

Roll No.-----

प्रश्नपुस्तिका क्रमांक
Question Booklet No.

O.M.R. Serial No.

K-1376

B.Sc. (Biotech.) (Sixth Semester) Examination, 2025-26

(NEP)

(BBT6004)

BIOINFORMATICS

Paper Code

BBT6004

(To be filled in the
OMR Sheet)

प्रश्नपुस्तिका सीरीज
Question Booklet Series

A

Time : 1:30 Hours]

[Maximum Marks-75

Instructions to the Examinee :

1. Do not open the booklet unless you are asked to do so.
2. The booklet contains 100 questions. Examinee is required to answer 75 questions in the OMR Answer-Sheet provided and not in the question booklet. All questions carry equal marks.
3. Examine the Booklet and the OMR Answer-Sheet very carefully before you proceed. Faulty question booklet due to missing or duplicate pages/questions or having any other discrepancy should be got immediately replaced.

परीक्षार्थियों के लिए निर्देश :

1. प्रश्न-पुस्तिका को तब तक न खोलें जब तक आपसे कहा न जाए।
2. प्रश्न-पुस्तिका में 100 प्रश्न हैं। परीक्षार्थी को 75 प्रश्नों को केवल दी गई OMR आन्सर-शीट पर ही हल करना है, प्रश्न-पुस्तिका पर नहीं। सभी प्रश्नों के अंक समान हैं।
3. प्रश्नों के उत्तर अंकित करने से पूर्व प्रश्न-पुस्तिका तथा OMR आन्सर-शीट को सावधानीपूर्वक देख लें। दोषपूर्ण प्रश्न-पुस्तिका जिसमें कुछ भाग छपने से छूट गए हों या प्रश्न एक से अधिक बार छप गए हो या उसमें किसी अन्य प्रकार की कमी हो, तो उसे तुरन्त बदल लें।

(Remaining instructions on the last page)

(शेष निर्देश अन्तिम पृष्ठ पर)

1. Which of the following best describes the core definition of bioinformatics?
 - (A) The study of manual biological data entry
 - (B) An interdisciplinary field combining biology, computer science, mathematics and statistics
 - (C) A branch of medicine focused solely on surgical tools
 - (D) The study of bird and insect wing structures only
2. What is the primary difference between Homology and Analogy?
 - (A) Homology refers to shared function; Analogy refers to shared ancestry
 - (B) Homology refers to shared ancestry; Analogy refers to independent evolution
 - (C) They are identical terms used interchangeably
 - (D) Analogy is used only for DNA, while Homology is for proteins
3. What was the historical significance of Margaret Dayhoff's work in the 1960s?
 - (A) She coined the term "Bioinformatics"
 - (B) She founded the Protein Data Bank (PDB)
 - (C) She sequenced the entire human genome
 - (D) She created the first protein sequence database
4. Which organization manages the EMBL Nucleotide Database?
 - (A) NCBI in the USA
 - (B) DDBJ in Japan
 - (C) European Bioinformatics Institute (EBI)
 - (D) The Human Genome Project
5. In an EMBL entry, what does the "FT" field represent?
 - (A) Feature Table, containing annotations like coding regions
 - (B) Final Text
 - (C) File Transfer protocol
 - (D) Fast Transcription

6. What is the primary purpose of the Entrez system at NCBI?
- (A) To store only raw DNA sequences
 - (B) To archive retired databases like UniGene only
 - (C) To provide cloud computing for AWS
 - (D) To act as an integrated search tool linking databases like GenBank and PubMed
7. How do Orthologs differ from Paralogs?
- (A) Orthologs are within one species; Paralogs are across different species
 - (B) Orthologs are the same gene in different species; Paralogs are duplicated genes within a species
 - (C) Orthologs are only DNA; Paralogs are only proteins
 - (D) There is no difference between them
8. The “Bermuda Principles” (1996) were established to ensure:
- (A) Private ownership of genetic data
 - (B) The use of AlphaFold for prediction
 - (C) High costs for sequencing
 - (D) Rapid public sharing of genomic data
9. What does the term “Deep Homology” refer to?
- (A) Sequencing ancient fossils
 - (B) Conserved genes like Pax6 controlling eye development across different phyla
 - (C) High-throughput sequencing of the ocean floor
 - (D) The structural similarity of bird and insect wings

