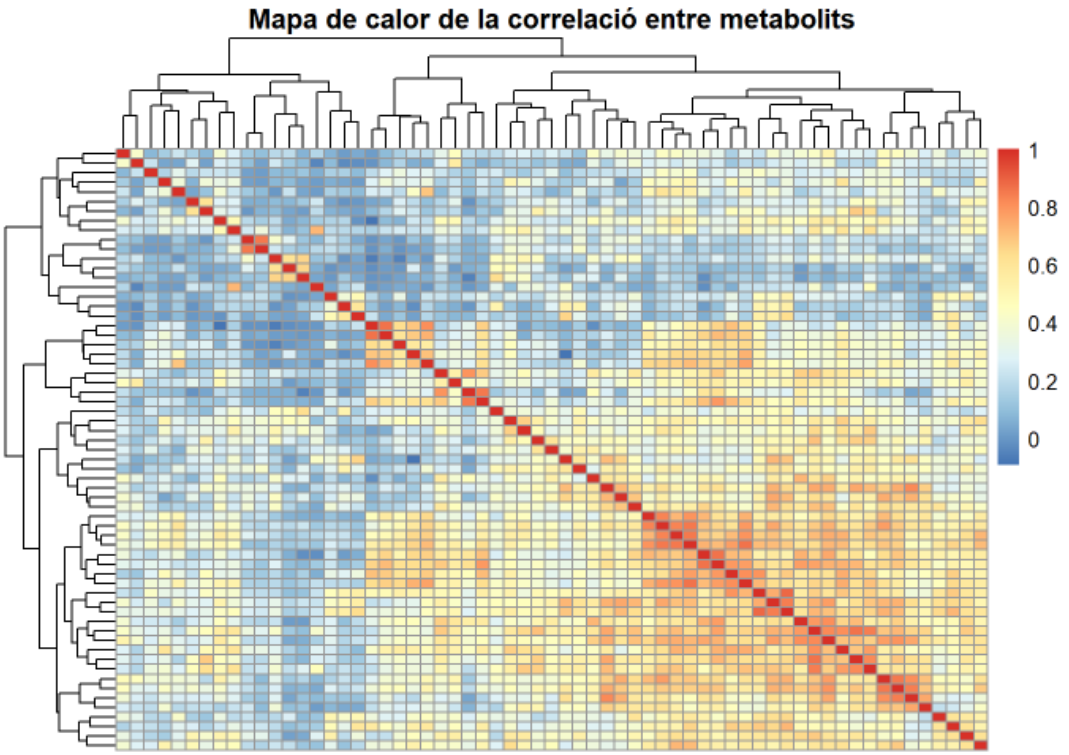
S'accedeix als noms reals dels metabòlits a través del SummarizedExperiment i s'obté que son: X1.6.Anhydro.beta.D.glucose, X2.Hydroxyisobutyrate, X3.Hydroxybutyrate, X3.Indoxylsulfate, Adipate, Alanine, Asparagine, Betaine, Carnitine, Citrate, Creatine, Creatinine, Dimethylamine, Ethanolamine, Formate, Fucose, Glutamine, Glycine, Hippurate, Histidine, Leucine, Methylamine, N.N.Dimethylglycine, Pyroglutamate, Pyruvate, Quinolinate, Serine, Succinate, Taurine, Threonine, Trigonelline, Trimethylamine.N.oxide, Tryptophan, Tyrosine, Valine, cis.Aconitate, trans.Aconitate, tau.Methylhistidine

Finalment, es representa el diagrama bidimensional (biplot) que permet mostrar les mostres i variables de forma simultània per poder visualitzar les seves relacions.



- Gràfic de tipus Heatmap.

Una altra tècnica exploratòria interessant és la visualització de la matriu de correlació entre els metabolits per veure les relacions entre les concentracions. Una forma de visualitzar-la és utilitzant la funció cor i un mapa de calor (heatmap): aquesta visualització permet identificar possibles grups de metabolits que tenen una correlació forta entre ells.



Les correlacions entre metabolits poden ser útils per agrupar metabolits relacionats entre si o, per exemple, per identificar vies metabòliques comunes. Es crea una taula que recull les correlacions superiors a 0.8 (Taula 4).

