

Task 7: Gene Set Analysis

Dongyuan Wu

2020-06-13

First of all, load data and packages what I need. The fetal brain, adult brain and adult liver datasets from roadmap epigenomics project with narrow peaks are downloaded using AnnotationHub. TxDb.Hsapiens.UCSC.hg19.knownGene is the hg19 annotation database will be used.

```
library(TxDb.Hsapiens.UCSC.hg19.knownGene)
library(AnnotationHub)
library(mygene)

deresult <- read.table("deresult.txt", header=TRUE)
degene <- subset(deresult, adj.P.Val < 0.05)

ah <- AnnotationHub()
ah <- subset(ah, species == "Homo sapiens")
fetal <- AnnotationHub::query(ah, c("EpigenomeRoadMap", "H3K4me3", "E081"))[[2]]
adult <- AnnotationHub::query(ah, c("EpigenomeRoadMap", "H3K4me3", "E073"))[[2]]
liver <- AnnotationHub::query(ah, c("EpigenomeRoadMap", "H3K4me3", "E066"))[[2]]

txdb <- TxDb.Hsapiens.UCSC.hg19.knownGene
txdb_gene <- genes(txdb)
```

Gene symbols were used to identify genes in previous tasks, while hg19 database uses gene entrez id. So I used mygene::queryMany() to transfer gene symbol into entrez id.

```
degene_entrez <- queryMany(degene$gene, scopes="symbol", fields="entrezgene",
                           species="human")$entrezgene
```

```
## Finished
## Pass returnall=TRUE to return lists of duplicate or missing query terms.
```

```
promoter <- promoters(txdb_gene[degene_entrez %in% txdb_gene$gene_id])

(fetal_perc <- length(subsetByOverlaps(fetal, promoter)) / length(fetal) * 100)
```

```
## [1] 30.70936
```

```
(adult_perc <- length(subsetByOverlaps(adult, promoter)) / length(adult) * 100)
```

```
## [1] 20.54484
```

```
(liver_perc <- length(subsetByOverlaps(liver, promoter)) / length(liver) * 100)
```

```
## [1] 17.05437
```

1. Are there changes in H3K4me3 between fetal and adult brain over promoters for genes differentially expressed between fetal and adult brain?

Yes. We can find there are 30.71% of DE genes in fetal brain while 20.54% of DE genes in adult brain, which means there are some genes expressed in fetal brain but not in adult brain.

2. Are promoters of genes differentially expressed between adult and fetal brain marked by H3K4me3 in liver?

Yes. There are 17.05% of DE genes in liver, which is different from adult and fetal brain.