

PEC 1 - Datos ómicos

[Code ▼](#)

Merce Merayo Pastor

2024-10-26

- 1 Resumen ejecutivo
- 2 Objetivos del Estudio
- 3 Materiales y Métodos
 - 3.1 Origen y naturaleza de los datos
 - 3.2 Herramientas bioinformáticas utilizadas
- 4 Resultados
 - 4.1 Carga de los datos y metadatos desde www.metabolomicsworkbench.org
 - 4.2 Preparación de los datos para el análisis
 - 4.2.1 Pasos durante la preparación de los datos
 - 4.2.2 Creación del objeto `SummarizedExperiment`
 - 4.3 Análisis multivariante, PCA y clustering
 - 4.3.1 Estructura de datos con los metabolitos
 - 4.3.2 Análisis de colinealidad
 - 4.3.3 Análisis de Clustering
 - 4.4 Análisis PCA
 - 4.4.1 Variabilidad explicada por componente
 - 4.4.2 Proyección de metabolitos en el espacio de PC
- 5 Discusión, limitaciones y conclusiones del estudio
- 6 Referencias
 - 6.1 GitHub

[Show](#)

1 Resumen ejecutivo

Este informe presenta un análisis del dataset de metabolómica correspondiente al estudio **ST003471** del repositorio **MetabolomicsWorkbench**, parte del proyecto **PR002130**. El estudio, titulado "*Gut Microbiota Regulates Stress Responsivity via the Circadian System - Mice Hippocampus*", examina cómo la microbiota intestinal influye en la regulación de la respuesta al estrés a través del sistema circadiano, utilizando como modelo el hipocampo de ratones.

La investigación original revela que el estrés y los sistemas circadianos están interconectados a través del eje hipotálamo-hipofisario-adrenal (HPA), el cual regula respuestas a estímulos externos. En este contexto, se descubrió que la microbiota intestinal actúa como un regulador clave de la función rítmica del eje HPA. Los hallazgos sugieren que el agotamiento de la microbiota altera significativamente tanto el transcriptoma como el metaboloma cerebral en regiones como el hipocampo y la amígdala, especialmente en vías relacionadas con la respuesta al estrés. Este efecto incluye una disrupción en los ritmos de secreción de glucocorticoides, lo que afecta las respuestas adaptativas al estrés en diferentes momentos del día. [1]

En el análisis actual, se emplean herramientas de bioinformática en **R**, incluyendo **Bioconductor** y los paquetes **SummarizedExperiment** y **mixOmics**, para realizar un análisis de componentes principales (PCA) y clustering, para poder explorar la variabilidad entre muestras y la identificación de posibles patrones de agrupamiento. Los resultados obtenidos ofrecen una visión de las principales fuentes de variación en los datos, lo que permite entender mejor cómo la microbiota intestinal podría modular las respuestas al estrés de forma circadiana.

Este trabajo demuestra la utilidad de las técnicas de análisis multivariante en bioinformática y cómo estas pueden contribuir a la comprensión de la influencia de la microbiota en la regulación de la respuesta al estrés en estudios metabolómicos.

2 Objetivos del Estudio

El objetivo de este análisis es realizar una exploración de los datos de metabolómica del hipocampo de ratones para:

1. Identificar y cuantificar las principales fuentes de variación entre las muestras mediante análisis de componentes principales (PCA), evaluando cómo estas variaciones podrían relacionarse con la influencia de la microbiota en la regulación de la respuesta al estrés de forma circadiana.
2. Comprobar si existe algún agrupamiento de interés entre muestras mediante técnicas de clustering, lo que permitiría detectar patrones en los metabolitos que puedan asociarse con las diferencias en los ritmos circadianos y la respuesta al estrés.
3. Utilizar herramientas bioinformáticas para proporcionar un análisis reproducible de los datos metabolómicos, con la finalidad de comprender la regulación circadiana de la respuesta al estrés relacionada con la microbiota.

Este estudio busca demostrar cómo los métodos multivariantes pueden mostrar patrones en los datos de metabolómica que reflejan interacciones biológicas complejas.

3 Materiales y Métodos

3.1 Origen y naturaleza de los datos

El dataset utilizado en este análisis proviene del estudio **ST003471** en el repositorio **MetabolomicsWorkbench**, que es parte del proyecto **PR002130** titulado “*Gut Microbiota Regulates Stress Responsivity via the Circadian System - Mice Hippocampus*”. Este estudio se centra en la relación entre la microbiota intestinal y la regulación de la respuesta al estrés en ratones, a través del sistema circadiano.

El conjunto de datos incluye medidas de abundancia relativa de diversos metabolitos en el hipocampo de ratones, una región del cerebro implicada en el procesamiento de la respuesta al estrés. Las muestras fueron recolectadas a lo largo de diferentes momentos del día para observar cómo la microbiota podría modular las respuestas al estrés de manera circadiana. Estos datos permiten una exploración de las vías metabólicas y cómo varían en función de la presencia o ausencia de microbiota, ofreciendo una ventana para entender la ritmicidad y la influencia de factores externos en los sistemas de respuesta al estrés. [1]

3.2 Herramientas bioinformáticas utilizadas

Se utilizaron varias herramientas en el entorno de programación **R**, incluyendo:

- **Bioconductor**: plataforma que proporciona acceso a paquetes bioinformáticos específicos para el análisis de datos ómicos.
- **SummarizedExperiment**: paquete empleado para organizar los datos metabolómicos y sus metadatos en un solo contenedor, permitiendo una manipulación estructurada y coherente.
- **mixOmics**: utilizado para llevar a cabo análisis de componentes principales (PCA) y clustering, métodos multivariantes clave para explorar patrones de agrupamiento en los datos.

Estas herramientas han sido necesarias para estructurar y analizar los datos de forma reproducible, además de explorar la variabilidad de las muestras y visualizar patrones relacionados con la respuesta al estrés y el sistema circadiano.

