

SEM	SET	PAPER CODE	TITLE OF THE PAPER
II	2014	14PBT2109	BASIC BIOINFORMATICS

**SECTION - A****Answer all the questions:****30 × 1 = 30****Choose the correct answer:**

- The term “Genomics” was coined by \_\_\_\_\_.  
a) Thomas Roderick                      b) Andrew Fire  
c) Craig                                      d) Margaret
- The Human Genome Initiative is started in \_\_\_\_\_.  
a) 1988                      b) 1998                      c) 2008                      d) 2013
- The PCR reaction is described by \_\_\_\_\_.  
a) Kary Mullis                      b) Thomas Roderick  
c) Craig                                      d) Margaret
- \_\_\_\_\_ is the development and execution of practices and procedures that properly manage the full data lifecycle needs of an enterprise.  
a) Data Resource Management                      b) data pre-processing  
c) Knowledge discovery                      d) data ware housing
- RCSB PDB is a \_\_\_\_\_ database  
a) Protein                      b) RNA                      c) DNA                      d) carbohydrates
- In Data mining, the identification of unusual data records or data errors that require further investigation is called \_\_\_\_\_.  
a) Anamoly detection                      b) classification  
c) clustering                      d) regression

7. In Flat file format, the LOCUS field provides \_\_\_\_\_.  
a) sequence length                      b) molecule type  
c) division                                  d) all the above
8. \_\_\_\_\_ is a curated protein sequence database which provides a high level of annotations, a minimal redundancy and high level of integration with other databases.  
a) SWISS – PROT                      b) PIR  
c) PROSITE                              d) SCOP
9. Pubmed is \_\_\_\_\_ databases.  
a) literature                              b) bibliography  
c) both a & b                              d) EST
10. \_\_\_\_\_ is a collection of single nucleotide variations, microsatellites, and small-scale insertions and deletions.  
a) dbSNP                                  b) dbEST  
c) dbGSS                                  d) Pubmed
11. The Pfam database contains information about protein \_\_\_\_\_.  
a) domains                                b) sequence  
c) structure                                d) all the above
12. \_\_\_\_\_ is a database of human genes and genetic disorders.  
a) OMIM                                  b) OMIB  
c) OMIA                                  d) OMIN
13. \_\_\_\_\_ is a software tool developed by the NCBI for submitting and updating entries to the GenBank sequence database.  
a) Sequin                                  b) BLAST  
c) FASTA                                  d) WEBIN
14. Prokaryotic genes can be predicted in its genome by identifying \_\_\_\_\_.  
a) Pribnow box                      b) RNA encoding genes  
c) SiRNA encoding genes              d) miRNA encoding genes

15. Vector screening system is done to remove the \_\_\_\_\_ from the submitting DNA sequences.
- a) cloned plasmid sequences    b) protein sequences  
c) RNA sequences                      d) none of the above
16. \_\_\_\_\_ trees illustrate the relatedness of organisms without representing common ancestor of the descendants.
- a) rooted              b) unrooted      c) both a & b              d) nodes
17. Which software program can read electron density maps in the structure of proteins?
- a) Swiss – Pdb Viewer    b) Rasmol      c) Chime    d) QIIME
18. \_\_\_\_\_ search protein database using a translated nucleotide query.
- a) tblastn                                      b) blastx  
c) blastn                                      d) blastp
19. Indel is a term which explains the \_\_\_\_\_ of bases in the DNA of an organism.
- a) insertion and deletion              b) inversion  
c) substitution                              d) addition
20. The ENV division of GenBank accommodates non-WGS sequences obtained via environmental sampling methods in which the source organism is \_\_\_\_\_.
- a) unknown                                      b) single known culture  
c) mixed culture                              d) none of the above
21. Needleman – Wunsch algorithm utilizes \_\_\_\_\_ in the alignment of nucleotide / protein sequences.
- a) dynamic programming              b) Levenshtein distance  
c) greedy algorithm                      d) none of the above
22. \_\_\_\_\_ tool is used to translate nucleic acid sequences.
- a) five frame translation              b) six frame translation  
c) four frame translation              d) ten frame translation

23. \_\_\_\_\_ is widely used for multiple sequence alignment of nucleotide and protein sequences.  
a) clustal      b) tblastn      c) blastx      d) tblastx
24. \_\_\_\_\_ includes insertions or deletions of a single base pair of DNA.  
a) point mutation      b) missense mutation  
c) frame shift mutation      d) all the above
25. Function of genes can be predicted at the RNA level by \_\_\_\_\_.  
a) microarray      b) ChIP      c) SAGE      d) all the above
26. \_\_\_\_\_ describe the 3-dimensional structure of every protein encoded by a given genome.  
a) structural genomics      b) expression proteomics  
c) functional proteomics      d) functional genomics
27. \_\_\_\_\_ is the method of creating a transcriptome without the aid of a reference genome.  
a) *De novo* transcriptome assembly      b) Reference transcriptome  
c) *Ab initio* transcriptome assembly      d) none of the above
28. Transcriptome is the subject that deals with the analysis of \_\_\_\_\_.  
a) mRNA      b) rRNA      c) tRNA      d) all the above
29. SAGE can be expanded as \_\_\_\_\_.  
a) serial analysis of gene expression  
b) serial analysis of genome expression  
c) support analysis of gene expression  
d) serial amplified gene expression
30. Assembler is a assembly software that has been used to generate \_\_\_\_\_.  
a) transcriptome      b) metabolome  
c) proteome      d) genome

\*\*\*\*\*