Medical Bioinformatics

Project

SS15

Stefan Jacob & Zarin Shakibaei

Outline

1. Introduction
2. Data
3. Methods
4. Results
5. Conclusion
6. Introduction
7. Data

Two different data archives have been downloaded.

The first archive comes from the Genome Characterization Center (GCC) and contains txt files with gene samples and their corresponding expression level (log2 lowess normalized). The second archive which is provided by Biospecimen Core Resource(BCR) contains carefully cataloged tissue and sampleinformation with important medical information about the patient. Each text file from the BCR contains other information e.g. the gender of the patient or the treatment, used drug and some information about the tumor (size, site, location, etc.)

Among the biotab txt files, eight contained useful information which have the following names:

nationwidechildrens.org\_biospecimen\_cqcf\_luad.txt

nationwidechildrens.org\_clinical\_cqcf\_luad.txt

nationwidechildrens.org\_biospecimen\_tumor\_sample\_luad

nationwidechildrens.org\_clinical\_cqcf\_luad.txt

nationwidechildrens.org\_clinical\_follow\_up\_v1.0\_luad.txt

nationwidechildrens.org\_clinical\_omf\_v4.0\_luad.txt

nationwidechildrens.org\_clinical\_patient\_luad.txt

nationwidechildrens.org\_clinical\_radiation\_luad.txt