We are starting the finishing this week!

1. Double click on the StartXWin-Cygwin icon (

). This opens a graphical version of Cygwin.
2. You will see a command line starting with a "$". This is a prompt where you type in your input.
3. Enter "ssh –Y xxx@199.107.197.20", and then the enter key. Note that the ssh must be lower case and the Y must be uppercase.
4. When asked for the password, type it in.
5. You will see "xxx@biology1 ~ $". The green text (xxx@biology1 is the name of the account and computer. The blue text is the directory and the prompt ($).
6. Each person will have a different project. The data we will finish are in the directories 120-D14, 130-G02, 145-P18, 150-G05, 175-P12, 180-B08, or 190-K08. To change the directory, type "cd xxxx", where the xxxx is your project directory, and then enter.
7. Now enter "cd edit_dir", and then enter.
8. To start consed, type "consed &", and then enter. The program will open.

Checking the data

9. Go to Assembly view to see how Consed assembled your data.
10. If you have lines below your contigs, there are inconsistencies. In most cases, the two sequences are just a little farther apart than they should be, so it isn't a worry.

    For those with more difficult contigs, some of the reads may have been misassembled. The misassemblies usually occur because of repeats that confuse the assembler. You may have to pull these reads out and/or reassemble the contigs.
11. Follow the directions in the "Strategies for Finishing" handout.
12. When calling the additional reads, follow the directions in the "Calling additional reads" handout. However, pick them from the **clone** (fosmid), not the subclone.
13. Make primers to close your gaps. There should be one primer from side of the gap, pointing into the gap. Check to see why the gap may be there. In some cases, there are dinucleotide repeats (e.g., ATATATATAT). These are hard to get through. When you call the chemistry for the reads, call both the 4:1 and dGTP chemistries for these.
14. Check the far left and right ends of the contigs for low quality bases. If you have just a few (20 or so), convert them to X's (in the trace window). If you have many, design a primer to improve the quality in that region. Since the fosmids overlap by quite a bit, not having good sequence for those 20 or so bases isn't important.
15. Improve the other regions that have low quality regions. You usually only need to call one read for these, and usually just the Big Dye chemistry.
16. When you are finished, save the latest ace file to windows. You can do this using WinSCP3. Drag the file to a windows drive or directory, such as the T drive. The ace file is the project name, such as 280-G02, then *.fasta,screen.ace.x*, where x is the highest

numbered ace file. For example, “*280-G02.fasta.screen.ace.2*”. It is in the edit_dir directory of your project.

17. To end the Linux session, enter “logout”.
18. To exit Cygwin, enter “logout”.

Submitting reads to Washington University

19. To submit the reads, log into the GEP project management system (gep.wustl.edu).
20. Click on “Order finishing reactions”.
21. Select the appropriate project and the ace file with all the oligo tags.
22. Click “Submit”.
23. Customize the reactions by selecting the correct chemistries. Call the reads with both 4:1 and dGTP chemistries **if there are dinucleotide repeats in the region that will be sequenced**. Otherwise you would normally call with BigDye reactions (the normal type).
24. Click on “Use selected oligos”.
25. Click on “Confirm reaction order”.
26. Confirm that the correct data is listed and then select “Submit order”.
27. Save the output page containing the names of the reactions for later use.