10. Which of the following is a primary focus of the GenBank database?
- (A) Comprehensive nucleotide sequences
 - (B) Protein 3D structures
 - (C) Tissue expression profiles exclusively
 - (D) Manual phylogenetic trees
11. If a researcher needs to find the specific 3D structure of a protein, which database established in 1971 should they use?
- (A) GenBank
 - (B) Protein Data Bank (PDB)
 - (C) UniGene
 - (D) KEGG
12. A scientist wants to compare a new gene sequence against all known sequences to find similarities. Which tool should they use?
- (A) CRISPR
 - (B) Microarrays
 - (C) RNA-Seq
 - (D) BLAST
13. To search for a complete mitochondrial genome specifically in humans using Entrez, which query is most appropriate?
- (A) Human OR mitochondrial DNA
 - (B) Human mitochondrial DNA AND complete genome
 - (C) BRCA1
 - (D) NM_000546

14. Which method is most effective for a scientist trying to reconstruct the evolutionary history (phylogeny) of a species?
- (A) Using analogy-based traits
 - (B) Phylogenetic analysis and sequence alignment
 - (C) Using only the UniGene archive
 - (D) Manual data entry
15. If you are looking for the raw sequence data in a GenBank entry, which section of the record do you scroll to?
- (A) ACCESSION
 - (B) DEFINITION
 - (C) ORIGIN
 - (D) LOCUS
16. A developer is creating a tool to design guide RNAs for gene editing. Which biological advancement will their tool primarily support?
- (A) Next-Generation Sequencing
 - (B) CRISPR
 - (C) The Atlas of Protein Sequence
 - (D) Needleman-Wunsch algorithm
17. If you need to retrieve the TP53 gene sequence using a specific NCBI accession number, which identifier would you use?
- (A) AC123456
 - (B) Hs.123
 - (C) NM_000546
 - (D) AlphaFold

18. Which step should a user take first when using GenBank via the Entrez system to find a human gene?
- (A) Select “Nucleotide” database
 - (B) Download the FASTA file
 - (C) Filter by mRNA
 - (D) Click on the Feature Table
19. A researcher is studying vertebrate forelimbs (human arm, bat wing, whale flipper). This study is an application of:
- (A) Analogy
 - (B) Horizontal Gene Transfer
 - (C) Structural Homology
 - (D) Metagenomics
20. When using the EMBL database, which unique tool allows for browsing sequence data?
- (A) BLAST
 - (B) ENA Browser
 - (C) MyNCBI
 - (D) LinkOut
21. What is the relationship between GenBank, EMBL, and DDBJ?
- (A) They are competing companies that do not share data
 - (B) DDBJ is the only one that uses BLAST
 - (C) GenBank is for humans, while EMBL is only for plants
 - (D) They are part of the INSDC and share sequence data daily

22. Why might “Convergent Evolution” be considered a challenge when determining homology?
- (A) It produces similar structures (like squid vs. human eyes) that do not share a common ancestor
 - (B) It speeds up the BLAST algorithm too much
 - (C) It only occurs in prokaryotes
 - (D) It requires cloud computing to analyze
23. Comparing the Smith-Waterman and Needleman-Wunsch algorithms, what is the primary functional difference?
- (A) Smith-Waterman is for 3D structures; Needleman-Wunsch is for sequences
 - (B) Smith-Waterman is for local alignment; Needleman-Wunsch is for global alignment
 - (C) One is manual; the other is AI-driven
 - (D) There is no difference
24. Which of these best illustrates the “transition from manual data analysis to computational methods” mentioned in the text?
- (A) The use of computers for phylogenetic trees and sequence analysis in the 1960s
 - (B) The creation of the human genome in 2003
 - (C) The retirement of UniGene in 2019
 - (D) The use of paper ledgers for gene tracking
25. Identify the outlier in the context of “Molecular Homology”:
- (A) Orthologs
 - (B) Paralogs
 - (C) DNA sequence similarities
 - (D) Analogous structures like bird vs. insect wings

