

Name:

Are you a **graduate** or **undergraduate** student? Please circle one.

Bioinformatics Take Home Test #1

-Due 9/13/13

(This is an open book exam based on the honors system -- you can use notes, lecture notes, online manuals, and textbooks.

Teamwork is not allowed, write down your own answers, do not cut and paste from webpages.

If your answer uses a citation, give the source of the quoted text. Do not cut and paste from the source.

You may always turn your quiz in by emailing it to your Professor and TA. If you are healthy, you may come to class and turn in a hard copy of your quiz instead. If you are sick, you may not turn in a hard copy of your quiz and are required to submit via email.

Point values are given in **red**.)

1. Give a short "definition" of the scientific discipline that is labeled as Bioinformatics. (less than 30 words) **(1pt)**

2. "Define" life, in fifteen words or less. (You may use another 15 words to point out the shortcomings of the definition you gave.) **(1pt)**

3. Are viruses alive? Please include 2-4 points of evidence to back up your answer (simply a yes or no will not suffice). **(2pt)**

4. How many peptides (short proteins) of 12 amino acids in length are possible, given that there are 20 possible amino acids? For your answer only consider the principles of combinatorics and ignore possible incompatibilities between amino acids) **(1pt)**

5. What is homology? **(1pt)**
 - a. The concept on which the Turing machine is based, aiming at the creation of artificial life
 - b. When two proteins share a function, such as nucleotide binding, they also share sequence similarities, because of the limited size of protein space, i.e. there is no other sequence that could carry out that function, so all the sequences with that function are homologous

- c. Similarity due to shared ancestry, i.e. both got the homologous trait from a common ancestor
 - d. Shared sequence similarity based on convergent evolution, i.e. the ancestor did not have it
 - e. A difference found because of diverging evolutionary paths since the last common ancestor
6. Can a protein be 23% homologous to another protein? (1pt)
- a. Only in cases of domain shuffling
 - b. Only if they shared a common ancestor in the Turing machine *in silico*
 - c. Only if they share 23% sequence identity
 - d. Yes (please explain)-

 - e. No (please explain)-

If you choose yes or no, make sure to include one or two sentences to explain your reasoning.

7. Do all proteins that evolved from the same ancestral protein have significant similarity in their primary sequence? Please explain your answer in three lines or less. (1pt)
8. Did all proteins that show significant similarity in their primary sequence evolve from the same ancestral protein (when no simple repeats are present)? (1pt)
9. If two proteins share a similar structure and can be aligned in the Swiss Protein DataBank Viewer (that program we've been using in class of Fridays) so that the structures overlap perfectly, are these two proteins usually considered homologous? (1pt)
- a. Yes
 - b. No
 - c. Can't tell- not enough information given

Give a short justification of your reasoning:

10. If two proteins are homologous, will their structures usually be alignable with the Swiss Protein DataBank Viewer? (1pt)
- a. Yes
 - b. No
 - c. Can't tell- not enough information given

Give a short justification of your reasoning:

11. What can you do using the Swiss Protein data bank file viewer, aka Deep View? List at least 2 possibilities (1pt)

12. Which structural elements make up the secondary structure of proteins? (1pt)

- a. Hydrogen bonds, Van der Waals interactions, and disulfide bridges
- b. Multiple protein chains interacting to form one macromolecule
- c. Alpha helices, beta sheets, and loops
- d. Nucleotide binding motifs, protein channels, hydrophobic domains, and other like motifs
- e. Alpha corkscrews, beta barrels, and delta turns

13. Which structural elements are often represented as only slightly curved yellow arrows? (1pt)

- a. Beta barrels
- b. Loops
- c. Beta sheets
- d. Alpha corkscrews
- e. Hydrogen bonds
- f. Covalent bonds (peptide bonds and disulfide bonds)
- g. Delta turns
- h. Alpha helices
- i. Nucleotide binding motifs

14. How many types of subunits (note: this asks for the different types of subunits, not the number of subunits) form the hexamer (the head) of nucleotide binding subunits in the F1 ATPase? (0.5pt)

In the catalytic cycle, do all of the catalytic subunits work in the same phase of the catalytic cycle at the same time? (0.5pt)

Are the different ATP binding subunit types that form the hexamer of the F1-ATP synthase homologous to each other? (1pt)

What may the ancestral ATPase have looked like (hint, consider what we know about the ancient gene duplication)? (1pt)

15. Under which conditions does evolution by natural selection occur? List at least 3 (2pt)

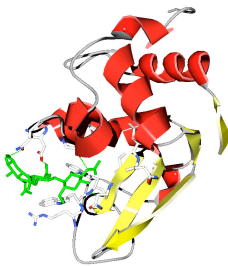
16. Is it possible to create a computer program to mimic evolution by means of artificial selection? If so, what can this type of exercise tell us? If not, why not? (2pt)

17. What is the Gaia hypothesis? (1pt)

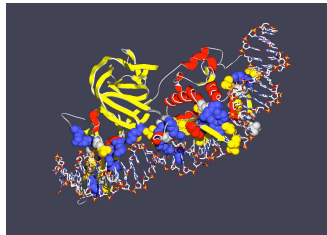
- a. A cute little demonstration showing that the plants on the Earth could control the planet's temperature simply by selection for black or white daisies (also known as the daisy world)
- b. The hypothesis that all life on Earth descended from a common ancestor
- c. The hypothesis that explains large glaciers and ice ages by positing a runaway cold-house, where the Earth gets colder and colder because the increased glaciers reflect more light from the surface
- d. The hypothesis that Mars would have cooled faster than the Earth and therefore been a much more habitable place for life much earlier, so that life could have arisen on Mars first and traveled to Earth on meteorites
- e. The hypothesis that the entire Earth is the unit of life and the entire biosphere is alive, because almost no single species can exist in complete isolation from other forms of life. It also includes the notion that the ecosystems on the Earth are protected by negative feedback loops that help maintain homeostasis.
- f. It is the hypothesis that the zircon crystals found in 3.8 billion year old rocks were produced by ancient life over 4 billion years ago. The zircon crystals are formed by running water and the bias in carbon isotope ratios indicates the presents of life.

18. We have worked with 3 enzymes in the Friday computer labs, which are shown below. Please identify them: (1pt)

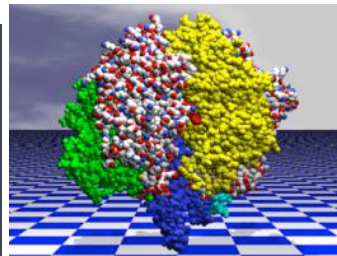
a.



b.



c.



For **Graduate Students:** (answer 2 out of 3) (2pts each)

1. Who was Alan Turing, and why should you care? At least three unique points of relevance are required.
2. In Thomas Mann's except from Dr. Faustus, he describes mushrooms, tiny trees, and algae meadows. Are these things alive and how do they relate to the question of what is life and how it arose?
3. In the article, Defining 'life', by Cleland and Chyba, what do they conclude is the underlying cause of the difficulty with defining life?

Extra credit:

1. The plot at http://web.uconn.edu/gogarten/mcb221_2007/chapter1/01-01.jpg contains a curve giving the increase of DNA sequence available in the databanks. According to the

depicted data, how long will it take on average for the number of nucleotide sequences deposited in genbank to double? (Rough estimation is ok)

2. According to this diagram, how many nucleotide sequences would you expect in Genbank in the Spring of 2013 and Spring of 2014?
3. About how many sequences and how many nucleotides were in the traditional part of Genbank in Spring of 2013?