Forming and editing contigs in GeneStudio

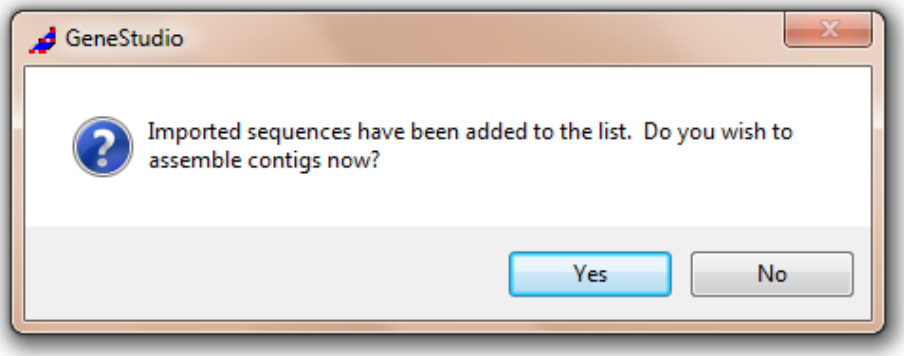
Before we run analyses on the sequences that we just produced we must first combine the forward and reverse sequences. This is called forming a contig. If you remember we used forward and reverse primers in the PCR reactions. We also sequenced the PCR products with the forward and reverse primers separately. In producing these two sequences we can align them (contig) and be sure that the base calls are correct. This tutorial will allow you to form and edit a contig from the sequences you produced. We can then go on to analyze these sequences.

Step 1: Open GeneStudio

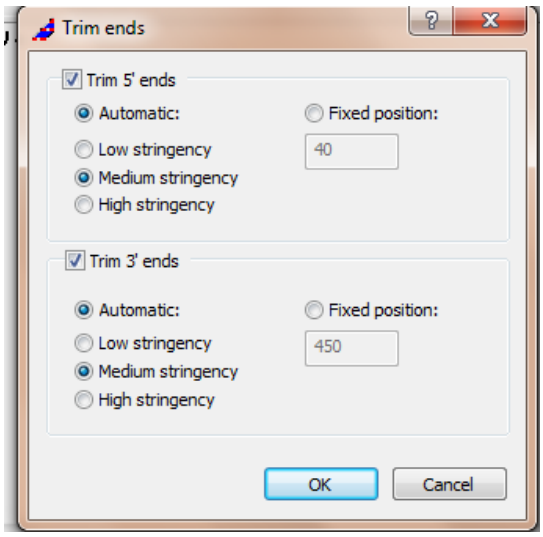
Step 2: On the right side of the window click the Contig editor button

Step 3: Click the File tab and then click "Import sequences…."

- Navigate to the file that contains your raw sequences. In this case they should be in the file folder "Raw_Sequences". Select the forward and reverse sequence of one sample (example: 1511L_A03 and 1511H_A01). In this case the specimen is 1511 and the forward product is 1511L and the reverse is 1511H. Once selected click ok to import. Click yes when this window pops up.



