

Practical NGS processing and Analysis

The aim of this practical is to extend the hands-on to the comparison of brain and liver. Use the same R wrappers and similar commands used in the hands-on and follow these steps:

1. Perform differential expression analysis between brain and liver using the EdgeR. Present results using a heatmap with hierarchical clustering in rows and columns and colored classification of differentially expressed genes (DEGs), i.e. overexpressed in brain versus liver and the other way around.
2. Perform gene ontology enrichment analysis of the two sets of DEGs using the command line wrapper of GOstats R package for biological processes. Plot results using any graphical representation and discuss results.
3. Analyze differential splicing using SUPPA between brain and liver for skipping exon, intron retention, mutually exclusive exon and alternative first exon. Plot top results using heatmaps. Different thresholds may be chosen for each event type.
4. Find H3K4me3 peaks shared by brain and liver and the ones exclusively found in each tissue using the narrow peaks found in /tutorial/results using bedtools intersect. Show results using a bar plot colored by the color code used during the hands-on. Palette is available at /tutorial/palettes/palTissue.txt. Any color may be chosen for shared peaks.
5. Create a BED file of 200bp up/downstream TSS of genes and overlap DEGs (step 1) with the 3 sets of H3K4me3 peaks classified in the previous step (4). Show three examples in the UCSC genome browser, including RNA-seq, ChIP-seq and ATAC-seq tracks. Ideally, one example of each peak set (i.e. shared peak, peak exclusively called in brain and peak exclusively called in liver). Discuss the integration of the three datasets in the TSS of the selected cases.
6. Show two examples of alternative first exons in the UCSC genome browser, including RNA-seq, ChIP-seq and ATAC-seq tracks. Discuss the integration of the three datasets in the TSS of the selected cases.

Delivery:

1. A report (pdf or html) with plots, description and discussion of the results of each step.
2. A readme.txt file with commands used to complete each step with a description of the aim. Commands should be reproducible and error-free.