Assignment: DNA sequence analysis

Fill this out **as you work through the data analysis**. You will need to include all of this information in your paper

1. After sequence reads were obtained from the miSeq sequencer, what data processing steps occurred before the sequences were given to you? (This information is in Lecture Video: DNA sequence analysis 1, ant Instructions: DNA sequence analysis). You'll need to add this information to your Materials and Methods section.

2. What herbal supplement(s) did you sequence? Did you get data from either of your samples?

3. **If you didn't get data** from either of your supplements, analyze data from another student. Use that supplement to answer the rest of the questions on this worksheet and to write your paper. Be sure to update your Materials and Methods so it is about this new supplement. What supplement(s) will you analyze today?

4. Look at the Krona graph. What species seem to be in your supplement? List both scientific and common names. Are the species identified with high confidence (green) or low confidence (red)? What is the percent abundance of each species (%)?

5. How many **high quality DNA sequence read pairs** were used in the analysis of your supplement(s)? Find this data on the Krona plot next to "**Count**" in the upper right hand corner. If there are <1000 high quality reads, your data might not actually be from the supplement you think you sequenced.

6. **For the rest of the questions, just use one supplement** (tell us which one). BLAST the consensus sequences from the supplement. Look at the first tab, showing results from your first consensus sequence. What subject species did your sequence hit? Are the top ~10 hits all from the same species? Are they at least from the same genus? Is there a good match between your sequence and the best hit sequences? If so, that is probably one of the species in your supplement. Based on the BLAST results, what species do you think your first consensus sequence came from?

7. Did you BLAST more than one consensus sequence? If so, check the other BLAST result tabs to see results from the other consensus sequences. For each sequence, record the species or genus that sequence is likely to have come from, based on your BLAST results. Record both the scientific and common names for these species.

8. Is there a good match between the BLAST search and RDP/Krona analysis? Describe

9. What are some reasons that BLAST and RDP/Krona might have given different results? (see the last page of the instructions)

10. Overall, does there seem to be a good match between the ingredients list on the bottle and the species identified from sequencing? Did you at least get sequences from some of the top ingredients on the bottle? Ingredients for supplements we provided here

11. Did you obtain sequences from taxa species that were not on the bottle's label? What are some possible explanations?

12. Did you obtain sequences from all the species that were on the bottle's label? What are possible explanations?

If you are unsure about any of your answers, please ask an instructor. Show your TA/CA that you have completed the handout before leaving lab and keep it to use in writing your paper