B. Sc. (Bioinformatics) (Sem. III) (CBCS) Examination

November / December – 2017

BI-301 : Bioinformatics Databases & Sequence Analysis

(New Course)

Time : **2½** Hours]   [Total Marks : **70**

Instructions : (1)   All questions are compulsory.

(2)   The right side figures indicate total marks of the question.

1   Attempt the following :   **14**

(a)   Answer the following short questions : (all compulsory)   **4**

(1)   Name of genetic disorder database.

(2)   Name of primary protein sequence database.

(3)   What is biomarker ?

(4)   Who provides automatic annotation databases for human, mouse, other vertebrate and eukaryote genomes?

(b)   Answer any **one** of the following questions.   **2**

(1)   What are the two challenges of data integration?

(2)   Supervised mining of "bigdata" without programming.

(c)   Answer any **one** of the following questions :   **3**

(1)   Supervised and unsupervised machine learning in Biology.

(2)   Data integration.

(d)   Answer any **one** of the following questions :   **5**

(1)   What are the cloud based resources in Bioinformatics?

(2)   What are the barriers for implementing data management?

2 Attempt the following : **14**

    (a) Answer the following short questions (All compulsory) **4**

        (1) Which sequences are included in JuncDB database?

        (2) RNAcentral is a database for coding RNA sequence. True or False. If false then correct it.

        (3) PROSITE offers tools for protein _____ and _____ .

        (4) The post translation in eukaryotes is facilitated by_____ and _____ proteins.

    (b) Answer any **one** of the following questions : **2**

        (1) Difference between coding DNA and non-coding DNA.

        (2) Explain KEGG pathway database.

    (c) Answer any **one** of the following questions. **3**

        (1) Explain Properties of proteins.

        (2) Write a note on transcription factors.

    (d) Answer any **one** of the following questions : **5**

        (1) Write a note on INSDC.

        (2) Protein structure database : PDB

3 Attempt the following : **14**

    (a) Answer the following short questions: (all compulsory) **4**

        (1) _____ are constructed for selected organisms by first clustering, then assembling EST and annotated gene sequences from GenBank.

        (2) What is /are the principles covered by BioGPS?

        (3) What is open reading frame (ORF)?

        (4) _____ is a field of biological research in which the genomic features of different organisms are compared.

(b) Answer any **one** of the following questions. **2**

(1) Into how many levels is the GOLD database classified? List out the levels.

(2) Explain in brief AmoebaDB

(c) Answer any **one** of the following questions. **3**

(1) What is the major purpose of GenoBase.

(2) Types of Signaling pathways. Explain.

(d) Answer any **one** of the following questions: **5**

(1) Types of protein -protein interactions.

(2) HMDB database.

4 Attempt the following : **14**

(a) Answer the following short questions: (All compulsory) **4**

(1) What is the other name of Global analysis of gene expression?

(2) Which techniques used to analyze gene expression?

(3) Which database classifies proteins into families and predicts the presence of domains and sites?

(4) _____ is the web resource designed for comparative and functional genomics in plants.

(b) Answer any **one** of the following questions : **2**

(1) Define gene expression data.

(2) Limitations of Microarray.

(c) Answer any **one** of the following questions : **3**

(1) What are the applications of PRINTS?

(2) Drug central

(d) Answer any **one** of the following questions : **5**

(1) Which are the drug-target databases?

(2) Plant DB p3db

5   Attempt the following :   **14**

    (a)   Answer the following short questions : (all compulsory)   **4**

        (1)   What is Multiple Sequence Alignment?

        (2)   What are gap penalties?

        (3)   What are the tools available for database searching for similar sequences?

        (4)   Multiple sequence alignment is used to detect regions of variability conservation in a family of proteins (true or false)

    (b)   Answer any **one** of the following questions :   **2**

        (1)   Domain

        (2)   Dot matrix

    (c)   Answer any **one** of the following questions :   **3**

        (1)   Steps for multiple sequence alignment.

        (2)   Progressive methods of multiple sequence alignment.

    (d)   Answer any **one** of the following questions :   **5**

        (1)   Align the following sequence

            Sequence 1   GGGGUUCGCUCA

            Sequence 2   AGAGGUUGCUCU

            Sequence 3   CGAGGCUGCUC

            Sequence 4   UGAGGCUCGCUCA

        (2)   Multiple sequence alignment as an extension of sequence pair alignment by dynamic programming.

———————