Grid Engine is the job queuing program installed on the Godel cluster. It accepts your jobs and tries to optimally distribute them in the cluster according to your specified options in a qsub script. Carlyle has prepared a sample qsub script that you can find in /usr/global, called “qsub\_test.sh.”

I’ve copied over his sample script into my directory /home/pxu/Bryan, modified it and renamed it qsub\_rnaseq.sh. You’ll notice that the first half of it consists of qsub options (preceded by **#$**,which signifies it as a parameter to be submitted to Grid Engine). The second half is a series of bash commands that Carlyle provided for debugging purposes.

Some important details on the qsub options (which you can find in the first 30 or so lines):

**#$ -e ./**

**#$ -o ./**

\* generates standard error and standard output files in your current working directory, you can change which directory you want them in, or comment it out if you don’t need it.

**#$ -N tophat**

\* specifies the label that will show up for the job, and the base-name for your error/output (in this case, it will show up as tophat.e##### and tophat.o#####).

**#$ -l mem\_free=33G**

\* forces the job to runonly on nodes that have at least 33 gigabytes of free RAM – you can increase/decrease as necessary. (Godel’s nodes have up to 100 gigs of RAM, I believe)

**#$ -pe smp 8**

\* specifies how many threads/slots to use; this number should match your “-p #” in your tophat and cufflinks commands

**#$ -m eas -M chimb@mymail.vcu.edu**

\* tells the cluster to send you a notification email upon the events specified (you can see the options in the comments above it). You should change the email to your own to receive notifications, and add/remove the options as you see fit (b e a s n).

You’ll also notice near the beginning of the bash code, I’ve included:

**source /home/pxu/Bryan/paths.sh**

\* this runs my shell script, paths.sh, which (if you take a look) simply appends all the necessary paths of bowtie, tophat, cufflinks etc. that are needed for RNA-seq analysis. If you update your software, you’ll need to edit this script to include the new path to the updated version(s).

To use this qsub script to submit a job to Grid Engine, simply enter (with the script in your cwd):

**qsub qsub\_rnaseq.sh “<command>”**

For example, to submit a tophat run:

**qsub qsub\_rnaseq.sh “tophat -p 8 –G Ss\_SK36.gff Ss\_S1\_thout Ss\_SK36 S1.fastq.gz”**

\* Note that the double quotes are required to encapsulate the entire command within one string, which is accepted and submitted to Grid Engine by the qsub script.

\* To check the status of your runs, simply enter **qstat** in your shell.

To submit a batch of jobs at the same time, simply create a batch file containing a list of the whole commands, one command per line.

For example, touch qsub\_tophat\_batch.bat and then list your commands in it:

**#! /bin/bash**

**qsub qsub\_rnaseq.sh “tophat -p 8 –G Ss\_SK36.gff Ss\_S1\_thout Ss\_SK36 S1.fastq.gz”**

**qsub qsub\_rnaseq.sh “tophat -p 8 –G Ss\_SK36.gff Ss\_S2\_thout Ss\_SK36 S2.fastq.gz”**

**qsub qsub\_rnaseq.sh “tophat -p 8 –G Ss\_SK36.gff Ss\_S3\_thout Ss\_SK36 S3.fastq.gz”**

Then you will need to set the file permissions to be executable. So in the shell, you would enter:

**chmod +x qsub\_tophat\_batch.bat**

Then simply run it to submit all the jobs you’ve listed:

**./qsub\_tophat\_batch.bat**