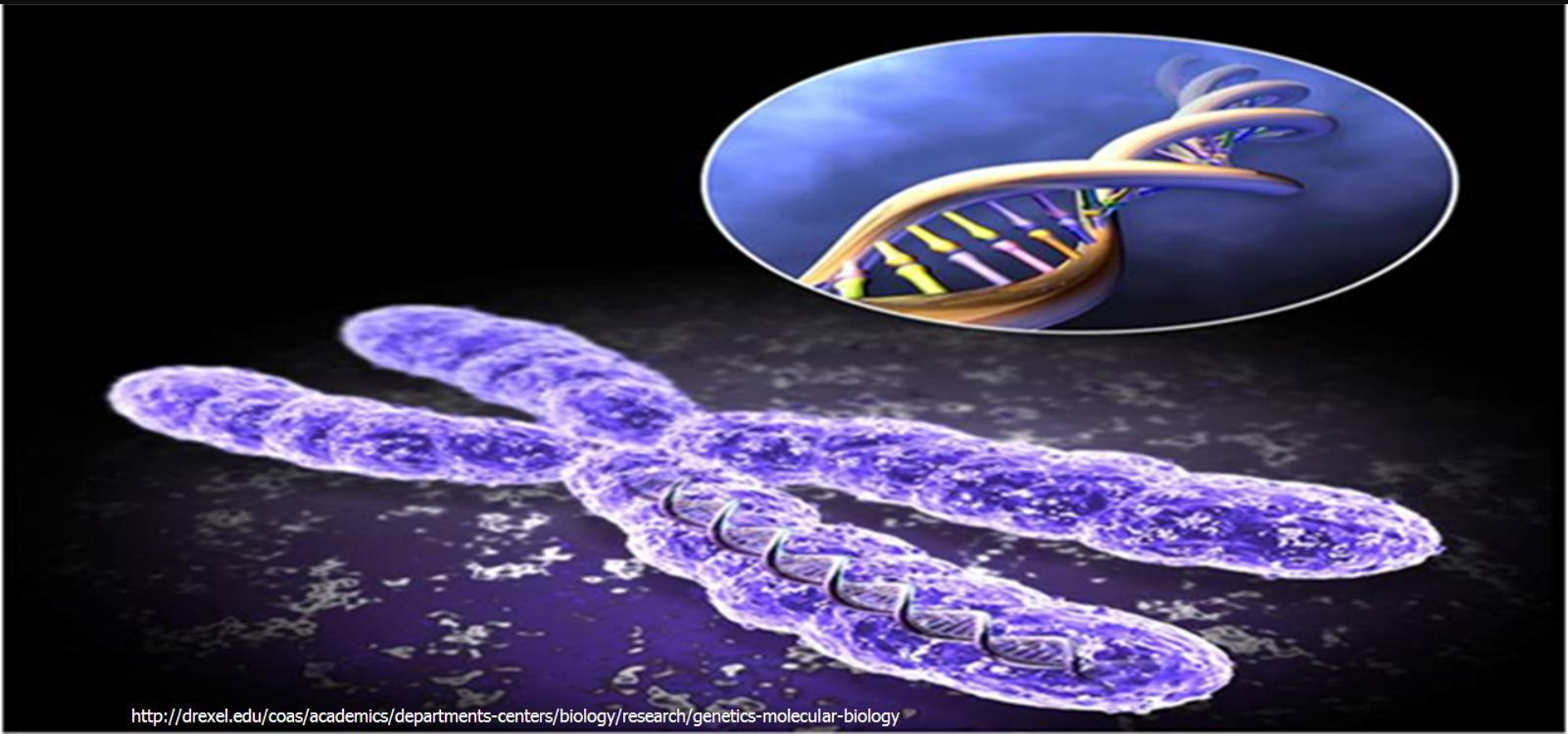
INTRODUCTION TO MOLECULAR BIOLOGY

SALWA HASSAN TEAMA



http://drexel.edu/coas/academics/departments-centers/biology/research/genetics-molecular-biology

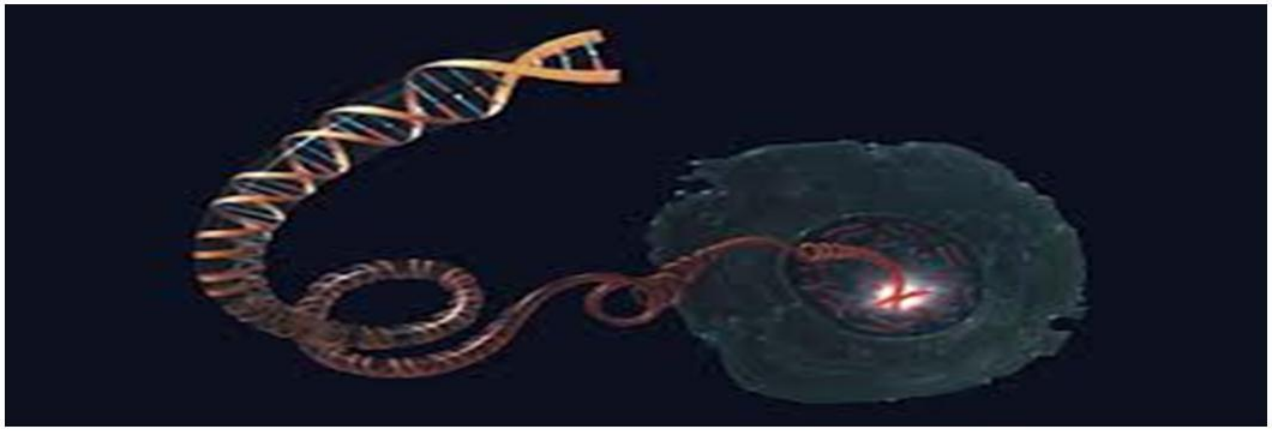
MOLECULAR BIOLOGY

- Molecular biology; the study of gene structure and functions at the molecular level to understand the molecular basis of hereditary, genetic variation,…. and the expression patterns of genes.
- The Molecular biology study the flow of information from DNA to RNA to protein.
- The Molecular biology field overlaps with other areas, particularly genetics and biochemistry.
- The Molecular biology allows the laboratory to be predictive in nature; events that occur in the future.

Source:http://www.differencebetween.info/difference-between-genome-and-dna

GENOME

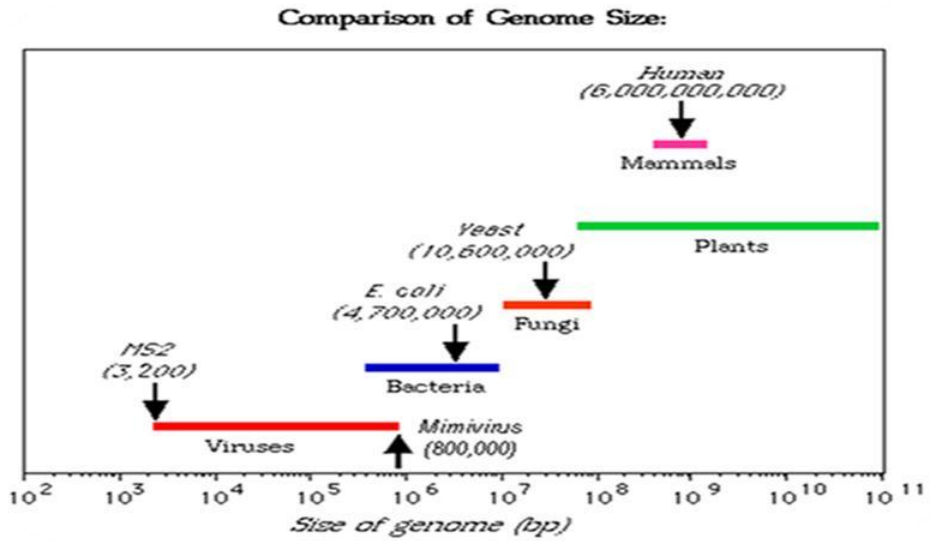
The genetic information within one mature cell of an organism
Encoded in the DNA (or for some viruses, RNA)



GENOME DATABASE

Organized in six major organism groups:

