Preliminary project decription

Group 3 - R in Bio Data Science

Angeliki Kallia Spentza -s212881   Daniel Pedrosa Martin -s213270

Javiera Laing Carmona -s196546   Klara Alicja Kotkowska -s213028

Mads Cort Nielsen -s120356

2022-04-07

rthritis

The original data comes from 236 RNA-seq synovial biopsy samples from the papers of Walsh et al. and Guo et al. as provided by Platzer et al. [1-3].

Links to datasets

Raw data 1: https://datadryad.org/stash/dataset/doi:10.5061/dryad.315v28r

Raw data 2: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89408

A rough workflow

Following tidying and wrangling of the raw data, a PCA, boxplot and heatmap of the gene-expression data will be done for each of the patient groups.

Libraries

Some of the libraries we expect to use are the following: `ggplot2`, `tidyr`, `patchwork`, `dplyr`, `broom`, `purrr`, `tibble`.

Aim

Our aim is to profile the patient groups based on the transcriptomic analysis. A biological analysis and conclusion will follow.

References

[1] Walsh, Alice M., et al. "Triple DMARD treatment in early rheumatoid arthritis modulates synovial T cell activation and plasmablast/plasma cell differentiation pathways." PLoS One 12.9 (2017): e0183928.

[2] Guo, Yanxia, et al. "CD40L-dependent pathway is active at various stages of rheumatoid arthritis disease progression." The Journal of Immunology 198.11 (2017): 4490-4501.

[3] Platzer, Alexander, et al. "Analysis of gene expression in rheumatoid arthritis and related conditions offers insights into sex-bias, gene biotypes and co-expression patterns." PloS one 14.7 (2019): e0219698.