

Vicario-Rubio-Elena-PEC1

Elena Vicario Rubio

2025-03-31

Índice

Informe PEC 1: Analisis del metaboloma fecal infantil	1
Abstract o resumen.	1
Objetivos.	1
Métodos.	2
Resultados.	3
Discusión.	8
Conclusiones.	8
Referencias.	8

Informe PEC 1: Analisis del metaboloma fecal infantil

Abstract o resumen.

Se analizó un dataset de metabolómica fecal (He et al., 2019) de bebés nacidos por parto vaginal y alimentados con leche materna (BF), fórmula estándar (SF) y fórmula experimental suplementada con la membrana de los glóbulos de grasa láctea (MfGM) bovina (EF) a los 2, 4, 6 y 12 meses. El objetivo fue estructurar los datos crudos del dataset usando una clase SummarizedExperiment de Bioconductor y realizar un análisis exploratorio para identificar las variaciones asociadas a la dieta y a la edad. Los datos de abundancia de los metabolitos y sus metadatos se integraron en la clase SummarizedExperiment, la cual se transformó logarítmicamente y se escaló para poder realizar el análisis de componentes principales (PCA) y gráficos de cajas y bigotes de los metabolitos más representativos. El análisis demostró como en las etapas tempranas hay una gran diferencia entre los tipos de alimentación mientras que a los 12 meses los perfiles convergen.

Objetivos.

Los objetivos de este estudio son organizar y dar estructura al dataset de metabolómica fecal de He et al., 2019 utilizando una clase SummarizedExperiment de Bioconductor para integrar los datos crudos de abundancia y los metadatos de las muestras junto a sus características; realizar un análisis exploratorio de los datos para identificar las principales variaciones en el metaboloma de los diferentes grupos de estudio (BF, SF, EF) a lo largo del primer año de vida, identificando los metabolitos clave que se diferencian entre los grupos a lo largo del tiempo y explorando los patrones metabólicos de los diferentes grupos.

Métodos.

El dataset se ha obtenido del GitHub “The Curated Gut Microbiome Metabolome Data Resource” del autor borenstein-lab, el cual contiene el estudio “Fecal microbiome and metabolome of infants fed bovine MFGM supplemented formula or standard formula with breast-fed infants as reference: a randomized controlled trial”, por He et al. (2019), publicado en Scientific Reports, <https://doi.org/10.1038/s41598-019-47953-4>. En el ensayo participaron 240 lactantes, todos nacidos por vía vaginal, 80 fueron alimentados con leche materna (BF), 80 fueron alimentados con una fórmula estándar (SF) y 80 fueron alimentados con una fórmula experimental suplementada con la membrana de los glóbulos de grasa láctea (MfGM) bovina (EF). De los cuales por diferentes exclusiones se quedaron 90 lactantes (15 niños y 15 niñas para cada grupo), los cuales fueron repartidos de forma aleatoria para seleccionar las muestras fecales de microbioma y metabolómica a los 2, 4, 6 y 12 meses. A partir de los 4 meses, se introdujo una suplementación de alimentos complementarios. Esto hizo que las muestras de los meses 4 y 6 se subdividieran en función de la suplementación de alimentos o no. Cada muestra fue secuenciada por amplicones de 16S para analizar la microbiota intestinal, también se realizó un perfilado metabolómico cuantitativo de los metabolitos presentes.

Se eligió para el preprocesamiento de los datos un SummarizedExperiment en vez de un ExpressionSet por su fácil manejo y extensibilidad, además de poder manejar múltiples matrices de datos, mientras que con ExpressionSet solo tenemos una forma de manejar, a través de “exprs”. También es más sencillo clasificar los metadatos a través de DataFrame de S4Vectors como colData y rowData que en los AnnotatedDataFrame.

Las herramientas utilizadas para el preprocesamiento, procesamiento y análisis exploratorio de los datos han sido el software R, en RStudio con Bioconductor, y sus paquetes asociados como son SummarizedExperiment, para dar estructura a los datos en crudo; PCAtools, para el PCA y su visualización; y S4Vectors. Además, se utilizaron paquetes de R como ggplot2, para los gráficos, dplyr y tidyr para la manipulación de los datos; matrixStatus, para los cálculos de matrices. Y también se utilizó Git y GitHub como control de versiones.