26. What was a direct impact of the Human Genome Project (HGP) on bioinformatics?
- (A) It led to the immediate retirement of GenBank
 - (B) It accelerated the development of tools like assembly algorithms and annotation pipelines
 - (C) It decreased the need for interdisciplinary collaboration
 - (D) It stopped the development of AI in biology
27. Analyze the role of “Multi-Omics Integration” in modern bioinformatics. It involves combining:
- (A) Only DNA and RNA
 - (B) Different versions of the BLAST tool
 - (C) Computer science and law
 - (D) Genomics, proteomics, and metabolomics
28. In the context of sequence databases, how does “Entrez” differ from “GenBank”?
- (A) GenBank is a search engine; Entrez is a database
 - (B) GenBank is the specific nucleotide database; Entrez is the integrated search system for multiple databases
 - (C) There is no difference; they are the same tool
 - (D) Entrez is only for proteins
29. What distinguishes “Next-Generation Sequencing” (NGS) from earlier sequencing methods?
- (A) It is slower and more expensive
 - (B) It enables high-throughput, cost-effective sequencing
 - (C) It is a manual process
 - (D) It does not require bioinformatics tools

30. What is the significance of “Synteny” in determining homology?
- (A) It refers to the conservation of gene order on a chromosome
 - (B) It is a tool for predicting protein folding
 - (C) It is the process of gene duplication
 - (D) It is a method for cloud storage
31. What is the primary difference between SWISS-PROT and TrEMBL?
- (A) SWISS-PROT is for DNA; TrEMBL is for proteins.
 - (B) SWISS-PROT is manually reviewed; TrEMBL is automatically annotated
 - (C) TrEMBL has higher data quality than SWISS-PROT.
 - (D) SWISS-PROT is unreviewed; TrEMBL is reviewed
32. Which database serves as the world’s primary repository for 3D structural data of biological macromolecules?
- (A) GenBank
 - (B) SWISS-PROT
 - (C) Protein Data Bank (PDB)
 - (D) TrEMBL
33. In an EMBL or SWISS-PROT entry, what does a “Stable Identifier” provide?
- (A) A list of all researchers who studied the protein
 - (B) A unique accession number for each protein
 - (C) The 3D coordinates of every atom
 - (D) The cost of sequencing the protein

34. Which technique measures the mass-to-charge (m/z) ratios of ionized peptides?
- (A) PCR
 - (B) Sanger Sequencing
 - (C) Microarrays
 - (D) Mass Spectrometry
35. What is a “Chromatogram” in the context of DNA sequencing?
- (A) A table of protein functions
 - (B) A file showing colored peaks representing nucleotides (A, T, G, C)
 - (C) A 3D model of a protein
 - (D) A gel showing DNA fragment sizes
36. What is the main limitation of the TrEMBL database?
- (A) It only contains human sequences
 - (B) It is too small to be useful
 - (C) Potential redundancy and errors due to lack of manual review
 - (D) It cannot be accessed online
37. The PDB ID “1ZNJ” refers to which specific protein example in the text?
- (A) Hemoglobin
 - (B) Human Insulin
 - (C) BRCA1
 - (D) Cytochrome c
38. Which blot is specifically used for the detection of RNA expression?
- (A) Southern Blot
 - (B) Western Blot
 - (C) Northern Blot
 - (D) Eastern Blot

39. What is the purpose of “Base calling” in bioinformatics?
- (A) Designing primers for PCR
 - (B) Identifying nucleotides from chromatogram peak height and shape
 - (C) Normalizing microarray data
 - (D) Folding a protein structure
40. Why are in silico tools like NEBcutter used in restriction digestion?
- (A) To sequence the DNA
 - (B) To simulate and predict DNA fragment sizes before the experiment
 - (C) To measure fluorescent intensity
 - (D) To identify post-translational modifications
41. A researcher has a protein accession number (P68871) and wants to find its oxygen transport function. Which database should they search?
- (A) PDB
 - (B) UniPort/SWISS-PROT
 - (C) NEBcutter
 - (D) ImageJ
42. If you need to visualize the atomic-level interactions between a drug molecule and a viral protein, which resource is most appropriate?
- (A) SWISS-PROT
 - (B) TrEMBL
 - (C) PDB
 - (D) PCR

