Session 5 Exercises

   de novo assembly of transcriptome sequences using iAssembler

Step 1. Log into the CAC Linux server. Create a session5 directory under your home directory on the CAC server. Copy all data files of the project to your session5 directory.

   mkdir session5

   cd session5

   cp -R /home/gfs08/jp86/ngw2010/session5/lecture2/* ./

   After you finish these steps, make sure you see the following files by typing "ls -l" followed by <u>Enter</u> key. "ls -l" command would give you the size of the file (in bytes), and last time it was modified.

   iAssembler-1.0b.x64 (directory containg the iAssembler program and the test data (test.fna))

   run_iAssembler.sh (a script for submitting a job to run iAssembler)

Step 2. Submit a job to run iAssembler.

   • Modify the first five lines in run_iAssembler.sh file.

      Modify the line "#PBS -A jp86_0004". Change "jp86_0004" to your own project name. For this project, v4 cluster (16G RAM) would work. Sometimes, the queue for the v4 cluster is very long, you might want to use v4-64g instead.

   • Submit the job by typing this command line followed by <u>Enter</u>.
     nsub run_iAssembler.sh
     The job would take a couple of minutes to finish. You can monitor the progress by "qstat" command. After the job is finishes, you will see a new folder created under the session5 directory: test.fna_output. The folder contains all the output files (all in txt format and can be opened by wordpad on Windows) from iAssembler assembly. Copy the folder to your local computer and examine the assembly results.