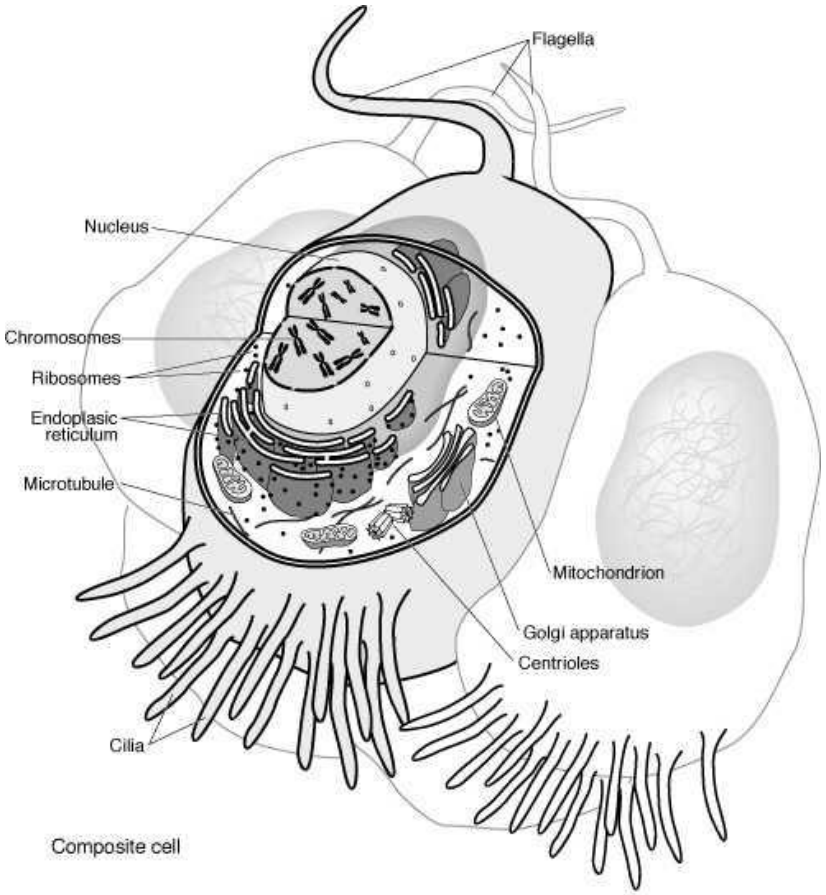
1. Biology in a nutshell

Cells and organisms

- Organisms
  - Unicellular: One single cell (or simpler)
    - Archaea
    - Prokaryotes: Bacteria, Yeast
    - Eukaryote: Protozoos
  - Pluricellular: Eukaryote cells
    - Different organitation levels (e.g. tissues)
    - Diversity in number, type and syze
  - Viruses are not properly organisms

Eukaryotes



- "Eu" good,
  "Karyo" nut or kernel →
- Presence of nucleous separated from del cytoplasm by the nucleous envelope.
- DNA: double-stranded, it is organized in chromosomes
- The cell contains other membrane-bound organelles
- Sexual reproduction is common

Key biomolecules

- The basic components of biological systems are
  - Sugars (carbohydrates)
  - Fats (lipids)
  - Nucleic acids
  - Proteins
- Sugars and lipids have no important role regarding "biological information" 💻
  Not discussed here

Nucleic acids and proteins

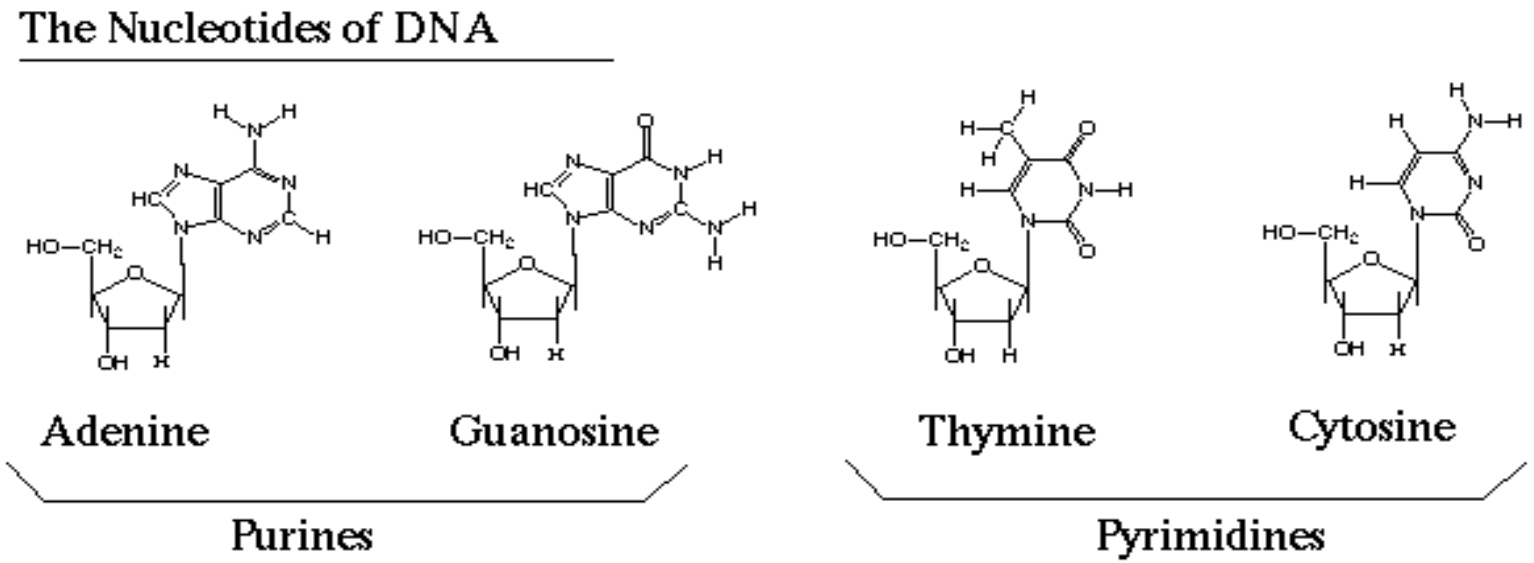
- Molecules that contain and transport information
  - DNA (4 different nucleotides)
    - Contains encoded biological information
  - RNA (4 different nucleotides)
    - Carries information from DNA to proteins
  - Proteins (20 different amino acids)
    - Function and structure of living beings

