

Informe PEC1

Ignacio Monasterio Perez

06-11-2024

Tabla de contenidos

1. Abstract

2. Objetivos del Estudio

3. Materiales y Métodos

3.1 Origen y naturaleza de los datos

3.2 Herramientas y Procedimiento

4. Resultados

4.1 Exploración inicial de los datos

4.2 Creación del Contenedor SummarizedExperiment

4.3 Acceso a los Datos Experimentales

4.4 Acceso a los metadatos

4.5 Análisis Comparativo: Cachexia vs Control

4.6 Análisis exploratorio

Boxplot comparativo de un metabolito

5. Discusión y Conclusiones

1. Abstract

En este informe se realiza un análisis exploratorio del dataset de cachexia para identificar las diferencias en metabolitos entre pacientes cachexic y controles. Los principales resultados incluyen...

2. Objetivos del Estudio

Los objetivos de este estudio son:

- Realizar una exploración del dataset de cachexia.
- Identificar patrones y diferencias significativas en los metabolitos entre grupos de pacientes.
- Proporcionar un análisis inicial para futuros estudios clínicos sobre cachexia.

3. Materiales y Métodos

3.1 Origen y Naturaleza de los Datos

Fuente de los datos: El dataset utilizado en este estudio proviene del repositorio de github proporcionado en el enunciado de la PEC: <https://github.com/nutrimetabolomics/metaboData/>.

Tipo de datos: Incluye variables de pacientes (Patient_ID, Muscle_loss) y 63 metabolitos medidos en pacientes cachexic y controles.

3.2 Herramientas y Procedimiento

Este análisis se realizó en R y RStudio, y el procesamiento de datos incluyó los siguientes pasos:

- Carga y Exploración de los Datos: Lectura del dataset y visualización de su estructura básica.
- Creación del Contenedor SummarizedExperiment: Almacenamiento de los datos en un objeto SummarizedExperiment para mantener organizados los datos y los metadatos.
- Análisis Comparativo: Cálculo de medias y varianzas de los metabolitos, así como comparación entre los grupos cachexic y control.
- Visualización: Uso de gráficos exploratorios para identificar diferencias significativas.

4. Resultados

4.1 Exploración inicial de los datos

```
# Carga y lectura de los datos
datos <- read.csv("C:/Users/Nacho/Downloads/human_cachexia.csv")

# Verificamos el tamaño y las columnas
dim(datos)
```

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

```
colnames(datos)
```

```
## [1] "Patient.ID" "Muscle.loss"
## [3] "X1.6.Anhydro.beta.D.glucose" "X1.Methylnicotinamide"
## [5] "X2.Aminobutyrate" "X2.Hydroxyisobutyrate"
## [7] "X2.Oxoglutarate" "X3.Aminoisobutyrate"
## [9] "X3.Hydroxybutyrate" "X3.Hydroxyisovalerate"
## [11] "X3.Indoxylsulfate" "X4.Hydroxyphenylacetate"
## [13] "Acetate" "Acetone"
## [15] "Adipate" "Alanine"
## [17] "Asparagine" "Betaine"
## [19] "Carnitine" "Citrate"
## [21] "Creatine" "Creatinine"
## [23] "Dimethylamine" "Ethanolamine"
## [25] "Formate" "Fucose"
## [27] "Fumarate" "Glucose"
## [29] "Glutamine" "Glycine"
## [31] "Glycolate" "Guanidoacetate"
## [33] "Hippurate" "Histidine"
## [35] "Hypoxanthine" "Isoleucine"
## [37] "Lactate" "Leucine"
## [39] "Lysine" "Methylamine"
## [41] "Methylguanidine" "N.N.Dimethylglycine"
## [43] "O.Acetylcarnitine" "Pantothenate"
## [45] "Pyroglutamate" "Pyruvate"
## [47] "Quinolate" "Serine"
## [49] "Succinate" "Sucrose"
## [51] "Tartrate" "Taurine"
## [53] "Threonine" "Trigonelline"
## [55] "Trimethylamine.N.oxide" "Tryptophan"
## [57] "Tyrosine" "Uracil"
## [59] "Valine" "Xylose"
## [61] "cis.Aconitate" "myo.Inositol"
## [63] "trans.Aconitate" "pi.Methylhistidine"
## [65] "tau.Methylhistidine"
```

```
# Vistazo a las primeras filas
head(datos)
```

```
## Patient.ID Muscle.loss X1.6.Anhydro.beta.D.glucose X1.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
## 4 NETL_005_V1 cachexic 154.47 52.98
## 5 PIF_115 cachexic 22.20 73.70
## 6 PIF_110 cachexic 212.72 31.82
## X2.Aminobutyrate X2.Hydroxyisobutyrate X2.Oxoglutarate X3.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
## 4 172.43 74.44 1199.91 555.57
## 5 15.64 83.93 33.12 29.67
## 6 18.36 80.64 47.94 17.46
## X3.Hydroxybutyrate X3.Hydroxyisovalerate X3.Indoxylsulfate
## 1 56.83 10.07 566.80
## 2 43.82 79.84 368.71
```

