Name:

Bioinformatics Take Home Exam #2   Due 10/04/15

Make sure each answer is only on one page, by using page breaks. Splitting an answer onto two pages leads to grading errors.
Do not write or type in font smaller than 12 point or write in cursive. Doing so will **lose** you **2 pts.**
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, first remove the instructions and extras (blank lines, alternative answers for multiple choice questions) from your document, so that only your answers, a minimal amount white space, and optionally the questions, are left. Indicate answer selection by highlighting in green (Yellow does not show)
Note on Late Quizzes: Late quizzes are an inconvenience and cannot be accepted at all after the answers have been released. If your quiz is submitted within the first 12 hours after the deadline, you will receive 5% off. Each additional 12 hours is an additional 5% off, up until the graded quizzes are returned or the answers released.
All questions worth 1pt

1. Assuming equal frequency of the different building blocks, two random protein sequences are on average _____ and nucleotide sequences are on average _____?
A. 20% identical and 5% identical.
B. 5% identical and 20% identical.
C. 20% identical and 40% identical.
D. 95% identical and 75% identical.
E. None of the above.

2. _____ sequences reach saturation before _____ sequences reach saturation, so _____ sequences can be used to look further back in time.
A. Nucleotide, protein, nucleotide
B. Protein, nucleotide, nucleotide
C. Nucleotide, protein, protein
D. Protein, nucleotide, protein
E. None of the above

3. The universe and the earth are approximately how old, respectively?
A. 20 billion years old and 500 million years old
B. 14 billion years old and 4.5 billion years old
C. 16 billion years old and 6 billion years old
D. 2 billion years old and 450 million years old
E. 1 million years old and 4500 thousand years old

4. **True/False** Life could have inhabited the Earth prior to 3.5 billion years ago and LUCA sometime before 3.5 billion years ago.

5. Who drew the first phylogenetic trees?
A. Lamarck
B. Darwin
C. Mayr
D. Henning
E. Baptiste

6. **True/False** BLINK, from NCBI, stands for Boolean-link and links to pre-computed Boolean searches.

7. Which biological processes are acquired traits capable of being passed on to offspring?
A. Long necks on giraffe, stripes on zebra, and spots on a cheetah
B. Epigenetic modifications and symbiotic gut Bacteria
C. Large antlers on deer, manes on male lions, and huge plum of feathers on peacocks.
D. Termite or ant fishing with sticks
E. All of the above.

8. **True/False**- Social Darwinism is correct in that charity has stopped evolution in our species.

9.**True/False**- Entrez is so effective because it has a humongous number of servers ready and waiting to compute your every request.

10.**True/False**- Entrez covers many databases simultaneously, but does so only slowly.

11. **True/False**- When inteins first begin to decay they lose the DNA-binding domain first, while the protein-binding domain must stay functional or it will destroy the function of the host proteins.

12.**True/False**- The finding that the ribosomal protein alone is responsible for the catalysis of translation is an argument against the RNA world hypothesis.

13. **True/False**- Having the same function IS necessary to have homology and without a shared function, homology can NOT exist.

14. **True/False** Among Site Rate Variation (ASRV) means that some sites will undergo multiple substitutions while other sites do not undergo any substitutions. Due to ASRV, protein and nucleotide sequences take longer to become saturated with substitutions than without ASRV.

15. What Boolean operations can be used in NCBI/Entrez searches?

16. Match the terms on the left with the definitions on the right- 6 pts

mRNA	1. The process of making a protein from an RNA template
tRNA	2. A molecular parasite that splices itself out at the RNA level
rRNA	3. A molecular parasite that splices itself out at the DNA level
transcription	4. The process of making RNA from DNA
replication	5. RNA that binds an A.A. & matches it with mRNA triplet
translation	6. A molecular parasite that splices itself out at the protein level
intein	7. An RNA copy of a gene, used in the process of making proteins
intron	8. Part of host gene left after RNA parasite is spliced out
exon	9. The host protein, which is spliced back together
extein	10. Host DNA left after molecular DNA parasite splices itself out
Non-existent thing put in to trip you up	11. RNA that makes up the ribosome and catalyzes protein synthesis
Another non-existent thing	12. Process of creating a new DNA molecule, from DNA strand