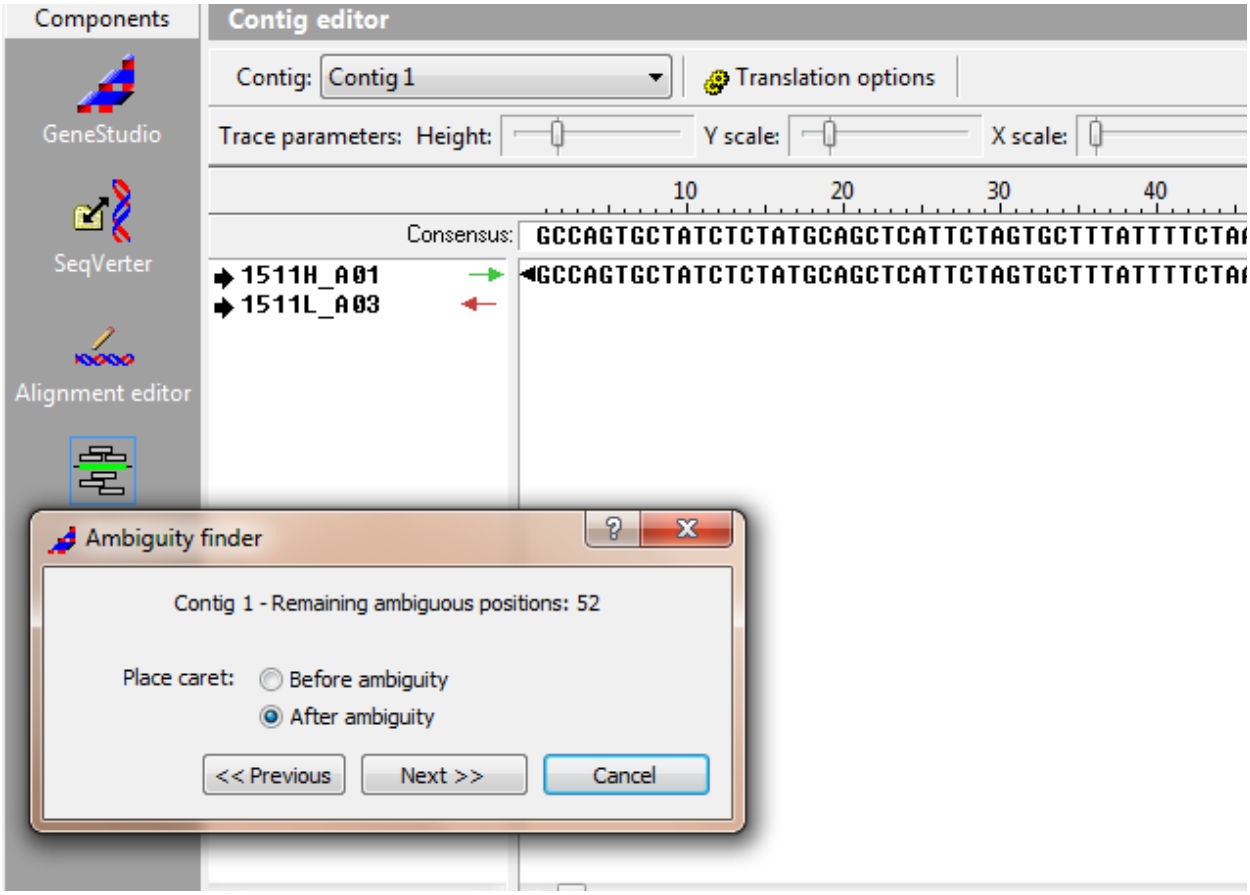
Then ensure that the following are checked in the next pop up window



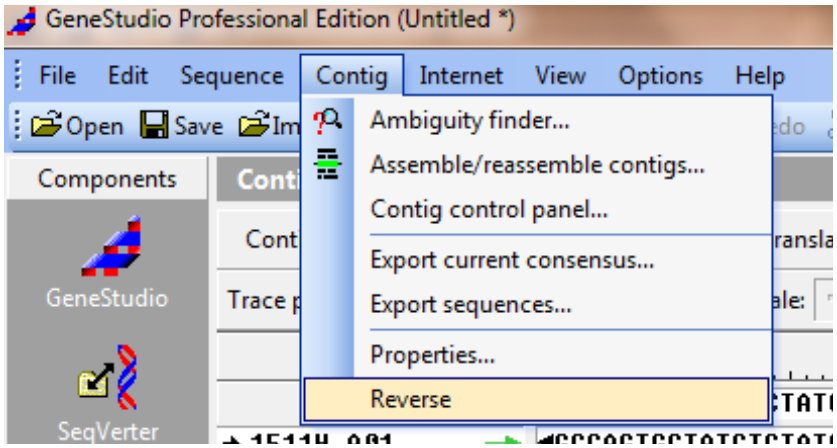
Click Yes and then click “Edit Ambiguities”

Step 4: Ensure that your alignment is not reversed

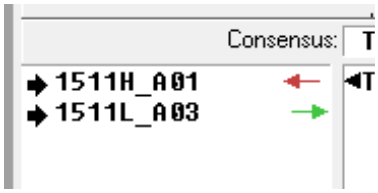
- Looking at the window you can see the consensus sequence and the two original sequence. Make sure that your forward primer is correct



In this case our reverse sequence (1511H) is going forward (green arrow) and the forward is the reverse (red arrow). To correct this go up to the Contig menu and click and then click reverse.