4 Resultados

4.1 Carga de los datos y metadatos desde www.metabolomicsworkbench.org

Show

Metabolite_name <chr>	RefMet_name <chr>
1 Factors	-
2 1-Methyl Adenosine	1-Methyladenosine
3 2-Oxoglutaric acid	2-Oxoglutaric acid
4 3-Methylhistidine	3-Methylhistidine
5 4-Guanidinobutyric acid	4-Guanidinobutanoic acid
6 4-Hydroxybenzaldehyde	4-Hydroxybenzaldehyde

6 rows | 1-3 of 6 columns

Show

En los datos recuperados hay información sobre los metabolitos considerados en el estudio además de un segundo fichero con todos los metadatos de información sobre el proyecto. Por otro lado, hay identificadores de las muestras que incluyen información sobre el tratamiento en el propio identificador. Adicionalmente, hay información sobre el origen anatómico de la muestras, en este caso del hipocampo, tipo de tratamiento y el y el tiempo *Zeitgeber* (*Zeitgeber_time*). Este último, en estudios circadianos, hace referencia al tiempo en función de un sincronizador ambiental, normalmente el ciclo luz/oscuridad, donde “ZT” se refiere a la hora en relación con el comienzo del periodo de luz (ZT11) [2].

4.2 Preparación de los datos para el análisis

Para el análisis, primero es necesario realizar algunos pasos de limpieza y transformación de los datos. En particular, vamos a eliminar columnas innecesarias y extraer los detalles de las muestras de forma estructurada.

Show

Se extraen los detalles de las muestras, como el origen y tipo de tratamiento, para poder utilizar esta información más adelante.

Creación de la matriz de metabolitos: En este paso, se transforman los datos de los metabolitos en una matriz en la que las filas representan cada metabolito y las columnas representan cada muestra y se convierten los valores a formato numérico para facilitar el análisis.

## [1] "1-Methyl Adenosine"	"2-Oxoglutaric acid"
## [3] "3-Methylhistidine"	"4-Guanidinobutyric acid"
## [5] "4-Hydroxybenzaldehyde"	"7-Methylguanine"
## [7] "Acetylcarnitine"	"Acetylcholine"
## [9] "Acetylhistidine"	"Adenine"
## [11] "Adenosine"	"Alanine"
## [13] "Aldopentose"	"AMP"
## [15] "Anserine"	"Arginine"
## [17] "Asparagine"	"Aspartic acid"
## [19] "Benzaldehyde"	"Betaine"
## [21] "Butyrylcarnitine"	"Carnitine"
## [23] "Carnosine"	"CDP-choline"
## [25] "CDP-ethanolamine"	"Choline"
## [27] "cis-Aconitic acid"	"Citrulline"
## [29] "CMP"	"Creatine"
## [31] "Creatinine"	"Cysteine"
## [33] "Cystine"	"Cytidine"
## [35] "Cytosine"	"Dihydroxyacetone-phosphate"
## [37] "Dulcitol/Mannitol/Sorbitol"	"Ethylmalonic acid/Glutaric acid"
## [39] "Fumaric acid"	"Galactonic acid/Gluconic acid"
## [41] "g-Aminobutyric acid"	"Glutamic acid"
## [43] "Glutamine"	"Glutathione"
## [45] "Glutathione oxidized"	"Glycine"
## [47] "GMP"	"Guanidoacetic acid"
## [49] "Guanine"	"Guanosine"
## [51] "Hexose"	"Histidine"
## [53] "Homoarginine"	"Hydroxyphenyllactic acid"
## [55] "Hypoxanthine"	"Imidazole"
## [57] "IMP"	"Indole-3-carboxyaldehyde"
## [59] "Inosine"	"Isocitric acid"
## [61] "Isoleucine"	"Isovalerylcarnitine"
## [63] "Leucine"	"Lysine"
## [65] "Malic acid"	"Methionine"
## [67] "Methionine sulfoxide"	"Mucic acid"
## [69] "N-Acetyl-alanine"	"N-Acetyl-arginine"
## [71] "N-Acetyl-aspartic acid"	"N-Acetyl-glucosamine"
## [73] "N-Acetyl-glutamic acid"	"N-Acetyl-glutamine"
## [75] "N-Acetyl-glycine"	"N-Acetyl-lysine"
## [77] "N-acetyl-methionine"	"N-Acetylneuraminic acid"
## [79] "N-Acetyl-ornithine"	"Nicotine amide"
## [81] "O-Phosphorylethanolamine"	"Ornithine"
## [83] "Pantothenic acid"	"Phenylalanine"
## [85] "Phosphocholine"	"Phosphoenolpyruvic acid"
## [87] "Pinene"	"Pipicolinic acid"
## [89] "Proline"	"Propionylcarnitine"
## [91] "Putrescine"	"Pyridoxal"
## [93] "Pyrogallol"	"Pyroglutamic acid"
## [95] "Ribulose-5-phosphate"	"Serine"
## [97] "Serine phosphate"	"Spermine"
## [99] "Sugar alcohol"	"Taurine"
## [101] "Threonic acid"	"Threonine"
## [103] "Trigonelline"	"Tryptophan"
## [105] "Tyrosine"	"UMP"
## [107] "Uracil"	"Uridine"
## [109] "Urocanic acid"	"Valine/5-Amino valeric acid"
## [111] "Valproic acid"	"Xanthine"

```
## [1] "ABX_62_ZT11" "ABX_63_ZT11" "ABX_64_ZT11" "ABX_65_ZT11" "ABX_87_ZT11"
## [6] "ABX_88_ZT11" "ABX_89_ZT11" "ABX_90_ZT11" "ABX_104_ZT17" "ABX_105_ZT17"
## [11] "ABX_106_ZT17" "ABX_107_ZT17" "ABX_108_ZT17" "ABX_119_ZT17" "ABX_121_ZT17"
## [16] "ABX_109_ZT23" "ABX_110_ZT23" "ABX_111_ZT23" "ABX_112_ZT23" "ABX_113_ZT23"
## [21] "ABX_114_ZT23" "ABX_115_ZT23" "ABX_116_ZT23" "ABX_72_ZT5" "ABX_73_ZT5"
## [26] "ABX_74_ZT5" "ABX_75_ZT5" "ABX_76_ZT5" "ABX_77_ZT5" "ABX_84_ZT5"
## [31] "ABX_86_ZT5" "CV_101_ZT11" "CV_102_ZT11" "CV_16_ZT11" "CV_17_ZT11"
## [36] "CV_18_ZT11" "CV_19_ZT11" "CV_20_ZT11" "CV_122_ZT17" "CV_123_ZT17"
## [41] "CV_66_ZT17" "CV_67_ZT17" "CV_68_ZT17" "CV_69_ZT17" "CV_70_ZT17"
## [46] "CV_71_ZT17" "CV_117_ZT23" "CV_118_ZT23" "CV_78_ZT23" "CV_79_ZT23"
## [51] "CV_80_ZT23" "CV_81_ZT23" "CV_82_ZT23" "CV_83_ZT23" "CV_10_ZT5"
## [56] "CV_11_ZT5" "CV_12_ZT5" "CV_13_ZT5" "CV_6_ZT5" "CV_7_ZT5"
## [61] "CV_8_ZT5" "CV_9_ZT5" "GF_32_ZT11" "GF_36_ZT11" "GF_41_ZT11"
## [66] "GF_42_ZT11" "GF_43_ZT11" "GF_45_ZT11" "GF_47_ZT11" "GF_48_ZT11"
## [71] "GF_91_ZT17" "GF_92_ZT17" "GF_93_ZT17" "GF_94_ZT17" "GF_95_ZT17"
## [76] "GF_96_ZT17" "GF_97_ZT17" "GF_98_ZT17" "GF_53_ZT23" "GF_54_ZT23"
## [81] "GF_55_ZT23" "GF_56_ZT23" "GF_57_ZT23" "GF_59_ZT23" "GF_60_ZT23"
## [86] "GF_61_ZT23" "GF_24_ZT5" "GF_25_ZT5" "GF_26_ZT5" "GF_27_ZT5"
## [91] "GF_28_ZT5" "GF_49_ZT5" "GF_51_ZT5" "GF_52_ZT5"
```

