

Algoritmi per la Bioinformatica

Giulio Pavesi
Dipartimento di Bioscienze
Università di Milano
giulio.pavesi@unimi.it

Orari del corso: Bioinformatica

- Contatti: Giulio Pavesi, Dip. Bioscienze,
Via Celoria 26, piano 2B, in fondo al
corridoio a destra, ricevimento
martedì/venerdì dalle 15.00 in poi (o per
e-mail)
- Sito del corso:
<http://www.beaconlab.it/>
seguite il link “Didattica” in alto a sinistra
- Lezioni :
 - Giovedì 10.15-13.15 (D.0.1) e venerdì 14.15-
16.15 (D.2.6)

Bioinformatica

=

Bio(logia)

∩

Informatica

(Scienza dell'informazione)

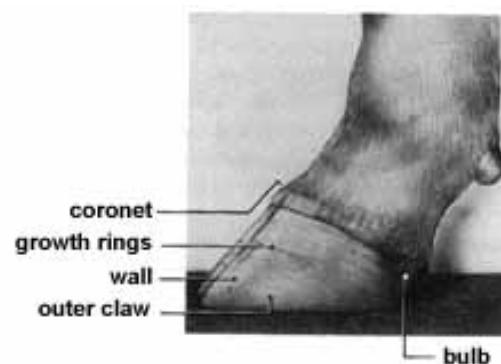
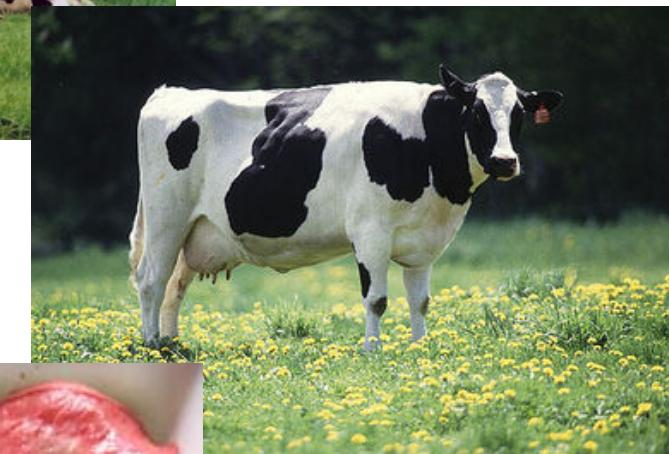
Information Science is NOT Computer Science

- Information science is a field primarily concerned with the:
 - analysis (**encoding and decoding**),
 - **classification**,
 - **manipulation**,
 - **storage**,
 - **retrieval**,
 - **transmission**,
- of information. (adapted from Wikipedia)

