Week 7

Computer Lab

SeqTrace & BLAST: Creating a consensus sequence & making sense out of it

**Background** The Sanger facility sends us forward and reverse reads… we need to:

1. cut off the crappy end parts,
2. “overlap” the trimmed sequences by matching up the base-pairs,
3. get a “consensus” sequence from the two reads

**How to open the SeqTrace program**

Go to your folder on the desktop

Got to the class software folder

Double click on SeqTrace

**Get your files on the desktop**

Download all your sequence files from the Google Drive. Make sure to unzip the downloaded folder.

(1 pair for every bug that you're working on; 27F and 1391R)

Example: For sample 94, download 008\_**94+27F**\_A02.ab1 001\_**94+1391R**\_H02.ab1

(Read number\_samplenumber+primer name\_plate position.ab1)

\*\*\*\*You can control click to highlight lots of files and right click to download all at once

Now go to SeqTrace

**Edit and create a consensus sequence with seqTrace**

1. Create new project (File → New Project) → Okay \*\*\*don’t change the defaults
2. To add files, go to “Traces” and click on “Add trace files”, then select the reads from the folder on the desktop (.ab1 files) you want to work with.
3. Click the arrow icon, in front of your REVERSE read file (1391R) (this designates it’s to be reverse in SeqTrace) make the arrow look like this: ←
4. Group your forward and reverse reads by highlighting both of them and clicking “Group selected forward/reverse files” (under “Traces”). **Name the group with the sample name!** Press Ok.
5. Under “Sequences” go to “Generate Finished Sequences” and click on “for all trace files”. (if it asks you to overwrite, say YES)
6. To view your consensus sequence, click on the read pair group and then click on the **magnifying glass icon** at the top of the page.

*The “Trace View” shows the quality scores, the chromatogram (trace) display, and the raw base calls from both the forward and reverse reads, as well as the consensus sequence. The consensus sequence is at the bottom.*

1. To export the consensus, go to “Sequences”, hover on “Export Sequences”, “From all Traces”

and then...

Find the desktop, then save.

Name it: **[lastname\_sample1\_sample2\_sample3].fasta**

**Example: Coil\_25\_26\_48.fasta**

*This will create a file containing the consensus sequence, which can then be used for analysis such as searching for closely related sequences using the BLAST program (Altschul, 1990) which can be used to identify the organism.*

**Fix up the first line of each FASTA sequence**

1. Open your Fasta file (saved to the desktop) using Notepad

2. Cleanup/Rename the 1st line to “>[sample#]” For example here is the before and after of a line in the sequence file:

Before:

>..\..\..\dnasequencingresults24765\003\_100+27F\_F02.ab1, ..\..\..\dnasequencingr

After

>100

**Copy your FASTA file back into the “Consensus” Google Drive Folder for consensus sequences (or we will lose it when you log out)**

**Also save your SeqTrace project and put it into the “SeqTrace” Google Drive Folder**

**BLAST your consensus Sequence**

1. Navigate to the Standard Nucleotide BLAST at NCBI:
2. <http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome>.Upload your consensus fasta file.
3. Check the box to exclude Uncultured/environmental sample sequences, since these will not be informative for identification. Be sure the nucleotide collection (nr/nt) is selected under database and click the “BLAST” button (Fig. 6).
4. Record your results i.e. list the **species** you observe on the first page of hits in your individual lab notebook - in order. In the compiled lab notebook, list only the first 4 (there might not be 4).

RDP - upload instructions

1. Go to<http://rdp.cme.msu.edu>.
2. Create an account.
3. Click on “my RDP/login”.
4. Upload the fasta file containing your 16S rDNA sequences.
5. Assign it a group name (your name is fine).