BIOLOGY 152

Bioinformatics Assessment

This assessment will be given at the beginning and end of the course as part of a National Science Foundation educational research project. It includes questions about concepts covered in BIO 152. If you are taking the assessment at the beginning of the course, you should not be alarmed if you do not know many of the answers and do not understand some of the terms used.

The results of this survey are anonymous. We are interested in aggregate data from the whole class. While participation is required, your performance will not affect your grade in BIO 152.

Before you begin, please use a #2 pencil to fill in the following information on the Scantron sheet:

1. **complete student ID number**
2. **fill in bubbles for student ID number**
3. **in "Additional Data" box write *Bio 152* under "Course" and fill in the month/year under "Date"**

Do NOT write on this packet.
Do NOT write your name or fill in name bubbles on the Scantron sheet.

Please respond to the following statements by indicating the extent to which you agree with the statement. Please use the Scantron sheet to record your response to each numbered statement using the following scheme:

A. strongly disagree
B. disagree somewhat
C. unsure whether you agree
D. agree somewhat
E. strongly agree

1. I am confident in my ability to use the science reference tools PubMed, NCBI Bookshelf, and OMIM.

2. I am confident in my ability to retrieve protein and nucleic acid sequence records using the Entrez search utility.

3. I am confident in my ability to read and evaluate the parts of a protein or nucleic acid sequence record.

4. I am confident in my ability to use BLAST to make comparisons between a protein or nucleic acid sequence and a database of sequences.

5. I know when to use BLASTN, BLASTP or BLASTX.

6. I am confident in my ability to explore various levels of protein structure using NCBI Structure/Cn3D.

7. I understand the significance of E scores when comparing a protein or nucleic acid sequence to a database of sequences.

8. I understand the difference between global and local sequence alignments.

9. I understand how to evaluate data from two-color microarray experiments.

10. I am confident in my ability to work with mathematical ratios.

11. I am confident in my ability to calculate probabilities.

12. I am confident in my ability to calculate logarithms.

13. I am confident in my ability to use Excel to work with quantitative data.

14. Studying bioinformatics helps strengthen my mathematical skills.

For the remaining questions, please select the lettered choice that provides the BEST answer to each question and record your choice on the Scantron sheet.

15. Which of the following on-line search utilities would you use to identify primary research articles about the *Drosophila* gene *Antennapedia*?

    A. BLAST
    B. Entrez
    C. NCBI Bookshelf
    D. OMIM
    E. PubMed

16. Which of the following would you use to identify protein sequence records from all species that have the word "Alzheimer's Disease" in their annotation?

    A. BLAST
    B. Entrez
    C. NCBI Bookshelf
    D. OMIM
    E. PubMed

17. In a nucleotide sequence record, the annotation "CDS" identifies:

    A. computer-derived sequences
    B. congenital defective sequenced alleles
    C. genomic DNA sequences
    D. sequence bounded by start and stop codons
    E. unverified sequence data

18. You have used Edman degradation to determine the amino acid sequence of a short peptide purified from an African Swallow. Which utility would you use to identify the human protein that is most similar to your purified peptide?

    A. BLASTN
    B. BLASTP
    C. BLASTX
    D. Entrez
    E. OMIM

19. In BLAST results, an E score of 0 indicates:

    A. the alignment of the two sequences has no gaps
    B. the query sequence returned no hits
    C. the wrong type of query sequence was entered
    D. there is no chance a random sequence would give the same result
    E. there is no chance the two sequences match

20. BLAST uses a local method of alignment instead of a global method of alignment because it:

    A. does not work in all countries in the world
    B. handles gaps much better
    C. is faster
    D. produces longer alignments
    E. works for both amino acid and nucleotide sequences

21. Which of the following could NOT be observed in a typical NCBI structure page for a metabolic protein of interest?

    A. any associated metal ions
    B. nucleotide sequence
    C. primary structure
    D. secondary structure
    E. the location of a specific amino acid

22. A two-color microarray experiment is performed and the scanned image shows varying levels of red intensity but no green signal. This is most likely due to:

    A. Failure to log transform the data
    B. Normal experimental procedure (i.e. this is what you would expect)
    C. Poor labeling of both starting DNA samples
    D. Poor labeling of both starting RNA samples
    E. Poor labeling of one of the samples

23. If a gene is shown to be induced four-fold in a microarray experiment, what would be the log2-transformed expression ratio?

    A. -4
    B. -2
    C. 0
    D. 2
    E. 4

24. Based on the signal intensity data from microarray scans provide below, which of the following genes display the same magnitude of change in expression?

Gene name	Red intensity	Green Intensity
A	50,000	10,000
B	10,000	20,000
C	50,000	20,000
D	5,000	25,000
E	20,000	5,000
F	5,000	50,000

A. A and D
B. A and E
C. B and C
D. D and E
E. D and F