43. To objectively measure the intensity of bands on a Western Blot image, which software should a scientist use?
- (A) MASCOT
 - (B) SEQUEST
 - (C) ImageJ
 - (D) BLAST
44. A lab is performing a qPCR experiment and needs to determine the relative abundance of a gene. Which data value from the output is essential for this calculation?
- (A) PDB ID
 - (B) Ct values
 - (C) Chromatogram peaks
 - (D) Accession number
45. Which database would a scientist use to quickly find millions of new, unreviewed protein sequences from a recently sequenced genome?
- (A) SWISS-PROT
 - (B) TrEMBL
 - (C) PDB
 - (D) KEGG
46. Before performing a PCR reaction, a student uses an in silico tool to ensure their primers do not form dimers or hit off-targets. This is an application of:
- (A) Protein identification
 - (B) Primer design
 - (C) Microarray normalization
 - (D) X-ray crystallography

47. If an experimentalist uses X-ray crystallography to determine a protein's shape, where would they deposit the resulting 3D coordinates?
- (A) SWISS-PROT
 - (B) GenBank
 - (C) PDB
 - (D) TrEMBL
48. A researcher is matching mass spectra against a database to identify a specific protein. Which tool are they likely using?
- (A) NEBcutter
 - (B) MASCOT or SEQUEST
 - (C) ImageJ
 - (D) UniProt Search
49. To study the expression of thousands of genes simultaneously in a cancer cell, which technique is most suitable?
- (A) Sanger Sequencing
 - (B) Restriction Digestion
 - (C) Microarrays
 - (D) Northern Blot
50. A scientist finds a "Reviewed" entry in UniProt. This means the information has been :
- (A) Generated by a computer only
 - (B) Manually curated by expert biologists
 - (C) Translated directly from EMBL without check
 - (D) Derived only from 3D structures

51. Which challenge is unique to Microarray data compared to Restriction Digestion?
- (A) Fragment prediction
 - (B) High false-positive rates due to multiple testing of thousands of genes
 - (C) The need for restriction enzymes
 - (D) Visualizing bands on a gel
52. What is the relationship between TrEMBL and SWISS-PROT entries?
- (A) SWISS-PROT entries are moved to TrEMBL after 10 years.
 - (B) They are entirely unrelated
 - (C) Some TrEMBL entries are reviewed and subsequently moved to SWISS-PROT
 - (D) TrEMBL only contains 3D structures found in SWISS-PROT
53. Analyze the data output of Mass Spectrometry. Why is PTM (Post-Translational Modification) analysis difficult?
- (A) PTMs do not change the mass of a protein
 - (B) PTMs shift the peptide mass, making the spectra complex to detect
 - (C) Mass spectrometry only works for DNA
 - (D) Databases like MASCOT cannot search for PTMs
54. Compare the “Data Quality” and “Redundancy” of SWISS-PROT vs. TrEMBL.
- (A) SWISS-PROT has lower “Redundancy” and higher quality
 - (B) TrEMBL has lower redundancy and higher quality
 - (C) Both have high redundancy and low quality
 - (D) There is no difference in redundancy between them
55. What bioinformatics problem is common to both Sanger sequencing (chromatograms) and Microarrays?
- (A) The need for primer design
 - (B) The requirement for normalization and correcting for background noise/errors
 - (C) Protein 3D visualization
 - (D) Mass-to-charge ratio calculations