[Show](#)

4.2.1 Pasos durante la preparación de los datos

Creación de la matriz de expresión de metabolitos y datos de las muestras: Se genera una matriz que organiza los valores de expresión de los metabolitos para cada muestra, junto con un `data.frame` que contiene información de cada muestra, como su tratamiento y otros metadatos.

[Show](#)

Limpieza y organización de la información de las muestras: La información de cada muestra se organiza en un formato compatible con Bioconductor . Esto incluye separar datos clave (origen, tratamiento y tiempo circadiano) y limpiar etiquetas innecesarias para facilitar el análisis.

[Show](#)

4.2.2 Creación del objeto `SummarizedExperiment`

Se construye un objeto `SummarizedExperiment` que integra los datos de expresión, información del proyecto, información de las muestras y detalles de los metabolitos.

[Show](#)

```
## class: SummarizedExperiment
## dim: 112 94
## metadata(2): info_muestras info_proyecto
## assays(1): counts
## rownames(112): 1-Methyl Adenosine 2-Oxoglutaric acid ... Valproic acid
##      Xanthine
## rowData names(1): Metabolite_name
## colnames(94): ABX_62_ZT11 ABX_63_ZT11 ... GF_51_ZT5 GF_52_ZT5
## colData names(2): ID_muestra info_muestra
```

Se accede a los metadatos sobre la información del proyecto:

[Show](#)

```
## $Project_Title
## [1] "Gut Microbiota Regulates Stress Responsivity via the Circadian System"
##
## $Project_Summary
## [1] "Metabolomics of adult mice Amygdala and Hippocampus. Samples collected from conventional, antibiotic-treated and ger
m-free mice across 4 timepoints of the day. The goal of the project is to investigate the role of the gut microbiota on the
circadian patterns of metabolites in the brain. Results indicated that metabolites involved in glutamate metabolism are alte
red across the day depending on microbial status."
##
## $Institute
## [1] "University College Cork"
##
## $Department
## [1] "Anatomy & Neuroscience"
##
## $PI_Name
## [1] "John Cryan"
##
## $Study_Title
## [1] "Gut Microbiota Regulates Stress Responsivity via the Circadian System - Mice Hippocampus"
##
## $Study_Summary
## [1] "Stress and circadian systems are interconnected through the hypothalamic-pituitary-adrenal (HPA) axis to maintain ap
propriate responses to external stimuli. Yet, the mechanisms of how such signals are orchestrated remains unknown. We now un
cover the gut microbiota as a regulator of HPA-axis rhythmic function. Additionally, microbiota depletion markedly disturbs
the brain transcriptome and metabolome in key stress response pathways in the hippocampus and amygdala across the day. This
is coupled with a dysregulation of the master circadian pacemaker in the brain that resulted in perturbed glucocorticoid rhy
thmicity. The resulting hyper-activation of the HPA-axis at the sleep/wake transition drives time-of-day-specific impairment
s of the stress response and altered stress-sensitive behaviors. Finally, using microbiota transplantation we confirm that d
iurnal oscillations of gut microbes underlie altered glucocorticoid secretion and identify L. reuteri as a candidate strain
for such effects. Together, our data offers compelling evidence that the microbiota regulates stress responsiveness in a cir
cadian manner and is necessary to respond adaptively to psychological stressors throughout the day."
```

[Show](#)

4.3 Análisis multivariante, PCA y clustering

El experimento consta de las medidas de 112 valores de diferentes metabolitos en 94 muestras. Para hacer un análisis multivariante, los metabolitos se consideran las variables y las muestras las observaciones, lo que permitirá hacer un análisis de correlación y clustering

Hasta este punto no se han detectado valores faltantes.

4.3.1 Estructura de datos con los metabolitos

Preparación del dataframe con los datos estructurados con los metabolitos como variables y las muestras como observaciones.