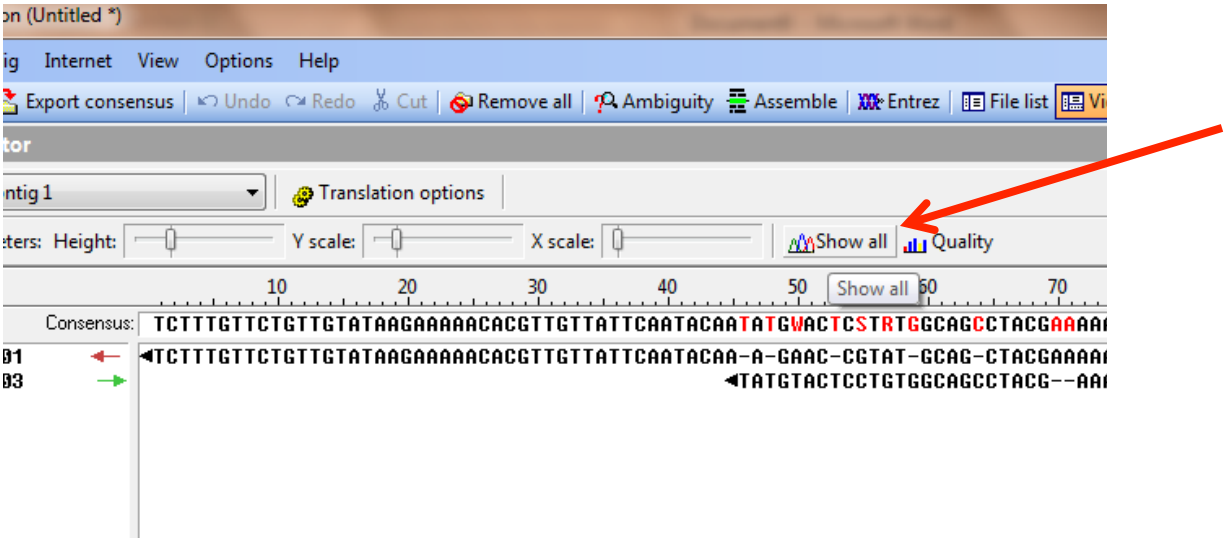


Now our sequences are aligned correctly. It is very important to make sure you know what the forward and reverse sequences should be.