56. Why is a digital image analysis tool (like ImageJ) necessary for Blots?
- (A) Because blots are qualitative only
 - (B) To objectively quantify band intensity from semi-quantitative data
 - (C) To sequence the protein in the band
 - (D) To predict the 3D structure of the protein on the membrane
57. Identify the outlier among these 3D structure determination methods:
- (A) X-ray crystallography
 - (B) Nuclear Magnetic Resonance (NMR)
 - (C) Cryo-Electron Microscopy (Cryo-EM)
 - (D) Southern Blotting
58. In the PDB ID “1A3N”, what does this 4-character code represent?
- (A) The gene name
 - (B) The unique PDB ID for a specific structure entry
 - (C) The sequence of the first four amino acids
 - (D) The resolution of the crystal
59. Which bioinformatics challenge is specifically associated with “Base calling errors”?
- (A) Microarrays
 - (B) Western Blots
 - (C) PCR Primer design
 - (D) Chromatograms from Sanger Sequencing
60. What is the common role of bioinformatics across all experimental techniques (PCR, MS, Sequencing)?
- (A) It replaces the need for laboratory experiments
 - (B) It creates the restriction enzymes used in digestion
 - (C) It only provides storage for 3D structures
 - (D) It acts as a bridge to transform raw experimental data into meaningful biological knowledge

61. Which of the following best defines Phylogeny Analysis?
- (A) The study of 3D protein folding in a single species
 - (B) The study of evolutionary relationships among biological entities based on genetic sequences
 - (C) The process of manual DNA sequencing
 - (D) The clinical diagnosis of infectious diseases
62. What is the primary difference between global and local sequence alignment?
- (A) Global alignment covers the entire length of sequences; local alignment finds the best matching regions
 - (B) Global alignment compares only two sequences; local alignment compares three or more
 - (C) Global alignment is for DNA; local alignment is for proteins
 - (D) There is no functional difference between them
63. In a phylogenetic tree, what does a “Node” represent?
- (A) A specific living species
 - (B) The length of time since evolution began
 - (C) A common ancestor or a branching point
 - (D) The total number of mutations
64. Which of the following describes a “Monophyletic” group?
- (A) A group containing an ancestor and some of its descendants
 - (B) A group that excludes the common ancestor
 - (C) A group of unrelated species that look similar
 - (D) A group containing an ancestor and all of its descendants
65. What is the purpose of “Bootstrapping” in phylogenetic analysis?
- (A) To increase the speed of the alignment algorithm
 - (B) To assess the statistical reliability or confidence of the tree branches
 - (C) To change the DNA sequence to RNA
 - (D) To identify the root of the tree

66. Which type of biological sequence uses Uracil (U) instead of Thymine (T)?
- (A) DNA
 - (B) RNA
 - (C) Protein
 - (D) Amino Acid
67. What does the “Molecular Clock” hypothesis suggest?
- (A) Mutations occur at a relatively constant rate over time
 - (B) All species evolved at exactly the same time
 - (C) DNA sequences never change
 - (D) Proteins fold faster than DNA replicates
68. A researcher wants to align three or more sequences to find highly conserved regions. Which tool is most appropriate?
- (A) BLAST
 - (B) Clustal Omega
 - (C) Needleman-Wunsch
 - (D) Smith-Waterman
69. If you are tasked with tracing human maternal ancestry, which type of DNA should you analyze?
- (A) Nuclear DNA
 - (B) Mitochondrial DNA
 - (C) Ribosomal RNA
 - (D) Bacterial DNA
70. A scientist has two very similar sequences of the same length and wants an end-to-end comparison. Which algorithm should they apply?
- (A) Smith-Waterman
 - (B) Maximum Likelihood
 - (C) BLAST
 - (D) Needleman-Wunsch

71. You have built a phylogenetic tree and need to identify a “Root” to provide evolutionary direction. What should you add to your analysis?
- (A) More sequences of the same species
 - (B) An Ingroup
 - (C) An Outgroup (a distantly related species)
 - (D) A gap penalty
72. Which software would you use specifically to visualize and annotate a completed phylogenetic tree?
- (A) MAFFT
 - (B) iTOL
 - (C) jModelTest
 - (D) RA_xML
73. If a study requires the most statistically robust tree building method that considers a specific model of sequence evolution, which should be chosen?
- (A) UPGMA
 - (B) Neighbor-Joining
 - (C) Maximum Likelihood
 - (D) Parsimony
74. What is the primary difference between the Neighbor-Joining (NJ) method and Maximum Likelihood (ML)?
- (A) NJ is character-based; ML is distance-based
 - (B) NJ is distance-based and fast; ML is character-based and computationally intensive
 - (C) NJ is only for proteins; ML is only for DNA
 - (D) There is no difference in the underlying logic