Nucleotides

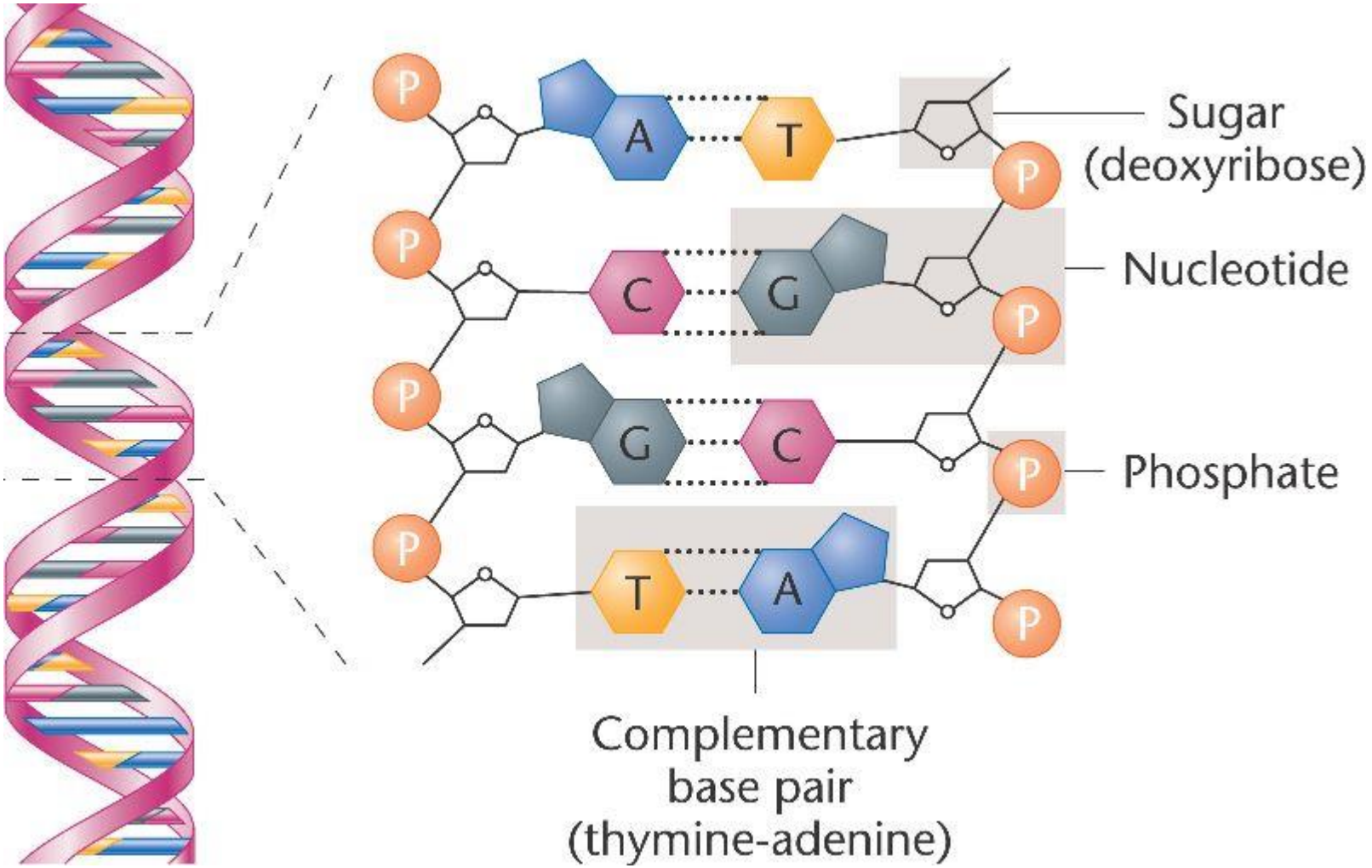
- Basic components of nucleic acids
- Consisting of
  - A sugar (Ribose or Deoxyribose)
  - A nitrogen base
  - A phosphate group
- In biological parlance we speak of "bases" instead of nucleotides to describe a string (3000bp = 3000 bp)

DNA nucleotides

Deoxyribonucleic Acid (DNA) contains four nucleotide bases.



The pairing between complementary bases



The primary structure of DNA

- Sequence of nuceotides
- Forms an unbranched polymer
- Organized in a double-stranded

```
atgaatcgta ggggtttgaa cgctggcaat
acgatgactt ctcaagcgaa cattgacgac
ggcagctgga aggcggtctc cgagggcgga
```

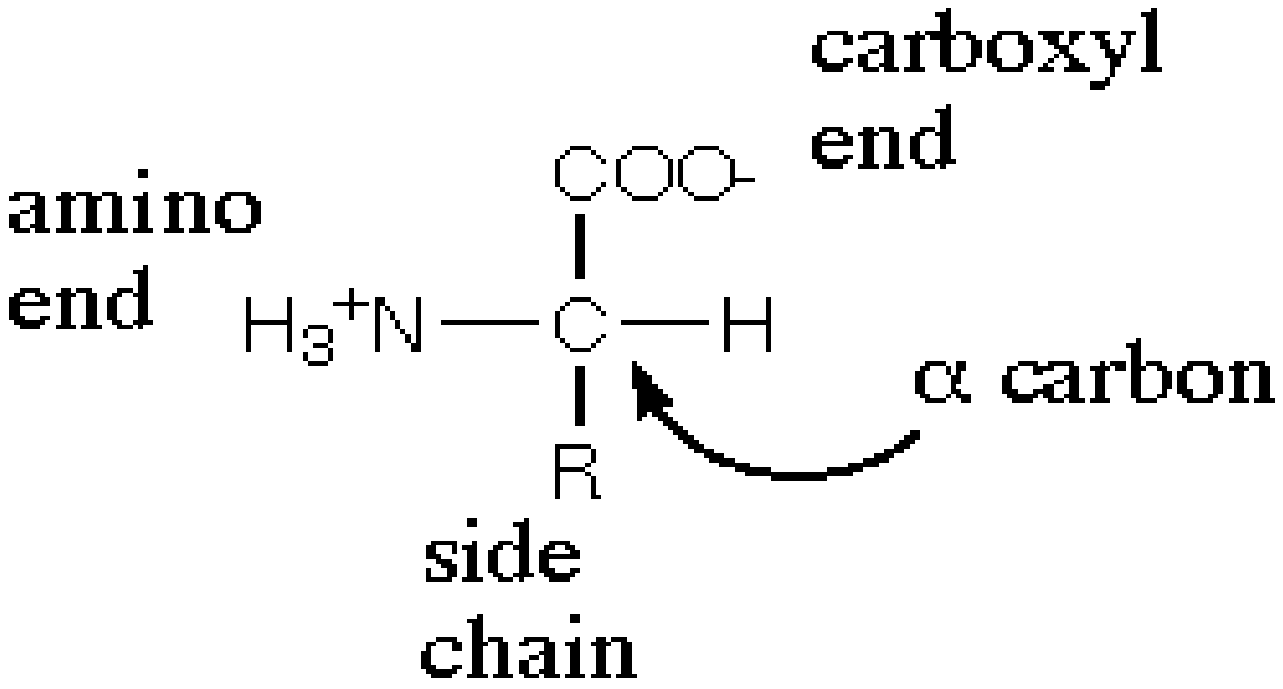
DNA vs RNA

- DNA is organized into a complementary double helix. RNA does not.
- One of the four bases are different
  - DNA💻 A, C, G, **T**
  - RNA: A, C, G, **U**
- Differ from the nucleotide sugar
  - DNA: Deoxyribose
  - ARN: Ribose