El preprocesado de los datos para la creación de SummarizedExperiment se hizo con los archivos mtb.tsv, metadata.tsv y mtb.map.tsv. Cargando los datos de abundancia y los metadatos de las muestras para su integración en una matriz de abundancia (assay) donde los metadatos de las muestras se reflejarán como “colData” y los métodos de los metabolitos se reflejarán como “rowData”.

En el análisis exploratorio de los datos, cargamos los datos ya preprocesados del SummarizedExperiment y hicimos una exploración básica para asegurarnos que los datos han sido correctamente cargados. A continuación, preparamos las matrices de datos para su análisis. Esta preparación consiste en una transformación logarítmica para reducir el sesgo y hacer comparables las varianzas, como hace el artículo original (He et al., 2019) y también se hizo un escalado de los datos para centrar la matriz. Para la visualización de los datos hicimos unos gráficos PCA Score con diferentes objetivos, ahondar en los metabolitos por edad y por grupo de estudio a través de los gráficos boxplot. Usamos la matriz logarítmica para poder ver las diferencias entre el mismo nivel, preparando los datos para las gráficas combinando los metadatos del estudio con los datos para poder extraer analizar la información relevante.

Resultados.

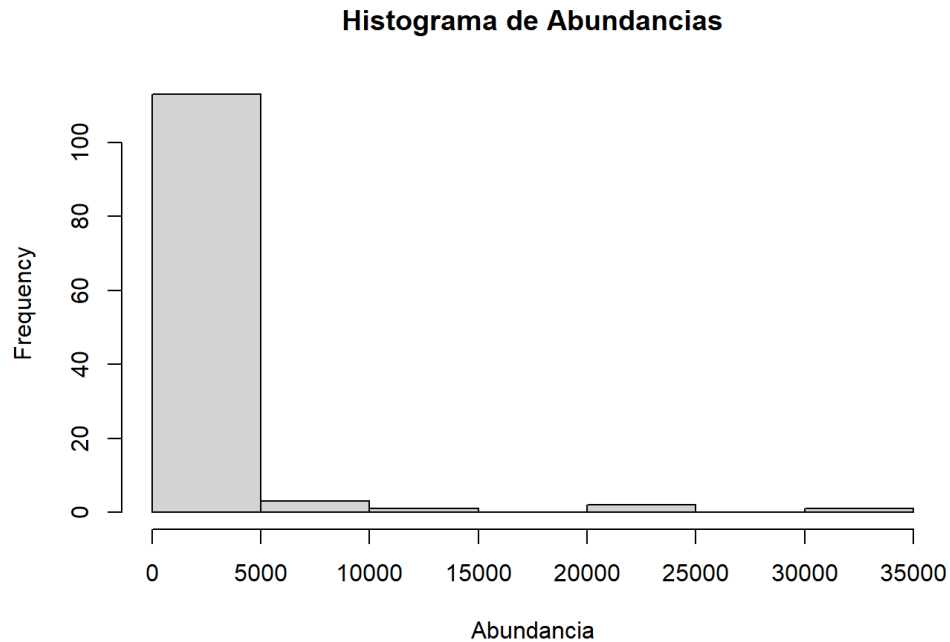


Figura 1: **Gráfica 1**

Histograma de abundancias La matriz obtenida tiene unas dimensiones de 120 filas por 277 columnas. En el análisis de Componentes principales podemos ver en el histograma de abundancias como la grafica antes de transformarla con el logaritmo es bastante plana y al transformarla cambia, esto se debe a que al reducir el sesgo podemos ver mejor las diferencias entre los datos (Gráfica 1 y 2) .

PCA Score Plot por Grupo y Tiempo (Gráfica 3) En este gráfico se muestra como se distribuyen las muestras según su perfil metabolómico. En el se muestra un solapamiento de casi todas las elipses, siendo relevante la distribución de los grupos de estudio. El grupo BF podemos observarlo en la parte superior del gráfico, mientras que los grupos EF y SF se agrupan en la parte inferior del gráfico, muy mezclado.

Esto puede demostrar que los bebés alimentados con leche materna y los alimentados con fórmula a a pesar de que la fórmula este suplementada, en los primeros meses tienen un metaboloma fecal muy diferente, y con el paso de los mese pareciéndose mas entre sí.

