

### (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](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:** La pregunta clave del estudio es cómo cambian los valores metabólicos en pacientes con sobrepeso u obesidad después de someterse a bypass gástrico o gastrectomía tubular. El objetivo es ver si estas cirugías afectan (o afectan diferente) sobre el metabolismo de los pacientes, lo que puede ayudarnos a entender mejor cuál es mejor o tiene menos efectos secundarios a largo plazo.

**Diseño Experimental:** Los pacientes son seleccionados según el tipo de cirugía que se les realiza. Se evalúan muchos parámetros metabólicos indicando el método utilizado (bypass gástrico o gastrectomía tubular). Es importante tener en cuenta que las diferencias en las características de los pacientes podrían influir en los resultados.

**Obtención de Datos Crudos:** En este estudio se utilizan técnicas de cromatografía líquida-espectrometría de masas y de espectrometría de masas en tándem para medir los análisis. Estas técnicas permite obtener multitud de datos de análisis. Esta información la sacamos del documento de AA\_Information\_S006. Es importante que las muestras sean tomadas correctamente, ya que sino, los resultados podrían no proporciona información del todo verdadera.

**Control de Calidad, Preprocesado y Normalización:** Un problema común con los datos es que pueden variar mucho entre los pacientes. Por ejemplo, algunas personas podrían tener mediciones muy altas o bajas por razones que no tienen que ver directamente con la cirugía. Para evitarlo, se pueden usar técnicas que ajusten los resultados para que las diferencias entre los pacientes no afecten tanto.

**Análisis Estadístico y Genes Diferencialmente Expresados:** Se deben realizar análisis estadísticos para comparar los resultados después de la cirugía. Algunas posibles pruebas para este caso son el t-test, ANOVA o Kruskal-Wallis Test.

**Análisis de Significación Biológica:** Se interpretan los datos evaluando la variación de los valores del análisis mediante pruebas estadísticas.

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 knirt 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.