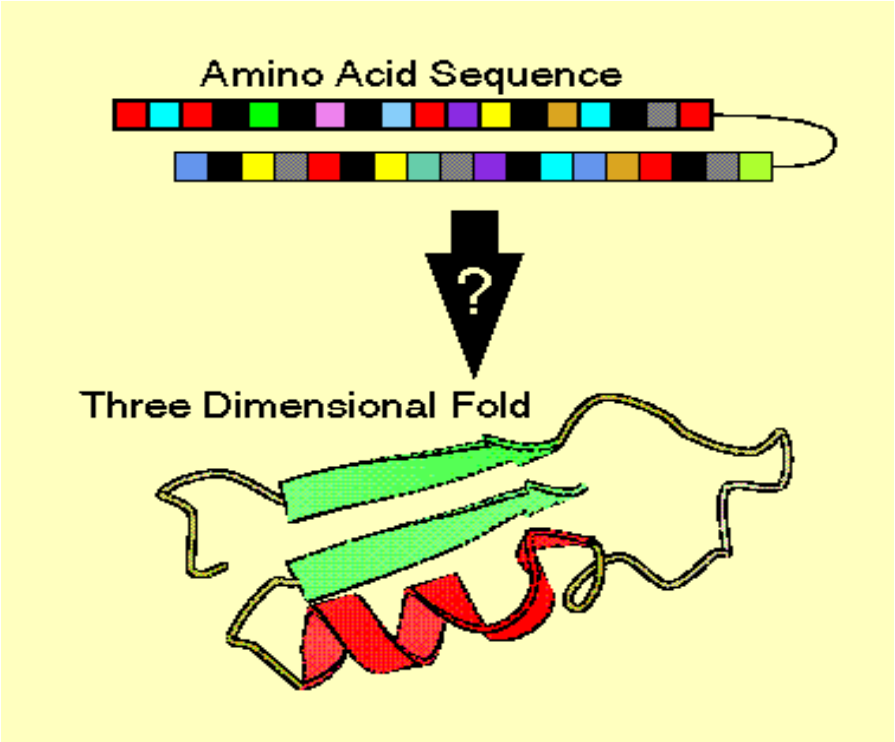
Proteins

- Amino acid sequence
- Forms an unbranched polymer
- There are 20 different amino acids (AA)
- The key function of proteins is in its three dimensional structure

Amino acids



Proteins “fold” into conformational structure

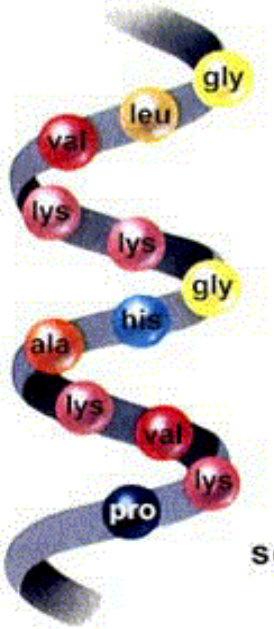


Protein's structure

Proteins fold into lowest to highest complexity level:

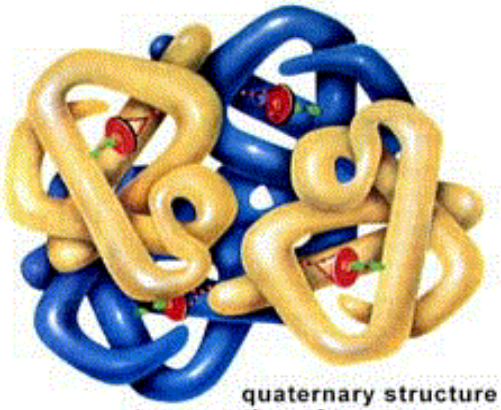
- Primary structure: amino acid sequence
- Secondary structure: regularly repeating local structures stabilized by hydrogen bonds.
- Tertiary structure: fold into 3-dimensional structures.
- Quaternary structure: structure formed by several protein molecules (protein complex).

primary structure
(amino acid sequence)



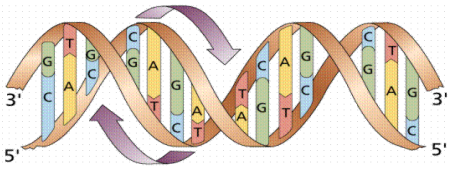
secondary structure
(α-helix)

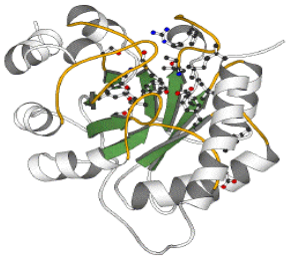
tertiary structure
(folded individual peptide)



quaternary structure
(aggregation of two or more peptides)