## 3	5.64		23.34		665.14			
## 4	175.91		25.03		411.58			
## 5	76.71		69.41		165.67			
## 6	31.82		35.16		183.09			
##	X4.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	
## 4	214.86	37.34	206.44	144.03	589.93	273.14	278.66	
## 5	97.51	407.48	44.26	15.03	1118.79	42.52	391.51	
## 6	132.95	81.45	14.44	25.28	237.46	157.59	66.69	
##	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	
## 4	200.34	13629.61	85.63	20952.22	1064.22	820.57	468.72	
## 5	84.77	854.06	105.64	6768.26	242.26	365.04	114.43	
## 6	40.04	1958.63	200.34	15677.78	614.00	459.44	314.19	
##	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
## 4	407.48	96.54	862.64	1685.81	5064.45	70.81	102.51	757.48
## 5	26.05	19.69	6836.29	432.68	395.44	26.58	52.98	1152.86
## 6	123.97	5.05	512.86	298.87	482.99	428.38	57.97	3568.85
##	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	
## 4	1043.15	223.63	37.71	368.71	103.54	290.03	48.91	
## 5	327.01	66.69	40.04	3640.95	101.49	122.73	27.94	
## 6	459.44	62.80	8.17	113.30	28.79	120.30	36.97	
##	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				
## 4	141.17	40.04	254.68	42.52				
## 5	5.31	46.06	45.60	74.44				
## 6	43.38	24.29	13.46	35.52				
##	Pyroglutamate	Pyruvate	Quinolinat	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
## 4	566.80	64.07	86.49	1248.88	144.03	111.05	837.15	1525.38
## 5	184.93	12.30	38.09	206.44	68.72	75.19	4.53	468.72
## 6	432.68	32.79	112.17	387.61	33.45	336.97	24.05	2059.05
##	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		
## 4	376.15	992.27	1450.99	235.10	323.76	30.57		
## 5	64.07	86.49	172.43	103.54	142.59	44.26		
## 6	105.64	862.64	880.07	239.85	127.74	29.67		
##	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
## 4 102.51 125.21 1863.11 247.15 75.94 249.64
## 5 160.77 186.79 101.49 749.95 98.49 84.77
## 6 36.97 89.12 287.15 129.02 121.51 399.41
## tau.Methylhistidine
## 1 160.77
## 2 130.32
## 3 83.93
## 4 254.68
## 5 79.84
## 6 68.72
```