Information Science and Biology

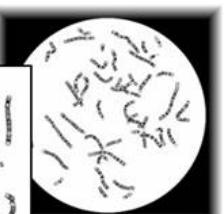
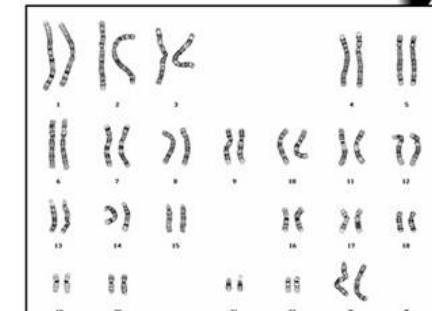
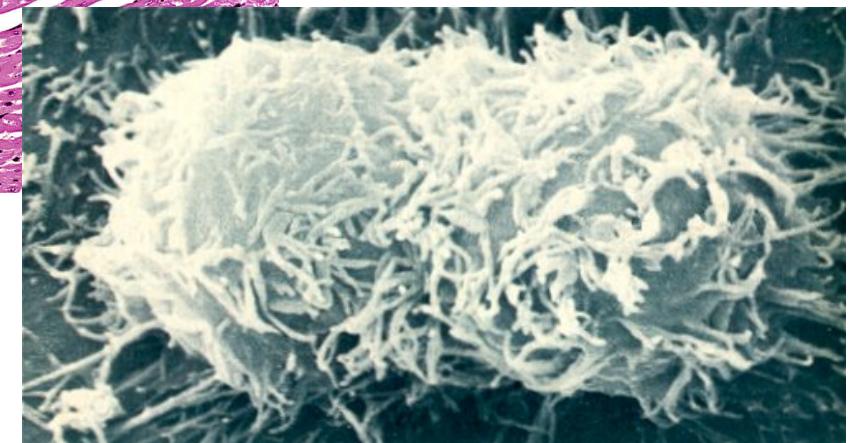
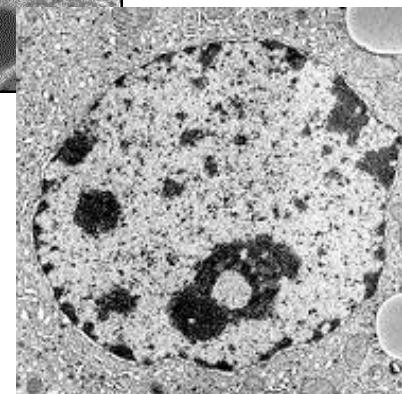
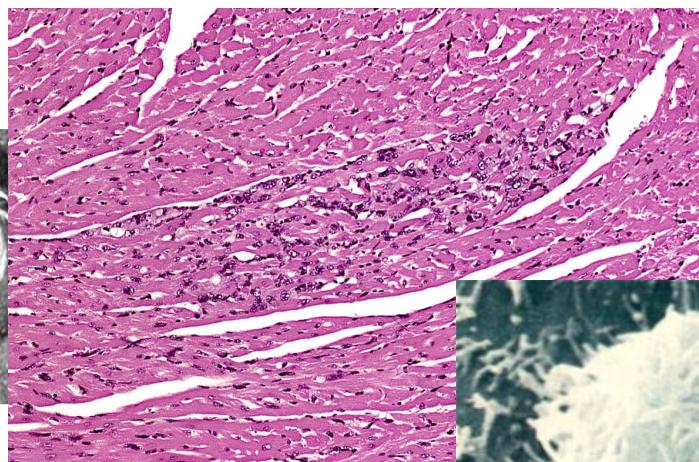
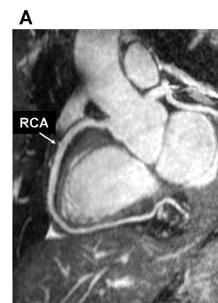
- Information science applied to the:
 - analysis (encoding and decoding),
 - classification,
 - manipulation,
 - storage,
 - retrieval,
 - transmission.
- of biological information (data).