The central dogma of molecular biology





DNA

transcription

mRNA

translation

Protein

CCTGAGCCAACTATTGATGAA

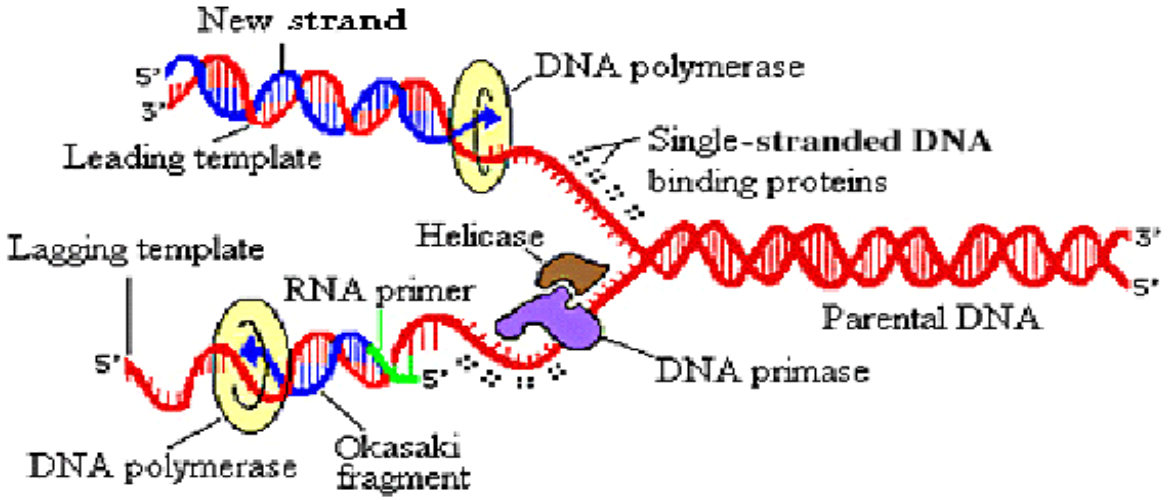
CCUGAGCCAACUAUUGAUGAA

PEPTIDE

From DNA to proteins

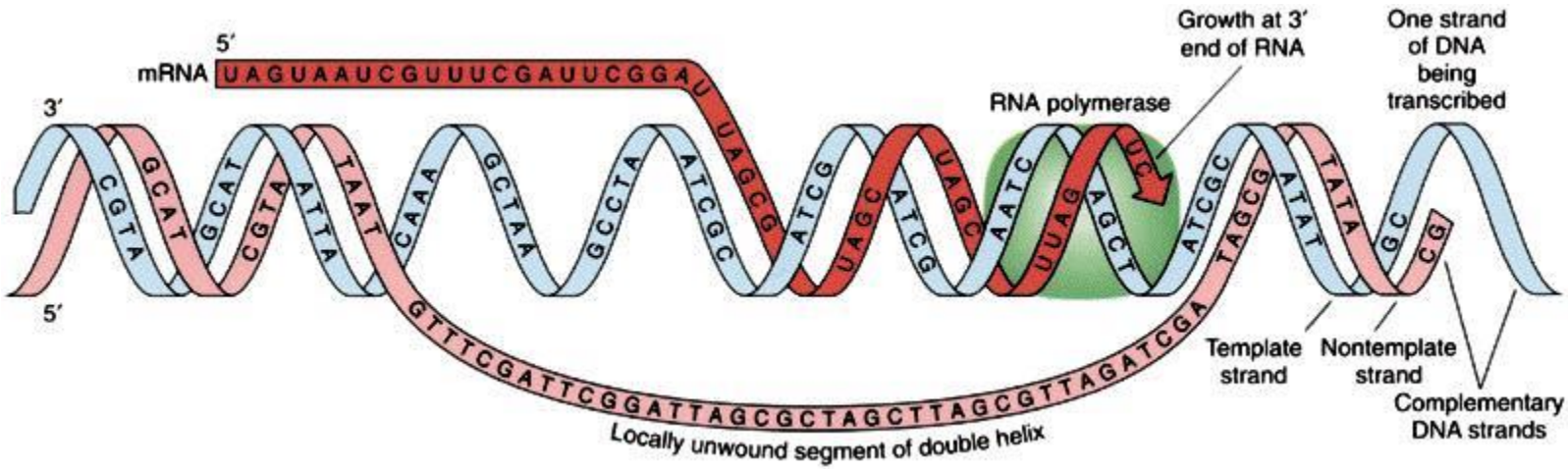
- DNA is replicated in a complex process involving many enzymes: **replication**
- DNA is copied in a string of complementary messenger RNA (mRNA): transcription **transcription**
- In eukaryotic cells, the mRNA is processed **splicing** eliminating coding fragments (“splicing”) and migrates from the nucleous to the cytoplasm.
- The mRNA carries coded information to ribosomes (ribosomal RNA) that "read" and perform protein synthesis: **translation**

1. Replication

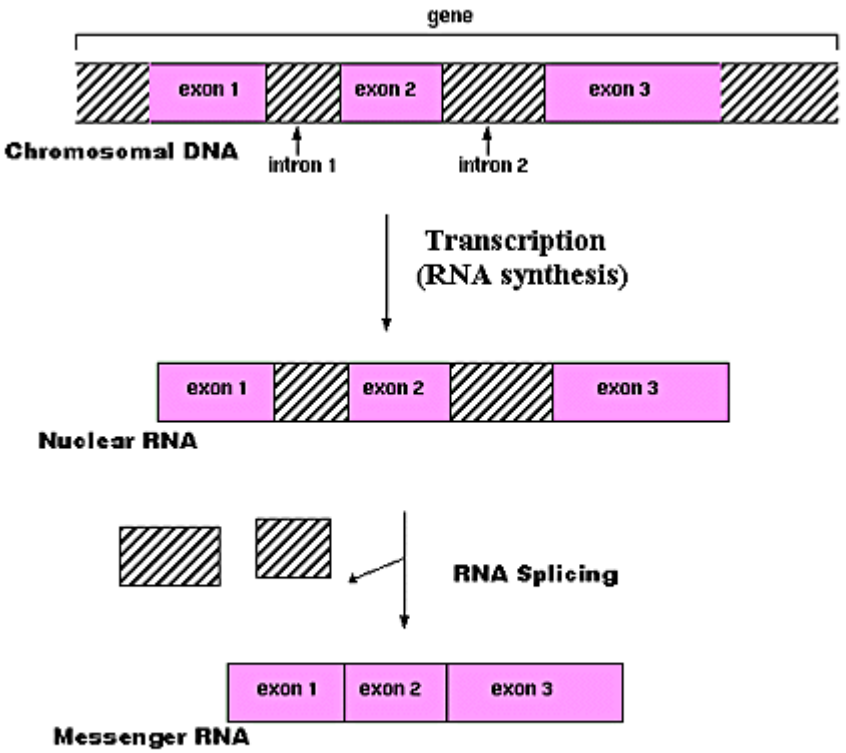


- Animations (1), (2), (3)