```
# Resumen estadístico
summary(datos)
```

```
## Patient.ID      Muscle.loss      X1.6.Anhydro.beta.D.glucose
## Length:77      Length:77      Min. : 4.71
## Class :character Class :character 1st Qu.: 28.79
## Mode :character Mode :character Median : 45.60
##                                     Mean :105.63
##                                     3rd Qu.:141.17
##                                     Max. :685.40
## X1.Methylnicotinamide X2.Aminobutyrate X2.Hydroxyisobutyrate X2.Oxoglutarate
## Min. : 6.42 Min. : 1.28 Min. : 4.85 Min. : 5.53
## 1st Qu.: 15.80 1st Qu.: 5.26 1st Qu.:15.80 1st Qu.: 22.42
## Median : 36.60 Median : 10.49 Median :32.46 Median : 55.15
## Mean : 71.57 Mean : 18.16 Mean :37.25 Mean : 145.09
## 3rd Qu.: 73.70 3rd Qu.: 19.49 3rd Qu.:54.60 3rd Qu.: 92.76
## Max. :1032.77 Max. :172.43 Max. :93.69 Max. :2465.13
## X3.Aminoisobutyrate X3.Hydroxybutyrate X3.Hydroxyisovalerate X3.Indoxylsulfate
## Min. : 2.61 Min. : 1.70 Min. : 0.92 Min. : 27.66
## 1st Qu.: 11.70 1st Qu.: 5.99 1st Qu.: 5.26 1st Qu.: 82.27
## Median : 22.65 Median : 11.70 Median : 12.55 Median : 144.03
## Mean : 76.76 Mean : 21.72 Mean : 21.65 Mean : 218.88
## 3rd Qu.: 56.26 3rd Qu.: 29.96 3rd Qu.: 30.27 3rd Qu.: 333.62
## Max. :1480.30 Max. :175.91 Max. :164.02 Max. :1043.15
## X4.Hydroxyphenylacetate Acetate Acetone Adipate
## Min. : 15.49 Min. : 3.49 Min. : 2.29 Min. : 1.55
## 1st Qu.: 41.68 1st Qu.: 16.28 1st Qu.: 4.95 1st Qu.: 6.11
## Median : 70.11 Median : 39.65 Median : 7.10 Median : 10.18
## Mean :112.02 Mean : 66.14 Mean : 11.43 Mean : 24.76
## 3rd Qu.:145.47 3rd Qu.: 86.49 3rd Qu.: 10.49 3rd Qu.: 19.11
## Max. :796.32 Max. :411.58 Max. :206.44 Max. :327.01
## Alanine Asparagine Betaine Carnitine
## Min. : 16.78 Min. : 6.69 Min. : 2.29 Min. : 2.18
## 1st Qu.: 78.26 1st Qu.: 20.49 1st Qu.: 28.79 1st Qu.: 14.44
## Median : 194.42 Median : 42.10 Median : 64.72 Median : 23.81
## Mean : 273.56 Mean : 62.28 Mean : 90.32 Mean : 52.09
## 3rd Qu.: 399.41 3rd Qu.: 89.12 3rd Qu.:127.74 3rd Qu.: 60.95
## Max. :1312.91 Max. :273.14 Max. :391.51 Max. :487.85
## Citrate Creatine Creatinine Dimethylamine
## Min. : 59.74 Min. : 2.75 Min. : 1002 Min. : 41.26
```

## 1st Qu.: 788.40	1st Qu.: 17.64	1st Qu.: 3498	1st Qu.: 142.59
## Median : 1790.05	Median : 44.26	Median : 7631	Median : 304.90
## Mean : 2235.35	Mean : 126.83	Mean : 8734	Mean : 358.17
## 3rd Qu.: 3071.74	3rd Qu.: 117.92	3rd Qu.:12333	3rd Qu.: 454.86
## Max. :13629.61	Max. :1863.11	Max. :33860	Max. :1556.20
## Ethanolamine	Formate	Fucose	Fumarate
## Min. : 16.12	Min. : 6.42	Min. : 5.70	Min. : 0.79
## 1st Qu.: 86.49	1st Qu.: 53.52	1st Qu.: 29.37	1st Qu.: 2.23
## Median : 204.38	Median : 95.58	Median : 61.56	Median : 4.10
## Mean : 276.26	Mean : 147.40	Mean : 88.67	Mean : 8.44
## 3rd Qu.: 407.48	3rd Qu.: 167.34	3rd Qu.:123.97	3rd Qu.: 7.85
## Max. :1436.55	Max. :1480.30	Max. :407.48	Max. :96.54
## Glucose	Glutamine	Glycine	Glycolate
## Min. : 26.84	Min. : 23.34	Min. : 38.09	Min. : 5.42
## 1st Qu.: 80.64	1st Qu.: 113.30	1st Qu.: 262.43	1st Qu.: 50.91
## Median : 210.61	Median : 225.88	Median : 528.48	Median :130.32
## Mean : 559.85	Mean : 306.87	Mean : 880.72	Mean :187.99
## 3rd Qu.: 407.48	3rd Qu.: 445.86	3rd Qu.:1096.63	3rd Qu.:267.74
## Max. :8690.62	Max. :1685.81	Max. :5064.45	Max. :720.54
## Guanidoacetate	Hippurate	Histidine	Hypoxanthine
## Min. : 7.03	Min. : 92.76	Min. : 14.15	Min. : 3.78
## 1st Qu.: 