[Show](#)

	1-Methyl Adenosine <dbl>	2-Oxoglutaric acid <dbl>	3-Methylhistidine <dbl>	4-Guanidinobutyric acid <dbl>
ABX_62_ZT11	3012420	31630	1053020	3646723
ABX_63_ZT11	2968448	55513	932342	3425795
ABX_64_ZT11	2903435	9389	1083327	3132739
ABX_65_ZT11	3068587	46558	961178	3396112
ABX_87_ZT11	3260243	85443	1033778	5249595
ABX_88_ZT11	2518544	139590	1103769	3970427

6 rows | 1-5 of 6 columns

4.3.2 Análisis de colinearidad

En este paso se calcula la matriz de correlación de los metabolitos para detectar posibles casos de colinearidad, lo cual sería relevante si en futuros estudios se quisiera realizar algún análisis de regresión.

[Show](#)

Tabla de correlaciones altas (>0.95) entre metabolitos que se podrían considerar colinearidad:

Correlaciones superiores a 0.95 entre metabolitos

Casos	Metabolito 1	Metabolito 2	Índice correlación
1	GMP	AMP	0.957
2	UMP	AMP	0.960
3	Ornithine	Arginine	0.975
4	AMP	GMP	0.957
5	IMP	GMP	0.969
6	UMP	GMP	0.990
7	Guanosine	Guanine	0.996
8	Guanine	Guanosine	0.996
9	Xanthine	Hypoxanthine	0.970
10	GMP	IMP	0.969
11	UMP	IMP	0.970
12	Leucine	Isoleucine	0.975
13	Phenylalanine	Isoleucine	0.958
14	Proline	Isoleucine	0.952
15	Valine/5-Amino valeric acid	Isoleucine	0.969
16	Isoleucine	Leucine	0.975
17	Methionine	Leucine	0.958
18	Phenylalanine	Leucine	0.976
19	Proline	Leucine	0.957
20	Valine/5-Amino valeric acid	Leucine	0.952
21	Proline	Lysine	0.950
22	Leucine	Methionine	0.958
23	Arginine	Ornithine	0.975
24	Isoleucine	Phenylalanine	0.958
25	Leucine	Phenylalanine	0.976
26	Proline	Phenylalanine	0.958
27	Isoleucine	Proline	0.952
28	Leucine	Proline	0.957
29	Lysine	Proline	0.950
30	Phenylalanine	Proline	0.958
31	AMP	UMP	0.960
32	GMP	UMP	0.990
33	IMP	UMP	0.970
34	Isoleucine	Valine/5-Amino valeric acid	0.969
35	Leucine	Valine/5-Amino valeric acid	0.952
36	Hypoxanthine	Xanthine	0.970

Esta tabla ayuda a identificar si existen metabolitos con relaciones muy fuertes, lo cual es útil para evaluar si se pueden reducir las variables o tener en cuenta la colinearidad en futuros análisis.

4.3.3 Análisis de Clustering

Hay dos enfoques para realizar un análisis de clustering:

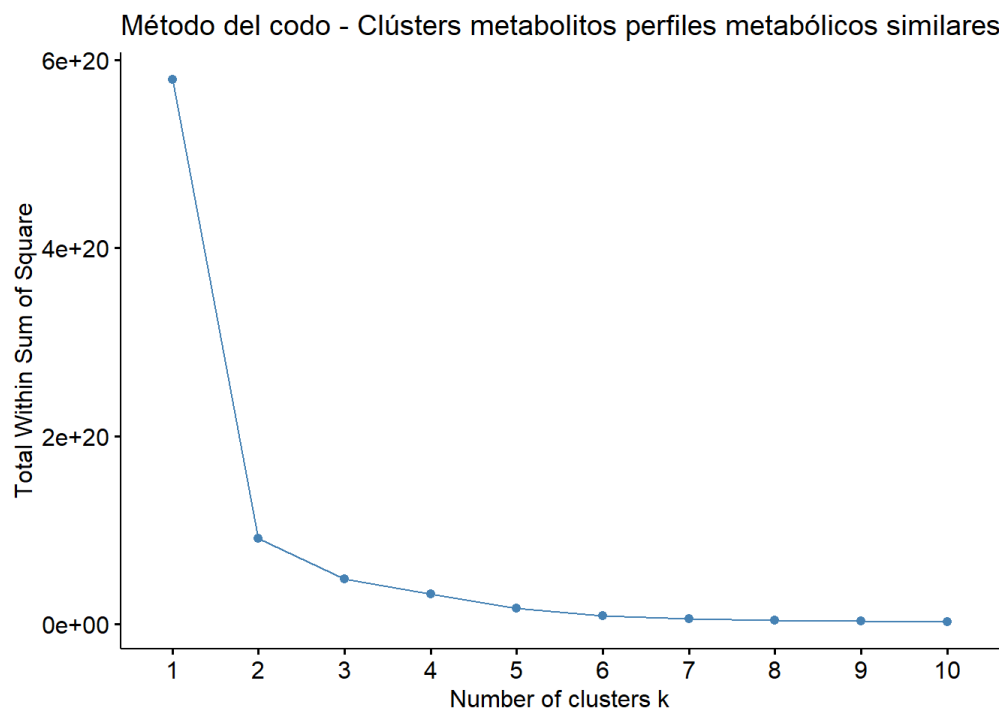
- Clustering de **muestras**: Para identificar si las muestras se agrupan en función de perfiles metabólicos similares.

- Clustering de **metabolitos**: Para identificar grupos de metabolitos que muestran patrones similares a lo largo de las muestras.

4.3.3.1 Clustering de muestras

Para determinar el número óptimo de clústers en las muestras se usa el método del codo.