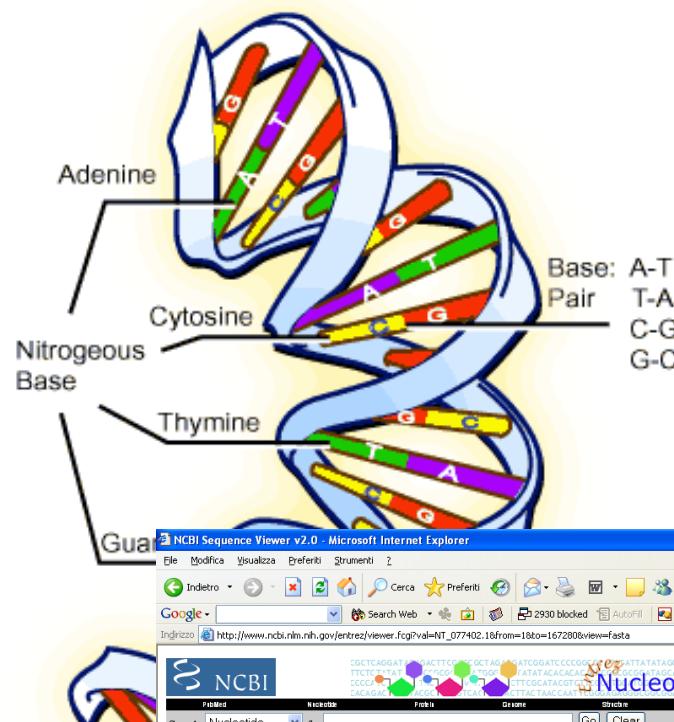
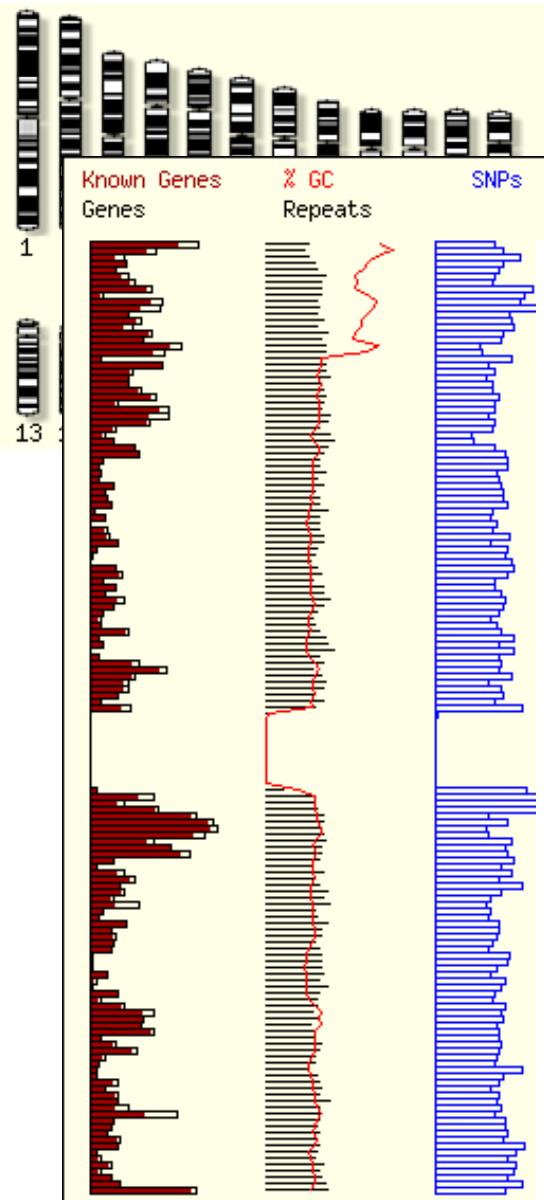
Biology: studying life