33.78	1st Qu.: 492.75	1st Qu.: 66.69	1st Qu.: 20.70
## Median : 64.72	Median : 1224.15	Median : 174.16	Median : 40.04
## Mean : 86.37	Mean : 2286.84	Mean : 292.64	Mean : 61.10
## 3rd Qu.:108.85	3rd Qu.: 2921.93	3rd Qu.: 419.89	3rd Qu.: 83.93
## Max. :561.16	Max. :19341.34	Max. :1863.11	Max. :265.07
## Isoleucine	Lactate	Leucine	Lysine
## Min. : 1.790	Min. : 7.32	Min. : 2.51	Min. : 10.49
## 1st Qu.: 3.900	1st Qu.: 35.52	1st Qu.: 9.12	1st Qu.: 30.27
## Median : 7.170	Median : 81.45	Median : 19.11	Median : 69.41
## Mean : 8.709	Mean : 158.46	Mean : 24.36	Mean :108.79
## 3rd Qu.:11.250	3rd Qu.: 139.77	3rd Qu.: 31.19	3rd Qu.:121.51
## Max. :40.040	Max. :3640.95	Max. :103.54	Max. :788.40
## Methylamine	Methylguanidine	N.N.Dimethylglycine	O.Acetylcarnitine
## Min. : 1.51	Min. : 1.70	Min. : 0.79	Min. : 1.23
## 1st Qu.: 5.26	1st Qu.: 4.26	1st Qu.: 7.03	1st Qu.: 3.94
## Median :14.73	Median : 7.85	Median : 21.98	Median : 11.47
## Mean :17.38	Mean : 15.32	Mean : 26.35	Mean : 19.73
## 3rd Qu.:24.05	3rd Qu.: 19.30	3rd Qu.: 40.04	3rd Qu.: 20.91
## Max. :52.46	Max. :141.17	Max. :120.30	Max. :254.68
## Pantothenate	Pyroglutamate	Pyruvate	Quinolinate
## Min. : 2.59	Min. : 21.33	Min. : 0.90	Min. : 5.21
## 1st Qu.: 11.13	1st Qu.: 68.72	1st Qu.: 4.85	1st Qu.: 26.58
## Median : 22.65	Median : 157.59	Median : 13.46	Median : 51.42
## Mean : 44.88	Mean : 211.45	Mean : 21.29	Mean : 66.44
## 3rd Qu.: 41.26	3rd Qu.: 301.87	3rd Qu.: 29.08	3rd Qu.: 87.36
## Max. :692.29	Max. :1064.22	Max. :184.93	Max. :259.82
## Serine	Succinate	Sucrose	Tartrate
## Min. : 16.12	Min. : 1.72	Min. : 6.49	Min. : 2.20
## 1st Qu.: 83.10	1st Qu.: 8.58	1st Qu.: 19.30	1st Qu.: 6.89
## Median : 142.59	Median : 30.88	Median : 40.85	Median : 12.94
## Mean : 197.69	Mean : 60.23	Mean : 113.23	Mean : 40.00
## 3rd Qu.: 270.43	3rd Qu.: 74.44	3rd Qu.: 94.63	3rd Qu.: 25.79
## Max. :1248.88	Max. :589.93	Max. :2079.74	Max. :837.15

##	Taurine	Threonine	Trigonelline	Trimethylamine.N.oxide
##	Min. : 17.81	Min. : 8.25	Min. : 10.07	Min. : 55.7
##	1st Qu.: 99.48	1st Qu.: 31.82	1st Qu.: 53.52	1st Qu.: 175.9
##	Median : 249.64	Median : 64.07	Median : 114.43	Median : 383.8
##	Mean : 525.12	Mean : 95.36	Mean : 270.44	Mean : 652.2
##	3rd Qu.: 665.14	3rd Qu.:137.00	3rd Qu.: 340.36	3rd Qu.: 735.1
##	Max. :4272.69	Max. :450.34	Max. :2252.96	Max. :5486.2
##	Tryptophan	Tyrosine	Uracil	Valine
##	Min. : 8.67	Min. : 4.22	Min. : 3.10	Min. : 4.10
##	1st Qu.: 21.33	1st Qu.: 23.57	1st Qu.: 11.94	1st Qu.: 12.18
##	Median : 46.99	Median : 60.34	Median : 27.39	Median : 33.12
##	Mean : 66.24	Mean : 81.76	Mean : 35.56	Mean : 35.67
##	3rd Qu.: 96.54	3rd Qu.:113.30	3rd Qu.: 44.26	3rd Qu.: 50.40
##	Max. :259.82	Max. :539.15	Max. :179.47	Max. :160.77
##	Xylose	cis.Aconitate	myo.Inositol	trans.Aconitate
##	Min. : 10.07	Min. : 12.94	Min. : 11.59	Min. : 4.90
##	1st Qu.: 29.96	1st Qu.: 36.23	1st Qu.: 30.27	1st Qu.: 12.43
##	Median : 50.40	Median : 129.02	Median : 78.26	Median : 26.84
##	Mean : 100.93	Mean : 204.22	Mean :135.40	Mean : 40.63
##	3rd Qu.: 89.12	3rd Qu.: 254.68	3rd Qu.:167.34	3rd Qu.: 57.40
##	Max. :2164.62	Max. :1863.11	Max. :854.06	Max. :217.02
##	pi.Methylhistidine	tau.Methylhistidine		
##	Min. : 11.36	Min. : 8.00		
##	1st Qu.: 67.36	1st Qu.: 27.39		
##	Median : 162.39	Median : 68.72		
##	Mean : 370.29	Mean : 89.69		
##	3rd Qu.: 387.61	3rd Qu.:130.32		
##	Max. :2697.28	Max. :317.35		

4.2 Creación del Contenedor SummarizedExperiment