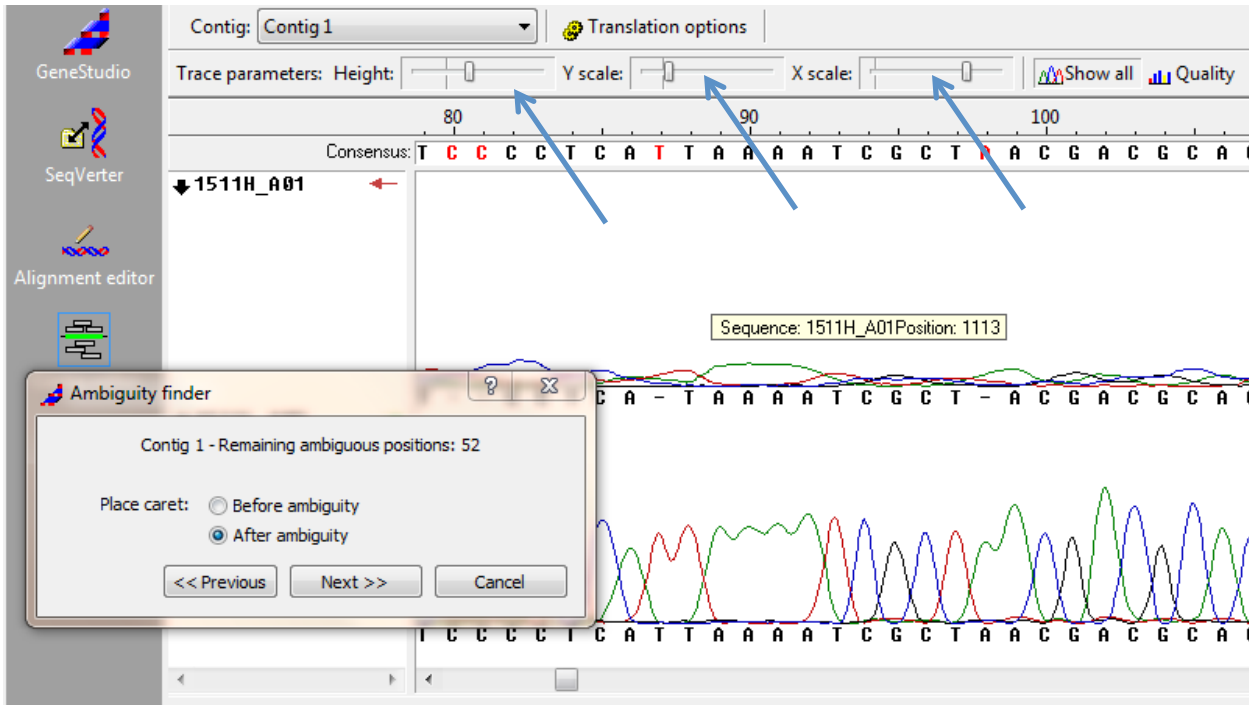


Step 5: Editing contig

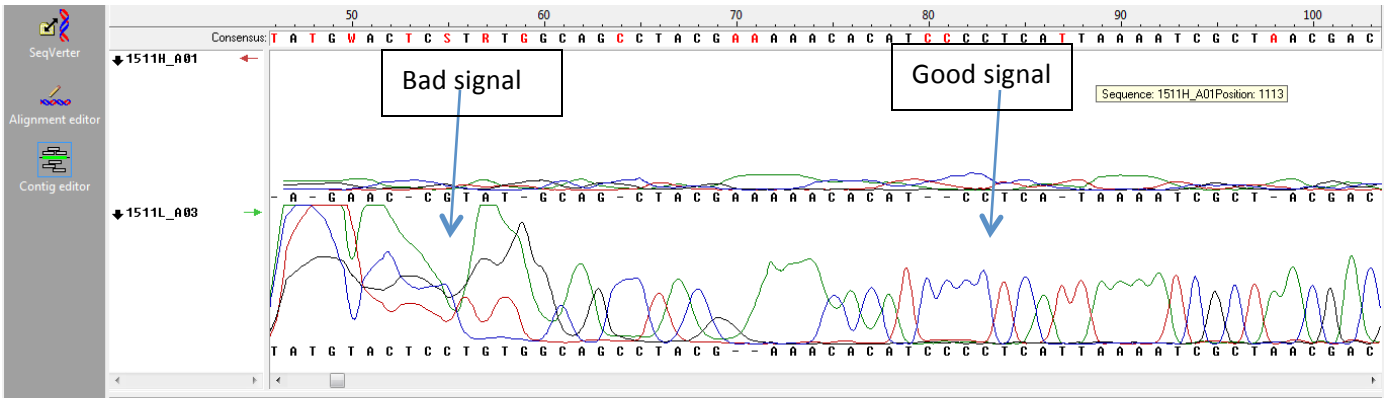
- Click show all, this will display the chromatograms of each sequence.



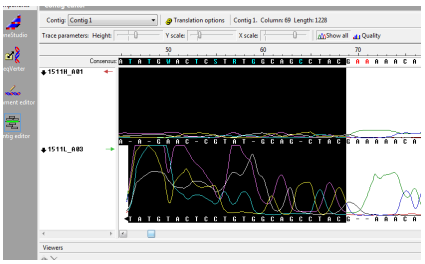
- We can make the chromatograms easier to see by adjusting the trace parameters.