75. Analyze the following group: A group includes a common ancestor and some, but not all, of its descendants. This is classified as:
- (A) Monophyletic
 - (B) Polyphyletic
 - (C) Paraphyletic
 - (D) Holophyletic
76. Why is “Model Selection” (e.g., using jModelTest) considered a critical step before building a tree?
- (A) It speeds up the alignment process
 - (B) It ensures the mathematical model fits the actual evolutionary patterns of the data
 - (C) It automatically roots the tree
 - (D) It eliminates the need for bootstrapping
77. How does an “Unrooted Tree” differ from a “Rooted Tree”?
- (A) An unrooted tree shows relationships but not the direction of evolution
 - (B) An unrooted tree is always more accurate
 - (C) A rooted tree does not have nodes
 - (D) Unrooted trees are only used for bacteria
78. In the context of sequence alignment, what is the impact of a “Gap Penalty”?
- (A) It rewards the algorithm for finding mutations
 - (B) It discourages the over-insertion of gaps to ensure biologically meaningful alignments
 - (C) It increases the length of the protein
 - (D) It forces two sequences to be identical

79. Evaluate the relationship between sequence similarity and homology:
- (A) High similarity always guarantees homology
 - (B) Homology is measured in percentages; similarity is not
 - (C) Similarity is a measurable quantity; homology is an inference of shared ancestry
 - (D) They are synonyms and used interchangeably
80. Which of the following best explains why “Trees are hypotheses”?
- (A) They are based on established historical records
 - (B) They cannot be changed once published
 - (C) They represent the best estimate of evolution based on available data and models
 - (D) They are not based on mathematical logic
81. What is the primary function of the Entrez system at NCBI?
- (A) To act as a standalone database for protein structures only
 - (B) To serve as a federated search engine that queries multiple databases simultaneously
 - (C) To replace the need for experimental DNA sequencing
 - (D) To provide a complex query language for legacy systems
82. Which of the following describes the “Ab Initio” approach to gene prediction?
- (A) It uses mathematical models to identify gene structures based on intrinsic DNA signals
 - (B) It relies solely on comparing new sequences to known databases
 - (C) It is used only for masking repetitive regions in the genome
 - (D) It requires RNA-seq data to function

83. In the context of genome annotation, what is the purpose of “Repeat Masking”?
- (A) To identify the specific function of a protein
 - (B) To translate DNA into RNA
 - (C) To increase the speed of a BLAST search
 - (D) To identify and “hide” repetitive sequences to prevent false gene predictions
84. What was the historical significance of the SRS (Sequence Retrieval System)?
- (A) It was the first tool to sequence the human genome
 - (B) It was an early integrated platform enabling cross-database queries
 - (C) It is the modern standard for AI-driven protein folding
 - (D) It is a tool specifically designed for CRISPR guide RNA design
85. Which of the following is a key characteristic of prokaryotic genomes compared to eukaryotes?
- (A) They contain numerous spliceosomal introns
 - (B) They are gene-dense with few introns
 - (C) They rely heavily on non-coding RNAs (ncRNAs)
 - (D) They are significantly larger and more complex
86. What does “Evidence-Based (homology) Prediction” involve?
- (A) Predicting genes using only neural networks
 - (B) Manually typing sequences into a Word document
 - (C) Using jModelTest to select evolutionary models
 - (D) Aligning a new genome against known proteins or transcripts to find matches
87. The tool “Apollo” is used in the genome annotation workflow for:
- (A) Automatic repeat masking
 - (B) Manual curation and refinement of gene models by biologists
 - (C) Statistical normalization of microarray data
 - (D) Measuring m/z spectra in mass spectrometry

