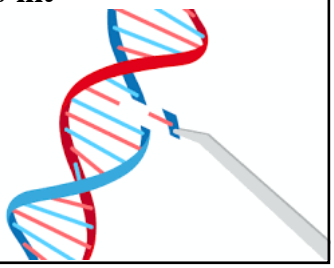
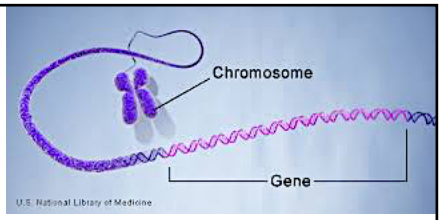
Bioinformatics I
From genes to proteins

Dr Manaf A Guma
University Of Anbar- College Of Applied Sciences-hit
Department Of Applied Chemistry

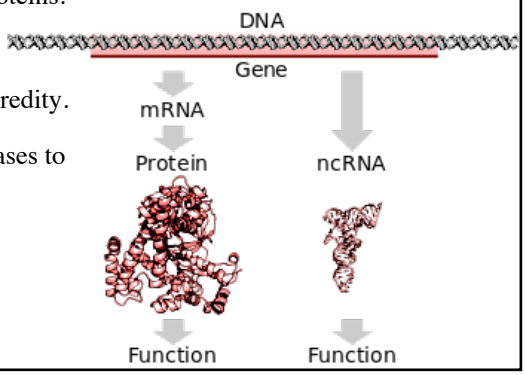


1

What is a gene?



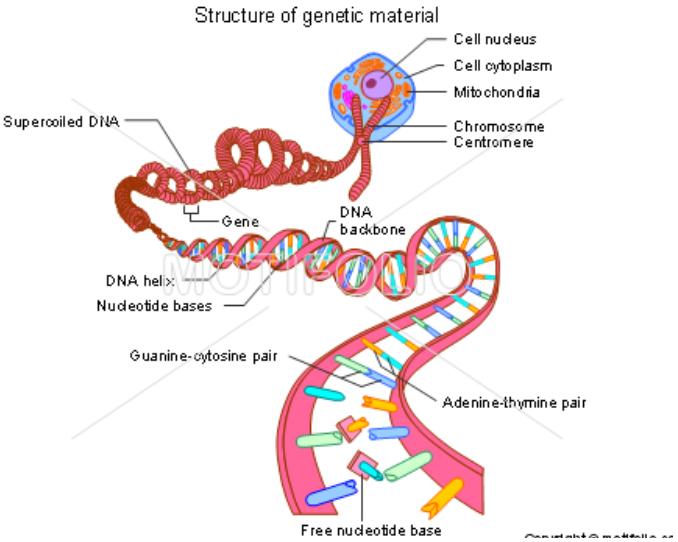
- **Genes** are made up of DNA.
- In biology, a **gene** is a sequence of nucleotides in DNA or RNA that encodes the synthesis of a gene product, either RNA or protein
- Some **genes** act as instructions to make molecules called proteins. However, many **genes** do not code for proteins.
- Gene: A **gene** is the basic physical and functional unit of heredity.
- In humans, **genes** differ in size from a few hundred DNA bases to more than 2 million bases.



2

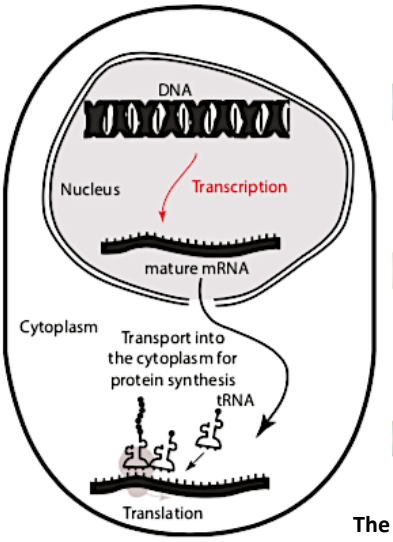
DNA is the Genetic Material

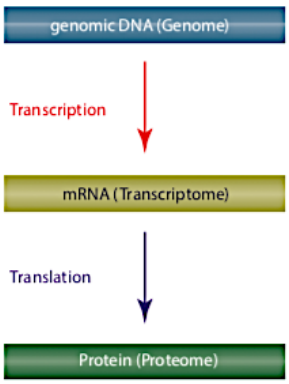
- ***DNA is a part of a Gene.***
- ***It is an (inheritance martial) that is transcribed into RNA (transcription), translated to a protein (translation in the ribosome).***
- ***In chromatin, DNA is tightly coiled around histones in the nucleus.***
- ***The process is called ''Central dogma of protein synthesis''.***



