

(1) Seleccionar un dataset de metabolómica:

1. Realizamos la descarga de datos, para lo que accedemos al repositorio de github, y en CODE descargamos en DownloadZip.
2. Desde el escritorio, elegimos la base 2018-MetabotypingPaper. Abrimos RStudio > File > New File > Markdown File.
3. Instalamos Bioconductor (copiando el link de la página web de Bioconductor). Como trabajo una versión antigua de R, no indico versión de Bioconductor.
4. También instalo BiocManager::install("SummarizedExperiment") porque sabemos que en el paso 2 lo necesitaremos.
5. En RStudio > ImportDataset > Escogemos las dos bases: DataValues_S013 donde tenemos los datos, y DataInfo_S013 donde están los metadatos. Antes de darle a import, copiamos el link, porque de la otra manera nos da error. Aunque los hemos importado igualmente.
6. Las nombramos como data_values y data_info respectivamente.
7. En el siguiente paso fijamos una semilla para generar números aleatorios en la muestra. Lo hacemos para practicar lo que hicimos en la tarea 1.1. Pero para seguir trabajando lo haremos con la base de datos

(2) Crear un contenedor SummarizedExperiment

8. Conseguimos crear el dataframe para las columnas (colData) y el datafram para las líneas (rowData).
9. Durante muchos intentos anteriores, nos daba problemas la longitud de las columnas de datos al intentar crear el SummarizedExperiment. Por eso escogemos una columna y una función en colData.
10. Ejecutamos la función Summarized Experiment.
11. Mostramos Summarized Experiment. Ha sido posible crear la matriz pero los nombres de las filas y columnas no aparecen, aparecen como 1,2,3,4 y 5. A pesar de que en rowData y colData aparecían bien. He intentado solucionarlo pero no he podido.

Para solucionar los diferentes problemas citados, hemos consultado esta bibliografía, aunque algunos no se han podido solucionar con éxito:

<https://www.youtube.com/watch?v=D8IVRmbMjyc>

https://www.bioconductor.org/help/course-materials/2019/BSS2019/04_Practical_CoreApproachesInBioconductor.html

(3) Realizar exploración de Dataset.

12. Para ello hemos visualizado y repasado los contenidos de la actividad 1.3. Y creamos un análisis específico para nuestro caso, siguiendo la línea de preguntas propuesta.

Pregunta Biológica: El interrogante principal es entender cómo varían los análisis metabólicos en pacientes de obesidad sometidos a dos tipos de cirugía: bypass y gastrectomía tubular.

Diseño Experimental: Se escogieron pacientes sometidos a cirugía, y se realizaron análisis de sus parámetros metabólicos. Esto permite evaluar los efectos de cada una de las dos cirugías y compararlas.

Obtención de Datos Crudos: Las mediciones se realizaron mediante cromatografía líquida-espectrometría de masas y espectrometría de masas en tándem. Podemos observarlo en el documento de AA_Information_S006.

Control de Calidad y Preprocesado: Es importante asegurar que las mediciones sean correctas y precisas, asegurando que las diferencias que existan se deban a los tratamientos y no a errores técnicos.

Análisis Estadístico y Genes Diferencialmente Expresados: Para evaluar los datos, se necesitan análisis estadísticos que identifiquen la variación entre los dos tipos de cirugía.

Análisis de Significación Biológica: Las diferencias en los parámetros ayudan a comprender mejor qué alteraciones metabólicas supone cada uno de los procedimientos y las posibles implicaciones a largo plazo.

Respuesta a la Pregunta Biológica: El estudio permite saber si hay diferencia entre las dos técnicas, y así poder escoger el mejor tratamiento para futuros pacientes.

Repositorio de GitHub: <https://github.com/AnaJuan7/Juan-Nicolas-Ana-PEC1.git>

Contiene:

- Contenedor de datos (.Rda)
- Código R.
- Metadatos en Markdown.
- No he adjuntado documento con datos de texto. NOTA: Al seleccionar knitr no puede crearse, inicialmente me daba solamente error en la función assay porque debe ser solamente números y he incluido variable con letras. Intentando corregirlo estoy creando otros problemas que no soy capaz de solucionar.