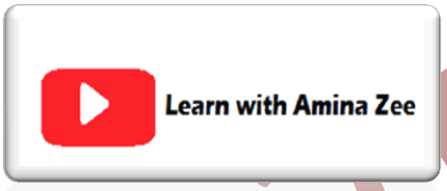
BIF 101 Introduction to bioinformatics:

Mid-term Short Notes

From Lec 1 to 88

Composed by Amina Zee



Question: 01

Describe the function of mitochondria?

Ans: Mitochondria are membrane-bound cell organelles (mitochondrion, singular) that generate most of the chemical energy needed to power the cell's biochemical reactions. Chemical energy produced by the mitochondria is stored in a small molecule called adenosine triphosphate (ATP).

Question: 02

What is the function of Golgi apparatus?

Golgi apparatus helps in,

- Covalent modification of proteins from ER,
- Sorting of proteins for
- Transport to other parts of the cell.

Question: 03

Define codon and anticodon?

The codon is unambiguous and specific means one codon can code for only one amino acid. Codons are tri-nucleotide units that present in mRNA and codes for a particular amino acid in protein synthesis.

Anticodon is tri-nucleotide units that present in tRNA. It is complementary to the codons in mRNA.

Question: 04

Define gene cloning?

Since 1970s, scientists have learned to isolate genes, then place them in new organisms and reproduce them by a set of techniques collectively known as gene cloning.

Technique of gene cloning:

Genetically Modified Organisms (GMOs):

The technique of Gene cloning led to the creation of a large number of genetically modified organisms with desirable characters.

Question: 05

What is Human genome project?

The Human Genome Project (HGP) was launched in 1990 and completed in 2003.The Human Genome Project (HGP) was **an international scientific research project** with the goal of determining the base pairs that make up human DNA, and of identifying, mapping and sequencing all of the genes of the human genome from both a physical and a functional standpoint.

Question: 06

What is Nucleic acid?

Nucleic acids are important group of biomolecules which are responsible for storage & transmission of hereditary information. Like proteins and polysaccharides, nucleic acids are also polymeric compounds. The repeating units in the nucleic acids are Nucleotides.

Question: 07

Describe chemical composition of DNA?

The DNA molecule consists of two strands that wind around one another to form a shape known as a double helix. Each DNA strand composed of three components:

- Deoxyribose
- Nitrogenous Base
- Phosphoric acid

And consist of **four bases, adenine (A), cytosine (C), guanine (G), and thymine (T)**.

Question: 08

Describe types of DNA?

There are four types of Deoxyribonucleotides.

- Deoxyadenosine
- Deoxythymidine
- Deoxyguanosine
- Deoxycytidine

These four deoxyribonucleotides make the structural units of DNA.

Question: 09

Differentiate between Nucleotide and nucleoside?

Nucleotide	Nucleotide
A molecule containing all these three components is called a nucleotide.	While a molecule without the phosphate group is called a nucleoside.
A nucleotide consists of a sugar molecule (either ribose in RNA or deoxyribose in DNA) attached to a phosphate group and a nitrogen-containing base. The bases used in DNA are adenine (A), cytosine (C), guanine (G), and thymine (T).	A nucleoside, composed of a nucleobase, is a pyrimidine (cytosine, thymine or uracil) or a purine (adenine or guanine), a five carbon sugar which is either ribose or deoxyribose.

Question: 10

Define Genetics?

Genetics is a branch of biology concerned with the study of genes, genetic variation, and heredity in organisms. Genes are made up of DNA. Some genes act as instructions to make molecules called proteins. However, many genes do not code for proteins. Every person has two copies of each gene, one inherited from each parent.

Question: 11

Name four sub-disciplines of Genetics?

- **Transmission (Classical) Genetics** (Deals with movement of genes and genetic traits from parents to offspring.)
- **Population Genetics** (Studies heredity in groups for traits determined by one or a few genes.)
- **Quantitative Genetics** (Studies group hereditary for traits determined by many genes simultaneously. Skin color, height, eye color)
- **Molecular Genetics (**Deals with the molecular structure and function of genes.)

Question: 12

Differentiate between genotype and phenotype?

Genotype	Phenotype
Genetic make-up of an organism is called Genotype. • Example of genotype includes Height. For an individual's gene makeup there is tall variety (T) and there is short variety (s). T and s are called the alleles. • Freckles or no freckles. Again the information that is passed from parent to child is carried in the cell of the genotype. .	Physical appearance of an organism is called Phenotype. Examples of phenotypes include height, wing length, and hair color. Phenotypes also include observable characteristics that can be measured in the laboratory, such as levels of hormones or blood cells.