3

The central dogma of molecular biology





The flow of information always proceeds from the genome to the proteome, not vice versa.

4

What does DNA / RNA mean ?

- **DNA is Deoxy ribo Nucleic Acid.**
- **RNA is RiboNucleic Acid.**
- **The name is based on the type of sugar molecules attached to the nucleic acid.**

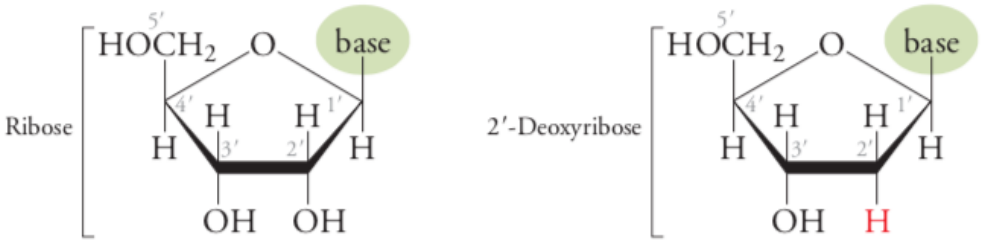
5

What is the units of the DNA ?

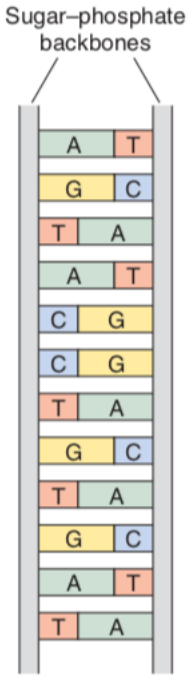
- **Nucleic acids** only four different types of structural units of DNA.
- DNA was known to contain chains of polymerized nucleotides—abbreviated A, C, G, and T—but these were thought to occur as simple repeating tetranucleotides.
- For example—ACGT-ACGT-ACGT-ACGT—
- The DNA structure ultimately elucidated by James Watson and Francis Crick in 1953 !
- Now: it is a Fingerprinting for each ''species''

6

- So, DNA contains the bases A, C, G, and T, whereas RNA contains A, C, G, and U.
- Linking a purine or a pyrimidine to a five-carbon sugar forms a **nucleoside.**
- *In DNA, the sugar is deoxyribose; in RNA, the sugar is ribose .*