3.4. Conclusions finals.

A través de l'anàlisi realitzat es poden extreure varies conclusions:

- Metabolits significatius: del total de metabòlits presents a les mostres, només 45 d'ells son significatius: M36, M40, M45, M57, M21, M43, M20, M7, M14, M27, M60, M59, M46, M54, M51, M38, M11, M24, M16, M15, M47, M52, M8, M23, M32, M4, M55, M9, M18, M13, M44, M3, M26, M31, M28, M50, M53, M63, M19, M22, M61, M41, M25, M17, M1.
- S'identifiquen potencials biomarcadors amb la comparació entre Cachexic_Mean i Control_Mean. Els 10 metabolits seleccionats son: M20, M31, M18, M26, M28, M53, M50, M21, M52, M27.
- Dels tres components principals seleccionats, el primer explica un 40.43% de la variància total, el primer i segon expliquen un 48.61% i els tres primers un 53.94%.
- Per als components principals (PC1, PC2 i PC3), s'han identificat els metabolits més influents en cada component. Alguns metabolits clau que apareixen amb alta influència en el PCA inclouen M20, M12, M35 (per PC1), M27, M49, M26 (per PC2), i M20, M26, M52 (per PC3).
- S'estableix un valor mínim d'influència per cada component, i s'obtenen els metabòlits menys rellevants. Els metabòlits poc rellevants comuns per als tres components principals son: M1, M2, M4, M5, M6, M7, M9, M10, M13, M14, M15, M16, M17, M18, M19, M20, M21, M22, M23, M24, M27, M28, M29, M30, M31, M32, M33, M36, M37, M38, M40, M42, M43, M44, M45, M46, M47, M48, M50, M51, M52, M53, M54, M55, M56, M57, M58, M59, M61, M62, M63.

- Els noms dels metabòlits poc significatius i que a més son poc rellevants per als 3 components principals son: X1.6.Anhydro.beta.D.glucose, X2.Hydroxyisobutyrate, X3.Hydroxybutyrate, X3.Indoxylsulfate, Adipate, Alanine, Asparagine, Betaine, Carnitine, Citrate, Creatine, Creatinine, Dimethylamine, Ethanolamine, Formate, Fucose, Glutamine, Glycine, Hippurate, Histidine, Leucine, Methylamine, N.N.Dimethylglycine, Pyroglutamate, Pyruvate, Quinolinate, Serine, Succinate, Taurine, Threonine, Trigonelline, Trimethylamine.N.oxide, Tryptophan, Tyrosine, Valine, cis.Aconitate, trans.Aconitate, tau.Methylhistidine.
- El gràfic interactiu 3D identifica quatre possibles valors atípics. Es tracta de les mostres: S2, S4, S5 i S27.
- S'identifiquen els metabòlits més relacionats entre sí mitjançant la matriu de correlació.

Enllaç al repositori GitHub creat:

https://github.com/M-Llorca/Llorca-Torres-Mar-PEC1.git

4. Apèndix: taules.

Taula 1:

Camp	Descripció	Valors
Study Description	Descripció de l'estudi	A metabolite concentration table from human urine samples with two groups
Patient.ID	Identificador del pacient	PIF_178, PIF_087, PIF_090, NETL_005_V1, PIF_115, PIF_110, NETL_019_V1, NETCR_014_V1, NETCR_014_V2, PIF_154, NETL_022_V1, NETL_022_V2, NETL_008_V1, PIF_146, PIF_119, PIF_099, PIF_162, PIF_160, PIF_113, PIF_143, NETCR_007_V1, NETCR_007_V2, PIF_137, PIF_100, NETL_004_V1, PIF_094, PIF_132, PIF_163, NETCR_003_V1, NETL_028_V1, NETL_028_V2, NETCR_013_V1, NETL_020_V1, NETL_020_V2, PIF_192, NETCR_012_V1, NETCR_012_V2, PIF_089, NETCR_002_V1, PIF_179, PIF_114, NETCR_006_V1, PIF_141, NETCR_025_V1, NETCR_025_V2, NETCR_016_V1, PIF_116, PIF_191, PIF_164, NETL_013_V1, PIF_188, PIF_195, NETCR_015_V1, PIF_102, NETL_010_V1, NETL_010_V2, NETL_001_V1, NETCR_015_V2, NETCR_005_V1, PIF_111, PIF_171, NETCR_008_V1, NETCR_008_V2, NETL_017_V1, NETL_017_V2, NETL_002_V1, NETL_002_V2, PIF_190, NETCR_009_V1, NETCR_009_V2, NETL_007_V1, PIF_112, NETCR_019_V2, NETL_012_V1, NETL_012_V2, NETL_003_V1, NETL_003_V2
Muscle.loss	Grup de pèrdua muscular (cachexic vs control)	cachexic, control, control
Metabolite	Nom complet del metabolit	X1.6.Anhydro.beta.D.glucose, X1.Methylnicotinamide, X2.Aminobutyrate, X2.Hydroxyisobutyrate, X2.Oxoglutarate, X3.Aminoisobutyrate, X3.Hydroxybutyrate, X3.Hydroxyisovalerate, X3.Indoxylsulfate, X4.Hydroxyphenylacetate, Acetate, Acetone, Adipate, Alanine, Asparagine, Betaine, Carnitine, Citrate, Creatine, Creatinine, Dimethylamine, Ethanolamine, Formate, Fucose, Fumarate, Glucose, Glutamine, Glycine, Glycolate, Guanidoacetate, Hippurate, Histidine, Hypoxanthine, Isoleucine, Lactate, Leucine, Lysine, Methylamine, Methylguanidine, N.N.Dimethylglycine, O.Acetylcarnitine, Pantothenate, Pyroglutamate, Pyruvate, Quinolinate, Serine, Succinate, Sucrose, Tartrate, Taurine, Threonine, Trigonelline, Trimethylamine.N.oxide, Tryptophan, Tyrosine, Uracil, Valine, Xylose, cis.Aconitate, myo.Inositol, trans.Aconitate, pi.Methylhistidine, tau.Methylhistidine