- Eukaryotes
- Bacteria
- Archaea
- Virus
- Viroids and
- Plasmid

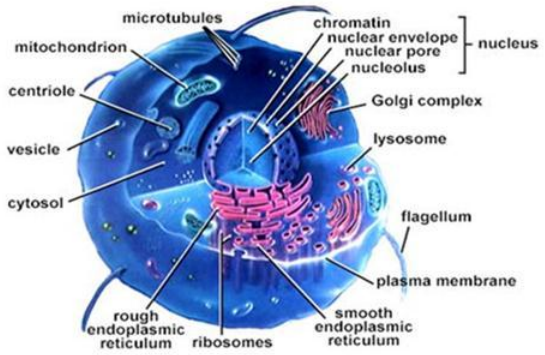


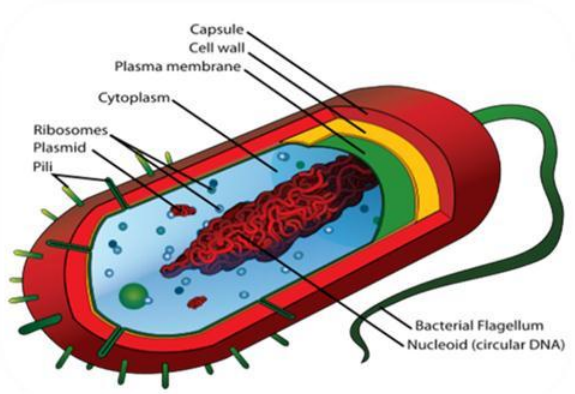
Source:http://www.19thpsalm.org/Ch06/

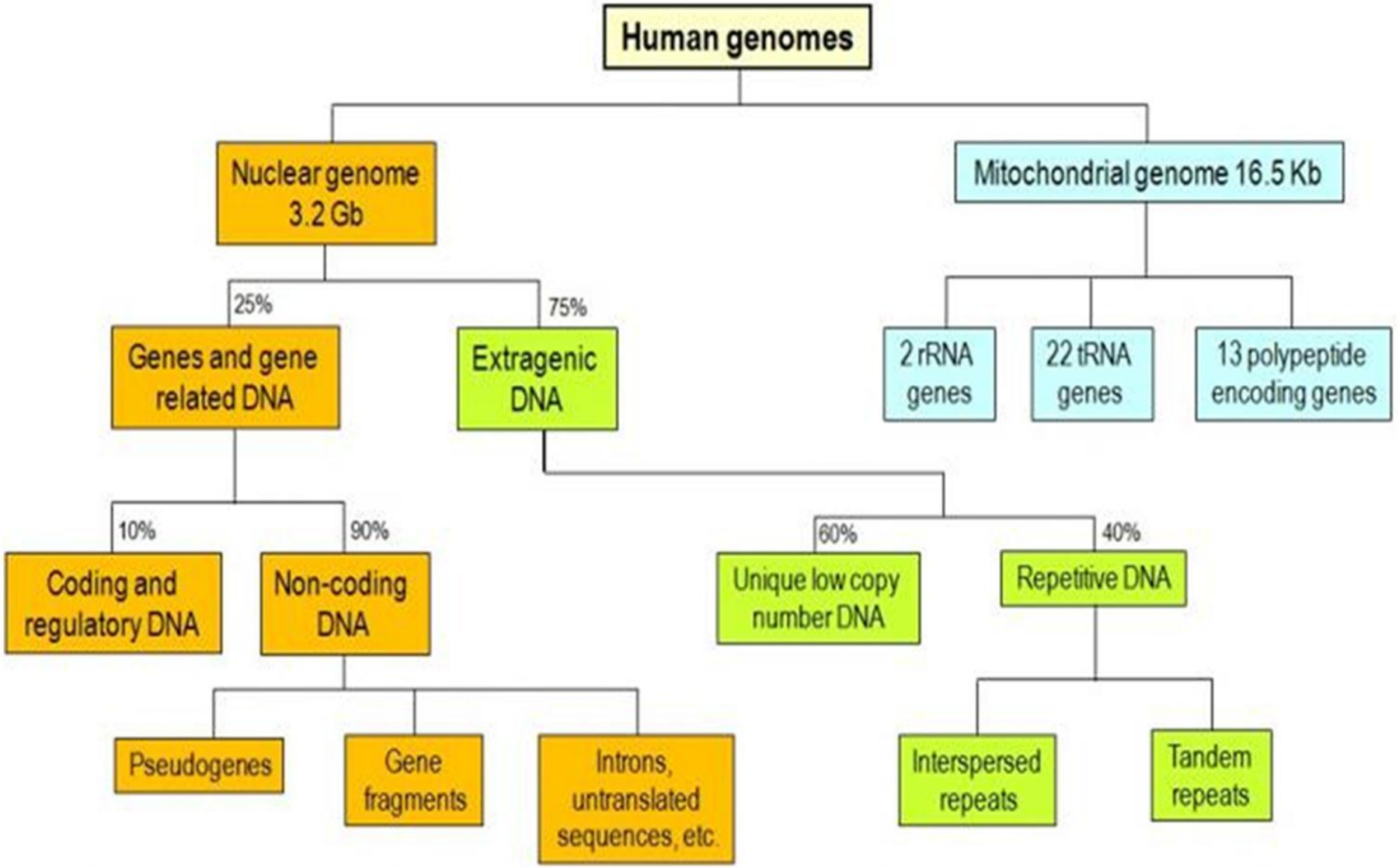
THREE DOMAIN OF LIFE

Eukaryotes, prokaryotes and archaea

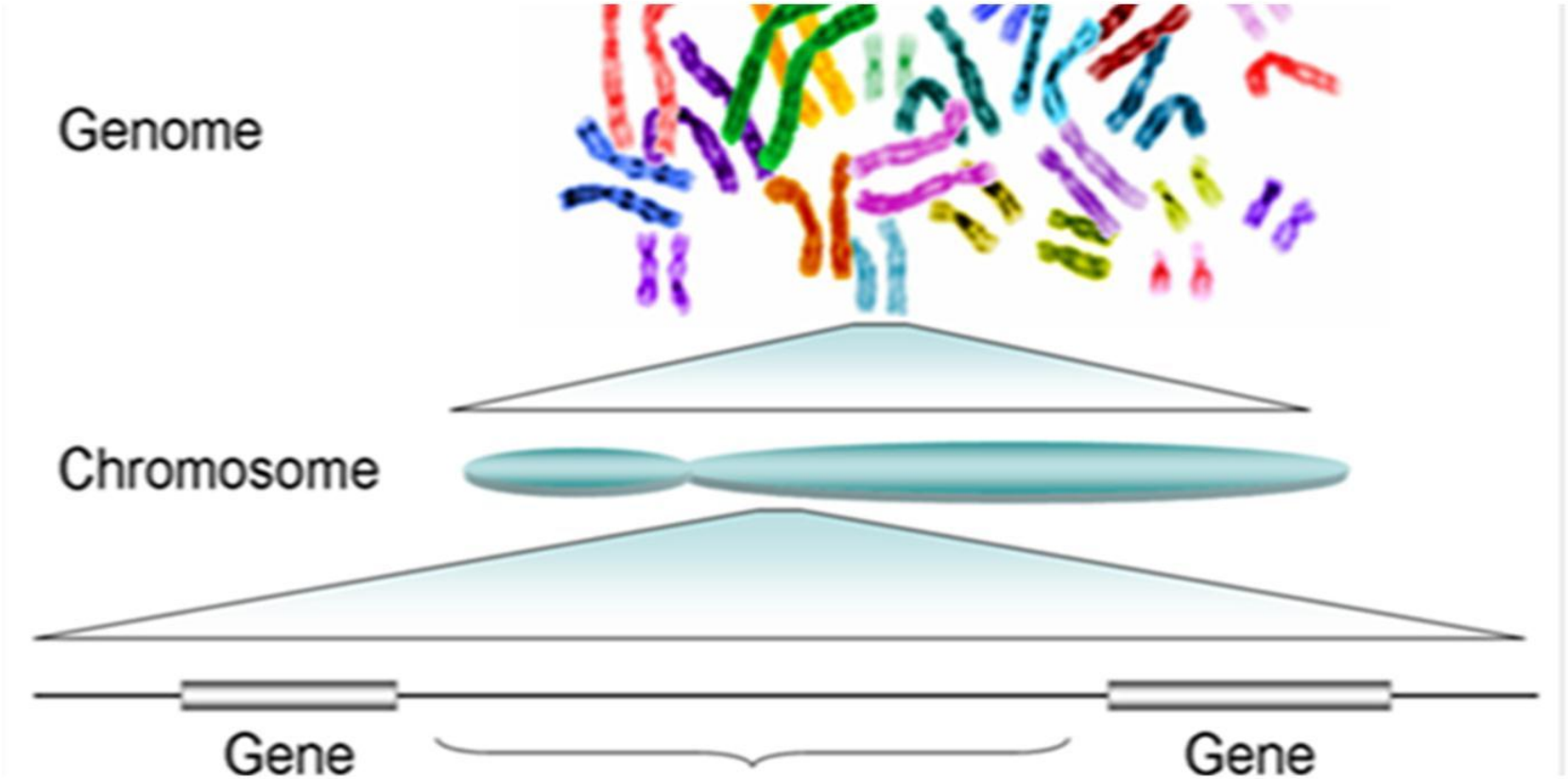
Eukaryotes are generally more advanced than prokaryotes