2. Transcription

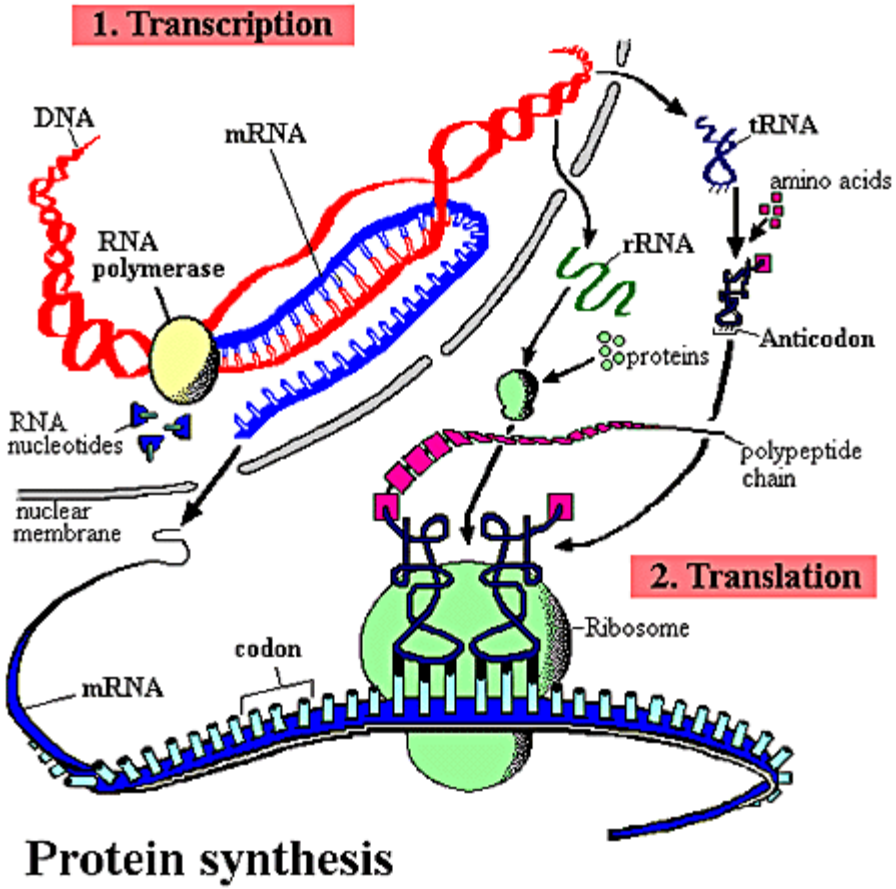


3. RNA processing or splicing



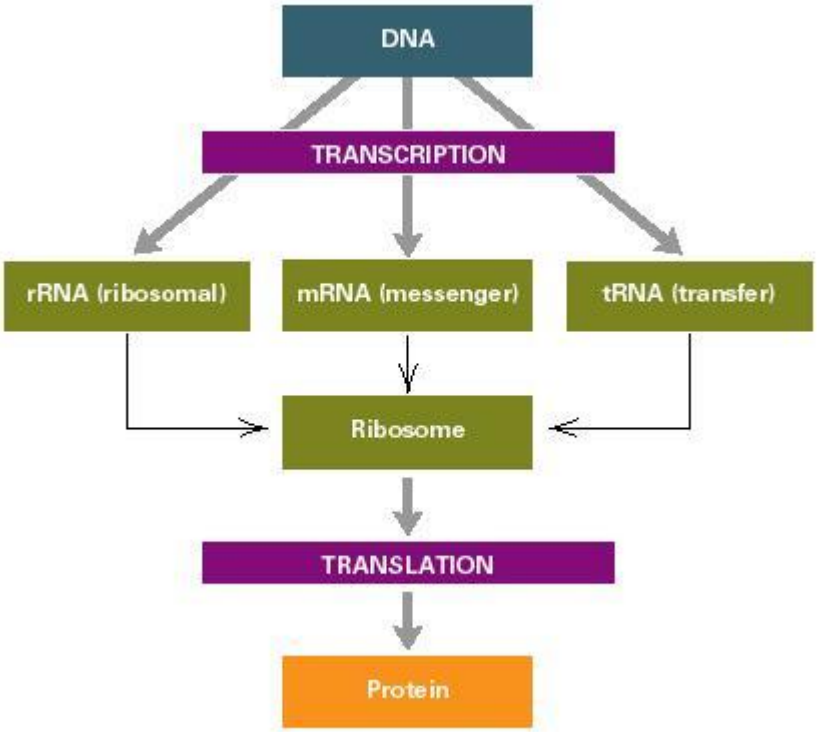
RNA synthesis and processing

4. Protein synthesis

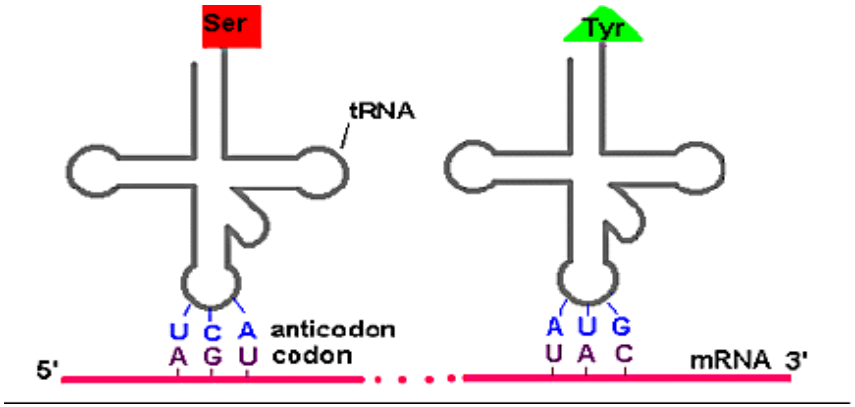


- The mRNA goes to the cytoplasm where it binds to ribosomes.
- Codon: mRNA information unit.
- The tRNA brings the complementary AA tRNA.
- The AA are bound to the protein to complete the sequence.
- Animations (1), (2)

Notice: 1 DNA vs. 3 RNA's !