- We begin by editing the beginning of the sequence. It usually takes a few base calls for the sequencing read to stabilize.

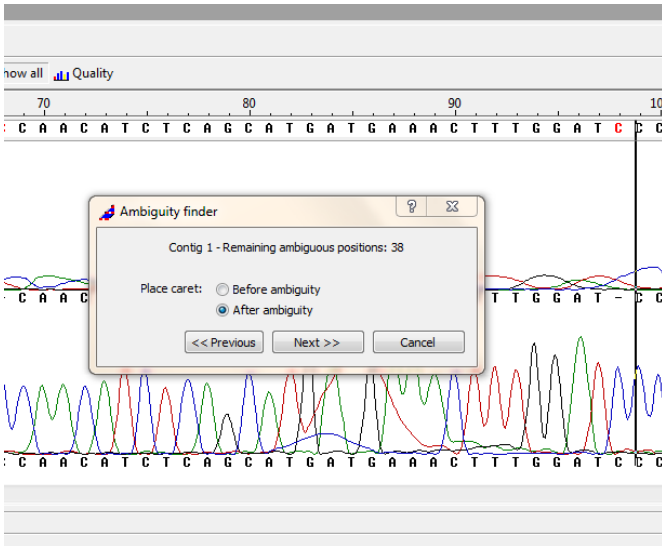


- Let us delete the first 15-20 bp of the forward sequenc by selecting those bases in the consensus.

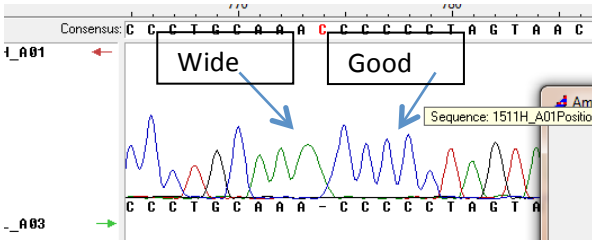


- Once selected hit delete.

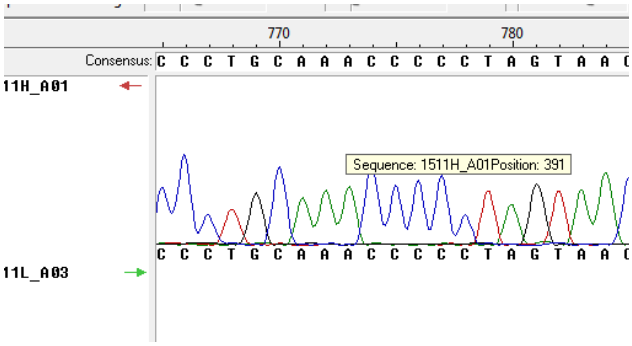
- Select the ambiguity tab and then click “next”



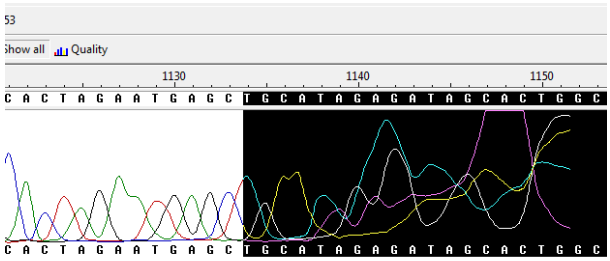
- Select the red C and type another c, this changes the c from red to blue. Click next, sometimes an additional base is called that is not really there. Notice the red c in the image below. Notice how wide the peak is compared to the surrounding sharp peaks.



- Select the red c and delete it.