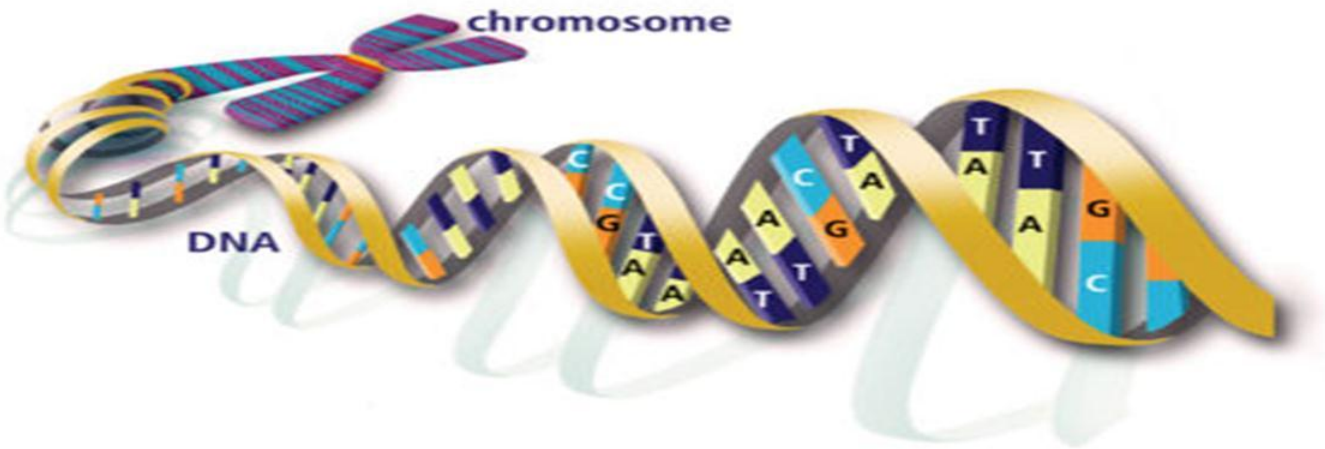

Human genomes
Nuclear genome 3.2 Gb
Mitochondrial genome 16.5 Kb
25%
75%
Genes and gene related DNA
Extragenic DNA
2 rRNA genes
22 tRNA genes
13 polypeptide encoding genes
10%
90%
60%
40%
Coding and regulatory DNA
Non-coding DNA
Unique low copy number DNA
Repetitive DNA
Pseudogenes
Gene fragments
Introns, untranslated sequences, etc.
Interspersed repeats
Tandem repeats

THE GENOME



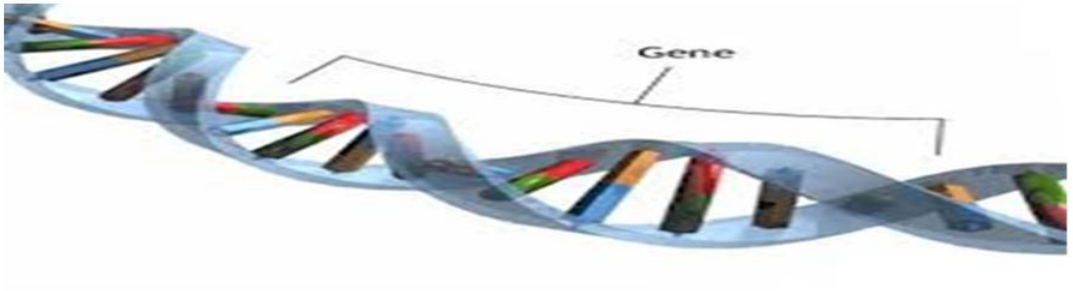
THE CHROMOSOME

- The storage place for all genetic information, the number of chromosomes varies from one species to another.



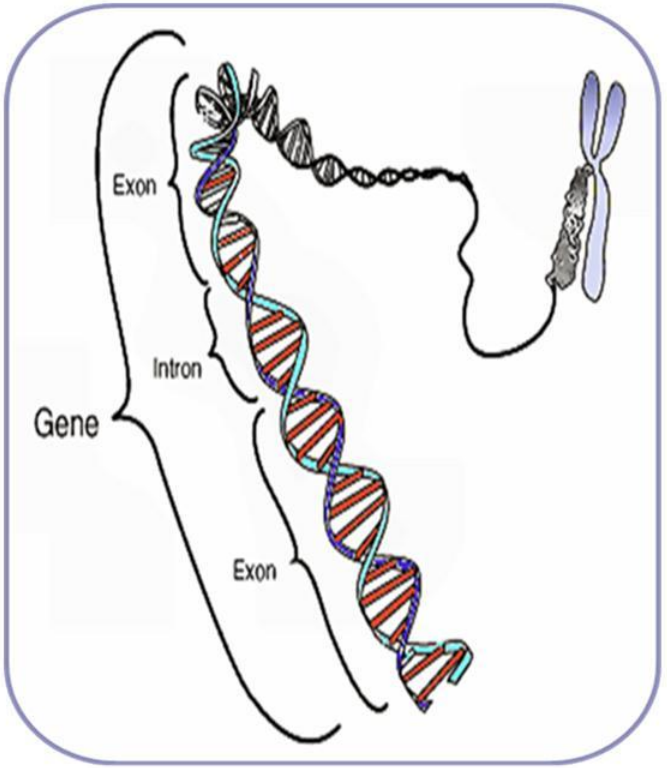
THE GENE

- The basic units of inheritance.
- It is a segment within a very long strand of DNA with specific instruction for the production of one specific protein.
- Located on chromosome on it's place or locus.
- **Allele;** A variant of the DNA sequence at a given locus. Each allele inherited from a different parent.



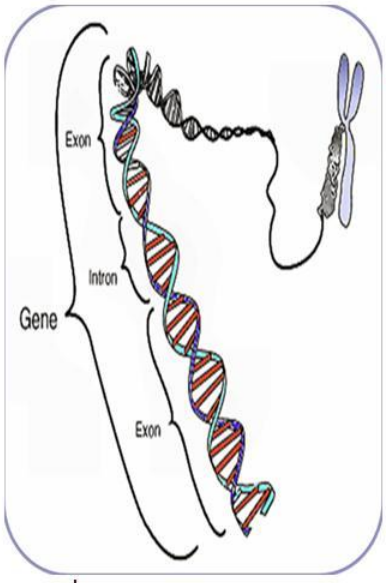
THE GENE

Most of the genes consist of; short coding sequences or exons are interrupted by a longer intervening noncoding sequence or introns; although a few genes in the human genome have no introns.



Source: National Human Genome Research Institute

THE GENE

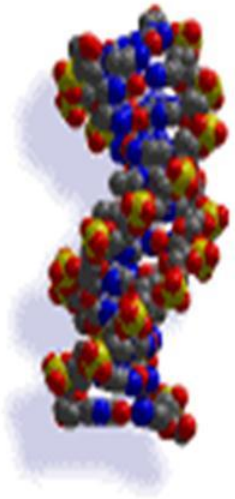


The actual number of genes contained in the human genome is not known. However, it has been estimated that the human genome contains about 30.000 essential genes.

THE GENE

The 30.000 or more protein encoding genes are scattered among 3 billion DNA base pairs 6×10^9 (diploid genome) among chromosomes.

DNA CONTENT



The molecule of DNA in a human chromosome ranges in size from 50×10^6 nucleotide pairs in the smallest chromosome up to 250×10^6 nucleotide pairs in the largest .

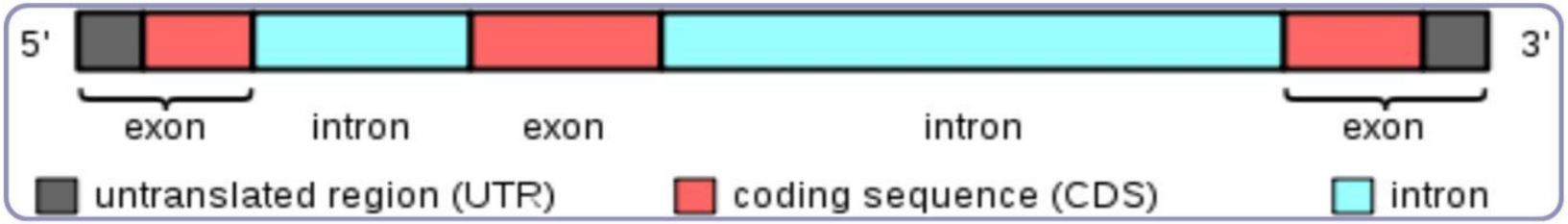
GENE STRUCTURE

Eukaryotic gene structure:

- The average gene 7-10 exons spread over 10-16kb of DNA.
- The gene must have (Exon; start signals; stop signals; regulatory control elements).

GENE STRUCTURE

The genes consist of; short coding sequences or exons are interrupted by a longer intervening noncoding sequence or introns; although a few genes in the human genome have no introns.



GENIC CODING DNA _(EXONS)

Coding DNA (Protein-coding sequences)

Coding DNA is defined as those sequences that can be transcribed into mRNA then translated into proteins.

NONCODING DNA

The exact amount and function of noncoding DNA that plays a role in cell physiology has been debated.

Noncoding DNA sequences within a genome that are not found within protein-coding exons.