```
# Separar los datos experimentales (metabolitos) desde la columna 3 en adelante
exp_data <- as.matrix(datos[, 3:ncol(datos)]) # Matriz de los metabolitos

# Transponer la matriz para que los metabolitos sean las filas y los pacientes las
# columnas
exp_data <- t(exp_data) # Transponemos la matriz

# Asignar nombres de filas (metabolitos) y nombres de columnas (Patient_ID)
rownames(exp_data) <- colnames(datos)[3:ncol(datos)] # Metabolitos como filas
colnames(exp_data) <- datos[, 1] # Pacientes como columnas (Patient_ID)

# Crear un dataframe para los metadatos, con Patient_ID como nombres de fila
metadata <- data.frame(Muscle_loss = datos[, 2]) # Solo la columna Muscle_loss
rownames(metadata) <- datos[, 1] # Asignamos Patient_ID como nombres de fila

# Verificamos que las dimensiones coincidan
if (ncol(exp_data) == nrow(metadata)) {
  print("Las dimensiones son correctas.")
} else {
  stop("Error: Las dimensiones no coinciden.")
}
```

```
## [1] "Las dimensiones son correctas."
```

```
# Crear el contenedor SummarizedExperiment  
library(SummarizedExperiment)
```

```
## Warning: package 'matrixStats' was built under R version 4.3.3
```

```
## Warning: package 'S4Vectors' was built under R version 4.3.2
```

```
## Warning: package 'GenomeInfoDb' was built under R version 4.3.3
```

```
se <- SummarizedExperiment(  
  assays = list(counts = exp_data), # Datos experimentales (transpuestos)  
  colData = metadata # Metadatos (Muscle_loss con Patient_ID como rownames)  
)
```

