**to download last lecture’s images, use the following commands from your terminal or command prompt (not logged into ssh):**

*on linux:*

scp epp622@160.36.207.XX:/home/epp622/Documents/EPP622/projects/<your\_project\_folder> /qc\_pipeline/reads/qc/\*.png /path/on/local/machine/

*on mac:*

scp scp epp622@160.36.207.XX:/home/epp622/Documents/EPP622/projects/<your\_project\_folder> /qc\_pipeline/reads/qc/\*.png ~/path/on/local/machine/

*on pc* (works with putty if pscp installed to path):

pscp epp622@160.36.207.XX:/home/epp622/Documents/EPP622/projects/<your\_project\_folder> /qc\_pipeline/reads/qc/\*.png c:\path\on\local\machine\

**connect to the lab computers via ssh (use the computer with your existing project directory if possible)**

ssh epp622@160.36.207.XX

the project directory is organized as follows:

snp\_pipeline

reads

Key\_SNPcall.txt

subs\_Beauregard.fastq

subs\_Covington.fastq

subs\_DM004.fastq

subs\_NewKawogo.fastq

subs\_Tanzania.fastq

refgenomes

trifida\_chr\_15.fasta

scripts

GBS\_align.sh

GBS\_call.sh

GBS\_format.sh

GBS\_index.sh

SNP\_Filtering.R

tools

picard (directory containing complete toolset)

vcftools (directory containing complete toolset)

GenomeAnalysisTK.jar

**find your project directory and copy the snp\_pipeline folders into it**

cd Documents/EPP622/projects/<your\_project\_folder>

cp -r ../../snp\_pipeline/ ./

**from the scripts folder, run a series of shell scripts:**

cd snp\_pipeline/scripts

bash GBS\_index.sh

-creates an index of our reference genome using Burrows-Wheeler transform

bash GBS\_align.sh

-creates a project folder 'batatas' where all alignment files will be stored

-pulls a list of sample names we will map from the 'KeySNPcall.txt' file from the 'reads' directory

-uses awk to write a more involved shell script ('presnpcalling.sh') for mapping our reads to the reference

bash presnpcalling.sh

-bwa aligns each of our samples to the indexed reference genome creating a .sam file

-picard tools organizes and indexes our .sam file, also creates .bam file

-samtools indexes our .bam file

bash GBS\_call.sh

-GATK snp calling and filtering

bash GBS\_format.sh

-pulls only the columns of interest for use in R graph production

**visualization:**

(to create graphs based on our snp discovery, change the following R script to include your project directory)

nano SNP\_Filtering.R (change the #### portion of line 11 to your project folder name, then save by hitting ctrl + o, hit enter, then crtl + x to exit)

Rscript SNP\_Filtering.R