

Gene Set Analysis

The Bioconductor package AnnotationHub() is used for this gene set analysis. The human hg38 annotation database is downloaded and saved as a TxDB object.

The differentially expressed genes from the statistical analysis was loaded and the gene ids were converted to entrez ids using the biomaRt package.

The annotation hub package was initialized and the respective narrow peak data was downloaded.

The differentially expressed gene list was loaded and the promoter() function was used to generate a genomic range object that contains all promoter ranges of the differentially expressed genes in adult and fetal brain. The subsetByOverlap() function is used to compute the percentage overlapped between the Differentially expressed genomic range object with the fetal, adult and liver genomic ranges.

Answer to question 1:

Yes, there are changes in H3K4me3 between fetal brain and adult brain samples. This is supported by the following:

- 1) There is considerable difference between the means of the peaks' width between adult brain and fetal brain samples (658 vs 1332). This suggests that there are more peaks in the adult samples, and it is consistent with the differences in overall sum of peaks
- 2) I found that about 15% of differentially expressed genes overlapped with the fetal brain and about 20% of differentially expressed genes are found in adult brain.

Answer to question 2:

No, changes in H3K4me3 between fetal brain and adult brains is not marked by H3K4me3 in the liver. This is supported by the following:

- 1) There is comparable mean widths between fetal brains and adult brain, suggesting that there are more peaks in the adult brain sample compared to the liver sample. Of the total sum of peaks of adult brain, only 13% (3840 / 29134) overlap with peaks in the liver sample