**A Few Basic UNIX Commands**

pwd Stands for Present Working Directory. Tells you where you are.

ls Short for “list.” Lists files and folders in the present directory.

cd [dirname] Changes directory to the directory named.

**Assembling With Newbler (aka GS De Novo Assembler)**

1. In the virtual machine, select **GS De Novo Assembler** from the menu bar on the left.
2. Select a directory and name for the new assembly project.
3. Go to the project tab and use the **plus** sign to add reads, either .sff files (GS Reads tab) or .fastq files (FASTA and FASTQ Reads tab).
4. Go to the **Parameters** tab, then select the **Output** sub-tab, and select “Complete Consed Folder” under the “Ace format” field.
5. Click the **Start** button on the right to begin the assembly.

**Selecting a Subset of an .sff File**

1. Open a terminal in your Virtual Machine, and navigate to the folder containing the master .sff file.
2. Run the following command:

sffinfo –a [name of master sff] | head –n [number of reads desired] > [name of output file]

Example:

sffinfo –a Giles.sff | head –n 30000 > Giles30kReads.txt

1. Next run this command:

sfffile –i [name of output file from Step 2] –o [name of new sff file] [name of master file]

Example:

sfffile –i Giles30kReads.txt –o Giles30k.sff Giles.sff

**Running consed**

1. Open a terminal in your Virtual Machine.
2. Navigate to the edit\_dir of an assembled project.
3. Type consed and press Enter.
4. Select the desired ace file, and click Open.

**Finding Weak Areas**

1. Use consed’s low quality consensus flag.

In the *Aligned Reads* window for a particular contig, go to **Navigate 🡪 Low Consensus Quality**. A new window titled “Low Consensus Quality” will open. You can double click a line in this window to go to that position, or use the buttons at the bottom of that window to jump between examples.

**NOTE:** You can change what is considered “Low Quality”. The default cutoff value is 25. To change this number, go to the *Consed Main Window* and select **Options 🡪 General Preferences**. In the Preferences window that opens, change the first value “Threshold for Low Consensus Quality.” For example, to be more stringent, change this value to 40, then go back to the *Aligned Reads* window and try **Navigate 🡪 Low Consensus Quality** again.

1. Run aceUtil.

In your SEA Virtual Machine, open a terminal. Type “aceUtil” (case-sensitive) and press Enter. When the small window opens, select **File 🡪 Set Input File** and navigate to the ace (consed) file you’d like to check. Wait for a moment, and a contig name should appear in the “contig” field. Then go to **File 🡪 Set Output File** and navigate to the same directory where the input file occurs, and name your output file to be the same as the input but increment the final digit by 1. Then select **Preferences 🡪 Settings** and set the discrepancy cutoff to 25%, and the low coverage value to 12 (or use your preferred values here!). Then go to **File 🡪 Analyze**. Wait until the Terminal displays how many discrepancies and low coverage regions were found, then select **File 🡪 Close**. Open the new ace file in consed, open the aligned reads window for the selected contig, then go to **Navigate 🡪 Tags 🡪 comment**.

1. Scan the genome by eye, screen by screen.

**Picking Primers in consed**

1. In consed, open the *Aligned Reads* window for the contig in question. Navigate to the area you’d like to fix.
2. Right-click the consensus at the weak area. From the dropdown menu that appears, select either the 3rd or 4th option that appears (**Primer from Clone**…NOT **Primer from Subclone**).
3. Look at the list of primers that appears, and select your favorite.
4. The consensus will be tagged with a yellow line indicating the primer position, and a red dot indicating the 3’ end of the primer.
5. There are many options that can be changed relating to how consed picks primers. Go to the *Consed Main Window* and select **Options 🡪 Primer picking preferences** to view and/or change these parameters.

**Adding Sanger Reads to a consed Assembly**

1. Place the chromatogram (.ab1) files in the project’s chromat\_dir.
2. Make a new file with the extension .fof (for **f**ile **o**f **f**ilenames) and save it in the project’s edit\_dir. This file should contain the exact names of the Sanger read files, one per line.
3. From the edit\_dir, type consed and press Enter to launch consed.
4. In the *Consed Main Window*, click the **Add New Reads** button.
5. From the **Files** pane on the right, select your .fof file.
6. Click **OK**.
7. Once the process has finished, save the ace file.