Noncoding DNA are never represented within the amino acid sequence of expressed proteins

NONCODING DNA

Intron
untranslated regions of mRNA
Genes for noncoding RNA (e.g. tRNA and rRNA)
Regulatory DNA sequences
Pseudogenes
Repetitive DNA sequences
Sequences related to Transposable genetic elements

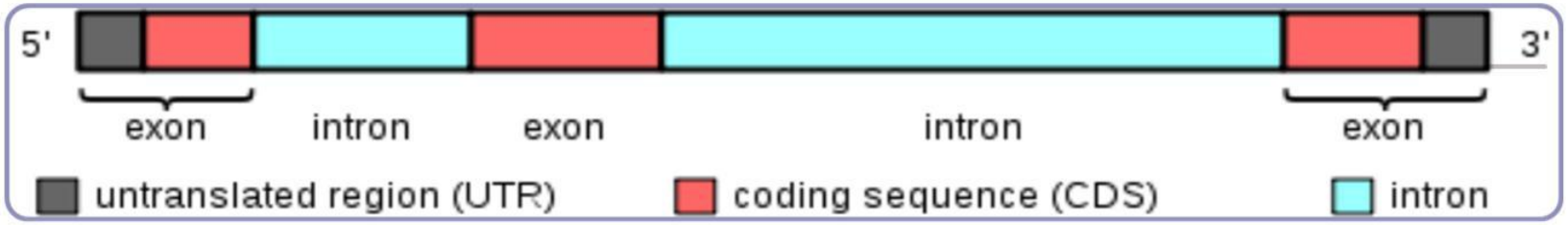
GENEIC NONCODING DNA

Intron

untranslated regions of mRNA

Genes for noncoding RNA (e.g. tRNA and rRNA)

Regulatory DNA sequences



Introns

Noncoding DNA sequences within genes (introns).

Noncoding DNA sequences within a gene that are found within protein-coding exons.

Noncoding DNA are never represented within the amino acid sequence of expressed proteins

GENES FOR NONCODING RNA

Genes for noncoding RNA include tRNA, ribosomalRNA, microRNA, snRNAand other non-coding RNA genes including about 60,000 long non coding RNAs (lncRNAs)

Genes for noncoding RNA molecules play many essential roles in cells, especially in the many reactions of proteinn synthesis and RNA processing.

Genes for noncoding RNA exact number in the human genome is yet to be defined, many of them are argued to be non-functional.

EXTRAGENEIC NONCODING DNA

- Pseudogenes
- Repetitive DNA sequences
- Sequences related to Transposable genetic elements
- Regulatory DNA sequences

PSEUDOGENES

Pseudogenes are inactive copies of protein-coding genes, often generated by gene duplication, that have become nonfunctional through the accumulation of inactivating mutations

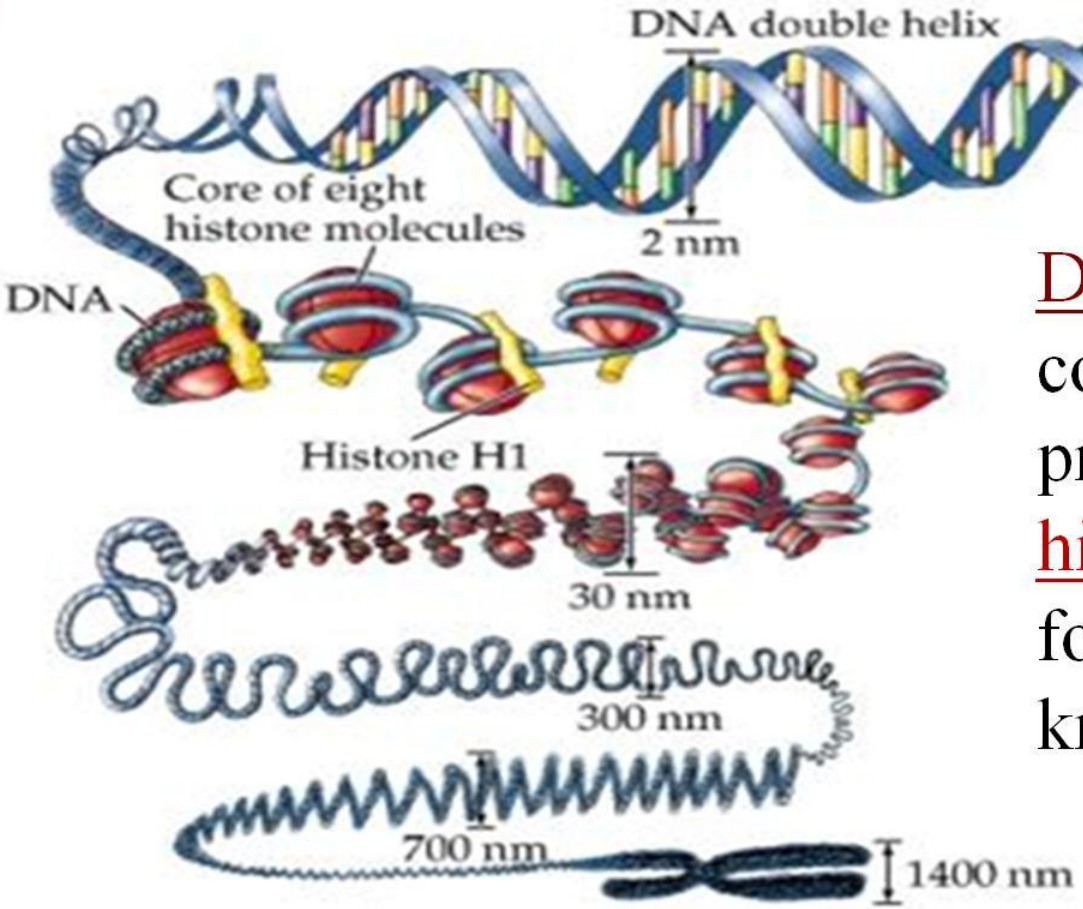
REPETITIVE DNA SEQUENCES

Repetitive DNA sequences comprise approximately 50% of the human genome

Major categories of **repeated sequence** or **repeats**:

- Tandem repeats: copies adjacent to each other, either directly or inverted
  - Satellite DNA
  - Minisatellite
  - Microsatellite
- Interspersed repeats (interspersed nuclear elements)
  - Transposable elements
    - SINEs (**S**hort **I**nterspersed **N**uclear **E**lements)
    - LINEs (**L**ong **I**nterspersed **N**uclear **E**lements)

DNA ORGNIZATION



DNA molecules complexed with other proteins, especially histones proteins to form a substance known as chromatin.

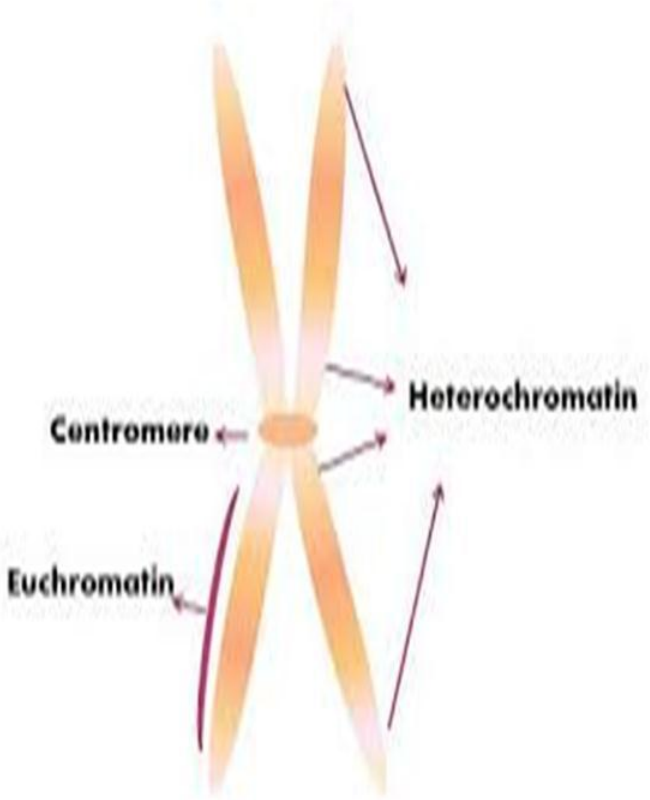
EUKARYOTIC CHROMATIN

Euchromatin

Less condensed region;
Transcriptionally active

Heterochromatin

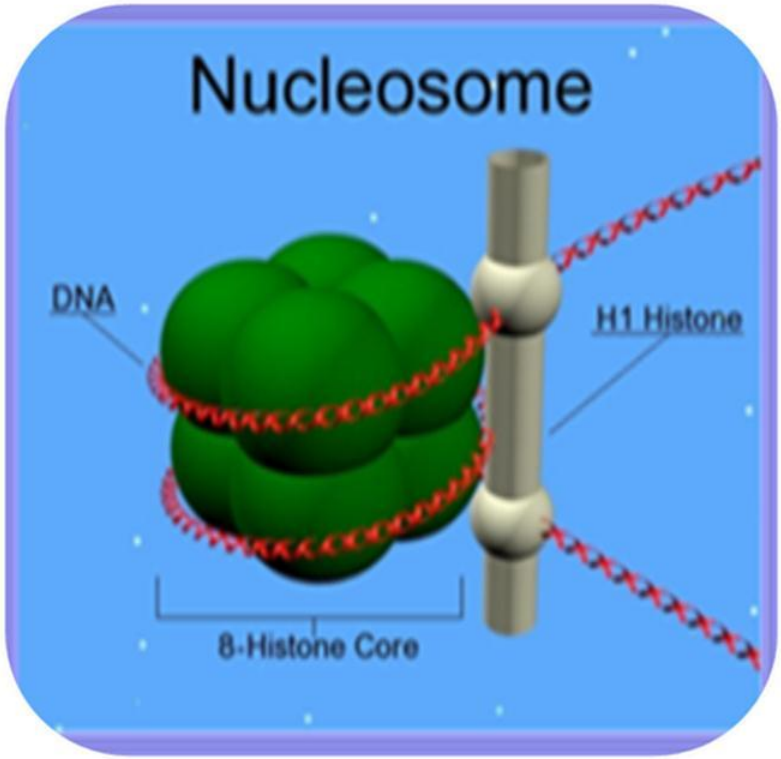
Densely condensed region
Transcriptionally inactive



Eukaryotic Chromatin Folding involves structures called **nucleosomes.**

Chromatin folding is the 30nm fiber

Chromatin condensation appears to involve the formation of a radial loop structure in eukaryotic chromosomes.