```
# Revisar el contenedor  
se
```

```
## class: SummarizedExperiment  
## dim: 63 77  
## metadata(0):  
## assays(1): counts  
## rownames(63): X1.6.Anhydro.beta.D.glucose X1.Methylnicotinamide ...  
## pi.Methylhistidine tau.Methylhistidine  
## rowData names(0):  
## colnames(77): PIF_178 PIF_087 ... NETL_003_V1 NETL_003_V2  
## colData names(1): Muscle_loss
```

4.3 Acceso a los Datos Experimentales

```
# Extraer la matriz de datos experimentales  
metabolitos <- assay(se)  
  
# Calcular la media de cada metabolito (filas)  
metabolito_means <- rowMeans(metabolitos)  
  
# Visualizar los metabolitos con mayor media  
head(sort(metabolito_means, decreasing = TRUE))
```

```
##          Creatinine          Hippurate          Citrate  
##          8733.9718          2286.8377          2235.3460  
##          Glycine Trimethylamine.N.oxide          Glucose  
##          880.7174          652.1569          559.8445
```

```
# Calcular la varianza de cada metabolito  
metabolito_vars <- apply(metabolitos, 1, var)  
  
# Visualizar los metabolitos con mayor varianza  
head(sort(metabolito_vars, decreasing = TRUE))
```



```
##          Creatinine          Hippurate          Citrate
##          41959525.4          8238722.0          4694015.7
##          Glucose          Glycine Trimethylamine.N.oxide
##          1923603.6          901916.6          828372.4
```

4.4 Acceso a los metadatos

```
# Extraer los metadatos
metadata <- colData(se)

# Contar el número de pacientes cachexic y control
table(metadata$Muscle_loss)
```

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

```
# Comparar medias de un metabolito (ejemplo con el primer metabolito) entre grupos
tapply(metabolitos[1, ], metadata$Muscle_loss, mean)
```

```
## cachexic control
## 128.68894 69.50533
```

4.5 Análisis Comparativo: Cachexia vs Control

```
# Comparar las medias de todos los metabolitos entre cachexic y control
group_means <- apply(metabolitos, 1, function(met) {
  tapply(met, metadata$Muscle_loss, mean)
})

# Mostrar las primeras comparaciones
head(group_means)
```

```
##          X1.6.Anhydro.beta.D.glucose X1.Methylnicotinamide X2.Aminobutyrate
## cachexic          128.68894          70.56426          23.669149
## control          69.50533          73.15500          9.528333
##          X2.Hydroxyisobutyrate X2.Oxoglutarate X3.Aminoisobutyrate
## cachexic          43.23766          183.11043          100.2747
## control          27.87100          85.51733          39.9110
##          X3.Hydroxybutyrate X3.Hydroxyisovalerate X3.Indoxylsulfate
## cachexic          29.260638          27.60638          265.1577
## control          9.898667          12.31267          146.3763
##          X4.Hydroxyphenylacetate Acetate Acetone Adipate Alanine
## cachexic          119.82255 85.63298 13.34638 34.817872 347.5911
## control          99.79867 35.60467 8.42000 8.993333 157.5840
##          Asparagine Betaine Carnitine Citrate Creatine Creatinine
## cachexic 75.39085 112.25298 64.62213 2720.853 174.91340 10722.140
## control 41.74900 55.97033 32.44367 1474.719 51.50433 5619.175
##          Dimethylamine Ethanolamine Formate Fucose Fumarate Glucose
```