88. A researcher wants to find all related publications and protein structures for a specific gene using one interface. Which tool should they use?
- (A) Entrez
 - (B) Repeat Masker
 - (C) Augustus
 - (D) HISAT2
89. If you are tasked with masking repetitive regions in a newly sequenced genome, which software would you run?
- (A) GeneMark-ES
 - (B) StringTie
 - (C) BLASTp
 - (D) RepeatMasker
90. A scientist has gathered RNA-seq data and wants to align it to a reference genome to improve annotation. Which tool is appropriate for this alignment?
- (A) MAKER2
 - (B) HISAT2 or STAR
 - (C) RepBase
 - (D) SRS
91. To validate that a predicted gene actually codes for a known protein, a researcher should run which process?
- (A) A BLASTp search against a database like UniProt
 - (B) A repeat masking session with RepBase
 - (C) An “Ab Initio” prediction with GeneMark
 - (D) A manual count of G-C content

92. Which tool would be best for integrating various lines of evidence (Ab Initio, RNA-seq, and homology) into a final gene set?
- (A) Apollo
 - (B) RepBase
 - (C) MAKER2
 - (D) PubMed
93. If a researcher is studying “Pan-Genomics,” what are they likely doing?
- (A) Sequencing a single gene in one person
 - (B) Annotating multiple genomes of a species to capture structural variation
 - (C) Building a 3D model of insulin
 - (D) Using Western blots to detect protein expression
94. How does the legacy SRS system differ from the modern Entrez system?
- (A) SRS used a complex query language; Entrez uses a user-friendly federated search with precomputed links
 - (B) SRS was for DNA; Entrez is for proteins
 - (C) SRS is AI-driven; Entrez is manual
 - (D) There is no difference; they are the same tool
95. What is a major challenge when predicting genes in eukaryotic genomes that is less prevalent in prokaryotes?
- (A) High gene density
 - (B) The presence of complex intron-exon structures and alternative splicing
 - (C) Lack of stop codons
 - (D) Absence of repetitive DNA
96. Analyze the role of RNA-seq in the annotation workflow. Why is it considered to “drastically improve accuracy”?
- (A) It provides direct experimental evidence of which parts of the genome are actually transcribed
 - (B) It masks, repetitive regions automatically
 - (C) It replaces the need for any “Ab Initio” tools
 - (D) It converts protein sequences back into DNA

97. Which of the following best describes the relationship between “Ab Initio” and “Homology-Based” prediction?
- (A) They are mutually exclusive and cannot be used together
 - (B) They both rely entirely on RNA-seq data
 - (C) Ab Initio is for proteins; Homology is for DNA only
 - (D) They are complementary; Ab Initio uses mathematical models, while Homology uses existing database knowledge
98. Why might a “Novel Gene” represent a challenge for genome annotation?
- (A) It is too small to be seen on a gel
 - (B) It lacks homology (matches) in existing databases, making it hard to validate
 - (C) It cannot be sequenced by NGS
 - (D) It always contains repetitive elements
99. Evaluate the “Emerging Trend” of Deep Learning in bioinformatics. What is its specific application in Unit IV?
- (A) Designing PCR primers
 - (B) Creating phylogenetic trees
 - (C) Storing data in the cloud
 - (D) Using neural networks for more accurate gene prediction (e.g., DeepGenome)
100. In a standard workflow, why is “Repeat Masking” performed before running an Ab Initio gene predictor like Augustus?
- (A) To make the genome sequence longer
 - (B) To prevent the gene predictor from misidentifying repetitive elements as actual genes
 - (C) Because Augustus can only read masked DNA
 - (D) To identify the organism’s genus and species

Rough Work / रफ कार्य

Rough Work / रफ कार्य

4. Four alternative answers are mentioned for each question as – A, B, C & D in the question booklet. The candidate has to choose the correct answer and mark the same in the OMR Answer-Sheet as per the direction :

Example :

Question :

Q. 1 (A) ● (C) (D)

Q. 2 (A) (B) ● (D)

Q. 3 (A) ● (C) (D)

Illegible answers with cutting and over-writing or half filled circle will be cancelled.

