HUMAN PHYSIOLOGY LABORATORY EXERCISE
Metabolism:
Cytochrome C in Humans Compared to Other Species Using Bioinformatics

INTRODUCTION: We live in a human-centered world, but there is a whole range of life beyond what we can see or even imagine. All life forms undergo metabolic processes to obtain energy. The way organisms metabolize substances, however, vary depending upon their ancestry and the environmental conditions they live in. For example, many eukaryotes including humans generate energy from sugars mainly through aerobic cellular respiration and live on land. Bacteria, on the other hand, can metabolize through various mechanisms such as nitrogen fixation, photosynthesis, aerobic and anaerobic respiration and may live in water. They can use different sources for energy such as single carbons, nitrogen and sulfur.

A molecule found common to bacteria and eukaryotes is the protein cytochrome C. For all organisms, it aids in metabolism by helping transfer electrons from one electron carrier to another. In humans, cytochrome C is one of the proteins in the electron transport chain in the inner mitochondrial membrane that transports the electron to oxygen, the final electron acceptor. Cytochrome C consists of about 100 amino acids. In different species, the amino acid sequences for cytochrome C is similar but not identical.

In this lab, we will use bioinformatics to analyze similarities and differences in amino acid sequences for cytochrome C in two species of bacteria and three species of eukaryotes (2 humans, chimpanzees, and honey bees) to understand their ancestry and cytochrome C function.

Purpose:
1. Understand and use bioinformatics with the programs BLAST and Seaview.
2. Compare and contrast amino acid sequences of cytochrome C in humans, other animals and bacteria.
3. Relate amino acid sequences of cytochrome C from various species to genome evolution & protein function.

Hypotheses:
1. If humans share common ancestry, then organisms more closely related to humans will share more common amino acid sequences in the cytochrome C protein.
2. If humans share similar metabolism function with other organisms, then these organisms will have a more common amino acid sequence in the cytochrome C protein as humans.
3. If an individual has a genetic disease with their cytochrome C protein, then this individual's amino acid sequence for cytochrome C will differ than that of a healthy individual.

MATERIALS AND METHODS:
Materials
Computer with internet connection
BLAST (Basic Local Alignment Search Tool) website: http://blast.ncbi.nlm.nih.gov/Blast.cgi
Seaview program
Amino acid sequences for subjects
Description of subjects

Methods:
Exercise 1: Use BLAST to identify accuracy of given amino acid sequences. Fill in Table 1.
1. Use the BLAST website: http://blast.ncbi.nlm.nih.gov/Blast.cgi and instructional video: http://youtu.be/yv7Tbo3XyOw

2. Under "Basic BLAST", choose the "protein blast" program to run because we are using amino acid sequences.
3. Enter your amino acid sequence for search
   a. In the "enter query sequence" box on the upper left corner, copy and paste the amino acid sequence for the organism you want to test. For example, for the *Unknown*, enter: MGDIEKGKKIFVQKCAQCHTVEKGGKHKTGPNLHGLFGRKTGQAVGFSYTDANKNK GITWGEETLMEYLENPKKYIPGTKMIFAGIKKKSERADLIAYLKKATNE
   b. Scroll down to the bottom of the page and click "BLAST". Wait for your results to show. Note: Results show the species and protein with the best match at the top of list.
4. Fill in information in Table 1 using the first match shown under the "description" tab.
   a. Fill in the species identified located in [brackets].
   b. The protein is listed before the species.
   c. For "fraction of AA sequences matched" and "identity value"
1. click on the first match.
      2. look at the first set of data on the top of page. Find the 4th column labeled "identities". Here is the fraction of AA matched (e.g. 105/105) and percentage for the identity value (e.g. 100%).
5. Repeat for the remaining organisms. Check for species match, protein, fraction of AA match and identity match.