Taula 2:

Estadístiques dels Metabolits

Metabolite	TotalMissing	TTestPvalue	Cachexic_Mean	Control_Mean
M1	0	0.035	128.69	69.51
M10	0	0.527	119.82	99.80
M11	0	0.002	85.63	35.60
M12	0	0.273	13.35	8.42
M13	0	0.008	34.82	8.99
M14	0	0.000	347.59	157.58
M15	0	0.004	75.39	41.75
M16	0	0.002	112.25	55.97
M17	0	0.032	64.62	32.44
M18	0	0.005	2720.85	1474.72
M19	0	0.020	174.91	51.50
M2	0	0.944	70.56	73.16
M20	0	0.000	10722.14	5619.17
M21	0	0.000	453.58	208.68
M22	0	0.020	326.77	197.13
M23	0	0.005	187.56	84.48
M24	0	0.002	108.60	57.44
M25	0	0.026	10.92	4.55
M26	0	0.009	827.22	140.96
M27	0	0.000	391.41	174.43
M28	0	0.013	1069.38	585.15
M29	0	0.049	219.27	138.98
M3	0	0.008	23.67	9.53
M30	0	0.117	97.62	68.74
M31	0	0.010	2875.73	1364.24
M32	0	0.005	364.23	180.47
M33	0	0.242	67.09	51.71
M34	0	0.109	9.66	7.22
M35	0	0.057	217.63	65.75
M36	0	0.000	31.26	13.56
M37	0	0.335	121.28	89.23
M38	0	0.001	21.22	11.36
M39	0	0.203	17.36	12.13

Metabolite	TotalMissing	TTestPvalue	Cachexic_Mean	Control_Mean
M4	0	0.005	43.24	27.87
M40	0	0.000	34.49	13.60
M41	0	0.024	25.56	10.60
M42	0	0.609	39.94	52.62
M43	0	0.000	270.29	119.26
M44	0	0.007	26.87	12.57
M45	0	0.000	83.75	39.32
M46	0	0.001	245.83	122.26
M47	0	0.004	79.63	29.84
M48	0	0.066	150.02	55.58
M49	0	0.376	47.23	28.68
M5	0	0.159	183.11	85.52
M50	0	0.014	655.72	320.52
M51	0	0.001	118.23	59.52
M52	0	0.003	359.64	130.69
M53	0	0.016	820.34	388.67
M54	0	0.001	81.82	41.83
M55	0	0.004	100.74	52.01
M56	0	0.539	37.51	32.49
M57	0	0.000	45.58	20.13
M58	0	0.134	129.29	56.51
M59	0	0.001	276.03	91.72
M6	0	0.102	100.27	39.91
M60	0	0.000	181.84	62.64
M61	0	0.020	48.81	27.81
M62	0	0.100	441.55	258.64
M63	0	0.017	105.67	64.65
M7	0	0.000	29.26	9.90
M8	0	0.003	27.61	12.31
M9	0	0.005	265.16	146.38

Taula 3:

Variables més importants per a PC1, PC2 i PC3

PC1	PC2	PC3
M20	M12	M35
M27	M49	M26
M22	M39	M60
M15	M41	M34
M51	M11	M11
M57	M25	M36
M14	M3	M16
M59	M8	M57
M46	M31	M52
M24	M46	M61