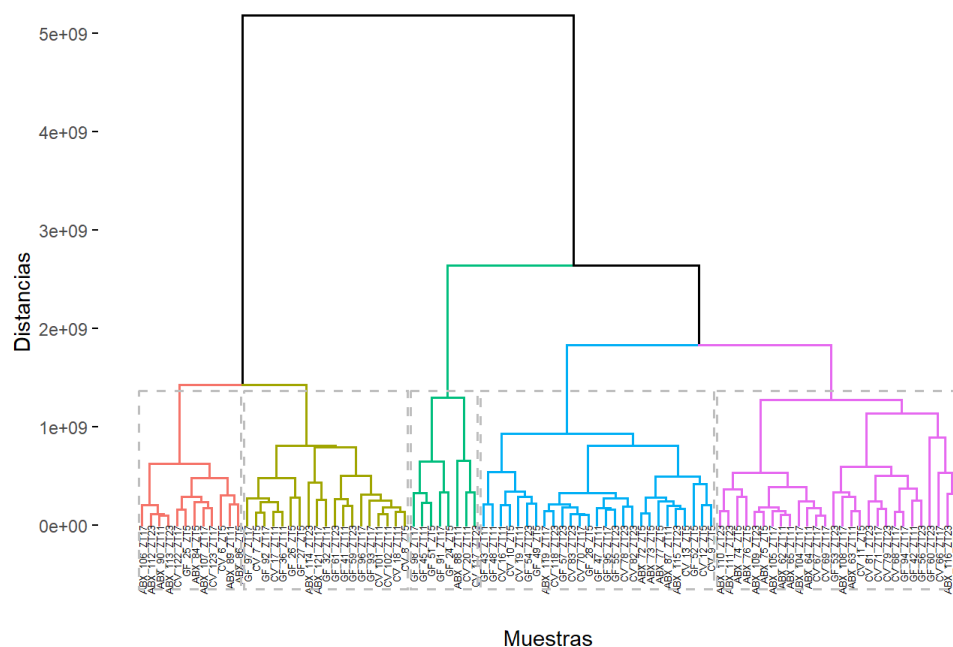
Show



El gráfico hace pensar que a partir de 5 clústers no se observa una reducción significativa en la variabilidad dentro de los grupos.

Show

Dendrograma de agrupamiento jerárquico aglomerativo de muestras - ward.D2



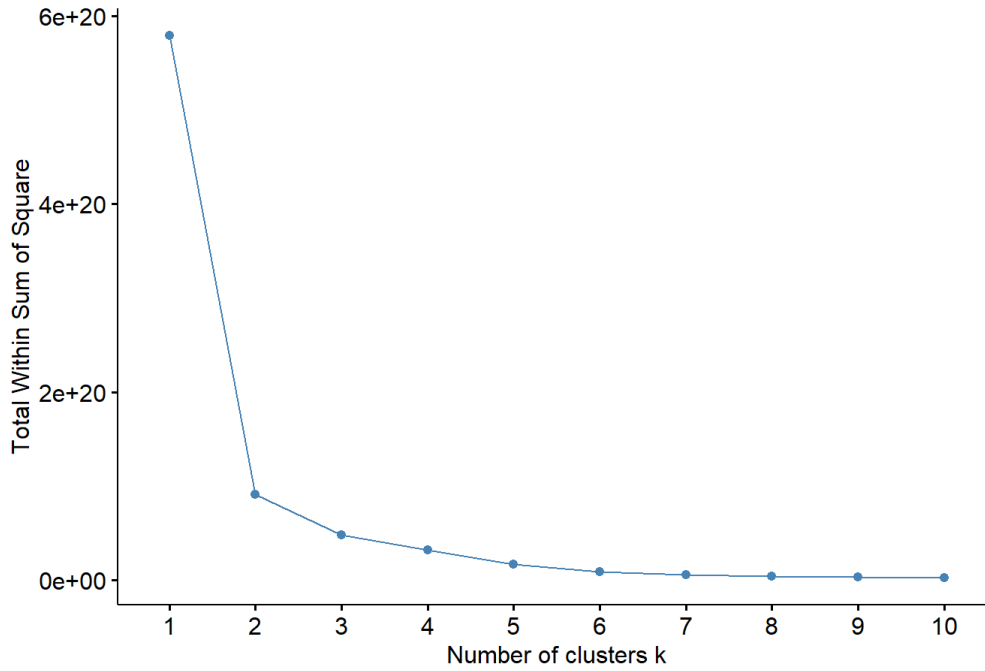
En el dendrograma no se observa una separación evidente únicamente por tratamiento o Zeitgeber_time, donde las letras del identificador de la muestra ABX, GF y CV indican el tipo de tratamiento.

4.3.3.2 Clustering de metabolitos

Para identificar la agrupación óptima de metabolitos que presentan patrones similares a lo largo de las muestras y establecer el número de clústers se usa también el método del codo.

Show

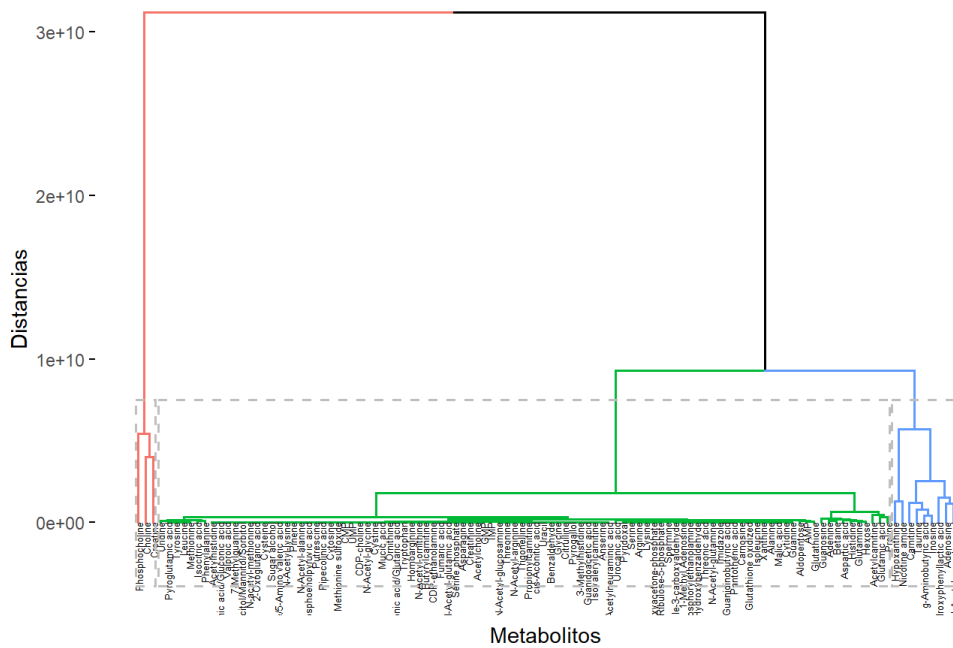
Método del codo - Clústers de metabolitos con patrones en muestras



En este caso parece que con 3 clústers es suficiente para capturar la mayor parte de la estructura de variación en los datos.