Question: 13

Differentiate between Dominant and recessive trait?

Dominant Trait	Recessive trait
When one characteristic expresses itself over the other i.e. round over wrinkled. . Having a widow's peak (a V-shaped hairline near the forehead) is dominant.	The trait that does not show through in the first generation i.e. wrinkled. For example, having a straight hairline is recessive

Question; 14

What is meant by genomics?

Genomics is the study of all of a person's genes (the genome**)**, including interactions of those genes with each other and with the person's environment.

Question: 15

Describe prokaryotic genome?

Prokaryotes are simple organisms.

Prokaryotes are the organisms whose Genetic material (DNA) is not enclosed in a nuclear membrane, there is no membrane bound organelles.

First prokaryotic Genome sequenced was that of Hemophilus influenzae. It paved the way for sequencing of many other organisms.

Question: 16

Describe Eukaryotic genome?

Eukaryotes have larger genomes. Eukaryotic genomes are composed of one or more linear DNA chromosomes. The number of chromosomes varies widely from Jack jumper ants and an asexual nematode, which each have only one pair, to a fern species that has 720 pairs. It is surprising the amount of DNA that eukaryotic genomes contain compared to other genomes.

Question: 17

DNA Sequencing	RNA Sequencing	Protein Sequencing
DNA sequencing is a laboratory technique used to determine the exact sequence of bases (A, C, G, and T) in a DNA molecule. The DNA base sequence carries the information a cell needs to assemble protein and RNA molecules. DNA sequence information is important to scientists investigating the functions of genes.	RNA-seq (RNA-sequencing) is a technique that can examine the quantity and sequences of RNA in a sample using next-generation sequencing (NGS). It analyzes the transcriptase, indicating which of the genes encoded in our DNA are turned on or off and to what extent.	Protein sequencing is the practical process of determining the amino acid sequence of all or part of a protein or peptide. This may serve to identify the protein or characterize its post-translational modifications**.**

Question: 18

What is meant by central Dogma?

The central dogma states that information in nucleic acid can be perpetuated or transferred, but the transfer of information form into protein is irreversible.

• DNA -> RNA -> Protein Synthesis

Question: 19

What is Transcription and translation?

Transcription	Translation
Transcription is the process by which the information in a strand of DNA is copied into a new molecule of messenger RNA (mRNA). The newly formed mRNA copies of the gene then serve as blueprints for protein synthesis during the process of translation.	Translation is the process by which the genetic code contained within a messenger RNA (mRNA) molecule is decoded to produce a specific sequence of amino acids in a polypeptide chain. The key components required for translation are mRNA, ribosomes, and transfer RNA (tRNA).

Question: 20

Describe types of RNA?

There are mainly three types of Ribonucleic acids (RNAs) present in the cells of living organisms.

- **Messenger RNA (mRNA) :**
  It is the type of RNA that carries genetic information from DNA to the protein biosynthetic machinery of the ribosome. It provides the templates that specify amino acid sequences in polypeptide chains.
- **Transfer RNA (tRNA) :**
  Transfer RNAs serve as adapter molecules in the process of protein synthesis. They are covalently linked to an amino acid at one end.
- **Ribosomal RNA (rRNA):**
  Ribosomal RNAs are components of ribosomes. rRNA is a predominant material in the ribosomes constituting about 60% of its weight. It has a number of functions to perform in the ribosomes.

Question: 21

Describe structure of RNA?

RNA consists of ribose nucleotides (nitrogenous bases appended to a ribose sugar) attached by phosphodiester bonds, forming strands of varying lengths. The nitrogenous bases in RNA are adenine, guanine, cytosine, and uracil, which replace thymine in DNA.

Question: 22

What are ribosomes?

The ribosome is the macromolecular machine that directs the synthesis of proteins. The ribosome is larger and more complex than the minimal machinery required for DNA or RNA synthesis.

The machinery for polymerizing amino acids is composed of at least three RNA molecules and more than 50 different proteins, with an overall molecular mass of >2.5 MDa.

Question: 23

How peptide bonds are formed?

Peptide bonds are formed by a biochemical reaction that extracts a water molecule as it joins the amino group of one amino acid to the carboxyl group of a neighboring amino acid. The linear sequence of amino acids within a protein is considered the primary structure of the protein

Question: 24

What are the binding sites of ribosomes for RNA?

Ribosomes are composed of two subunits, one small and one large.

Four binding sites are located on the ribosome, one for mRNA and three for tRNA. The three tRNA sites are labeled P, A, and E. The P site, called the peptidyl site, binds to the tRNA holding the growing polypeptide chain of amino acids.