```
## cachexic      453.5806      326.7721 187.56447 108.59915 10.92191 827.2189
## control      208.6833      197.1253  84.48333  57.44467  4.55200 140.9580
##              Glutamine  Glycine Glycolate Guanidoacetate Hippurate Histidine
## cachexic    391.4104 1069.3779 219.2696      97.62426 2875.73 364.2323
## control    174.4273 585.1493 138.9837      68.73967 1364.24 180.4723
##              Hypoxanthine Isoleucine Lactate Leucine Lysine Methylamine
## cachexic     67.08702  9.660851 217.63191 31.26170 121.28234 21.21638
## control     51.71433  7.218000  65.74833 13.55667  89.22933 11.36000
##              Methylguanidine N.N.Dimethylglycine O.Acetylcarnitine Pantothenate
## cachexic      17.36468      34.48979      25.56447      39.94404
## control      12.12833      13.59667      10.59800      52.62267
##              Pyroglutamate Pyruvate Quinolinolate Serine Succinate Sucrose
## cachexic     270.2923 26.86553  83.74723 245.8298  79.62894 150.02447
## control     119.2580 12.56633  39.32400 122.2630  29.83600  55.57967
##              Tartrate Taurine Threonine Trigonelline Trimethylamine.N.oxide
## cachexic    47.23468 655.7200 118.23319  359.6377      820.3406
## control    28.67600 320.5223  59.51867  130.6870      388.6690
##              Tryptophan Tyrosine Uracil Valine Xylose cis.Aconitate
## cachexic    81.82404 100.7423 37.51362 45.58255 129.28915 276.0255
## control    41.83300  52.0140 32.49333 20.13267  56.50933  91.7240
##              myo.Inositol trans.Aconitate pi.Methylhistidine tau.Methylhistidine
## cachexic    181.83766  48.81404      441.5532      105.66766
## control     62.64133  27.80933      258.6400      64.65033
```

Si el experimento tiene información adicional sobre cómo se recopilieron los datos, se puede almacenar en el componente `metadata()`.

```
# Añadir información al componente metadata
metadata(se) <- list(experiment_info = "Datos de metabolitos de pacientes con cachexia")

# Ver la metadata
metadata(se)
```

```
## $experiment_info
## [1] "Datos de metabolitos de pacientes con cachexia"
```

4.6 Análisis exploratorio

```
library(ggplot2)
```

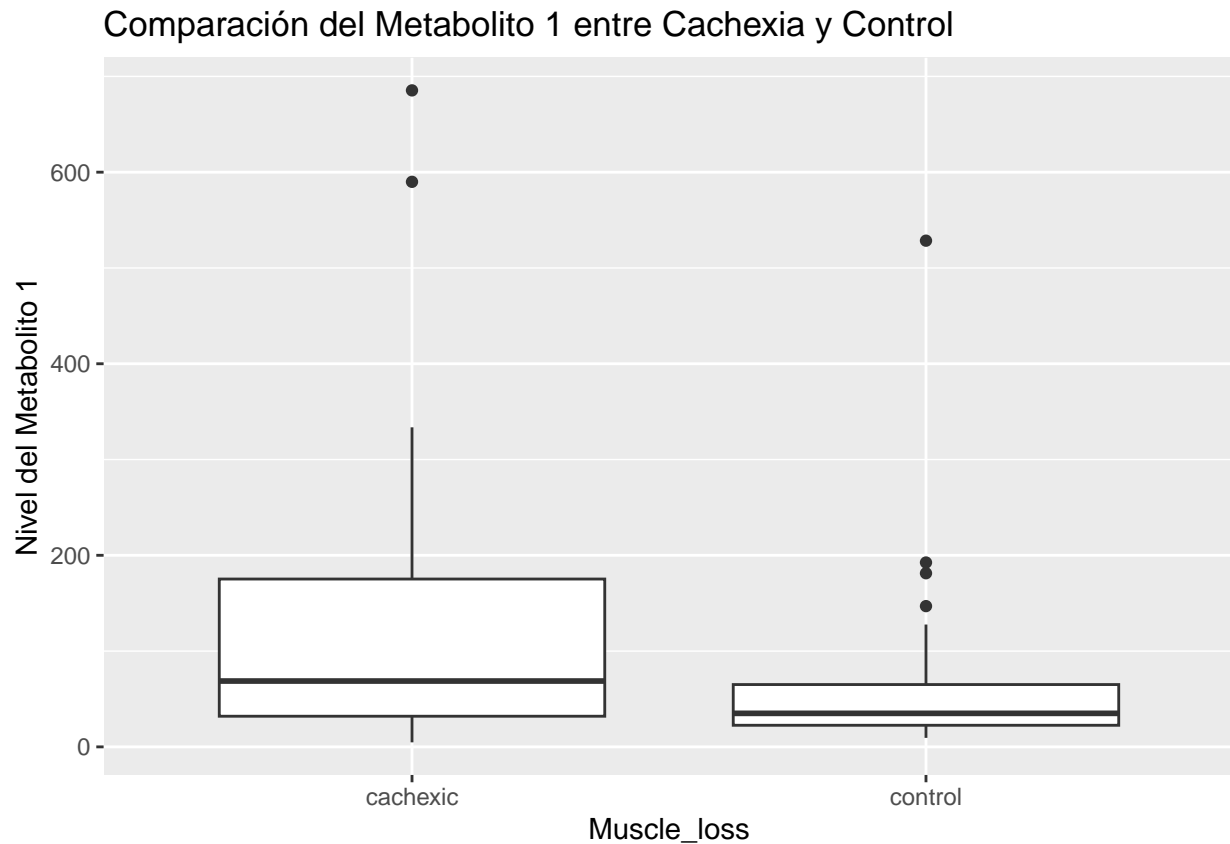
```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

