Hippocampus TPM Counts

Kennedi Todd

2/8/2021

Set working directory

Load packages

```r
library(data.table)   # fread()
library(rtracklayer)  # to import gtf file
```

```
## Loading required package: GenomicRanges

## Loading required package: stats4

## Loading required package: BiocGenerics

## Warning: package 'BiocGenerics' was built under R version 4.0.5

## Loading required package: parallel

##
## Attaching package: 'BiocGenerics'

## The following objects are masked from 'package:parallel':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, parApply, parCapply, parLapply,
##     parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##     Filter, Find, Map, Position, Reduce, anyDuplicated, append,
##     as.data.frame, basename, cbind, colnames, dirname, do.call,
##     duplicated, eval, evalq, get, grep, grepl, intersect, is.unsorted,
##     lapply, mapply, match, mget, order, paste, pmax, pmax.int, pmin,
##     pmin.int, rank, rbind, rownames, sapply, setdiff, sort, table,
##     tapply, union, unique, unsplit, which.max, which.min

## Loading required package: S4Vectors

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:data.table':
##
##     first, second

## The following object is masked from 'package:base':
```

```
##
##      expand.grid

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

## The following object is masked from 'package:data.table':
##
##      shift

## Loading required package: GenomeInfoDb

## Warning: package 'GenomeInfoDb' was built under R version 4.0.5
```

Get info for downloading files.

```r
data.files <- read.csv(file = "tbi_data_files.csv",header = TRUE, sep = ",")
donor.info <- read.csv(file = "DonorInformation.csv", header = TRUE, sep = ",")

# Subset hippocampus data
data.files <- data.files[data.files$structure_acronym == "HIP",]

# get donor metadata by joining tables based on donor_id
df.hip <- merge(data.files, donor.info, by = "donor_id")

# Exract info
links.hip <- df.hip$gene_level_fpkm_file_link
donorid.hip <- df.hip$donor_id
rnaid.hip <- df.hip$rnaseq_profile_id
donorsex.hip <- df.hip$sex
```