5. Each question carries equal marks. Marks will be awarded according to the number of correct answers you have.
6. All answers are to be given on OMR Answer Sheet only. Answers given anywhere other than the place specified in the answer sheet will not be considered valid.
7. Before writing anything on the OMR Answer Sheet, all the Instructions given in it should be read carefully.
8. After the completion of the examination candidates should leave the examination hall only after providing their OMR Answer Sheet to the invigilator. Candidate can carry their Question Booklet.
9. There will be no negative marking.
10. Rough work, if any, should be done on the blank pages provided for the purpose in the booklet.
11. To bring and use of log-book, calculator, pager and cellular phone in examination hall is prohibited.
12. In case of any difference found in English and Hindi version of the question, the English version of the question will be held authentic.

Impt. On opening the question booklet, first check that all the pages of the question booklet are printed properly. If there is any discrepancy in the question booklet, then after showing it to the invigilator, get another question booklet of the same series.

4. प्रश्न-पुस्तिका में प्रत्येक प्रश्न के चार सम्भावित उत्तर— A, B, C एवं D हैं। परीक्षार्थी को उन चारों विकल्पों में से एक सही उत्तर छॉटना है। उत्तर को OMR आन्सर-शीट में सम्बन्धित प्रश्न संख्या में निम्न प्रकार भरना है :

उदाहरण :

प्रश्न :

प्रश्न 1 (A) ● (C) (D)

प्रश्न 2 (A) (B) ● (D)

प्रश्न 3 (A) ● (C) (D)

अपठनीय उत्तर या ऐसे उत्तर जिन्हें काटा या बदला गया है, या गोले में आधा भरकर दिया गया, उत्तर निरस्त कर दिया जाएगा।

5. प्रत्येक प्रश्न के अंक समान हैं। आपके जितने उत्तर सही होंगे, उन्हीं के अनुसार अंक प्रदान किये जायेंगे।
6. सभी उत्तर केवल ओ. एम. आर. उत्तर-पत्रक (OMR Answer Sheet) पर ही दिये जाने हैं। उत्तर-पत्रक में निर्धारित स्थान के अलावा अन्यत्र कहीं पर दिया गया उत्तर मान्य नहीं होगा।
7. ओ. एम. आर. उत्तर-पत्रक (OMR Answer Sheet) पर कुछ भी लिखने से पूर्व उसमें दिये गये सभी अनुदेशों को सावधानीपूर्वक पढ़ लिया जाये।
8. परीक्षा समाप्ति के उपरान्त परीक्षार्थी कक्ष निरीक्षक को अपनी OMR Answer Sheet उपलब्ध कराने के बाद ही परीक्षा कक्ष से प्रस्थान करें। परीक्षार्थी अपने साथ प्रश्न-पुस्तिका ले जा सकते हैं।
9. निगेटिव मार्किंग नहीं है।
10. कोई भी रफ कार्य, प्रश्न-पुस्तिका के अन्त में, रफ-कार्य के लिए दिए खाली पेज पर ही किया जाना चाहिए।
11. परीक्षा-कक्ष में लॉग-बुक, कैलकुलेटर, पेजर तथा सेल्युलर फोन ले जाना तथा उसका उपयोग करना वर्जित है।
12. प्रश्न के हिन्दी एवं अंग्रेजी रूपान्तरण में भिन्नता होने की दशा में प्रश्न का अंग्रेजी रूपान्तरण ही मान्य होगा।

महत्वपूर्ण : प्रश्नपुस्तिका खोलने पर प्रथमतः जाँच कर देख लें कि प्रश्न-पुस्तिका के सभी पृष्ठ भलीभाँति छपे हुए हैं। यदि प्रश्नपुस्तिका में कोई कमी हो, तो कक्षनिरीक्षक को दिखाकर उसी सिरिज की दूसरी प्रश्न-पुस्तिका प्राप्त कर लें।