

Your Name:

Bioinformatics Take Home Test #1 –Due 9/19/16

Please submit a hardcopy of your answers (with or without the questions), if possible. If you have an emergency and cannot submit a quiz in person, email it in by the start of class on the due date. If you do so, only email your answers.

Late quizzes will not be accepted after the answers have been released.

All questions worth 1pt if not otherwise indicated

1. Give a one sentence description of the field known as Bioinformatics.
2. Bioinformatics does NOT include which of the following activities?
 - A. Analyzing genome data using computers
 - B. Figuring out a protein structure from X-ray crystallography
 - C. Detecting homologs using primary sequence similarity
 - D. Detecting homologs using secondary sequence similarity
3. ATP, GTP, NAD, NADP, FMN, and FAD are example of what type of cofactor?
 - A. Nucleotide or modified nucleotide
 - B. Lipid or modified Lipid
 - C. Amino acid or modified amino acid
 - D. oligosacharides
4. The cofactors listed in question 3 provide support for which hypothesis regarding the early evolution of life?
 - A. DNA World
 - B. Compartmentalized Cells First
 - C. RNA World
 - D. Protein World
 - E. Metabolism First
5. Proteins that evolved from the same ancestral protein
 - A. ALWAYS have significant similarity in their primary sequence.
 - B. Cannot have different functions
 - C. Can diverge so that they have only limited homology.
 - D. Do not necessarily retain detectable similarity in primary sequence.
 - E. Will ALWAYS have detectable levels of primary sequence similarity
6. What might be the reason for proteins that show significant similarity in their primary sequence (when no simple repeats or low complexity regions are present) to be homologous.
 - A. Sequence space is so big that stumbling onto a significantly similar sequence by chance is very unlikely

- B. This is a trick question, because NOT all currently known proteins that share significant similarity are homologous
- C. There is a direct link between sequence similarity and percent homology
- D. Homologous sequences share a function and there is only a small region of protein space that can accomplish any one function
- E. All of the above

7. Which of the following is NOT considered strong evidence for homology?

- A. The arrangement of secondary structure elements in space is the same or very similar.
- B. Significant primary sequence similarity
- C. Significant primary sequence identity
- D. Identical function
- E. None of the above

8. Which of the following is the BEST working definition of life?

- A. A contained metabolic system capable of evolution with heredity.
- B. Cells. Life could only exist in a form compartmentalized by a lipid bilayer.
- C. Self-sustained metabolic system that does not require input from any other living system.
- D. An interacting web with intricate feedback loops to ensure homeostasis.
- E. Anything that can reproduce itself perfectly with no errors.
- F. All of the above are equally good answers.

9. What is homology?

- A. The creation of artificial life by any means.
- B. A difference found because of diverging evolutionary paths since the last common ancestor.
- C. Similarity due to shared ancestry, i.e. both got it from a common ancestor.
- D. When two proteins share a function, such as nucleotide binding, they also share sequence similarities, because of the limited size of protein space.
- E. Shared sequence similarity based on convergent evolution, i.e. the ancestor did not have it.

10. Can a protein be 84% homologous to another protein?

- A. Yes.
- B. Mostly no, with the exception of cases of domain shuffling.
- C. Yes, if they share 84% sequence identity.
- D. No, without exception.

11. The Swiss Protein data bank file viewer, aka Deep View canNOT be used to do which of the following things?

- A. Visualize where a nucleotide cofactor binds.
- B. Compare the structures of two sequences.
- C. Color a structure by how closely it matches another.
- D. Detect homology with 100% certainty.
- E. Visualize the conformational change that the ATPsynthase undergoes during its catalytic cycle.