17. Inteins are composed of which of the following domains? Choose 2.
A. Self-splicing domain
B. Walker motif
C. Nucleotide binding domain (GRASP)
D. Hydrolase domain
E. Helix-turn-Helix DNA binding domain
F. Homing endonuclease domain

18. What are the functions of the two domains that are present in full inteins? Choose 2.
A. Splices itself out of the host protein, putting the host protein back together
B. Cleave uninfected DNA, so that the molecular parasite will spread
C. Creates a channel through the lipid bilayer, to allow molecules to pass through
D. Binds ATP
E. Binds a nucleotide
F. Cleaves carbohydrates

19. Which of the following are databases available through the NCBI aka Entrez? Circle all that apply-

A. BioProject (formerly Genome Project)
B. Bookshelf
C. Database of Genome Survey Sequences (dbGSS)
D. GenBank
E. Genome Reference Consortium (GRC)
F. NCBI Help Manual
G. Nucleotide Database
H. Protein Database
I. PubMed Central (PMC)
J. Taxonomy
K. All of the above and many many more.

20. Sequences that do not show significant similarity-
A. are not homologous
B. might never-the-less be homologous
C. are homologous

21. If the following searches were conducted in PubMed for articles, what would the searches return? Please draw Venn diagrams to illustrate your answers (i.e. depict each of the individual searches as a circle). 2pts.
A. Gogarten J NOT Gogarten JP

B. Gogarten JP AND Doolittle WF

C. Gogarten J OR ATPsynthase

D. (Gogarten JP OR Swithers K) AND Inteins

22. What does the abbreviation NCBI stand for and why is this site important in the field of bioinformatics? Limit your answer to 30 words or less.

23. What is the definition of homology?
A. The common ancestor of all cellular life
B. Similarity due to shared ancestry, i.e. both got it from a common ancestor
C. The percent of time the A.A. residues of two sequences are similar
D. Similarity due to convergent evolution
E. Has the same function
F. How similar two sequences are and degrades over time with sequence evolution

24. **True/False**- Gut Bacteria can be passed on to children,but sharing these symbionts between communities is less frequent.

25. What is the Black queen hypothesis?
A. Life is like an arms race, where all lifeforms have to run faster and faster just to stay in place.
B. Gaia favors cooperation, which is the driving force behind multi-species communities being interdependent.
C. Leaky goods can be taken up by all members in a community, so selection for streamlined genomes will result in all members of a community producing only a subset of the required leaky goods.
D. DNA based organisms took over from the RNA world, after DNA was created by a virus in an act of genome warfare.
E. None of the above.

26. **True/False**- The Modern Synthesis does not give any weight to the effects of mutations themselves.

27. How might mutual aid be selected for?
A. Trick question: it cannot be selected for, because even if a stingy species is going extinct, it cannot decide to stop being stingy.
B. When cooperation between species results in more offspring for both, the entire community thrives and spreads.
C. When cooperation is between close relatives, like siblings, helping each other survive leads to some of your own genes being passed on.
D. When one bird helps defends another bird's nest, that second bird will remember and return the favor when the first bird is under attack.
E. B-D.

Extra credit question--
1.   Two random nucleotide sequences with equal frequencies of A, G, T, and Cs without alignment have an average percent identity of 25%. How would the average percent identity

change, if the frequencies for the nucleotides are not equal. Use composition with 24%G 24%C and 26%A, 26%T as an example.

2.   Pushing the boundaries of life 2 pt
a.   Which of the following features of life as we know it is inescapable and will surely be found in all alien life discovered?
A. DNA
B. RNA
C. The central dogma (DNA →RNA →Proteins)
D. Ester linked lipid bilayers
E. Parasites
F. All of the above

b.   If we found a virus that brought its own ribosome with it into the host cell and used only its own protein manufacturing equipment, would that virus be alive? Why (in 30 words or less)?