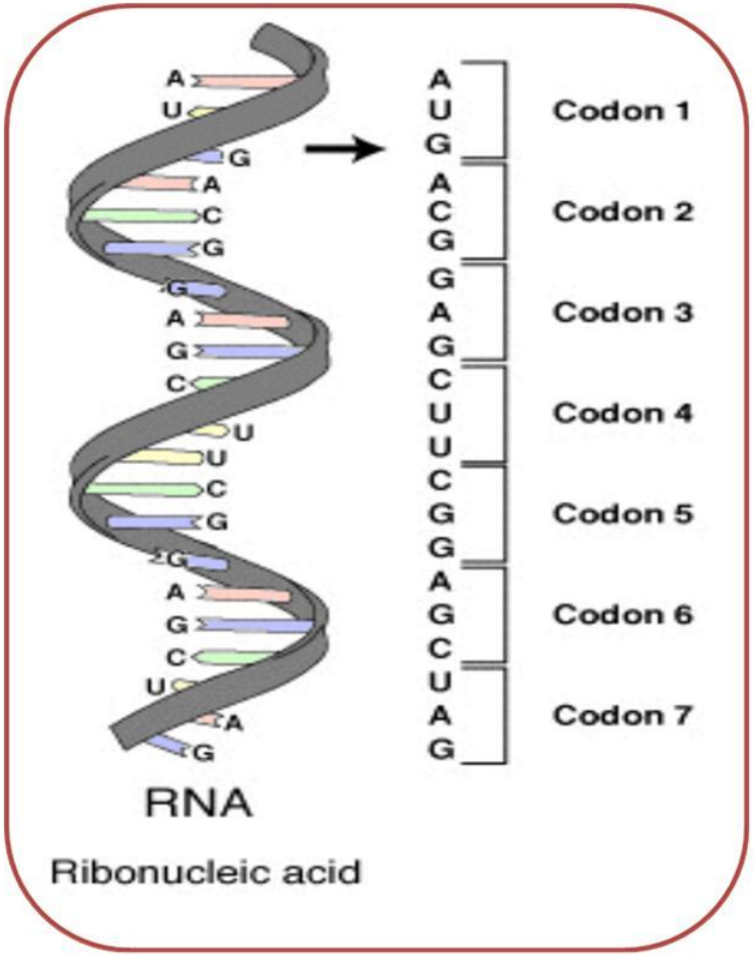


DNA ORGANIZATION

- The nucleus of eukaryotic cells contains DNA, histones, and non-histone proteins. These three elements combine to form the chromatin fibers of the nucleus.
- Four histones, H2A, H2B, H3 and H4, combine by twos to form the core particle of nucleosomes, the basic structural units of chromatin, which have a core of histones around which the DNA winds.
- The remaining histone, H1, ties DNA to the nucleosome and possibly also sets up linkages that wind nucleosome chains into chromatin fibers.

THE GENETIC CODE

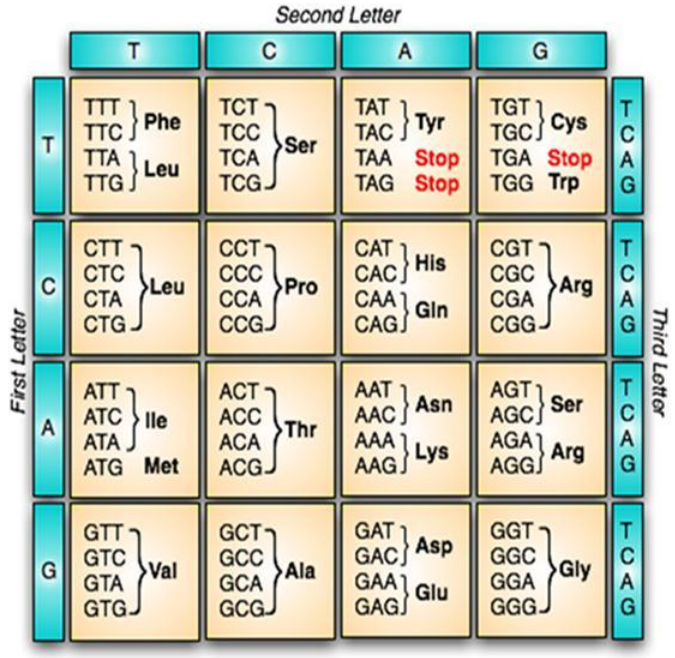
- The sequence of codons in the **mRNA** defines the primary structure of the protein.
- Three nucleotides in mRNA (codon) specify one amino acid in a protein.



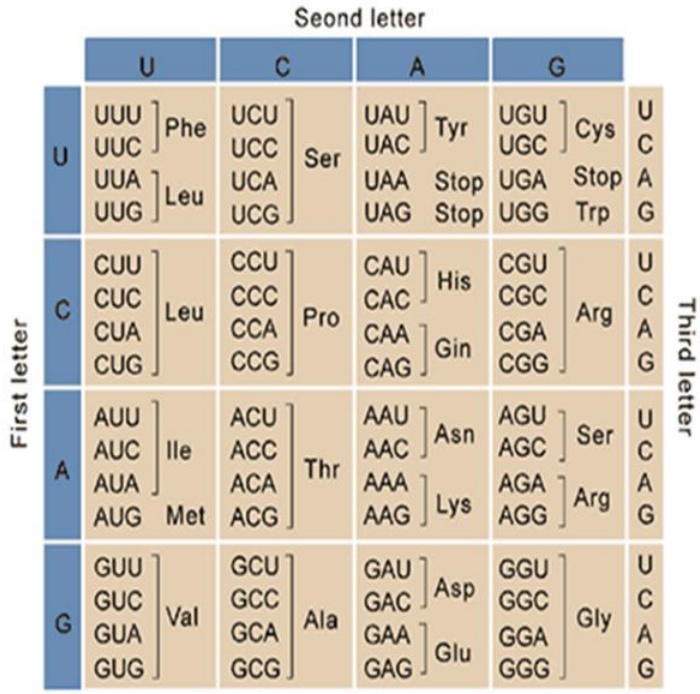
THE GENETIC CODE

- The purine and pyrimidine bases of the DNA molecule are the letters or alphabet of the genetic code.
- Series of codons in part of a mRNA molecule. Each codon consists of three nucleotides.
- 64 different combination of bases ; 61 of them code for 20 amino acid (AA); the last 3 codon (UAG,UGA,UAA) don not code for amino acids , they are termination codons.
- The sequence of codons in the mRNA defines the primary structure of the protein.
- Degenerate, specific, no gaps, non overlapping, almost universal,…

THE GENETIC CODE



DNA Codon



RNA Codon

The Mitochondria

Mitochondria is a membrane enclosed organelle found in most eukaryotic cells. These organelles range from 1-10 micrometers (μm) in size.

Mitochondria generates most of the cell's supply of adenosine triphosphate (ATP).

Mitochondria is involved in a range of other processes, such as signaling, cellular differentiation, cell death, as well as the control of the cell cycle and cell growth.

Mitochondria has been implicated in several human diseases, including mental disorders, cardiac dysfunction, and may play a role in the aging process.

Mitochondria has its own DNA.

The mitochondrial DNA is only 16 kb in length (less than 0.03 % of the length of the smallest nuclear chromosome) and encodes only a few dozen genes. Although the products of these genes function in mitochondria, the majority of proteins found in mitochondria, product of nuclear genes.

.

MITOCHONDRIAL DNA

Human cells have hundreds of mitochondria, each containing a number of copies of a small circular molecule, mitochondrial chromosome (Mitochondrial DNA).