Boxplot comparativo de un metabolito

```
# Crear un dataframe para ggplot2
metabolito_df <- data.frame(
  Patient_ID = colnames(metabolitos),
  Muscle_loss = metadata$Muscle_loss,
  Metabolito1 = metabolitos[1, ] # Primer metabolito (por ejemplo)
```

```
)

# Crear un boxplot para comparar entre cachexic y control
ggplot(metabolito_df, aes(x = Muscle_loss, y = Metabolito1)) +
  geom_boxplot() +
  labs(title = "Comparación del Metabolito 1 entre Cachexia y Control",
       y = "Nivel del Metabolito 1")
```



La diferencia en los niveles de “Metabolito 1” entre ambos grupos es bastante clara:

- Pacientes con Cachexia: Tienen un rango de niveles de metabolito mucho más amplio y valores atípicos elevados. Esto podría indicar una posible relación entre niveles elevados de este metabolito y la presencia de cachexia.
- Pacientes en el Grupo Control: Muestran niveles de “Metabolito 1” más bajos y menos variables, lo que podría sugerir que niveles altos de este metabolito son menos comunes o inexistentes en personas sin cachexia.

5. Discusión y Conclusiones

La comparación de medias sugiere una fuerte asociación entre niveles elevados de ciertos metabolitos y la presencia de cachexia. Los pacientes con esta condición muestran consistentemente niveles más altos en múltiples metabolitos en comparación con el grupo control. Esto podría interpretarse como una señal de que la cachexia está relacionada con alteraciones significativas en el metabolismo de los pacientes, reflejadas en concentraciones elevadas de varias sustancias, muchas de las cuales están involucradas en el metabolismo energético y proteico.

Estos resultados refuerzan la hipótesis inicial: la cachexia podría provocar un cambio en el perfil metabólico, posiblemente debido a un aumento en el catabolismo (degradación) de nutrientes o a alteraciones en los procesos de síntesis. La diferencia de estos niveles podría ser útil como biomarcador para el diagnóstico o monitoreo de cachexia, aunque sería necesario realizar pruebas estadísticas adicionales (como un test t o ANOVA) para determinar si estas diferencias son estadísticamente significativas. Además, la evaluación de la variabilidad y la identificación de posibles valores atípicos podrían aportar más precisión en el uso clínico de estos metabolitos como indicadores de la progresión de la enfermedad.

En resumen, este análisis comparativo subraya el impacto de la cachexia en los niveles de numerosos metabolitos, lo que sugiere una afectación sistemática en los pacientes con esta condición en comparación con aquellos sin ella.

Repositorio en GitHub

https://github.com/ignaciomonasterio/Monasterio-Perez-Ignacio_PEC1.git