

First, you will need to log into the AWS and follow:

Day7_installing_Rsubread.docx

Day7_featurecounts_worksheet.docx

We are on the AWS because Day7_featurecounts_worksheet.docx requires the bam files.

Then you will need to log off the AWS to do

Day7_differential_expression_worksheet.doc

It is easier to use Deseq2 on your own computer with Rstudio and the counts file you need for input is very small.

After this you have two options.

1. I'm interested in learning what to do with gene lists
2. I'm interested in learning how to get R to run on a supercomputer with a input and output file.

See below what to do for each of those options.

1. I'm interested in learning what to do with gene lists

After you get done with Deseq2 you will have a list of genes that is significantly different. You could also make a pvalue ranked list of all genes.

You could use those gene lists for Go analysis or GSEA analysis. Check out these worksheets for how.

The worksheet:

Genelist_to_go.pdf shows you how to use go.

And you can find out how to use GSEA here.

<https://github.com/Dowell-Lab/codeclub/blob/master/gsea/gsea.md>

In general I think GSEA is more robust.

2. I'm interested in learning how to get R to run on a supercomputer with a input and output file.

Go to the scripts folder of day 7 and look inside the routin directory. It has another way to count using feature counts. What are the pros and cons of this other way?