PCA Score Plot por Alimentación (Gráfica 4) En este gráfico se muestran como se distribuyen as muestras según su alimentación. Podemos observar como todas las elipses están solapadas entre sí, diferenciando el tamaño de la elipse roja (2 meses), la cual es grande, mientras que la elipse de 12 meses es más compacta y pequeña, lo que nos puede decir que la población es mas homogénea. Esto concuerda con el articulo ya que los bebés a los 12 meses no tienen grandes diferencias en el microbioma intestinal.

PCA Score Plot por Edad (Gráfica 5) En este gráfico se muestran como se distribuyen según la edad de los bebés agrupando todas las dietas, en el se puede observar como todas las elipses están solapadas entre sí, destacando como cambia el perfil metabolomico a lo largo del primer año de vida. Viendo que con 2 meses los datos se agrupan a la derecha del gráfico, ocupando gran parte del dentro, mientras que a los 12 meses, los datos se agrupan mas hacia abajo del gráfico y están mas juntos.

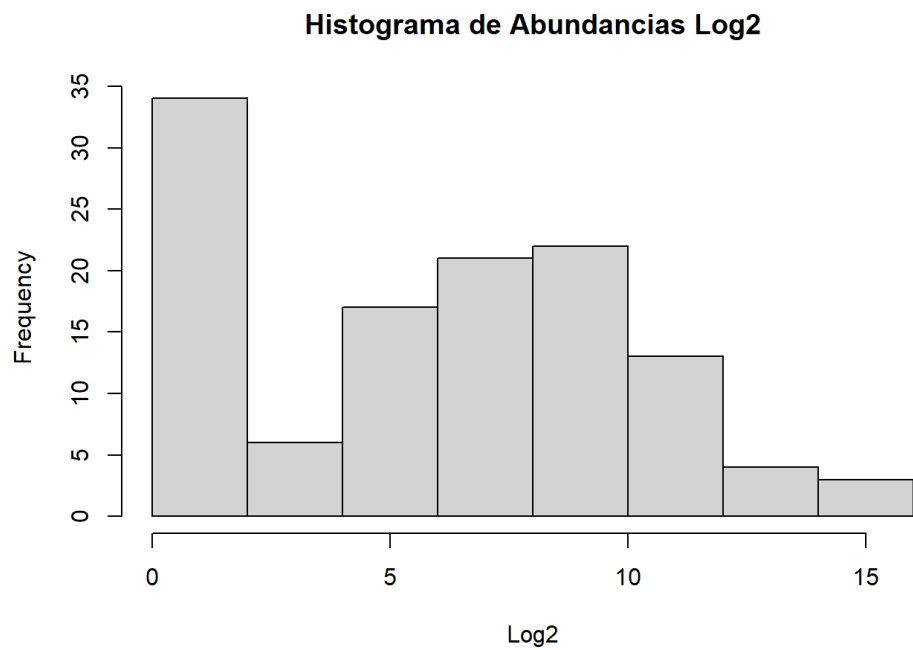


Figura 2: Gráfica 2

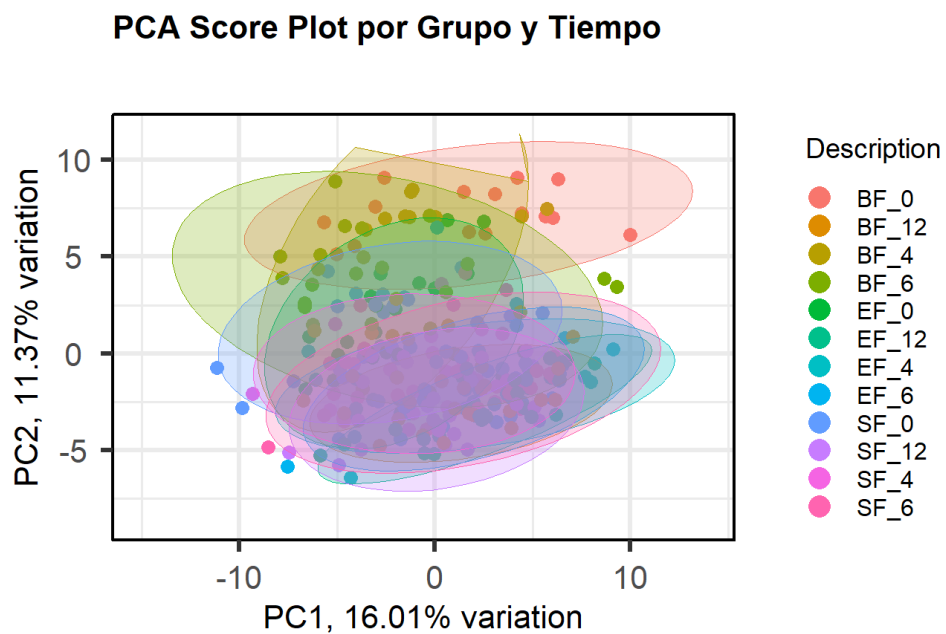


Figura 3: Gráfica 3

PCA Score Plot por Alimentación

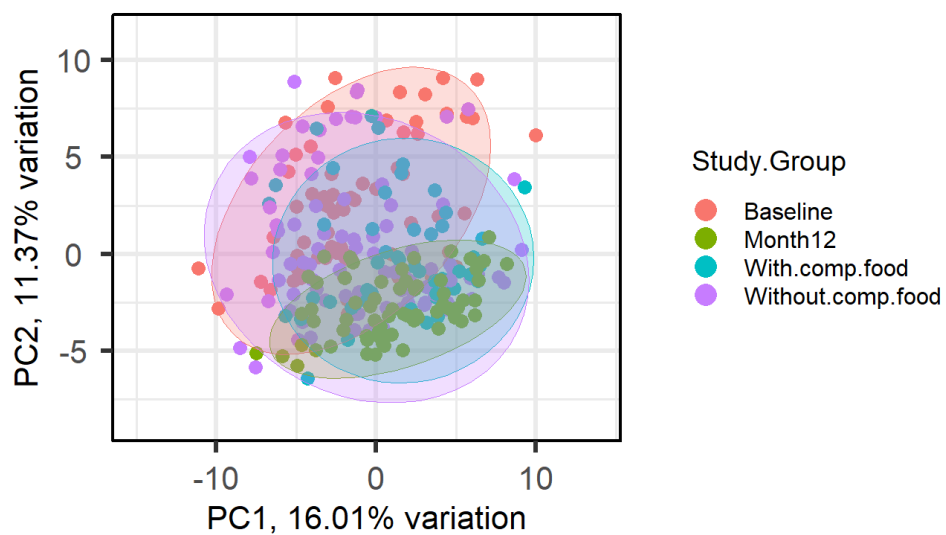


Figura 4: Gráfica 4

PCA Score Plot por Edad

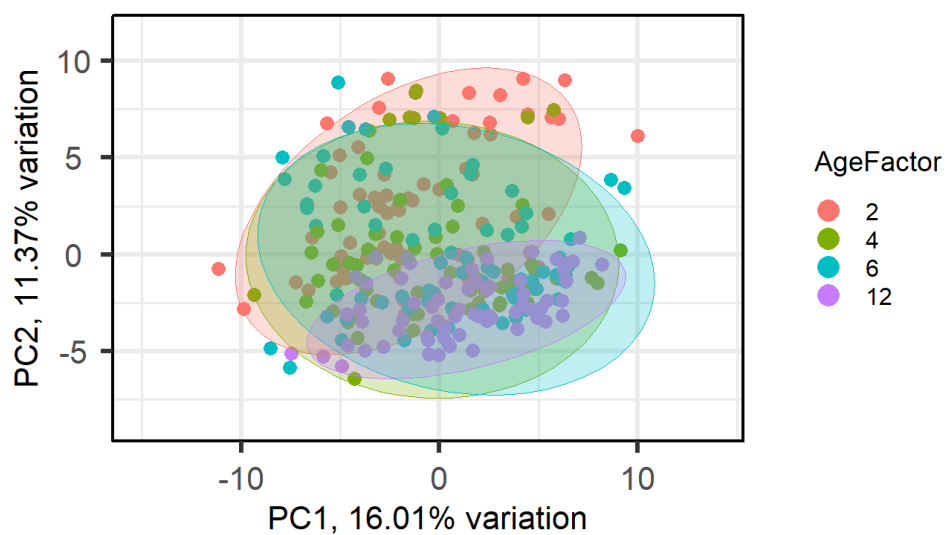


Figura 5: Gráfica 5

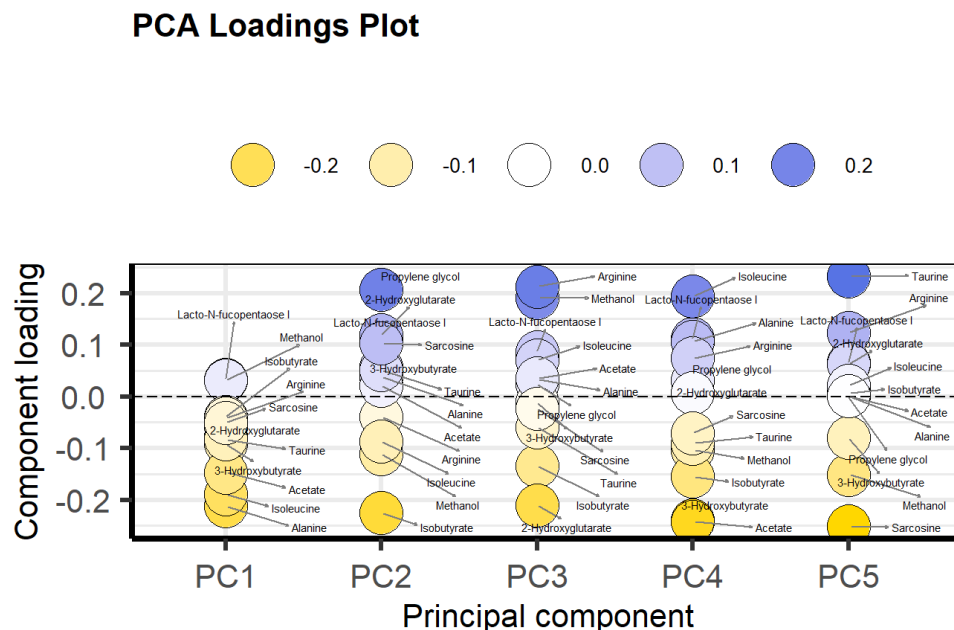


Figura 6: Gráfica 6

Metabolitos que más contribuyen a PC1 y PC2 (Gráfica 6) En este gráfico se muestran cuales son los principales metabolitos que más contribuyen en cada Componente Principal. Dichos metabolitos están relacionados con la dieta del bebé a lo largo de su primer año de vida.

Comparación de metabolitos por grupo de estudio y por edad (Gráfica 7 y 8) En estos gráficos comparamos las cantidades de metabolitos en las distintas dietas y edad que se muestran en el estudio.

En ellos se puede interpretar como el butirato, el isobutirato y el isovalerato tienen una tendencia positiva con la edad, siendo el nivel más bajo a los 2 meses y creciendo hasta alcanzar su máximo a los 12 meses. En este grupo de metabolitos también podemos observar como el paso del tiempo es mucho más relevante para el aumento de sus concentraciones que el tipo de suplementación con comida de la fórmula o la leche materna en los meses intermedios (4-6 meses). Esto concuerda con la maduración del microbioma.

Mientras que, en metabolitos como el lactato y el succinato se muestra una tendencia descendente con la edad, siendo los niveles más altos a los 2 meses y los niveles más bajos a los 12 meses. En el caso del succinato si que se aprecia nivel un poco más bajos con la suplementación con comida de la fórmula o la leche materna en los meses intermedios (4-6 meses). Esto también confirma el descenso de los procesos de fermentación temprana, para dar paso a mejores digestiones de los bebés.