- Now the peaks are even. Continue to the next ambiguity until you are near the end of the sequence. Once at the end we can delete the unclear base calls

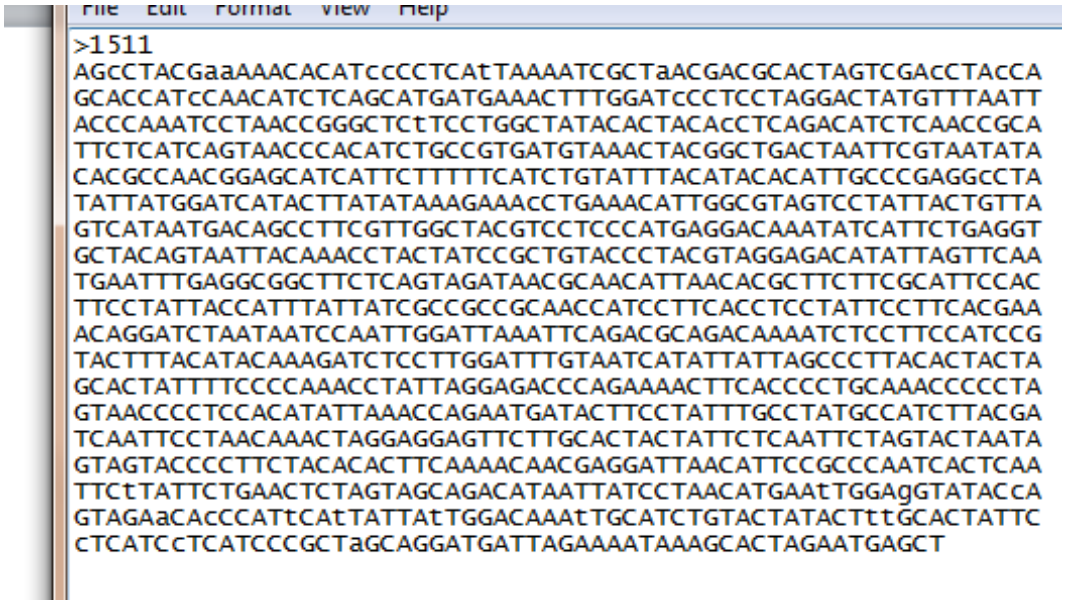


- Scroll through the whole consensus alignment to double check you have corrected all ambiguities.

Step 6: Renaming and exporting consensus sequence.

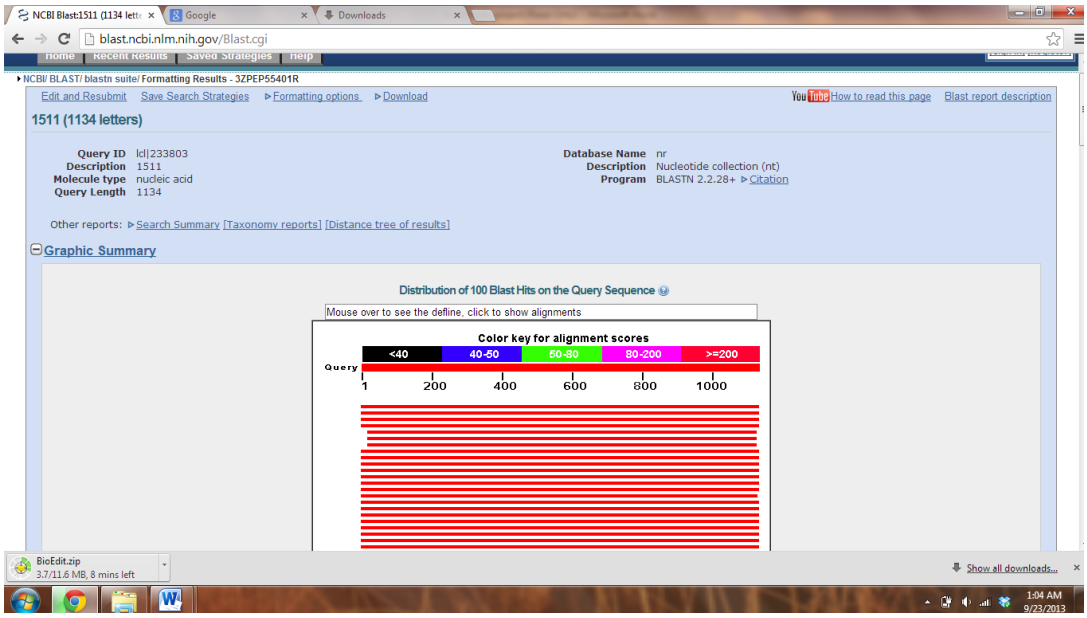
- Go to the contig tab and select properties
  - Change the name to the specimen number (in this case 1511)