Question: 25

What is meant by initiation Factor?

In prokaryotes, translation initiation is controlled by three initiation factors: IF1, IF2, and IF3. Both IF1 and IF2 are involved in positioning the initiator tRNA in the partial P site of the 30S subunit, while the GTPase activity of IF2 signals the beginning of translation elongation.

Question: 26

How translocation took place in large subunits?

These movements are coordinated within the ribosome and are collectively referred to as translocation.

- The initial steps of translocation are coupled to the peptidyl transferase reaction.
- Once the growing peptide chain has been transferred to the A-site tRNA, the A- and P-site tRNAs have a preference to occupy new positions in the large subunit.
- The 3’ end of the Asite tRNA is bound to the growing polypeptide chain and prefers to bind in the P-site of the large subunit.
- The now deacetylated P-site tRNA is no longer attached to the growing polypeptide chain and prefers to bind in the E-site of the large subunit.
- In contrast, at this time, the anticodons of these tRNAs remain in their initial location in the small subunit bound to the mRNA.

- Thus, translocation is initiated in the large subunit before the small subunit, and the tRNAs are said to be in “hybrid states”.

Question: 27

What is Amino acid?

An amino acid is **an organic molecule** that is made up of a basic amino group ($-NH_2$), an acidic carboxyl group ($-COOH$), and an organic R group (or side chain) that is unique to each amino acid. The term amino acid is short for α-amino [alpha-amino] carboxylic acid.

Question: 28

Define proteins and describe its functions?

Proteins are large biomolecules and macromolecules that comprise one or more long chains of amino acid residues.

- **Functions:**

Protein has many roles in your body. It **helps repair and builds your body's tissues**, allows metabolic reactions to take place and coordinates bodily functions. In addition to providing your body with a structural framework, proteins also maintain proper pH and fluid balance.

Questions: 29

Write a short on structure of proteins?

Within a protein, multiple amino acids are linked together by peptide bonds, thereby forming a long chain.

- **Primary Structure** Sequence of amino acids bonded by peptide linkages.
- **Secondary Structure** α helices (hydrogen bonds present between amino acid residues), β pleated sheets (hydrogen bonds present between amino acid residues).
- **Tertiary Structure** Generated by bending and folding of the polypeptide chain.
- **Quaternary structure** Arrangement of multiple folded protein molecules in a multi-subunit complex.

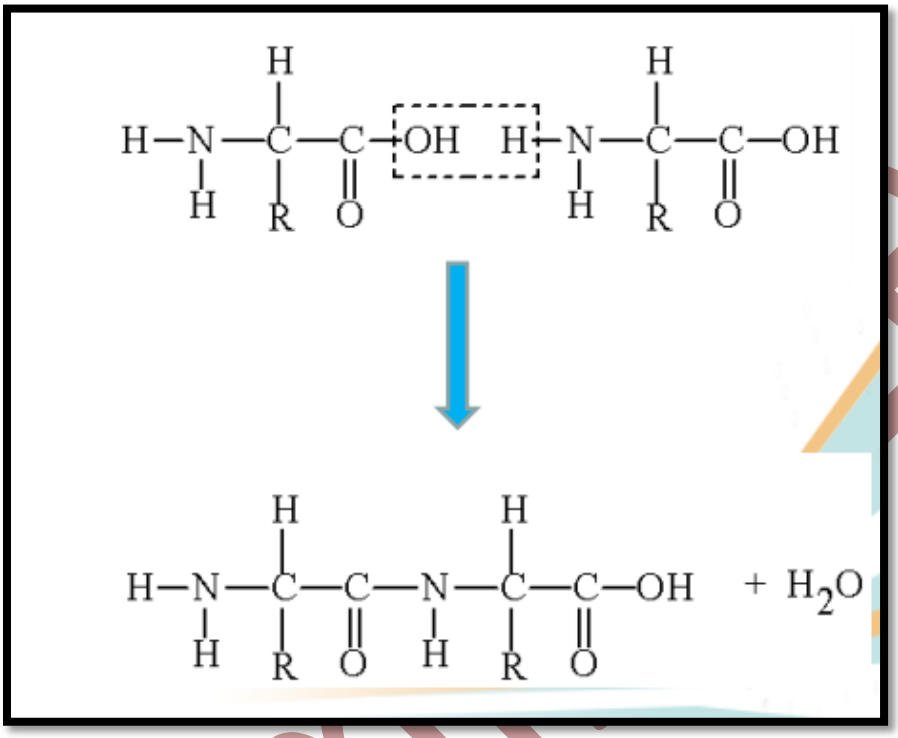
Question: 30

What is chemical composition of proteins?

