Ocular disease mechanisms elucidated by genetics of human fetal retinal pigment epithelium gene expression

Lab Journal Theme07 - Gene Expression Analysis

Lisa Hu
414264
Bio-Informatica
Hanzehogeschool Groningen, ILST
Marcel Kempenaar
8 March 2022

Contents

Exploratory Data Analysis

For decompressing the data, run the code chunks in the Rmd file that deem fit for your situation:

- If you downloaded the data from the official site: Run coded chunks `decompress-data` and `create-dataset`.
- If you want to use the dataset delivered with the project: Run the `decompress-dataset` code chunk.

After decompressing the data, the data can be read:

```r
## Read the dataset
dataset <- read.table("./gene_count.txt", sep="\t", header=TRUE)
## Indices for dataset
glucose.data <- seq(2, 47, 2)
galactose.data <- seq(3, 48, 2)

## Sample of the dataset
pander(dataset[0:5, 0:5])
```

Gene	X1_glucose	X1_galactose	X2_glucose	X2_galactose
__alignment_not_unique	0	0	0	0
__ambiguous	73052	71663	90130	114748
__no_feature	6143654	3901459	4560099	10675855
__not_aligned	0	0	0	0
__too_low_aQual	0	0	0	0

' To create a visual representation of the data, boxplots can be created:

```r
boxplot(log2(dataset[glucose.data]+1), main="Glucose", names=seq(1, 23),
        xlab="Sample", ylab="log2(Count)", outline=FALSE, cex.axis=0.6, cex.lab=0.8)
```
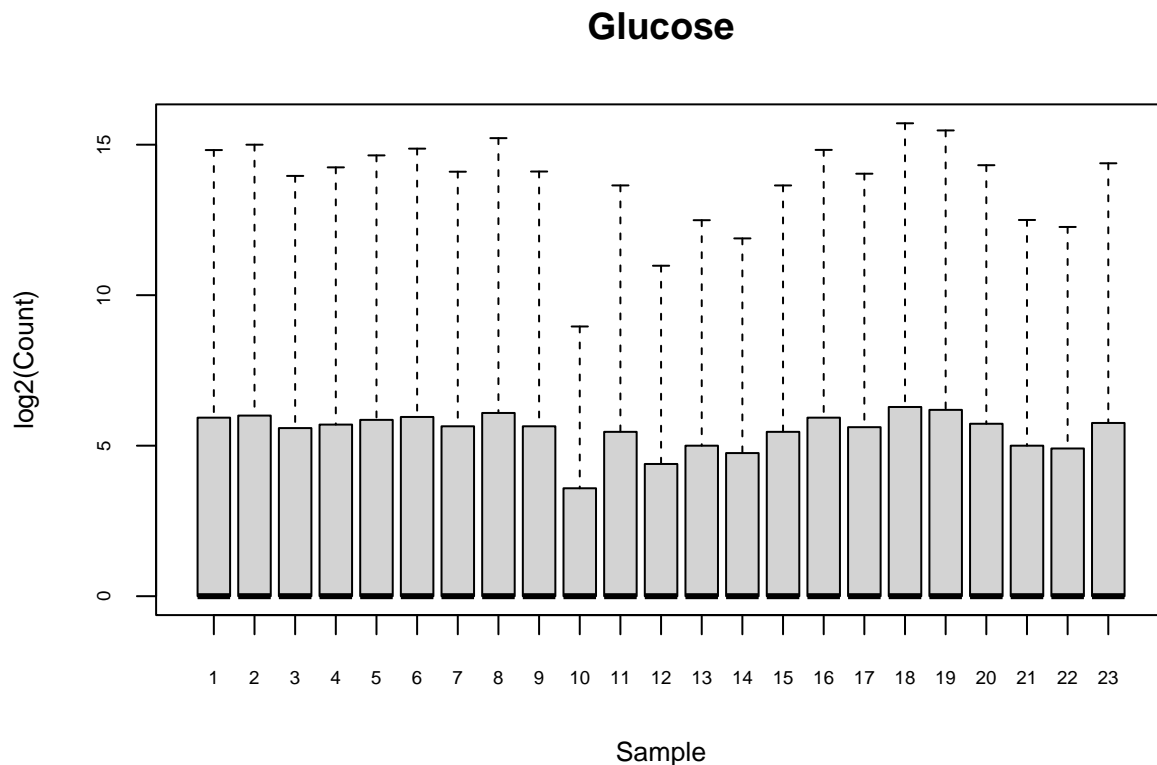
```
boxplot(log2(dataset[galactose.data]+1), main="Galactose", names=seq(1, 23),
        xlab="Sample", ylab="log2(Count)", outline=FALSE, cex.axis=0.6, cex.lab=0.8)
```

Galactose