

Gene Expression Profiles, Revisited

BIOF2014

Data

Gene expression profiles have been generated by RNA sequencing across different normal tissues and tumour samples. These datasets are available from the GTEx and TCGA projects.

We will work with the TCGA expression dataset.

Download `expr_pancan_ccg.rds` from Moodle.

Questions

1. Use the sample ID to determine its tissue type, e.g. tumour vs. normal tissue, tumour type, tissue type, etc.
2. Design a statistical model in Stan to compare the expression of a specific gene in a tumour type against the normal tissues that the tumour type is derived from.
3. Design a statistical model in Stan to compare the expression of a specific gene in a tumour sample against its matched adjacent normal tissue (e.g. paired analysis).

Resources

1. https://docs.gdc.cancer.gov/Encyclopedia/pages/TCGA_Barcodes/
2. <https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/tissue-source-site-codes>
3. <https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/sample-type-codes>