MITOCHONDRIAL DNA

Mitochondrial DNA contains 37 genes, essential for normal mitochondrial function

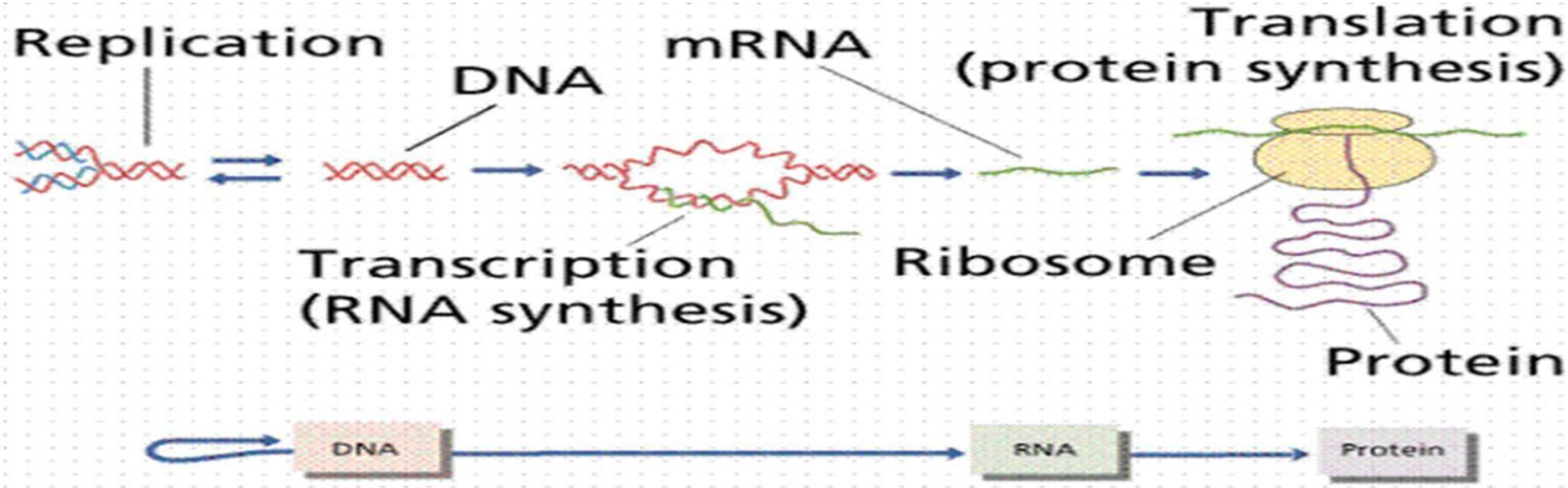
Genes provide instructions for making enzymes involved in oxidative phosphorylation and genes provide instructions for making molecules called transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs).

Although the products of these genes function in mitochondria, the majority of proteins found in mitochondria, product of nuclear genes.

CENTRAL DOGMA OF MOLECULAR BIOLOGY

DNA molecules serve as templates for either complementary DNA strands during the process of replication or complementary RNA during the process of transcription.

RNA molecules serve as a template for ordering amino acids by ribosomes during protein synthesis.



RNA

Three major classes of RNA:

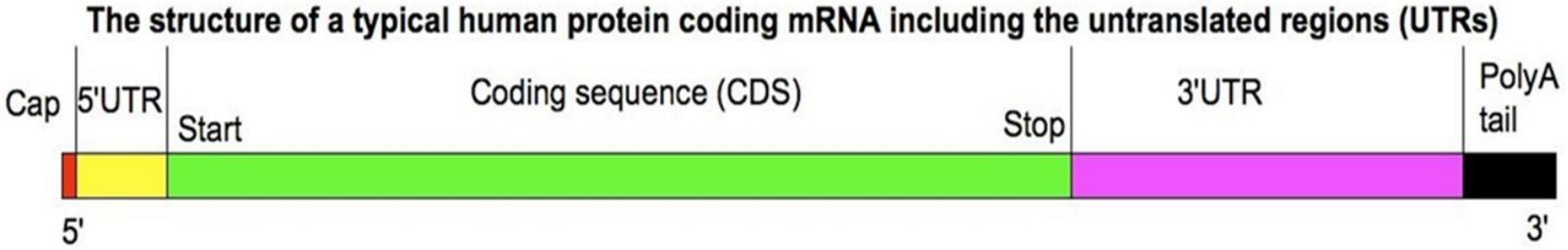
- Messenger (mRNA),
- Transfer (tRNA) and
- Ribosomal (rRNA)

Minor classes of RNA include:

- Small nuclear RNA,
- Small nucleolar RNA,
- ……

mRNA

- Transcripts of structural genes that encode all the information necessary for the synthesis of a polypeptide/s of protein.
- Intermediate carrier of genetic information; deliver genetic information to the cytoplasm where protein synthesis takes place.
- The 5' terminus is capped by 7 methyguanosine triphosphate.
- Synthesis of the poly (A) tail involves cleavage of its 3' end and then the addition of about 200 adenine residues.



tRNA and rRNA

tRNA

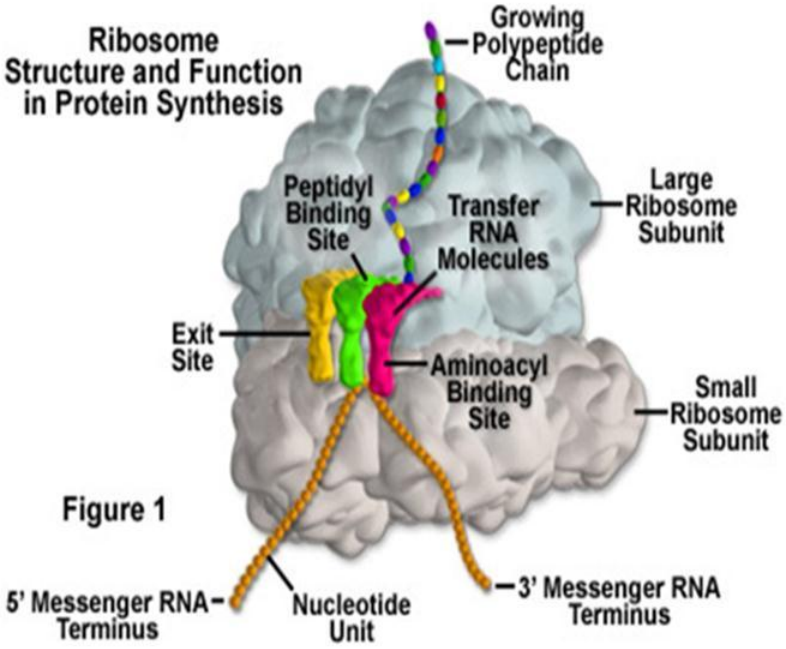
- Small molecules consisting approx. 70-80 nucleotides.
- All tRNAs share a common secondary structure represented bya coverleaf.
- They have four- base paired stems defining three stem loops(the D loop, anticodon loop, and T loop) and the acceptor stem to which amino acids are added in the charging step.
- tRNA molecules that carry amino acids to the growing polypeptide.

rRNA

- The RNA component of the ribosome.
- Provide a mechanism for decoding mRNA into amino acids and interacts with tRNAs during translation by providing peptidyl transferase activity

RIBOSOMES

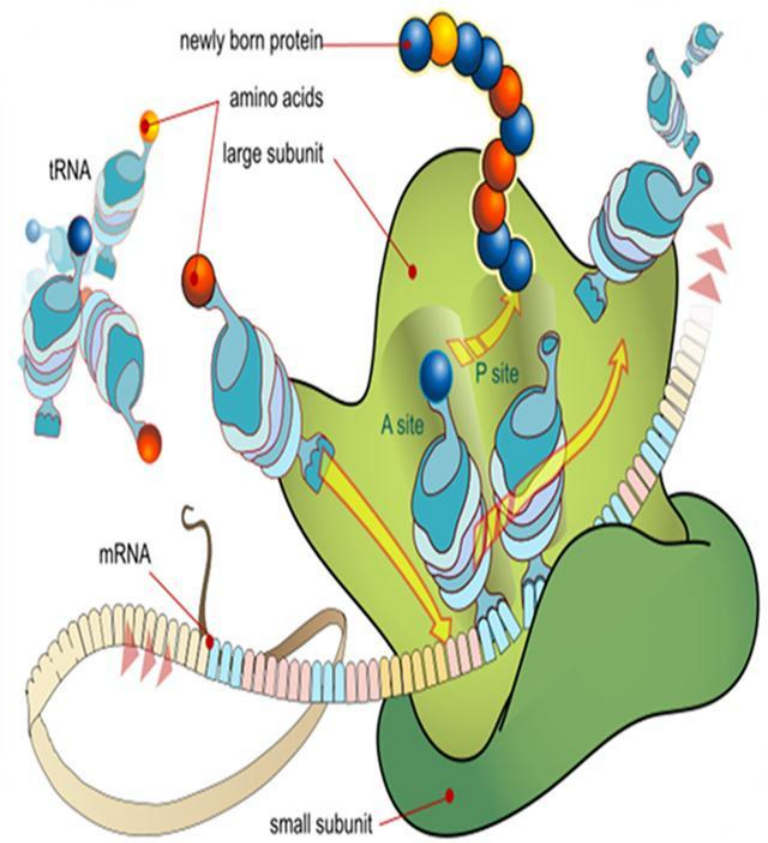
- Factory for protein synthesis.
- Composed of ribosomal RNA and ribosomal proteins; known as a Ribonucleoprotein (RNP).
- Translate messenger RNA (mRNA) to build polypeptide chains using amino acids delivered by transfer RNA (tRNA).