The genetic code



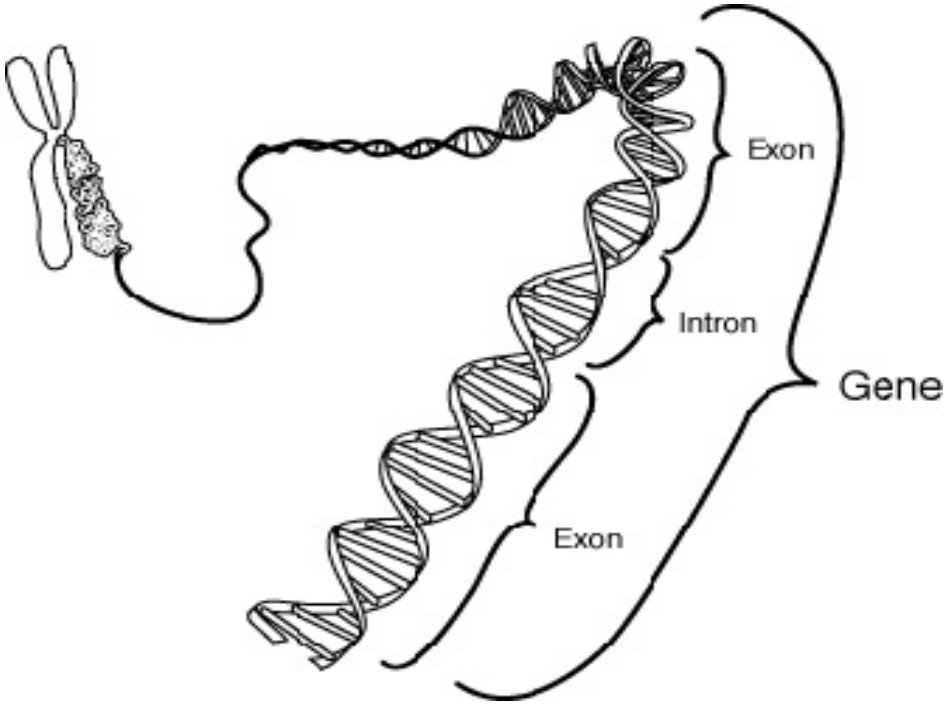
2nd base in codon

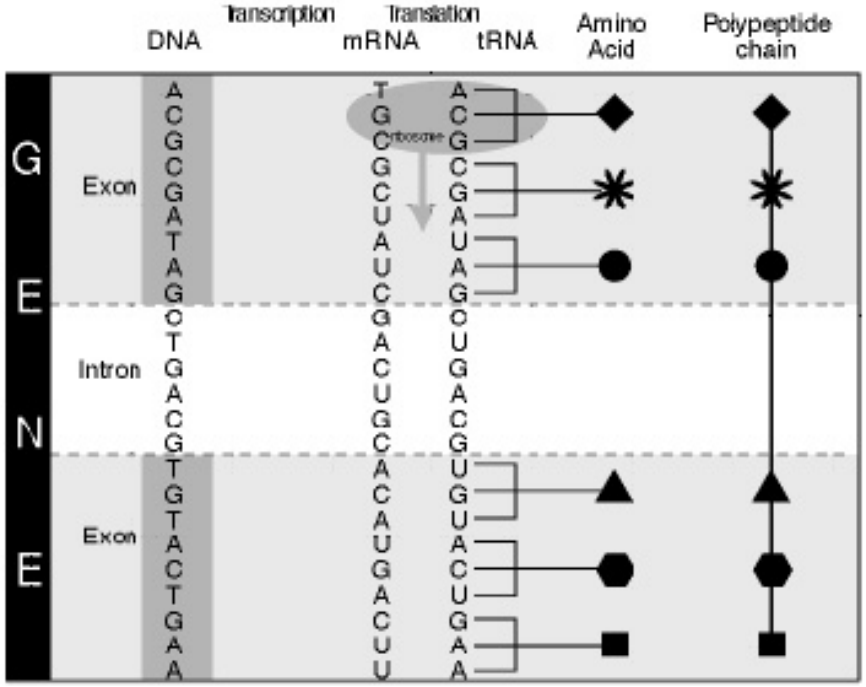
1st base in codon	U	C	A	G	3rd base in codon
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

What is a gene? (pre ENCODE)

- A gene used to be defined as ...
  - The functional and physical unit of heredity,
  - transmitted from one generation to their offsprings,
  - consisting of DNA fragments,
  - with the information needed for the synthesis of a specific protein (most of them...)

Gene components





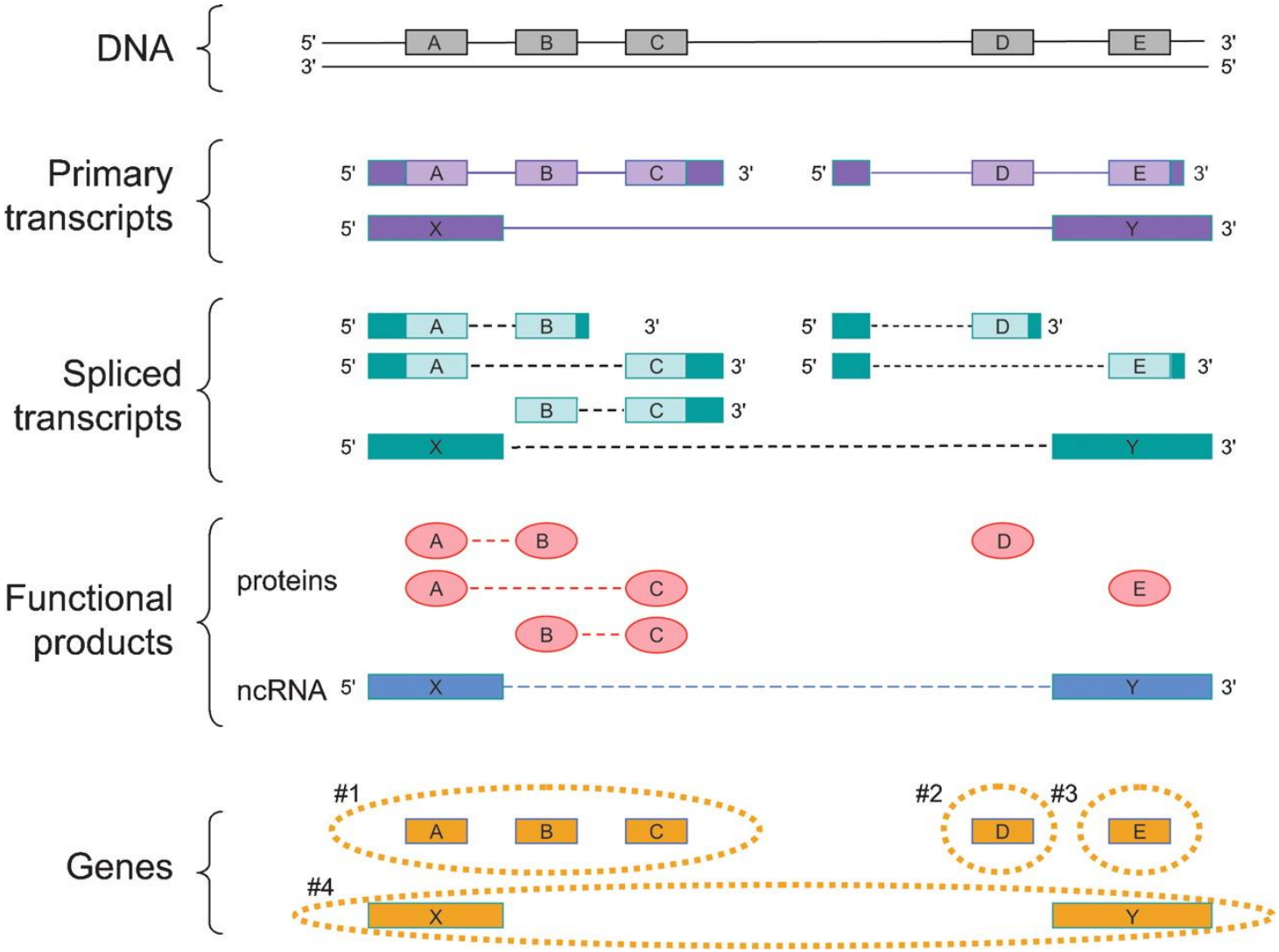
The ENCODE project

In September 2003 The National Human Genome Research Institute (NHGRI) launched a public research consortium named ENCODE, the Encyclopedia Of DNA Elements

Its goal was "*to carry out a project to identify all functional elements in the human genome sequence*"

The findings of the ENCODE project have led to re-consider foundamental ideas such as the definition of gene

What is a gene (post-ENCODE)



Gene expression

- Genes can be “turned on” or “off”
- When a gene is turned on we say it is *expressing* or *being expressed*
  - *The central dogma implies that when a gene is expressed it is, at least transcribed*
  - *That is: expression → mRNA*
- Each cell expresses only a fraction of its genes
- Remaining genes are repressed ("off")