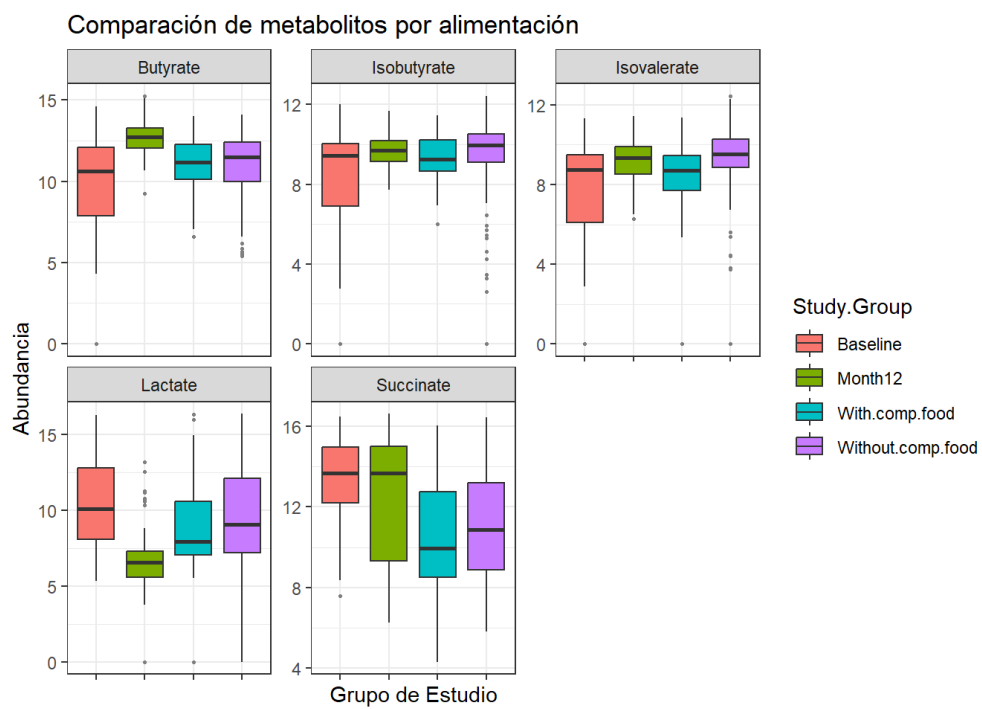


Figura 7: Gráfica 7

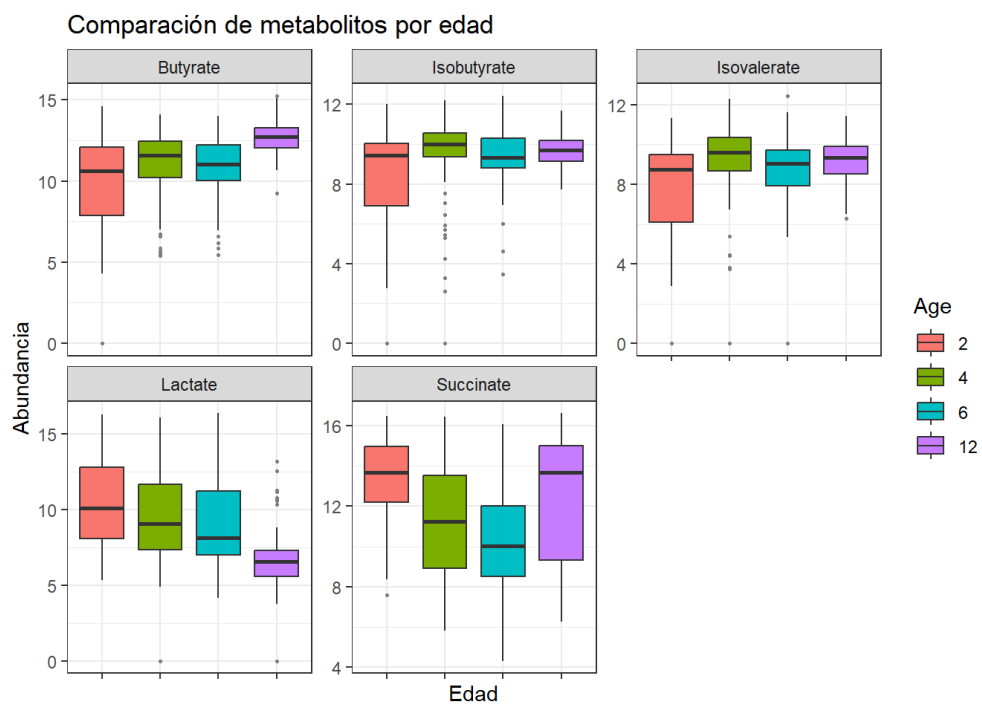


Figura 8: Gráfica 8

Discusión.

El análisis de los datos nos ha proporcionado una visión general de como evoluciona el metaboloma de un bebé en el primer año de vida según su alimentación. El factor más determinante según el PCA es la edad, la cual moldea el perfil metabolómico fecal. Esto se puede ver ya que las muestras de 2 meses se distinguen fácilmente de las de 12 meses, quedando en el medio las de 4 y 6 meses, lo que nos refleja una significativa progresión de la maduración metabólica intestinal.

