

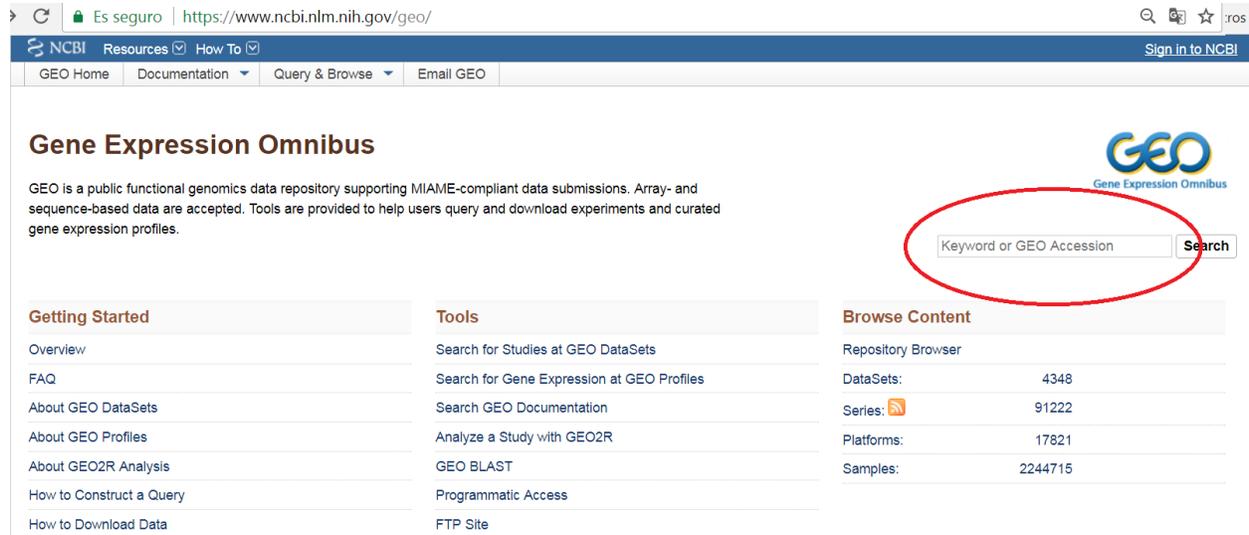
Examen “Análisis Estadístico de datos Ómicos en RBioconductor”

Fecha entrega: 24 de Noviembre (término de semestre Escuela de Postgrado Facultad de Ciencias).

1. Ingresar al módulo de GEO de la página del NCBI:

<https://www.ncbi.nlm.nih.gov/geo/>

2. En el campo de búsqueda, ingrese una frase o palabra clave para buscar algún estudio de su interés (marcado con la elipse roja en la imagen adjunta). La idea es que busquen un trabajo en donde hayan realizado RNA-seq y se descarguen el archivo con los conteos para realizar un análisis de expresión diferencial.



The screenshot shows the NCBI Gene Expression Omnibus (GEO) website. The search bar is highlighted with a red circle. The page includes a navigation menu, a search bar, and a table of content statistics.

| Getting Started | Tools | Browse Content |
|--|--|------------------------------------|
| Overview | Search for Studies at GEO DataSets | Repository Browser |
| FAQ | Search for Gene Expression at GEO Profiles | DataSets: 4348 |
| About GEO DataSets | Search GEO Documentation | Series: 91222 |
| About GEO Profiles | Analyze a Study with GEO2R | Platforms: 17821 |
| About GEO2R Analysis | GEO BLAST | Samples: 2244715 |
| How to Construct a Query | Programmatic Access | |
| How to Download Data | FTP Site | |

3. Una vez que hayan descargado el archivo con los conteos, realizar un análisis de expresión diferencial utilizando al menos dos aproximaciones estadísticas diferentes (pueden ser los paquetes vistos en clases u otros disponibles para R) y comparar los resultados mediante un diagrama de Venn.

4. Fijar un criterio en función del foldchange y valor p ajustado (en caso de los métodos clásicos) o un valor de probabilidad a posteriori (en caso de los métodos bayesianos) para determinar si los genes presentan una expresión diferencial. Justificar los valores utilizados.

5. Utilizando los genes con expresión diferencial realice análisis de sobrerrepresentación (enriquecimiento) utilizando las bases de datos KEGG y Gene Ontology. Grafique los resultados utilizando visualizaciones como las entregadas por el paquete clusterProfiler.

6. Discuta e interprete brevemente los resultados obtenidos por usted en función al problema biológico involucrado en el estudio. Remarque las diferencias y similitudes encontradas por usted y el trabajo publicado de donde obtuvo los datos.

7. Realice un informe con una breve introducción, Resultados y discusión. Suba en un archivo comprimido los datos de conteo con el cual hizo los análisis, el script de R y el documento del informe.