Exercise 2: Use BLAST to compare amino acid sequences for cytochrome C between species. Fill in Table 2.
1. Use the BLAST website: http://blast.ncbi.nlm.nih.gov/Blast.cgi and instructional video: http://youtu.be/xY_FOkSfhu0
2. In your protein BLAST search page under "enter query search", click the box titled "Align two or more sequences." A second search box will appear underneath.
3. Enter the sequence of *Homo sapiens 1* in the second box titled "Subject". Enter all the sequences of the organisms that you want to compare it to in the first "Query" box. Be sure to hit return after each sequence when entering multiple organisms.
4. Click "BLAST"
5. Scroll down and find the "identities" match percentage between *Homo sapiens 1* and each other organism. Enter in the appropriate box in Table 2. Note: The two sequences are aligned next to each other. The differences in amino acids are identified with a blank or a "+" sign.

Exercise 3: Use Seaview to visually align cytochrome C amino acid sequences for all 6 organisms. Answer questions based on alignment.
1. Use instructional video: http://youtu.be/-XfF8sJaZ-c
2. Open Seaview.
3. Enter the sequences to be aligned in Seaview
   a. Go to Appendix 1. Amino Acid Sequences for Species
   b. Highlight ALL the text below the heading including the species name.
   c. Copy (CTRL+C) and paste (CTRL+V) into the window called "alignment drag zone". Note: you must paste by using CTRL + V on the keyboard, not with the mouse.
   d. Click "confirm"
4. Align the sequences
   a. Select the drop down menu called "align"
   b. Select "align all"
   c. Click "OK"
5. Answer corresponding questions in the results section.

RESULTS:

Exercise 1: Use BLAST to identify accuracy of given amino acid sequences.

Table 1. Protein Identification for Different Species

	Species	Protein	Fraction of AA sequences matched	Identity values (%)
Eukarya	1.			
	Homo sapiens 1: healthy			
	Homo sapiens 2: disease			
	Pan troglodytes			
	Apis mellifera			
Bacteria	*Pseudomonas*			
	Caldithrix abyssi			

1. A. What is the range of identity values in table 1 above?

   B. Do all your organisms above have the protein cytochrome C? Yes or No

   C. Based on your answers to A and B above, do your given sequences correctly identify the protein and organism? Explain using the data in table 1.

_____

Exercise 2: Use BLAST to compare amino acid sequences for cytochrome C between species.
Fill in the table below with the "identity value" for each pair of organisms.

Table 2. Amino Acid Identity Table for Cytochrome C

Cytochrome C	*Homo Sapiens 2*	*Pan troglodytes*	*Apis mellifera*	*Pseudomonas*	*Caldithrix abyssi*
Homo Sapiens 1					

2. A. <u>List</u> in order the species that have cytochrome C that is the most similar to least similar to humans.

   B.   What does this result imply about our evolutionary relationships and common ancestry with bacteria and other eukaryotes?

3. Based on the identity value of *Homo sapiens 1* vs *2*, do you think *Homo sapiens 2* has a genetic defect in its cytochrome C protein? Please explain.

4. What amino acid(s) are different between *Homo sapiens1* and *Pan troglodytes*? Do you think this difference has any effect on cytochrome C function? Briefly justify your answer.

Exercise 3: Use Seaview to visually align cytochrome C amino acid sequences for all 6 organisms.

Comparison of Cytochrome C sequence in six species Using SeaView

Use your alignment of cytochrome C sequences in Seaview to answer the following questions.

5. A. Which domain of life looks more diverse, bacteria or eukarya? _____

   B. Who is more related: the two bacteria or humans and bees? _____

   C. Briefly explain your answers to 5A and B by pointing out patterns from the data.

6. A. List the name and position of the amino acids that are identical in position in the cytochrome C sequence across all species investigated. The part of the genome that codes for these amino acids are referred to as "conserved regions" of the gene.

   B. What might be the function of these conserved regions for the cytochrome C protein?

7. Which species has the sequence that is the MOST different from the other species? Suggest a reason (based on protein function) why this species' cytochrome C sequence might differ the most. (*Hint: Look at the gaps (-) in alignment sequence and look at description of subjects.)

DISCUSSION/CONCLUSION:
Does your data support hypothesis #2? Evaluate based on patterns derived from the data.

WEBSITE:
http://www.coexploration.org/C-DEBI/toolkits_biology.html