- Go to File and select "Export current consensus"
  - Navigate to your edited sequences folder within your project files and set the file name as the specimen number (1511) and save the file as a FASTA file. Here is what that file will look like:

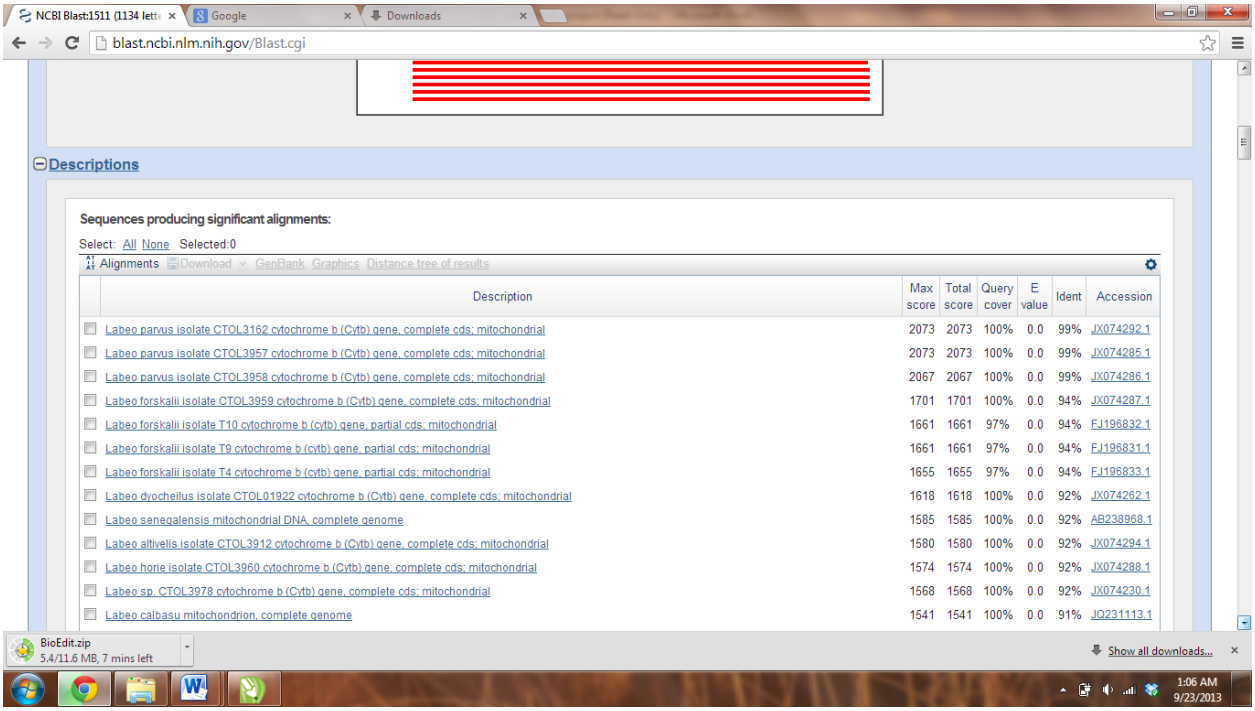


Step 7: BLASTing your sequence with NCBI genbank.

- Go to the BLAST website http://blast.ncbi.nlm.nih.gov
- Enter your sequence in the sequence query window. You can either upload your file or copy and paste it in the window. Make sure that the Nucleotide database is selected and hit BLAST
- You should see something like the following



- Scroll down and you will see a list of the most closely related sequences in the database.



- We see that *Labeo parvus* CTOL3162 is close to our query sequence. Since we have identified our specimen as *Labeo cylindricus* it seems like we have done everything correctly.
- Form and edit your remaining contigs and verify them in BLAST. Once completed with that we can combine all sequences in an alignment. We will work on that next.