Taula 4:

Correlacions > 0.8

Metabolit_1	Metabolit_2	Correlacio
M42	M2	0.8494645
M25	M5	0.8573558
M46	M7	0.8051757
M41	M12	0.8083981
M49	M12	0.8588489
M57	M14	0.8366967
M27	M15	0.8556964
M46	M15	0.8021854
M51	M15	0.8315868
M59	M18	0.8628231
M21	M20	0.8623471
M22	M20	0.8190004
M24	M20	0.8048195
M43	M20	0.8443799
M20	M21	0.8623471
M43	M21	0.8167289
M20	M22	0.8190004
M33	M22	0.8088955
M32	M23	0.8072242
M47	M23	0.8205849

Metabolit_1	Metabolit_2	Correlacio
M55	M23	0.8450884
M20	M24	0.8048195
M5	M25	0.8573558
M15	M27	0.8556964
M28	M27	0.8027463
M46	M27	0.8674403
M51	M27	0.8477669
M27	M28	0.8027463
M51	M28	0.8352743
M23	M32	0.8072242
M55	M32	0.8572686
M22	M33	0.8088955
M57	M36	0.8804148
M12	M41	0.8083981
M2	M42	0.8494645
M20	M43	0.8443799
M21	M43	0.8167289
M7	M46	0.8051757
M15	M46	0.8021854
M27	M46	0.8674403
M23	M47	0.8205849
M12	M49	0.8588489
M15	M51	0.8315868
M27	M51	0.8477669
M28	M51	0.8352743
M23	M55	0.8450884
M32	M55	0.8572686
M14	M57	0.8366967
M36	M57	0.8804148
M18	M59	0.8628231

5. Apèndix: Codi R.

Creació del SummarizedExperiment:

```
library(Biobase)
library(SummarizedExperiment)

expressionValues <- t(as.matrix(cachexia_data[, 3:65]))
# Es transposa la matriu per tenir els metabolits com files i les conc
entracions com columnes

# Metadades de les columnes: informació adicional sobre les mostres (c
olumnes 1 i 2)
colData <- data.frame(
  Patient.ID = cachexia_data$Patient.ID,  # ID del pacient
  Muscle.loss = cachexia_data$Muscle.loss,  # Pèrdua muscular
  row.names = paste0("S", 1:77)  # Nombre de las filas = Nombre de la
muestra
)

# Metadades de les files: es guarden els noms del metabolits
rowData <- data.frame(
  Metabolite = rownames(expressionValues)
)

# Per simplificar la matriu, es canvien els noms dels pacients (per id
entificació de la mostra) i dels metabolits:

colnames(expressionValues) <- paste0("S", 1:77)
# Es canvia el nom de les columnes on cada sample (S, mostra) correspo
n a un pacient
rownames(expressionValues) <- paste0("M", 1:nrow(expressionValues))
# Es canvia el nom de les files on cada una és un metabolit (M) difere
nt

# Es crea l'objecte SummarizedExperiment amb les dades preparades prèv
iament
se_cachexia <- SummarizedExperiment(
  assays = list(raw = expressionValues),
  colData = colData,
  rowData = rowData
)

# S'afegeix una nova metadada amb la descripció de lestudi
metadata(se_cachexia) <- list(study_description = "A metabolite concen
tration table from human urine samples with two groups")

# Es mostra l'objecte SummarizedExperiment creat:
se_cachexia
```

```
## class: SummarizedExperiment
## dim: 63 77
## metadata(1): study_description
## assays(1): raw
```

```
## rownames(63): M1 M2 ... M62 M63
## rowData names(1): Metabolite
## colnames(77): S1 S2 ... S76 S77
## colData names(2): Patient.ID Muscle.loss
```

La resta de codis es poden consultar al repositori GitHub. S'adjunta arxiu .Rmd amb la PAC completa, incloent la taula de les metadades acompanyades d'una breu descripció.

6. Bibliografía

https://web.ub.edu/web/actualitat/w/la-caquexia-es-un-sindrome-multiorganico-segun-un-articulo-de-nature-reviews-cancer-en-que-participan-expertos-de-la-ub

https://www.metaboanalyst.ca/docs/Format.xhtml

https://www.bioconductor.org/packages//release/bioc/vignettes/omicsViewer/inst/doc/quickStart.html

https://www.bioconductor.org/packages/devel/bioc/vignettes/SummarizedExperiment/inst/doc/SummarizedExperiment.html

https://mixomicsteam.github.io/mixOmics-Vignette/id_03.html#id_03:pca-vars