

Segunda Prueba de Evaluación Continua (PEC2) Análisis de datos Ómicos, (23 Diciembre 2020- 12 Enero 2021)

Enunciado

El objetivo de esta práctica es doble:

- Partiendo de datos preparados deberéis analizarlos siguiendo las pautas discutidas en los debates 3 y 4 y, es decir **a partir de los read counting**.
- En los datasets **no es necesario llevar a cabo ni alineamiento ni mapeo de las “reads” frente a un genoma de referencia**. Es decir **TODO EL PROCESO DE ANALISIS lo llevareis acabo a partir de los READ COUNTING (COUNT TABLE)**
- Una vez obtenidos los resultados deberéis redactar un informe con la estructura tradicional de un informe técnico-científico (ver sección “Informe del análisis”).

Los datos para el análisis (debéis escoger uno)

Elegir un dataset de los que contiene **PEC2datasets.zip**. La Información del de dichos datasets se encuentra en:

- PEC2datasets información.xls (Phenodata, sample names, variable labels, y el count_data)**
- Datasets contenidos en: PEC2datasets.zip**

“Pipeline” de análisis

Tal como habéis aprendido en la unidad de análisis de microarrays se debería seguir una serie de pasos ordenados. El proceso estándar consistirá en:

1. Identificar que grupos hay y a qué grupo pertenece cada muestra.
2. Control de calidad de los datos crudos (desde la vertiente estadística STAT world (clustering, PCA) no BIO world)
3. Normalización
4. Identificación de genes diferencialmente expresados
5. Anotación de los resultados
6. Comparación entre distintas comparaciones (si hay más de una comparación, ver que genes han sido seleccionados en más de una comparación)
7. Análisis de significación biológica (“Gene Enrichment Analysis”)

Podeis obviamente emplear EdgeR o DESeq2 o limma- Voom aquella que creáis conveniente ya que no "esta para nada cerrada esta cuestión"

Informe del análisis

Una vez realizado el análisis debéis redactar un informe exponiendo **qué habéis hecho, como lo habéis hecho y qué resultados habéis obtenido todo en un FICHERO RMARKDOWN**.

Como cualquier informe científico-técnico vuestro informe debería contener las partes siguientes:

1. **Abstract**, con un resumen breve de no más de cinco líneas.
2. **Objetivos**: Que se pretende con este estudio
3. **Materiales y Métodos** (tipo de experimento, diseño experimental, métodos que habéis utilizado en el análisis)
4. **Resultados**
Que se obtiene como resultado del análisis
5. **Discusión**
Que limitaciones consideramos que pueden haber en el estudio (si consideramos que hay alguna...)
6. **Conclusión**: NO HACE FALTA. Vuestro “rol” aquí es técnico. Como bioinformáticos se os presupondrá la capacidad de manejar la información biológica mediante los programas adecuados, pero ello no implica que debáis tener los conocimientos específicos que puede requerir la interpretación biológica de los resultados.

Segunda Prueba de Evaluación Continua (PEC2)
Análisis de datos Ómicos, (23 Diciembre 2020- 12 Enero 2021)

7. **Apéndice:** Podéis poner el código de R que hayáis utilizado en un apéndice con comentarios.

De cara a evaluar el trabajo deberéis entregar un fichero markdown como el que llevasteis a cabo en la PEC1 CON el fichero fuente (rnw o rmd) y su version “compilada” en formato pdf.

Este archivo debe contener vuestro nombre y apellidos en la cabecera (o en la primera línea del script de R). (p.ej. Apellido1Apellido2Nombre-informePEC2.zip).