12. Which elements make up the secondary structure of proteins?
- A. Hydrophobic domains, nucleotide binding motifs, and protein channels.
 - B. Beta barrels, alpha corkscrews, and delta turns.
 - C. Beta sheets, alpha helices, and loops.
 - D. Van der Waals interactions, hydrogen bonds, and disulfide bridges.
 - E. Multiple protein chains interacting to form one macromolecule.
13. How many **types** of subunits (note: this asks for the **different types** of subunits, encoded by separate genes, not the number of subunits) form the hexamer (the head) of nucleotide binding subunits in the F1 ATPase?
14. True/False In the catalytic cycle of the ATPase, all of the catalytic subunits work in different phases of the catalytic cycle at any one point in time.
15. True/False The different ATP binding subunit types that form the **hexamer** of the F1-ATP synthase are homologous to each other.
16. True/False The ancestral ATPase could have looked like a homohexamer with six identical subunits.
17. True/False It is NOT possible to create a computer program to mimic evolution by means of artificial selection, because computers programs are NOT capable of finding new solutions that a human has never thought of before.
18. What is the Gaia hypothesis?
- A. Earth's plants control the planet's temperature by selection for flower color
 - B. All life on Earth descended from ONE common ancestor
 - C. The unit of life is the entire Earth and the entire biosphere is alive, because no single species can exist in complete isolation. The ecosystems on the Earth are protected by negative feedback loops that help maintain homeostasis.
 - D. Mars cooled faster than the Earth and therefore was a more habitable place for life earlier. Life arose first on Mars and traveled to Earth on meteorites
 - E. Large glaciers and ice ages result from a runaway cold-house, where the Earth gets colder because glaciers reflect more light from the surface
 - F. Zircon crystals 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.
19. Evolution by natural selection requires which three things to occur?
- A. Variation among offspring, a niche, and competition for resources.
 - B. Heredity, excess offspring, and a niche.
 - C. Excess offspring, competition for resources, and variation among offspring.
 - D. A niche, a human to naturally select the best offspring, and oxygen.
 - E. None of the above.

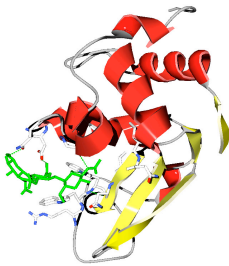
20. Which of the following is NOT part of the explanation for how complex functional molecules were assembled, despite the vastness of protein space?

- A. Gaia directs protein evolution, through negative feedback loops, to the correct region of protein space.
- B. There are multiple unrelated solutions for the same functionality, exemplified by the fact that there are non-homologous enzymes inhabiting completely different regions of protein space with the same function.
- C. An exact function does not need to be hit upon, because natural selection can take a protein with limited function and make it better.
- D. Similar structures have similar function, so there are entire regions of protein space occupied by homologs that all function equally well, or nearly so.
- E. Protein space is made slightly smaller by removing all of the possibilities that cannot be synthesized or they will clog up the ribosome.

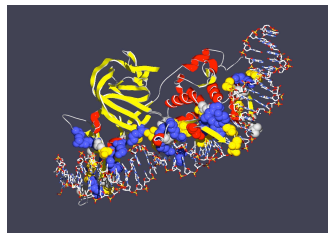
21. How many peptides (short proteins) of 70 amino acids in length are possible, given that there are 20 possible amino acids? For your answer only consider the principles of combinatorics and ignore possible incompatibilities between amino acids. How does this number compare to the estimated number of protons in the universe?

22. We have worked with 3 enzymes in the Friday computer labs, which are shown below. Please identify them:

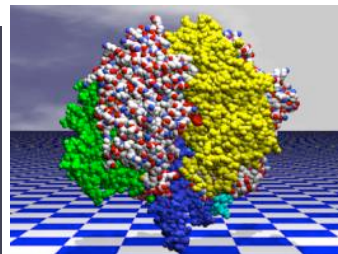
a.



b.



c.



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

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)

According to this diagram, how many nucleotide sequences would you expect in Genbank in the Spring of 2015 and Fall of 2016?

About how many sequences and how many nucleotides were in the traditional Genbank in the Fall of 2016? How many nucleotides were part of Whole Genome Sequences accessible through the NCBI

2. Who was Alan Turing, and why should you care? At least three unique points of relevance are required.