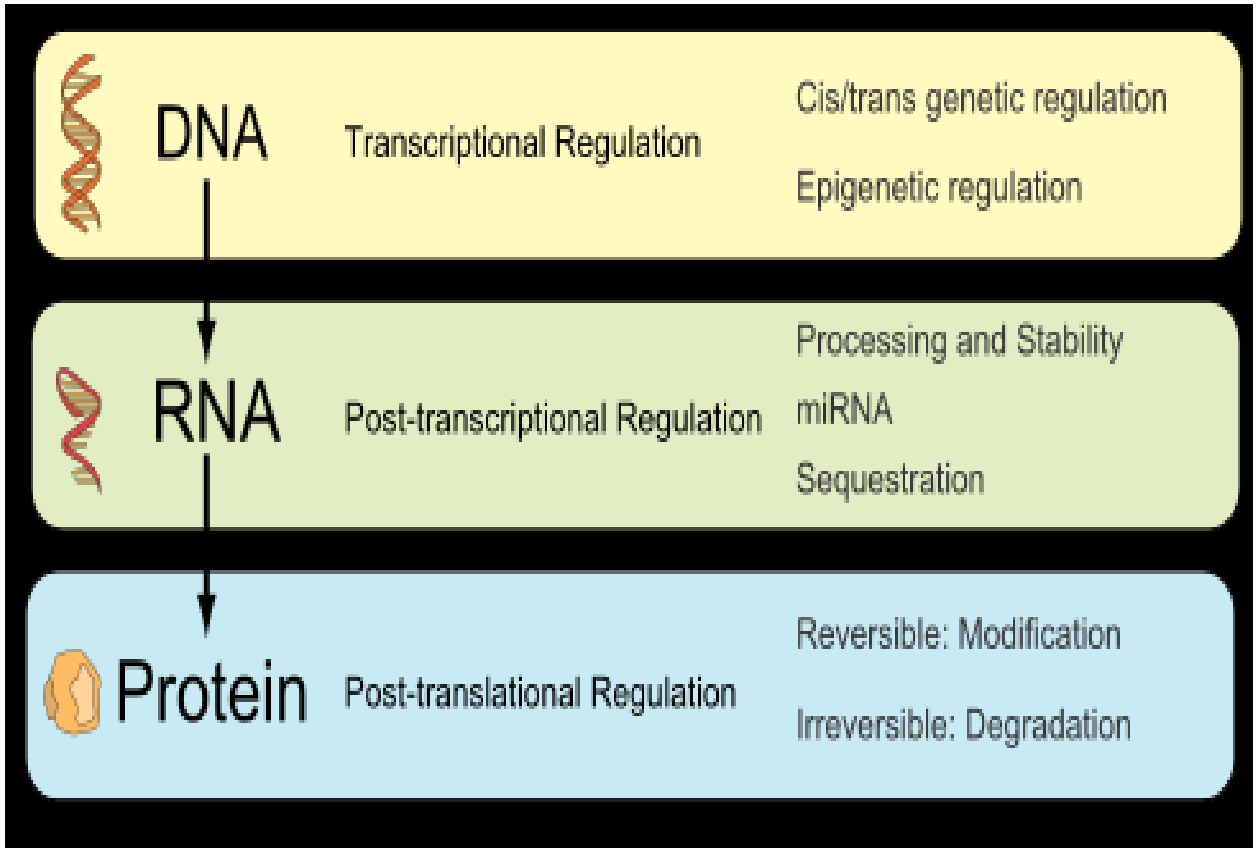
Regulation of gene expression

- Each cell expresses (or active or "on") only a fraction of their genes.
- Remaining genes are repressed ("off").
- The process consisting of activating some genes and inhibiting others is globally known as *gene regulation*.
- Gene regulation determines:
  - The appearance and different function of different cells types
  - The ability of some cells to react quickly to environmental changes

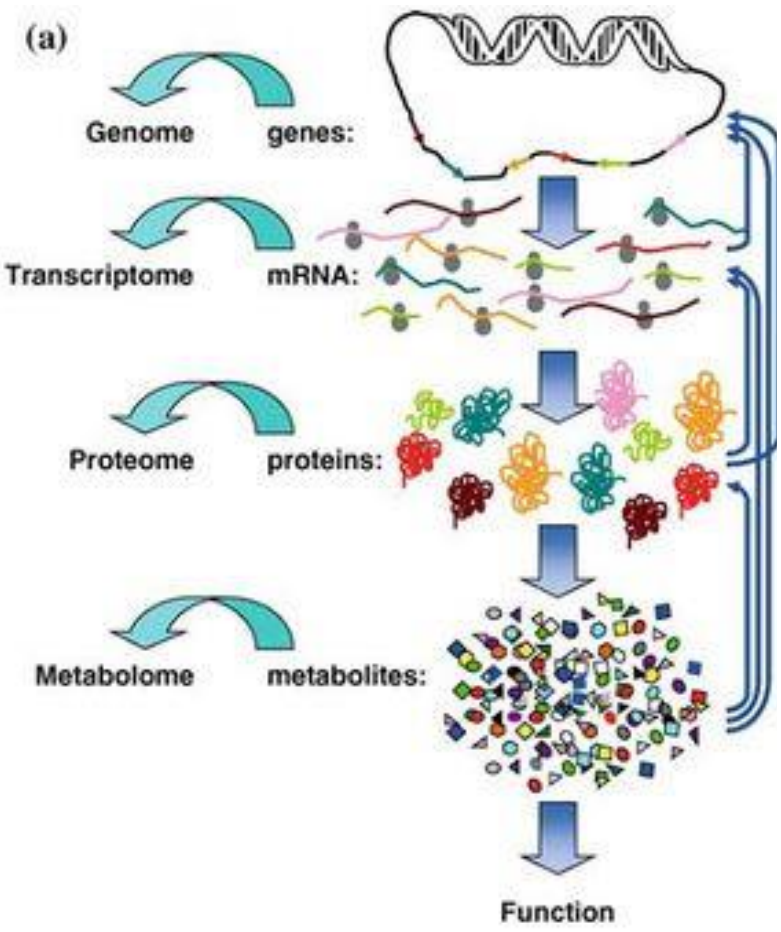
How are genes regulated?