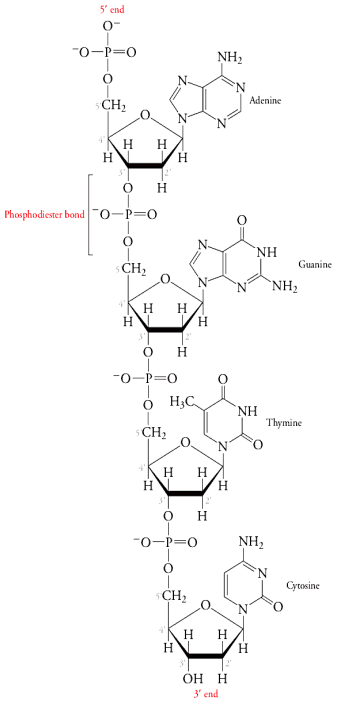


7



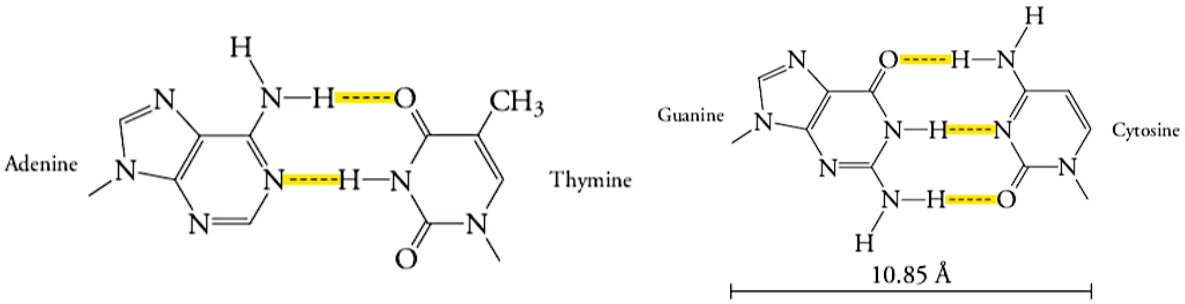
DNA is a double helix

- Nucleotides are linked by phosphodiester bonds form a polymer.
- There are hydrophobic interactions between stacked bases.
- The end of the polymer that carries a phosphate group attached to C5' is known as the **5' end,** and the end that carries a free OH group at C3' is the **3' end.**
- **The sequence is red from 5' end to 3' end.**



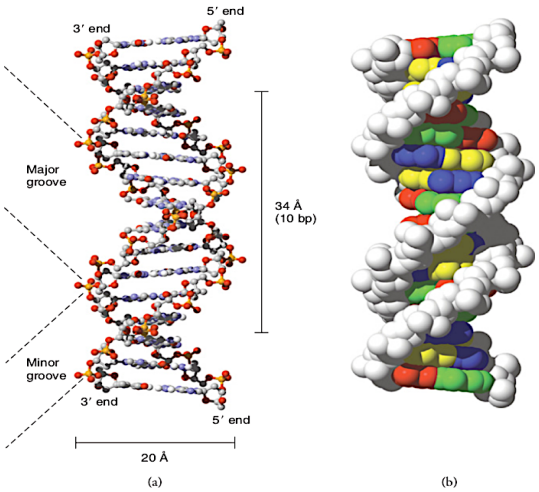
8

- DNA contains two polynucleotide strands whose bases pair through **hydrogen bonding** .
- Two hydrogen bonds link adenine and thymine, and three hydrogen bonds link guanine and cytosine:



9

- The diffraction (scattering) of an X-ray beam by a DNA fibre suggested a (helix)'' spiral'' with a repeating spacing of 3.4 Å.

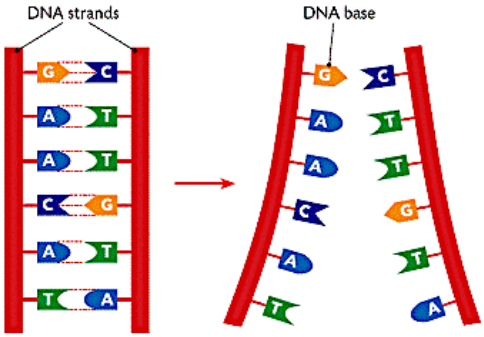


- *The size of a DNA is expressed in units of base pairs* ***(bp)*** *or kilo- base pairs (1000 bp,* ***(kb)****.*

Model of DNA. (a) Ball-and-stick model with atoms colored: C gray, O red, N blue, and P gold (H atoms are not shown). (b) Space-filling model with the sugar–phosphate backbone in gray and the bases color-coded: A green, C blue,G yellow, and T red.

10

Each DNA strand has a complementary strand.



Forward = Reverse Complement (Forward)

AGCTTCTAGTCGACTAGAAGCT

AGCTTCTAGTCGACTAGAAGCT

https://www.bioinformatics.org/sms/rev_comp.html
https://www.genscript.com/sms2/rev_comp.html

From Thomas Krahn's presentation.

