VEME *de novo* assembly practical session
2013-08-29

Assembly with Velvet on Galaxy

1. Log into UF HPC's Galaxy instance: http://galaxy.hpc.ufl.edu/
2. Get some data:
    a. Shared Data: Data Libraries: VEME2013: wine_yeast.100K.fq
3. NGS: QC and manipulation: **FASTQ Groomer**
    a. Input FASTQ quality scores type: Sanger
4. NGS: QC and manipulation:FastQC: **FastQC:Read QC**
    a. Use the defaults or add a title for easier reference later
    b. Notice poor quality at ends of reads
5. NGS:QC and manipulation:**FASTQ Quality Trimmer**
    a. Window size: 5
    b. Quality score: 30
    c. Rerun FastQC on trimmed dataset
6. NGS: Assembly:**Velvet**
    a. pre-process fastq data with velveth
    b. assembly with velvetg

VEME *de novo* assembly practical session
2013-08-29

Assembly with SHARCGS in batch mode

1. Log into the UF HiPerGator Supercomputer

   Use 'gator.hpc.ufl.edu' as the host name in the PuTTY connection dialog.

2. Copy the training dataset to your scratch space

   a. $ cd scratch

   b. $ $ cp -a /project/bio/training/VEME/ .

3. Optional: Modify the email address and set the email options for the batch system by editing the sharcgs.pbs file with nano or vim.

   a. Nano sharcgs.pbs

4. View the job script (Use pageUp and pageDown or arrow keys to traverse the file and “q” to exit the pager)

   a. $ less sharcgs.pbs

5. Submit the batch job

   a. qsub sharcgs.pbs

6. Check the job status (substitute your account's id for XYZ)

   a. $ qstat -u vemeXYZ

7. View the results of the run

   a. $ less

Extra: Assembly with Velvet in batch mode

1. View the velvet.pbs job script
2. Submit the velvet.pbs to the batch system
3. Check the job status for completion
4. View the results in the kmer21 and compare the results from runs with other kmer lengths in auto_kmer_XX directories
    a) $ cd kmer21
    b) $ less Log
    c) $ cd ../auto_kmer_15
    d) $ less Log
    e) $ cd ..
    f) $ tail */Log