Esta variación observada con el PCA coincide con las tendencias identificadas en los metabolitos, como la disminución del lactato y succinato, ya que son productos de la fermentación de los azúcares simples como la glucosa y la lactosa, y el aumento de los ácidos grasos de cadena corta ramificada como el butirato, isobutirato e isovalerato, que reflejan el cambio y la maduración de la microbiota intestinal hacia un mejor aprovechamiento de los alimentos. Todo estos resultados concuerdan con las observaciones del propio artículo de He et al. (2019).

El tipo de dieta también se manifestó como un factor influyente en los primeros meses de vida. El PCA score por Grupo y Tiempo mostró una diferenciación relativamente clara entre el grupo BF y los grupos alimentados con fórmula (SF y EF) durante los primeros meses (2 meses). Y como señalaron He et al. (2019), nuestro PCA tampoco mostró una clara diferenciación entre los grupos SF y EF.

Por último, destacar la convergencia de los perfiles metabolómicos a los 12 meses, a pesar de las diferencias de alimentación causadas por la leche materna/fórmula, por la introducción de alimentos a los 4-6 meses, homogeneizando y ayudando a la maduración del microbioma.

Conclusiones.

- El dataset de He et al., 2019 se pudo procesar fácilmente en un SummarizedExperiment, lo que demuestra lo útil que es para el procesamiento de datos ómicos.
- El análisis nos indicó que la edad es el factor más determinante para el metaboloma fecal durante el primer año de vida, como se muestra en el PCA y en los cambios de los metabolitos como el butirato, el isobutirato, el isovalerato, el lactato y el succinato.
- La alimentación inicial, ya sea materna o de fórmula, genera perfiles metabolómicos diferentes en los primeros meses de vida, que se atenúan con la introducción de la comida y convergen al año de vida.

Referencias.

- Repositorio de GitHub: <https://github.com/evadlerr/Vicario-Rubio-Elena-PEC1.git>
- He, X., Parenti, M., Grip, T., Lönnerdal, B., Timby, N., Domellöf, M., Hernell, O., & Slupsky, C. M. (2019). Fecal microbiome and metabolome of infants fed bovine MFGM supplemented formula or standard formula with breast-fed infants as reference: a randomized controlled trial. *Scientific reports*, 9(1), 11589. <https://doi.org/10.1038/s41598-019-47953-4>.
- Falcon, S., Morgan, M., & Gentleman, R. (s. f.). An Introduction to Bioconductor's ExpressionSet Class.
- Forbes, J. D., Azad, M. B., Vehling, L., Tun, H. M., Konya, T. B., Guttman, D. S., Field, C. J., Lefebvre, D., Sears, M. R., Becker, A. B., Mandhane, P. J., Turvey, S. E., Moraes, T. J., Subbarao, P., Scott, J. A., Kozyskyj, A. L., & Canadian Healthy Infant Longitudinal Development (CHILD) Study Investigators (2018). Association of Exposure to Formula in the Hospital and Subsequent Infant Feeding Practices With Gut Microbiota and Risk of Overweight in the First Year of Life. *JAMA pediatrics*, 172(7), e181161. <https://doi.org/10.1001/jamapediatrics.2018.1161>.
- Bode L. (2009). Human milk oligosaccharides: prebiotics and beyond. *Nutrition reviews*, 67 Suppl 2, S183-S191. <https://doi.org/10.1111/j.1753-4887.2009.00239.x>
- Riviére, A., Selak, M., Lantin, D., Leroy, F., & De Vuyst, L. (2016). Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Frontiers*

in microbiology, 7, 979. <https://doi.org/10.3389/fmicb.2016.00979>Mohan, R. et al. Effects of Bifidobacterium lactis Bb12 supplementation on body weight, fecal pH, acetate, lactate, calprotectin, and IgA in preterm infants. *Pediatr. Res.* 64, 418–422 (2008).

- He, X., Parenti, M., Grip, T. *et al.* Metabolic phenotype of breast-fed infants, and infants fed standard formula or bovine MFGM supplemented formula: a randomized controlled trial. *Sci Rep* **9**, 339 (2019). <https://doi.org/10.1038/s41598-018-36292-5>
- Lee, H., Zavaleta, N., Chen, SY. *et al.* Effect of bovine milk fat globule membranes as a complementary food on the serum metabolome and immune markers of 6-11-month-old Peruvian infants. *npj Sci Food* **2**, 6 (2018). <https://doi.org/10.1038/s41538-018-0014-8>