All the proteins are made up of twenty different types of amino acids. So these amino acids are called standard amino acids.

In a protein molecule, each amino acid residue is joined to its neighbor by a specific type of covalent bond which is called Peptide Bond.

Amino acids can successively join to form dipeptides, tripeptides, tetrapeptides, oligo peptides and polypeptides.



Question: 31

What is Genbank?

The GenBank sequence database is an open access, annotated collection of all publicly available nucleotide sequences and their protein translations.

The GenBank database is designed to provide and encourage access within the scientific community to the most up-to-date and comprehensive DNA sequence information.

Question: 32

Comparing Sequences:

Sequence alignment appears to be extremely useful in a number of bioinformatics applications.

By comparing sequences, we can get:

• Similarity

• Specific Differences

• Relationship

• Evolutionary Insight

Question: 33

What is meant by exact and inexact matching?

Exact Matching	**Inexact Matching**
Exact Matches include complete matching in terms of residues and their order	In exact matching includes partial matching in terms of residues and order with room for variations in both.

Question: 34

What is Sequence Alignment and its types?

The process of inexact matching while keeping in view the conserved residues is called sequence "Alignment".

There are two types of sequence alignment

- Pairwise sequence Alignment
- Multiple Sequence Alignment

Question: 35

What is Pairwise sequence Alignment?

Pairwise Sequence Alignment is **used to identify regions of similarity that may indicate functional, structural and/or evolutionary relationships between two biological sequences.**

Pairwise sequence alignment helps compare two sequences.

• It takes into consideration inexact matching

• Matches are colored and missing nucleotides are denoted by dot **'.'**

Question No: 36

Describe types of Pairwise Alignment?

Global and local pairwise alignments are two types of alignment and they can lead to different results

Local Alignment	Global Alignment
Local alignment finds the regions between two sequences that have the strongest similarity, excluding less similar regions. • It is used to Find similar domains, motifs, detecting distant homology	Global alignment, Maximizing sequence matches over the entire length of two sequences, by introducing gaps. • It is used to determine overall similarity, conservation.

Question; 37

What is identity sequence and Similarity sequence?

Identity Sequence	Similarity Sequence
Identity is the number of nucleotides or amino acids which match exactly between two biological sequences. Identity measurement is made on the shorter of the two sequences **Formula:** Identity = Num. of matches/Smaller Length × 100%	Sequence similarity is a measure of an empirical relationship between sequences. Similarity is the comparison between sequences calculated by using alignment approach.

Question: 38

What is Multiple Sequence Alignment?

Multiple Sequence Alignment (MSA) is generally the alignment of three or more biological sequences (protein or nucleic acid) of similar length.

For Multiple sequence alignment, We use progressive Alignment as Dynamic programming become expensive.

Question: 39

What is Progressive Alignment?

A progressive alignment method is described that utilizes the Needleman and Wunsch pairwise alignment algorithm iteratively to achieve the multiple alignment of a set of protein sequences and to construct an evolutionary tree depicting their relationship.

Question: 40

What is CLUSTAL?

It is an online tool to do multiple alignments. The CLUSTAL series of programs are widely used in molecular biology for the multiple Alignment of both nucleic acid and protein sequences and for preparing phylogenetic trees.

Question: 41

What is BLAST tool?

The Basic Local Alignment Search Tool (BLAST) **finds regions of local similarity between protein or nucleotide sequences**. The program compares nucleotide or protein sequences to sequence in a database and calculates the statistical significance of the matches.

Question: 42

Describe types of Blast?

BLAST n	Compare the nucleotide query against nucleotide database.
BLAST p	Compare the amino acid query against the protein database
BLAST x	Compare the nucleotide sequence against protein database.
t BLAST n	Compare the protein query against nucleotide sequence.
t BLAST x	Compare a protein query against the all six frames of a nucleotide sequence database.

Question: 43

What is FASTA tool?

FASTA is tool developed in 1988, the FASTA format is a text-based format for representing either nucleotide sequences or amino acid (protein) sequences, in which nucleotides or amino acids are represented using single-letter codes. The format also allows for sequence names and comments to precede the sequences**.**

Question: 44

Describe the types of FASTA?

t FAST x 35	Compare a protein sequence to DNA sequence database.
FAST y 35	Compare a DNA sequence to Protein sequence database.
FAST m 35	Compare ordered peptides to protein database.
FAST s 35	Compare unordered peptides to a protein sequence database.

"Work Hard, Achieve your Goal and make your own World"

Prepared and Composed by AMINA ZEE

Best of Luck

☺Thank You☺