RIBOSOMES

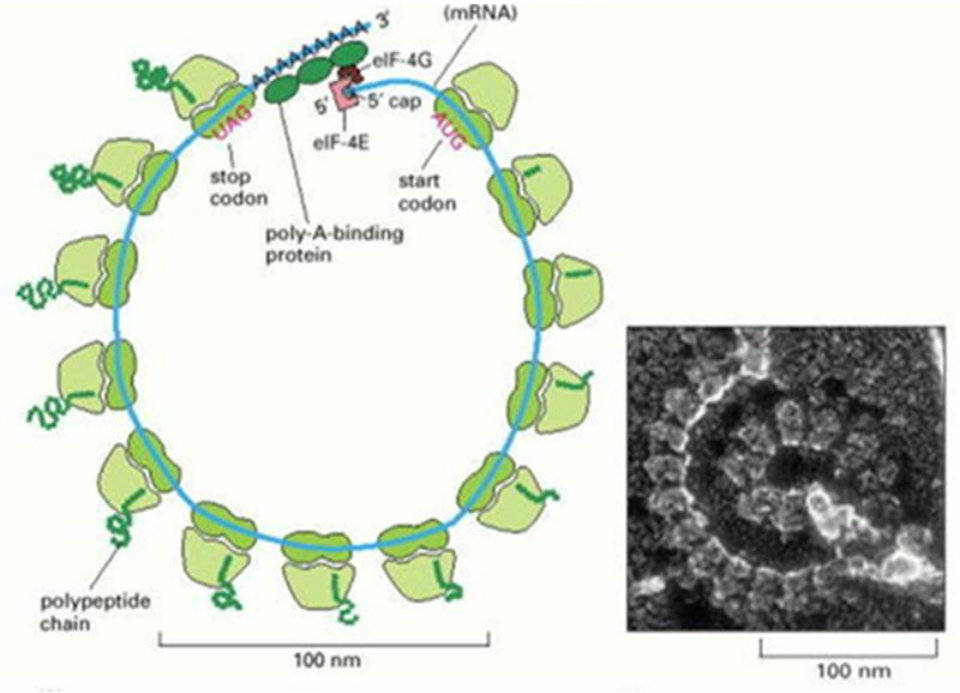
Eukaryotic ribosomes are larger. They consist of two subunits which come together to form an 80S particle

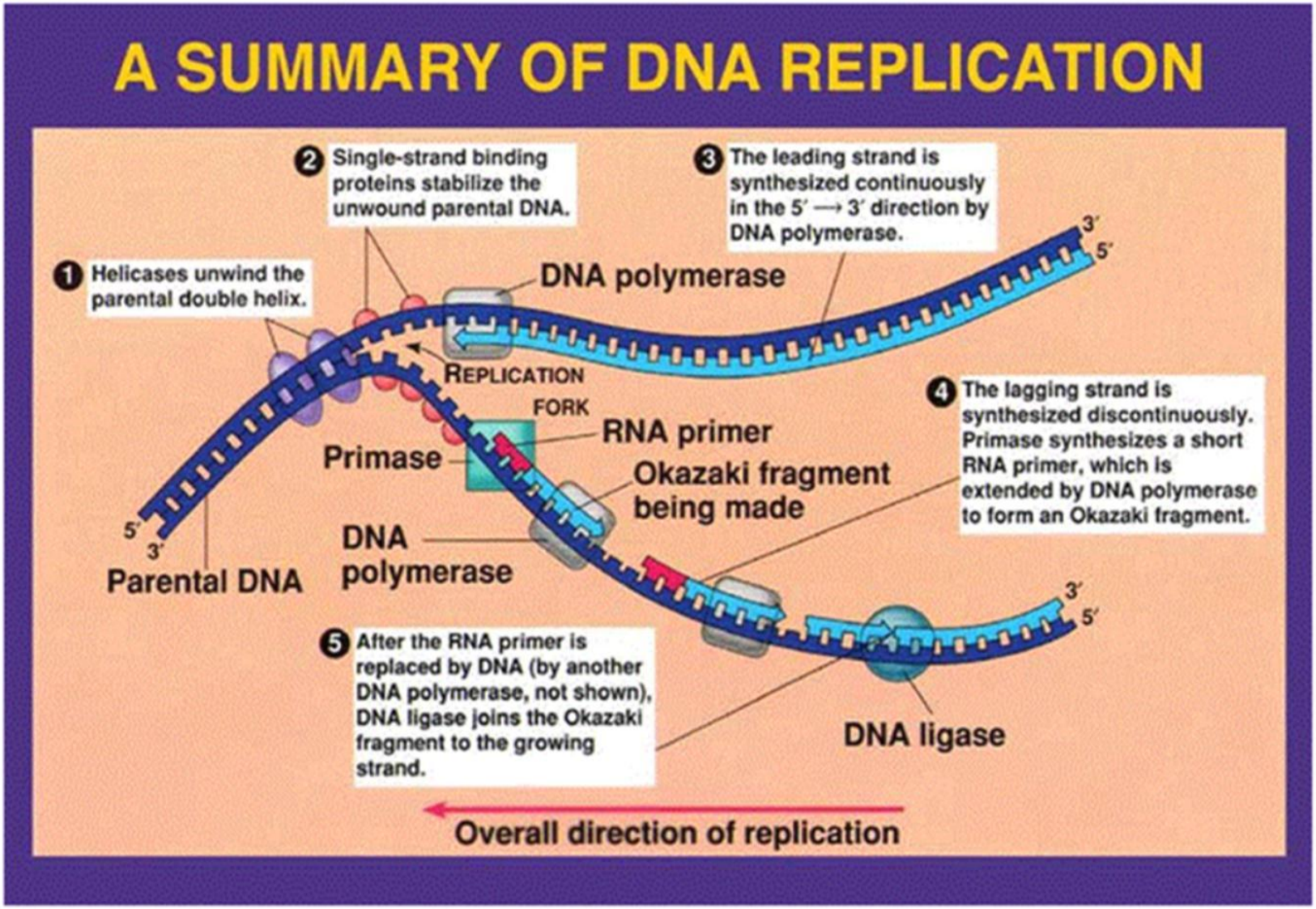
- 60S subunit holds (three rRNAs 5S, 5.8S, 28S and about 40 proteins).
- 40S subunit contains (an18S rRNA and about 30 proteins).



POLYSOMES

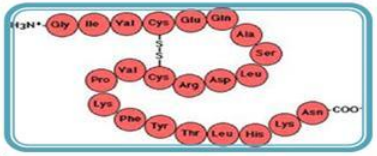
Most mRNA are translated by more than one ribosome at a time; the result, a structure in which many ribosomes translate a mRNA in tandem, is called a **polysomes.**



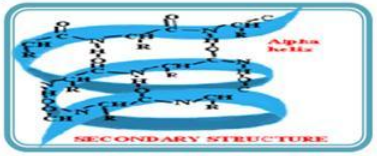


Source: http://serc.carleton.edu/microbelife/research_methods/genomics/replication.html

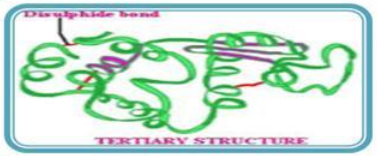
PROTEIN STRUCTURE



Primary structure: Formed by joining the amino acid sequence into a polypeptide.



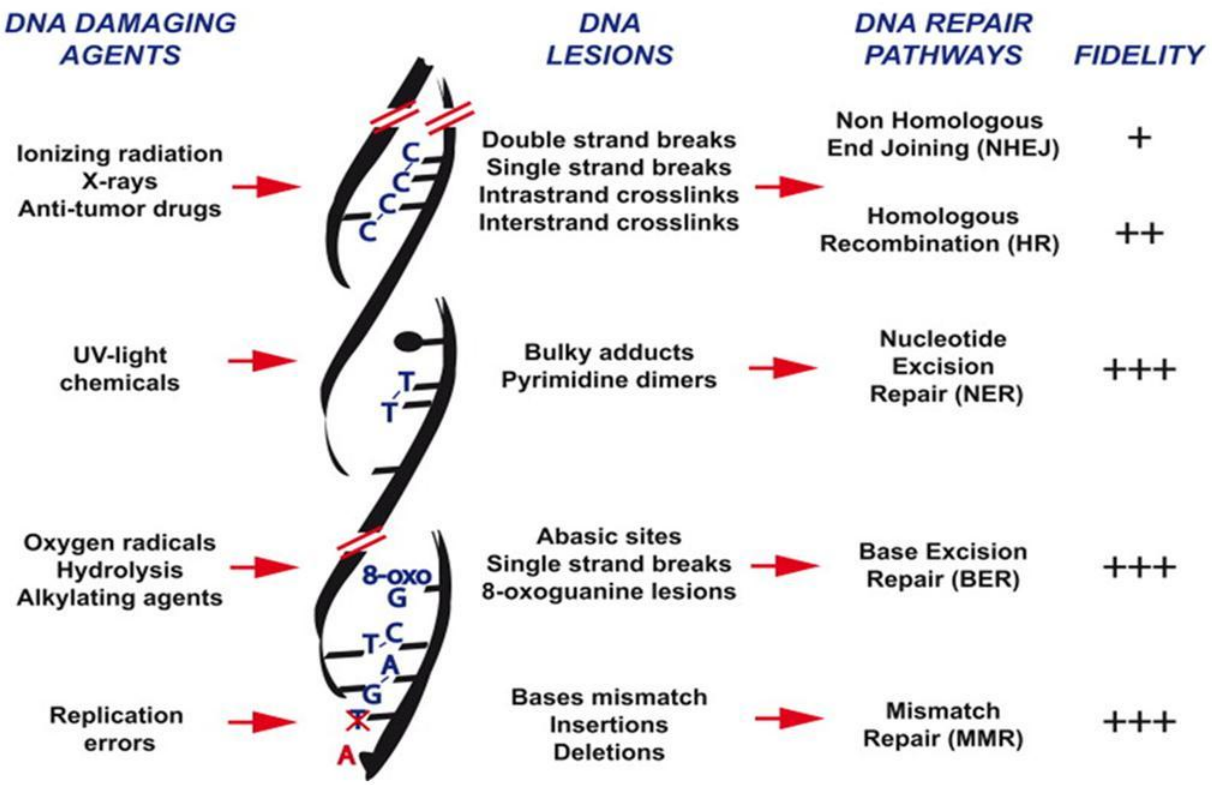
Secondary structure: Different conformation that can be taken by the polypeptide: alpha helix and strands of beta sheet.