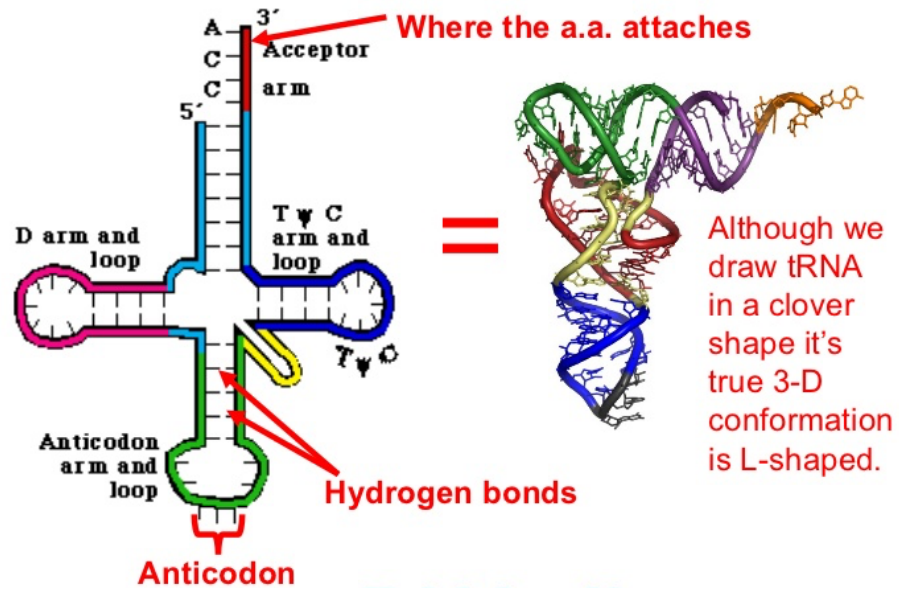
11

RNA is a single-stranded

- RNA, which is a single-stranded polynucleotide, has greater conformational freedom than DNA.
- What are the types of RNA?

1. The transcribed RNA is known as **messenger RNA (mRNA)** because it carries the same genetic message as the gene.
2. The mRNA is translated in the **ribosome,** a cellular particle consisting of protein and **ribosomal RNA (rRNA) (discussed later !).**
3. At the ribosome, small molecules called **transfer RNA (tRNA),** which carry amino acids, recognize sequential sets of three bases (known as **codons**)

12



A transfer tRNA molecule. This 76-nucleotide single-tranded RNA molecule folds back on itself so that base pairs

13

So far, what did we understand?

- There are two types of sequences which are constructed by a biomolecules:
- Nucleotides sequences:

1. RNA seq: such as `AUUGCCGGCUUUA`
2. DNA seq: such as `ATTGCCGGCTTTA`
3. Proteins sequences: : such as `ASMDAIKKKMQMLKLDKENALD`

- These two sequences are the base of bioinformatics that we are going to deal with.

14

Review ?

- Exercise 1.1
  What is the difference between the two polynucleotides DNA and RNA?
- Exercise 1.2
  DNA consists of two complementary nucleotide strands. Which base pairings are observed between these two nucleotide strands?
- Exercise 1.3
  What is the meaning of the terms genome, transcriptome, and proteome?
- Exercise 1.4
  The 20 naturally occurring amino acids are encoded by base triplets (codons) of the genetic code. Which consideration led to the discovery of the triplet codon organization of the genetic code?
  Exercise 1.5
- Build the genetic code of your name. If this is not possible, use the name CRICK.
- Exercise 1.6
- What is meant by the central dogma of molecular biology?
- Exercise 1.7
  What is meant by the term splicing, and how does this process contribute to the dis- crepancy between the relatively low number of genes in the human genome but the larger number of proteins actually produced?

15

Review ?

- Exercise 1.8
  Which amino acids show the following properties: (A) hydrophobic, polar, and small and (B) hydrophobic and aliphatic?
- Exercise 1.9
  In which direction is the primary structure of proteins read?
- Exercise 1.10
  Which structural elements can be found in the secondary structure of proteins?
- What is meant by the central dogma of molecular biology?
- Describe the differences between prokaryotes and eukaryotes cells.
- What are the functions of each particles in each in the cell ?
- Such as mitochondria? Cytoplasm? …etc

16

REVIEW ?

1. How does DNA encode genetic information and how is this information expressed?
2. What is the relationship between the nucleotide sequence in a gene and the amino acid sequence of a protein?
3. List some reasons why knowing a gene's sequence might be useful.
4. How has the DNA sequence changed and how does this
5. affect the encoded protein?

17