**Genomics cross-study analysis, data integration, exploratory data analysis, data mining**

**Keywords:**

data mining

cross study analysis

data normalization

batch-corrected RNA-seq

|  |
| --- |
| **Relevant journal publications** |
| |  | | --- | | **ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases**  BMC Genomics 2014 volume 15, Article number: 284  415 Citations  <https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-284> | |  |  |  | | --- | | **Google**  **enabling cross-study analysis of RNA-Sequencing data**  [https://www.google.com/search?lr=&as\_qdr=all&sxsrf=AleKk00QSjfjp\_4g\_LC4EyNExgXHpC1w4w:1606327451528&q=enabling+cross+study+analysis+of+rna-sequencing+data&sa=X&ved=2ahUKEwi-jZCJpJ7tAhVkw1kKHXaKCWMQ1QIoAHoECAUQAQ](https://www.google.com/search?lr=&as_qdr=all&sxsrf=ALeKk00QSjfjp_4g_LC4EyNExgXHpC1w4w:1606327451528&q=enabling+cross+study+analysis+of+rna-sequencing+data&sa=X&ved=2ahUKEwi-jZCJpJ7tAhVkw1kKHXaKCWMQ1QIoAHoECAUQAQ)  **Google**  **RNA-Seq data mining** | | **Massive mining of publicly available RNA-seq data from human and mouse**  Nature Communications volume 9, Article number: 1366 (2018)  96 Citations  <https://www.nature.com/articles/s41467-018-03751-6>  **Omics Playground: Explore Omics Data Freely**  <https://omicsplayground.readthedocs.io/en/latest/index.html>  Omics Playground is a comprehensive self-service platform platform for visualization, analytics and exploration of Big Omics Data. It allows users to apply a multitude of state-of-the-art analysis tools to their own data to explore and discover underlying biology in a short time.  The platform offers a unique combination of features that distinguishes it from the other analytics platforms currently available. We believe that data preprocessing (primary analysis) and statistical testing (secondary analysis) are now well established, and the most challenging task is currently data interpretation (tertiary analysis) that often takes the longest time but where actual insights can be gained. Therefore, Omics Playground focuses strongly on tertiary analysis while providing good support for secondary analysis.  Reanalyzing Public Datasets  To illustrate the use case of the Omics Playground, we reanalyzed different types of publics datasets, including microarray, bulk RNA-seq, single-cell RNA-seq and proteomic datasets to recapitulate the results.  <https://omicsplayground.readthedocs.io/en/latest/examples/examples.html> |      |  | | --- | | **Google**  **how to analyze geo datasets in r**  <https://www.google.com/search?lr=&as_qdr=all&sxsrf=ALeKk03S9orXF4hAyQbXO282dDPFBUHzAA:1609876679459&q=how+to+analyze+geo+datasets+in+r&sa=X&ved=2ahUKEwjyuvH8yYXuAhUOQ80KHZVNBEgQ1QIoAHoECAgQAQ> | | **Exploratory analysis of datasets obtained from GEO**  <http://www.biotechworld.it/bioinf/2015/10/22/exploratory-analysis-of-datasets-obtained-from-geo/>  bioinformatics > Exploratory analysis of datasets obtained from GEO  October 22, 2015  **Using the GEOquery Package**  <https://bioconductor.org/packages/release/bioc/vignettes/GEOquery/inst/doc/GEOquery.html>  September 21, 2014  **Retrieve and analyze a gene expression data set from NCBI GEO in R**  <https://www.youtube.com/watch?v=gC-WuK4BbJY>  Nov 2, 2012  **Analysing data from GEO - Work in Progress**  <https://sbc.shef.ac.uk/geo_tutorial/tutorial.nb.html>  30 Jun 2020 |  |  | | --- | | **comparing rna-seq datasets** | | **Broad Institute Gene Set Enrichment Analysis (GSEA)**  <https://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Using_RNA-seq_Datasets_with_GSEA> |  |  | | --- | | **pcaExplorer: an R/Bioconductor package for interacting with RNA-seq principal components**  BMC Bioinformatics 2019 volume 20, Article number: 331  17 Citations  <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-019-2879-1> | | **pcaExplorer**  <http://bioconductor.org/packages/release/bioc/html/pcaExplorer.html>  **pcaExplorer User Guide**  27 October 2020  <http://bioconductor.org/packages/release/bioc/vignettes/pcaExplorer/inst/doc/pcaExplorer.html>  **Up and running with pcaExplorer**  27 October 2020  <http://bioconductor.org/packages/release/bioc/vignettes/pcaExplorer/inst/doc/upandrunning.html>  data component (count matrix, experimental data, dds object, annotation) |   **BioStars**  **Question: RNA-seq data comparison across experiments**  <https://www.biostars.org/p/117451/>  **BioStars**  **Question: Comparing similarity of RNAseq datasets**  <https://www.biostars.org/p/334235/>   |  | | --- | | **Feature-based classification of human transcription factors into hypothetical sub-classes related to regulatory function**  BMC Bioinformatics volume 17, Article number: 459 (2016)  7 Citations  <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1349-2> | |