Tertiary structure: Result from folding the secondary structure components of the polypeptide into three-dimensional configuration.

Quaternary structure: complex of several protein molecules or polypeptide chains, usually called protein subunits, which function as part of the larger assembly or protein complex.

DNA DAMAGE/REPAIR



Common Tools and Techniques in Molecular Biology

- Nucleic acid fractionation
- Polymerase chain reaction
- Probes, Hybridization
- Vector, Molecular cloning
- Nucleic acid enzymes
- Microarray
- DNA sequencing
- Electrophoretic separation of nucleic acid
- Blotting techniques
- DNA: Southern blotting
- RNA: Northern blotting
- Protein: Western blotting
- Molecular cytogenetic analysis
- …..

Functional Genomics and Other Technologies

Genomics: Involves the sequencing of the complete genome, including structural gene, regulatory sequences and noncoding DNA segments in the chromosome of an organism and the interpretation of all the structural and functional implications of these sequences and of many transcripts and proteins the genome encode. Genomic information offers new therapies and diagnostic methods for treatment of many diseases.

Transcriptomics: is the systematic and quantitative analysis of all the transcript present in a cell or a tissue under a defined set of conditions (the transcriptome).

Proteomics: is the total set of proteins expressed from the genome of cell via transcriptome. It is the quantitative study of the proteome.

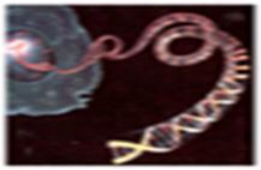
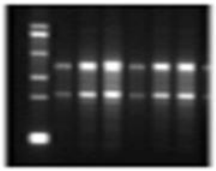
Metabolomics: is the study of all the small molecules, including metabolic intermediates (amino acids, nucleotides ,sugars,…) that exist within a cell. The metabolome provides a sensitive indicator of the physiological status of a cell and has potential uses in monitoring disease and its management.

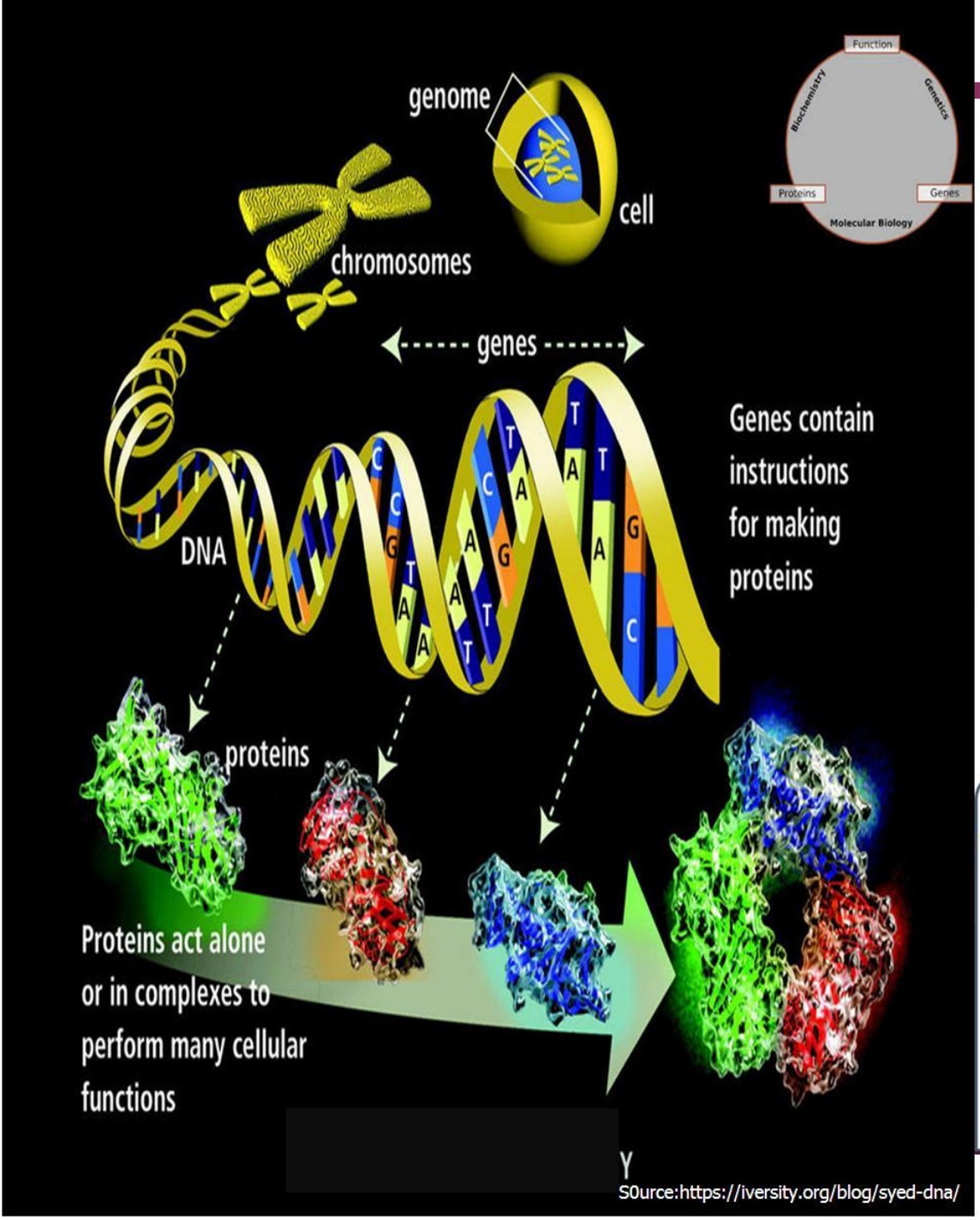
from genotype to phenotype

DNA => pre-mRNA => mRNA => => => => protein

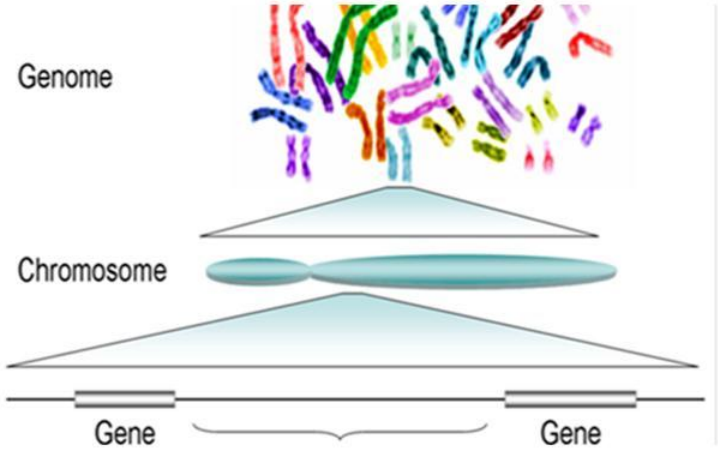
↓ transcription ↓ splicing ↓ translation ↓ post-translational modifications

genome transcriptome proteome

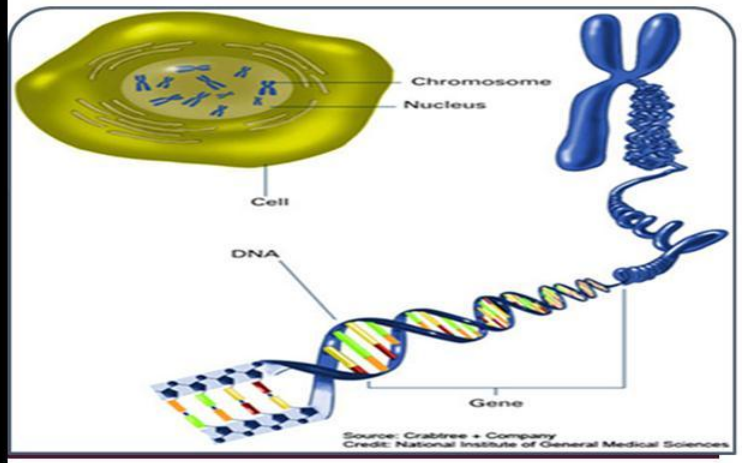





S0urce:https://iversity.org/blog/syed-dna/



Source:http://www.pgpstudy.org/genmaterial/genmatles2/celldnagenesreadpg7.htm



THANK YOU