Show

Dendrograma agrupamiento jerárquico aglomerativo Metabolitos (ward.D2)



Show

```
## grupos_metabolitos
## 1 2 3
## 100 9 3
```

Show

```
## [1] "Metabolitos en el grupo rojo:"
```

Show

```
## [1] "Choline" "Creatine" "Phosphocholine"
```


Show

```
## [1] "Metabolitos en el grupo azul:"
```

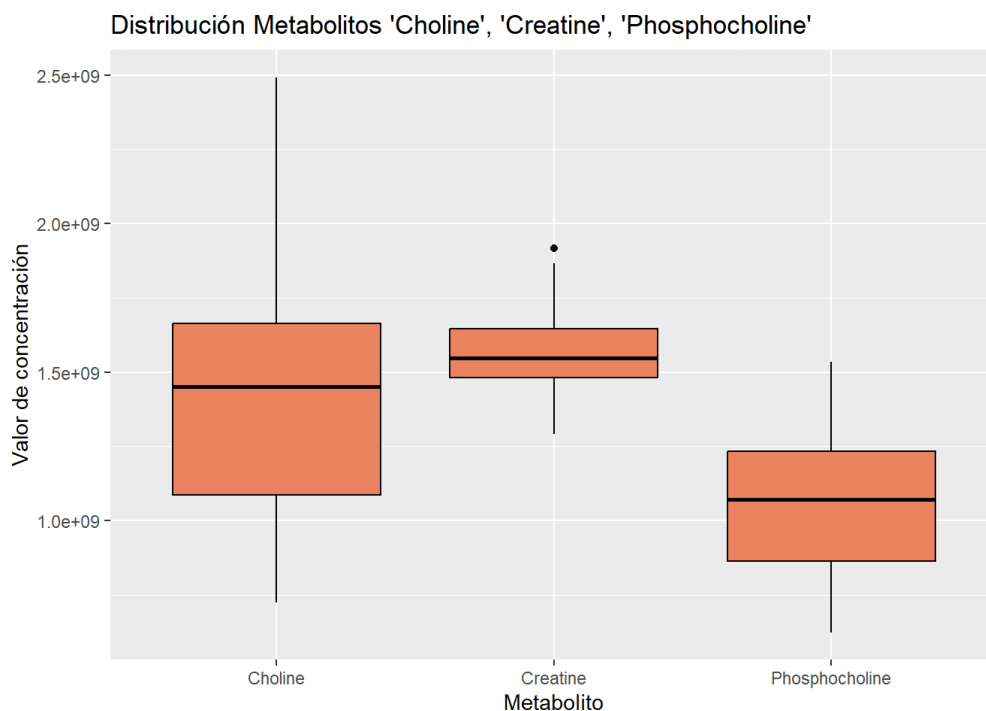
Show

```
## [1] "Adenosine"      "Carnitine"
## [3] "g-Aminobutyric acid" "Hydroxyphenyllactic acid"
## [5] "Hypoxanthine"    "Inosine"
## [7] "N-Acetyl-aspartic acid" "Nicotine amide"
## [9] "Taurine"
```

El clúster representado en rojo contiene 3 metabolitos (Choline, Creatine, Phosphocholine) que tienen un patrón de variación muy diferente en comparación con el resto de los metabolitos. Esto podría ser porque tienen funciones muy específicas y diferenciadas al resto de metabolitos.

Se analiza más en detalles si los metabolitos Choline, Creatine y Phosphocholine tienen distribuciones similares en términos de concentración.

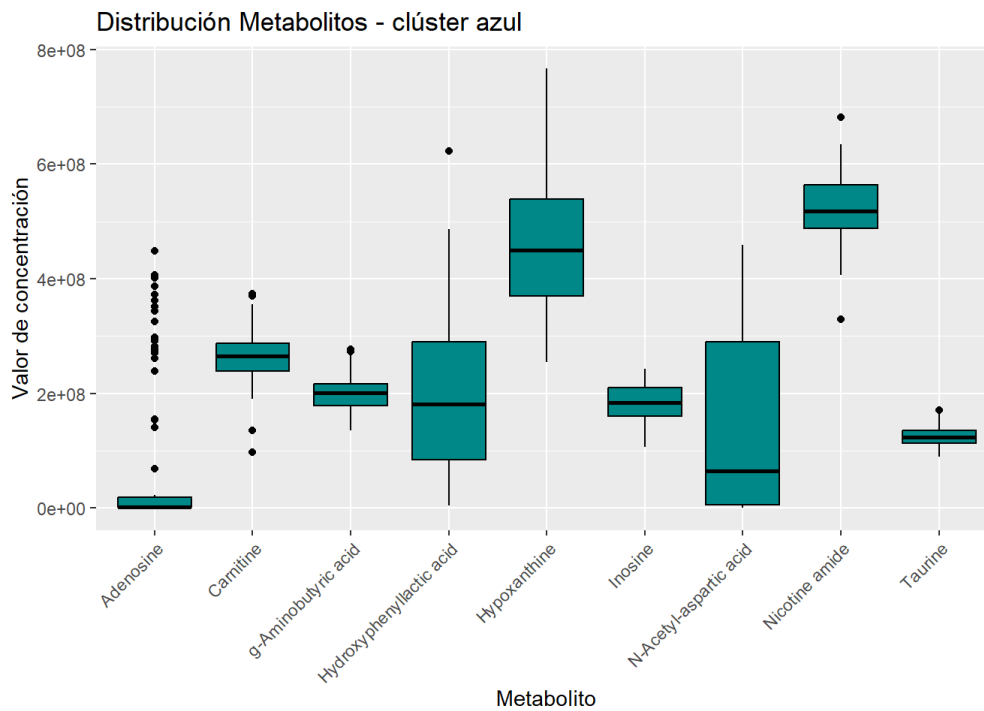
Show



Algo similar podría ocurrir en el clúster representado en azul, donde se han agrupado 9 metabolitos.