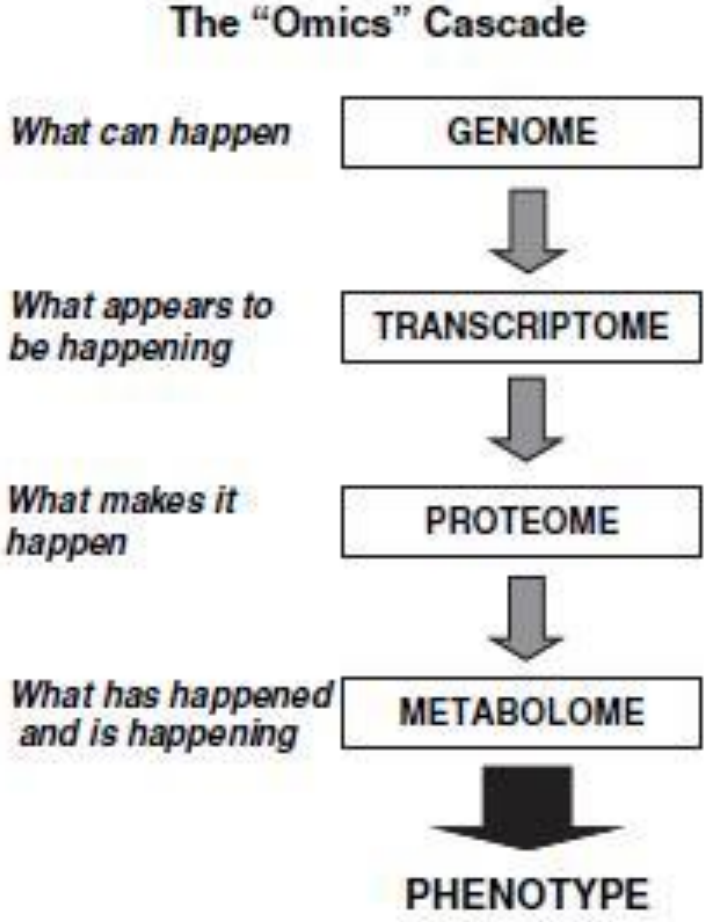
- Gene regulation can occur at any point in the process of expression but often occurs during transcription.
- Environmental signs or other cells activate proteins called *transcription factor*.
- They bind to the *regulatory regions* of genes, increasing or decreasing the level of transcription → They control the amount of gene product produced by the gene in every moment.

Forms of gene expression regulation



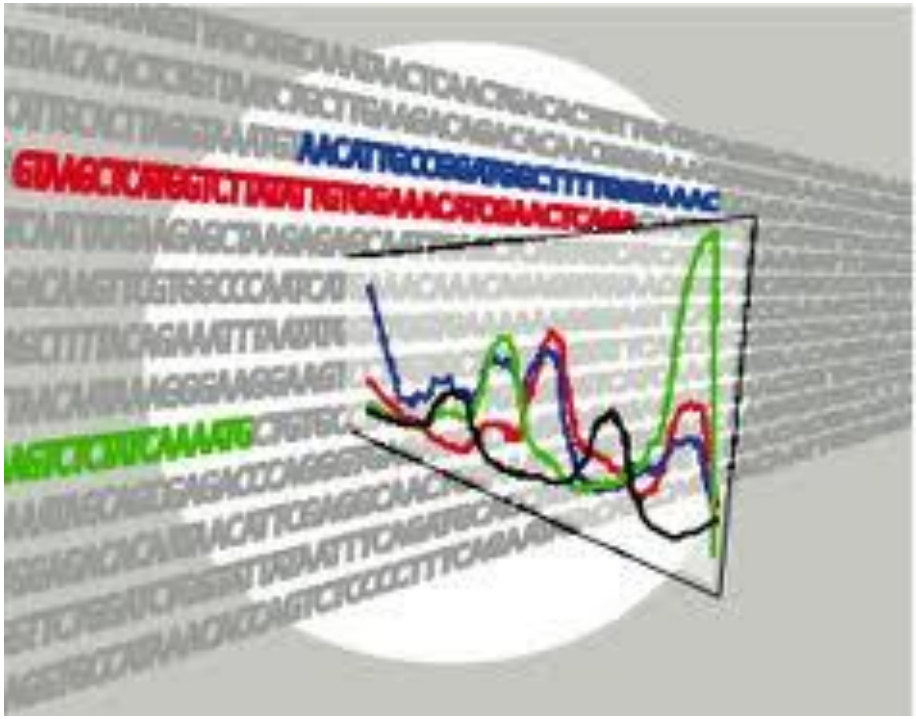
Put it altogether: *The OMICS cascade*





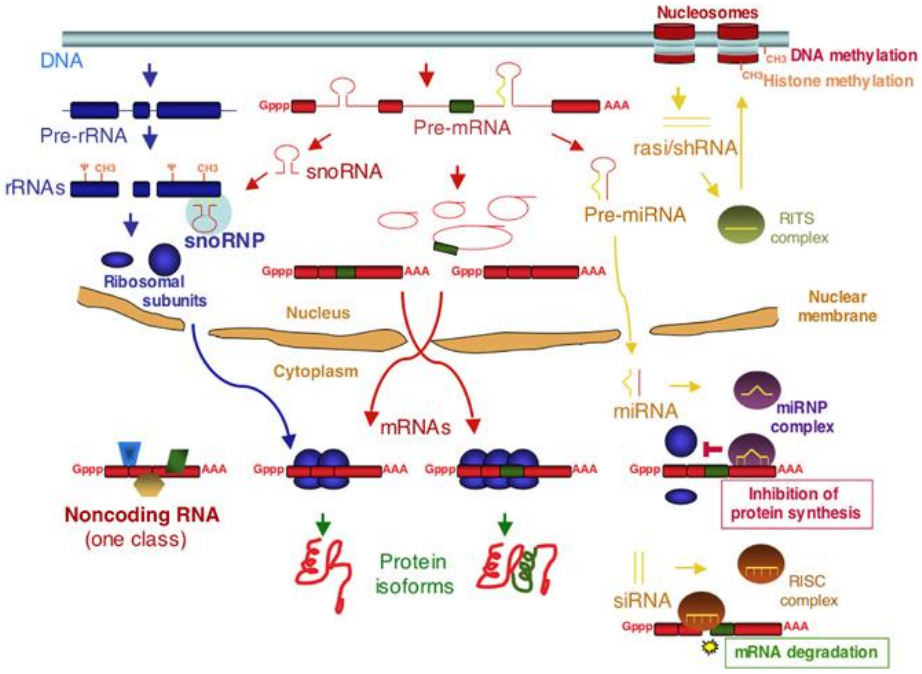
Genomics

- A genome is an organism's complete set of DNA, including all of its genes
- Genomics relies on *sequencing technologies* to study genomes and intragenomic phenomena



Transcriptomics

- The transcriptome is the set of all RNA molecules, in one or a population of cells.
- ***Transcriptomics***, examines expression levels of mRNAs in a given cell population, often using high-throughput techniques such as
  - RTqPCR
  - microarrays or
  - NGS.



‘omics’ are high throughput

- Most ‘omic’ approches generate huge quantities of data.
- The management, storage, analysis and interpretation of these high throughput datasets cannot be conceived without all type of computing and quantitative ressources
- ***Bioinformatics and computational biology are a must for omics sciences***