**Contents of RNA-Seq**

**Shell script files** (ending in.sh)

Gene\_counts

bam\_files

* Coverage
* Forward
* Reverse
* Gene\_counts

report\_fastq\_temp

GO

**Gene\_counts**

These contain differential gene expression tables and output performed using edgeR. “gene\_counts” contains the gene count table reflective of either, forward, or reverse strands. In the “dge\_output” folder there are Excel files with naming conventions where “temp” refers to 37 degree Celsius strain comparisons and “nontemp” refers to normal temperature condition comparisons. “All” refers to gene count tables that pertain to the forward or reverse strand, when mapping these reads to the reference, strandedness was taken into account. “fwdstrand” and “revstrand” refer to forward and reverse strands respectively, where gene count tables were only calculated using the designated strand.

Each tab refers to a mutant strain comparison. Columns look like this:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| (blank) | logFC | AveExpr | t | P.Value | adj.P.Val | B | Sample1 | Sample2 | Sample3 | Sample4 |

The columns to pay attention to are:

(blank) – this contains the systematic gene names

logFC – this reports the log2 fold-change

adj.P.Val – this contains the adjusted P-Value for the analysis

Sample1-4 – these are raw gene count values before transformation/normalization used for DGE analysis

**Bam\_files**

These are the aligned RNA-Seq data in binary format.

**Coverage**

“coverage” – within here are individual sample files that can be visualized using IGV or IGB

* Reverse\_strand – these are individual sample files on the reverse strand
  + Individual\_samples – each sample for a given strain
  + Mean\_coverage\_reversestrand – these are mean comparisons of strains (reverse strand)
    - Log\_mean\_comparisons – these look at log2 comparisons of the mutant strain compared to wildtype in IGV or IGB (reverse strand)
    - Subtract\_mean\_comparisons – these look at the difference between mutant strain and wildtype to be visualized in IGV or IGB (reverse strand
* Forward\_strand - these are individual sample files on the forward strand
  + Individual\_samples – these are each sample for a given strain
  + Mean\_coverage\_forwardstrand - these are mean comparisons of strains (forward strand)
    - Log\_mean\_comparisons - these look at log2 comparisons of the mutant strain compared to wildtype in IGV or IGB (forward strand)
    - Substract\_mean\_comparisons - these look at the difference between mutant strain and wildtype to be visualized in IGV or IGB (forward strand)
* Mean\_coverage\_bothstrands- these are mean comparisons of strains (either strand)
  + Individual\_samples – each sample for a given strain
  + Log\_mean\_comparisons - these look at log2 comparisons of the mutant strain compared to wildtype in IGV or IGB (either strand)
  + Substract\_mean\_comparisons- these look at the difference between mutant strain and wildtype to be visualized in IGV or IGB (either strand)

**report\_fastq\_temp**

This is the QC file generated from the sequencing facility

**GO**

This directory contains GO output from ontologizer. Genes where B-H adjust P-value < 0.0001 for each differential gene expression comparison (all strands, forward, and reverse) were fed into the Gene Ontology program. The output with “table-“ refers to the GO table where each row is a GO term and the “p.adjusted” column refers to the B-H adjusted P-value calculated by the program. Be sure to report that column when referring to significance of a particular term and **not** the “p.min” column. These files can be viewed if opened in Excel.

The output with “anno-“ comprises the list of significant genes as well as the associated GO terms that are listed in the “table-” file.

The naming scheme of each file is similar to what is used before:

Table = GO table

\_[strain]\_ = ex: mutant strain sir2, lsd1, lsd2 etc.

\_nontemp\_ or \_temp\_ = nontemp is normal temp conditions and temp is 37 degrees Celsius

\_all\_ or \_fwd\_ or \_rev\_ = both strands, forward, or reverse strand

Anno- = annotation file with the list of significant genes as well as the associated GO terms reflected in the table file