El gráfico muestra que *Choline* tiene la mayor variabilidad en concentración entre las muestras, mientras que *Creatine* presenta una distribución más constante con menor dispersión. *Phosphocholine* tiene una variabilidad intermedia, sugiriendo una respuesta algo similar a *Choline* pero más estable que *Creatine*.

Show



El gráfico muestra que *Hypoxanthine* y *N-Acetyl-aspartic acid* tienen alta variabilidad en sus concentraciones, mientras que *Adenosine* presenta valores bajos con bastantes outliers. Por otro lado, *g-Aminobutyric acid*, *Inosine* y *Taurine* son más estables entre muestras.

4.4 Análisis PCA

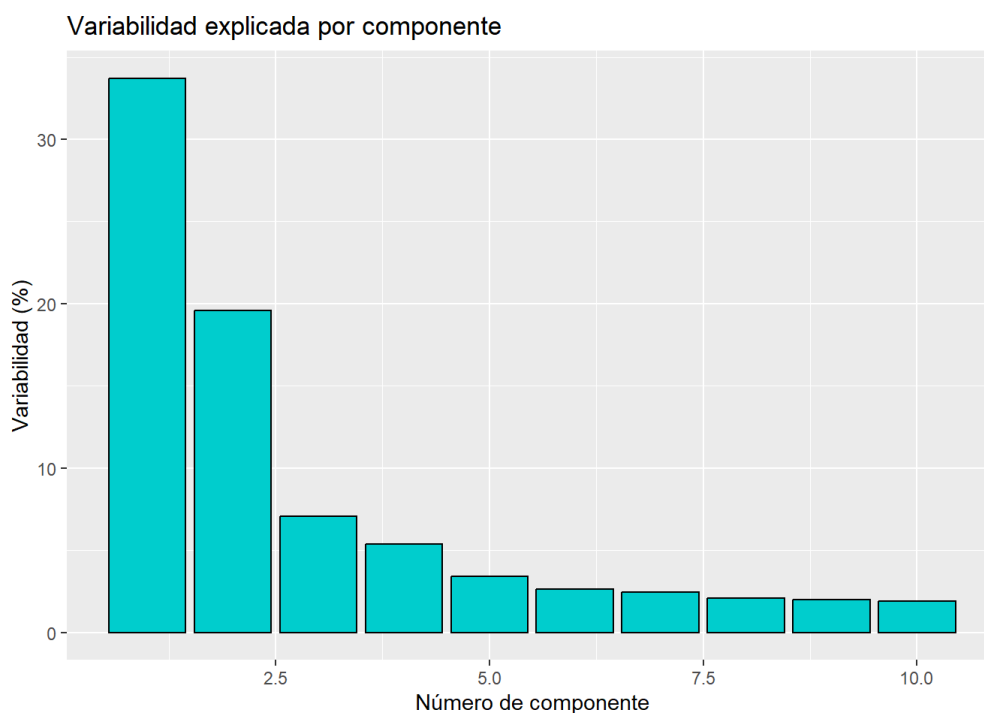
El objetivo del análisis de componentes principales (PCA) es reducir la dimensionalidad de los datos. Los primeros componentes principales capturan la mayor parte de la variabilidad y se facilita la visualización de patrones en las muestras y metabolitos.

4.4.1 Variabilidad explicada por componente

Se calcula la variabilidad explicada por cada componente para determinar cuántas componentes son necesarias para capturar una cantidad aceptable de la variabilidad total de los datos.

Como muestra el gráfico de barras, las dos primeras componentes principales (PC1 y PC2) explican una gran parte de la variabilidad: PC1 explica aproximadamente el 34% y PC2 el 20%.

Show



A continuación se muestran los valores de la variabilidad acumulada hasta el decimo componente principal.

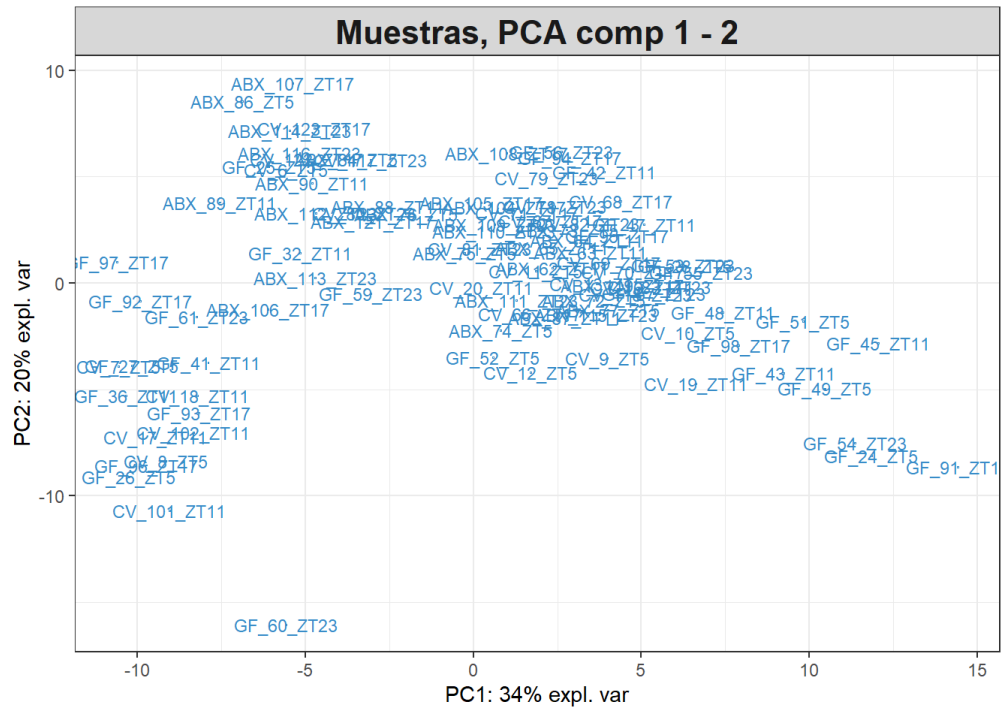
Show

##	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
##	0.3371549	0.5328263	0.6034837	0.6570484	0.6909815	0.7172116	0.7416258	0.7626408
##	PC9	PC10						
##	0.7824020	0.8013619						

Las 5 primeras componentes casi explica el 70% de la variabilidad de los datos, lo cual es suficiente para obtener una representación de las principales fuentes de variación en los datos.

El siguiente gráfico muestra la distribución de las muestras en el plano de las dos primeras componentes principales (PC1 y PC2). Cada punto representa una muestra etiquetada con su identificador, que incluye el tipo de tratamiento y el Zeitgeber time (ZT).

Show

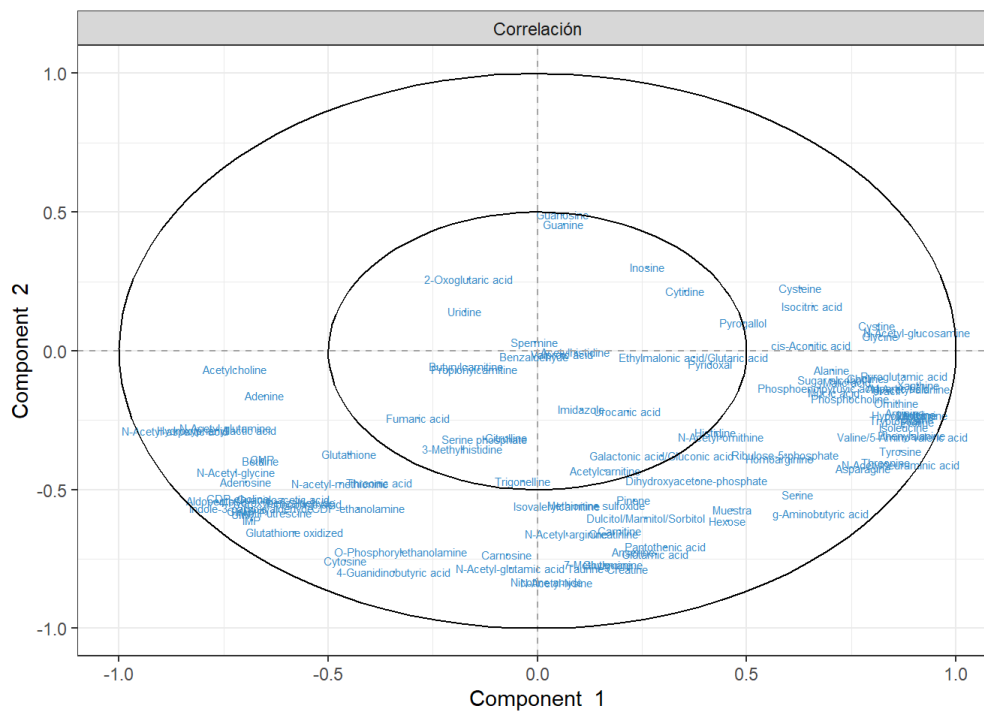


No se observa una clara separación de las muestras por tratamiento (ABX, GF, CV) ni por Zeitgeber time, lo que indica que la variabilidad explicada por las primeras dos componentes no está directamente asociada con estos factores de forma evidente.

4.4.2 Proyección de metabolitos en el espacio de PC

El siguiente gráfico muestra cómo se proyectan los metabolitos en el espacio definido por las dos primeras componentes principales. Cada punto representa un metabolito, y su posición muestra cómo contribuye a las componentes principales. Los metabolitos que se encuentran más cerca de los extremos (bordes) de los ejes PC1 y PC2 tienen una mayor influencia en la variabilidad explicada por esas componentes.

Show



5 Discusión, limitaciones y conclusiones del estudio

Los resultados del clustering hacen pensar que algunos metabolitos presentan patrones similares, y es posible que las respuestas estén relacionadas con el tratamiento o al ciclo circadiano. Si los metabolitos en un clúster están asociados a funciones o rutas metabólicas específicas, esto podría señalar áreas del metabolismo que responden conjuntamente a los factores experimentales. En el contexto de este estudio, estos patrones podrían reflejar cómo la microbiota y el ritmo circadiano influyen en algunos procesos metabólicos.

Los resultados del análisis PCA muestra agrupaciones de metabolitos con patrones de variación similares, sugiriendo posibles relaciones funcionales o rutas metabólicas comunes [3]. Metabolitos como *2-Oxoglutaric acid* y *Uridine* se proyectan en direcciones cercanas, lo que indica que podrían estar correlacionados, aunque no con un valor superior a 0,95, o responder a factores experimentales. Sin embargo, el PCA solo explica una parte de la variabilidad, por lo que no captura todas las posibles interacciones metabólicas influenciadas por el tratamiento y el tiempo circadiano.

Algunas limitaciones adicionales proceden de que este estudio es parte de un proyecto mayor sobre cómo la microbiota intestinal afecta los ritmos circadianos de metabolitos en el cerebro y su relación con el sistema de respuesta al estrés (eje HPA). Sin embargo, faltan datos sobre la regulación de glucocorticoides y la actividad rítmica del HPA, lo que limita el análisis directo de estas interacciones en este análisis [1].

6 Referencias

- [1] This data is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench, <https://www.metabolomicsworkbench.org> (<https://www.metabolomicsworkbench.org>), where it has been assigned Project ID PR002130. The data can be accessed directly via it's Project DOI: 10.21228/M89525 This work is supported by NIH grant, U2C- DK119886.
- [2] Dyar KA, Eckel-Mahan KL. Circadian Metabolomics in Time and Space. *Front Neurosci*. 2017 Jul 11;11:369. doi: 10.3389/fnins.2017.00369. PMID: 28744188; PMCID: PMC5504240.
- [3] Nyamundanda, G., Brennan, L. & Gormley, I.C. Probabilistic principal component analysis for metabolomic data. *BMC Bioinformatics* 11, 571 (2010). <https://doi.org/10.1186/1471-2105-11-571> (<https://doi.org/10.1186/1471-2105-11-571>)

6.1 GitHub

El código y los datos de este análisis están disponibles en el repositorio de GitHub (<https://github.com/mercemerayo/Merayo-Pastor-Merce-PEC1>).