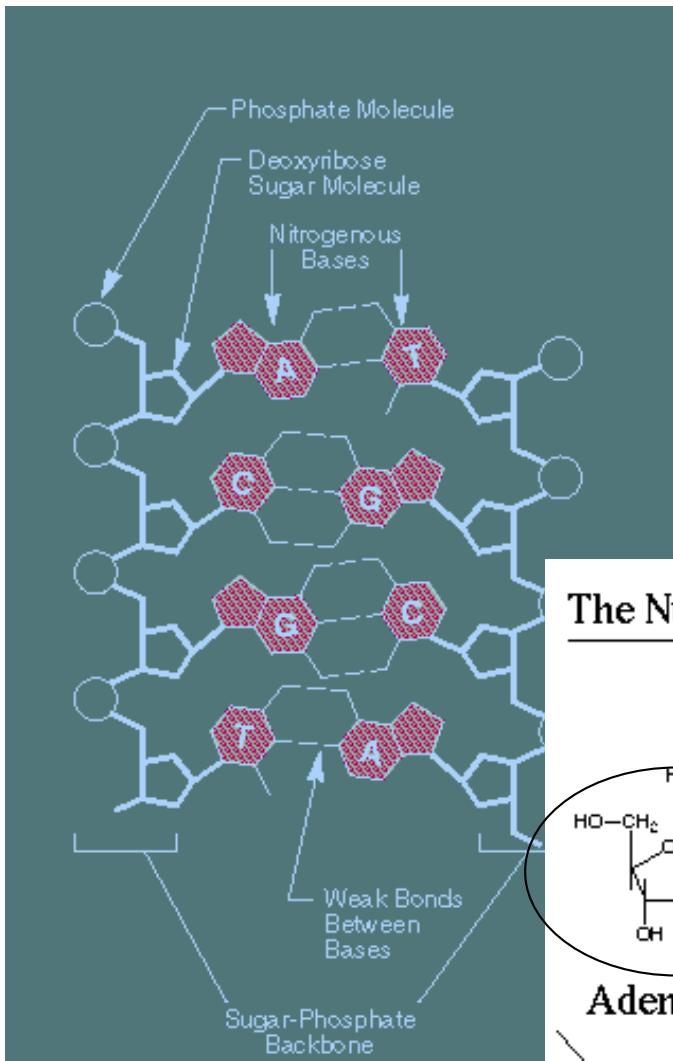
Biology: studying life

Figure 3

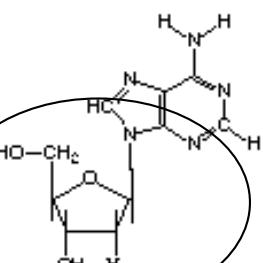




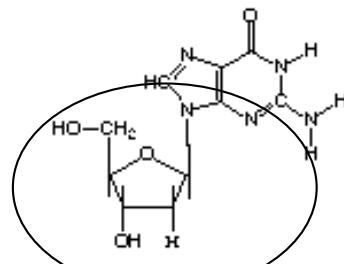
DNA



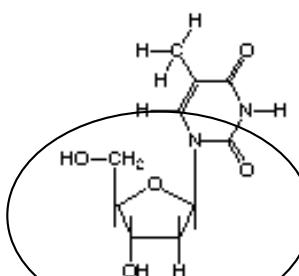
The Nucleotides of DNA



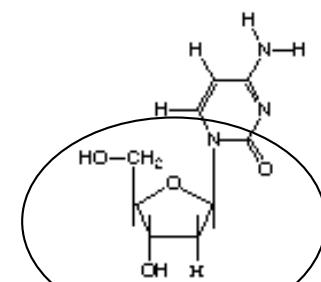
Adenine



Guanosine



Thymine

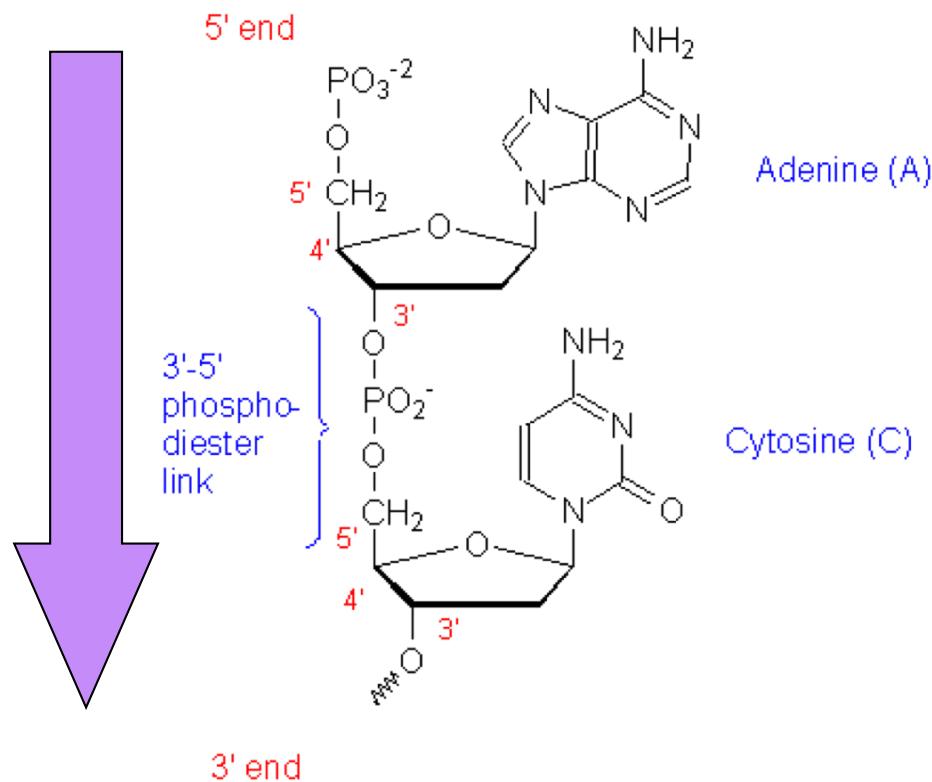


Cytosine

Purines

Pyrimidines

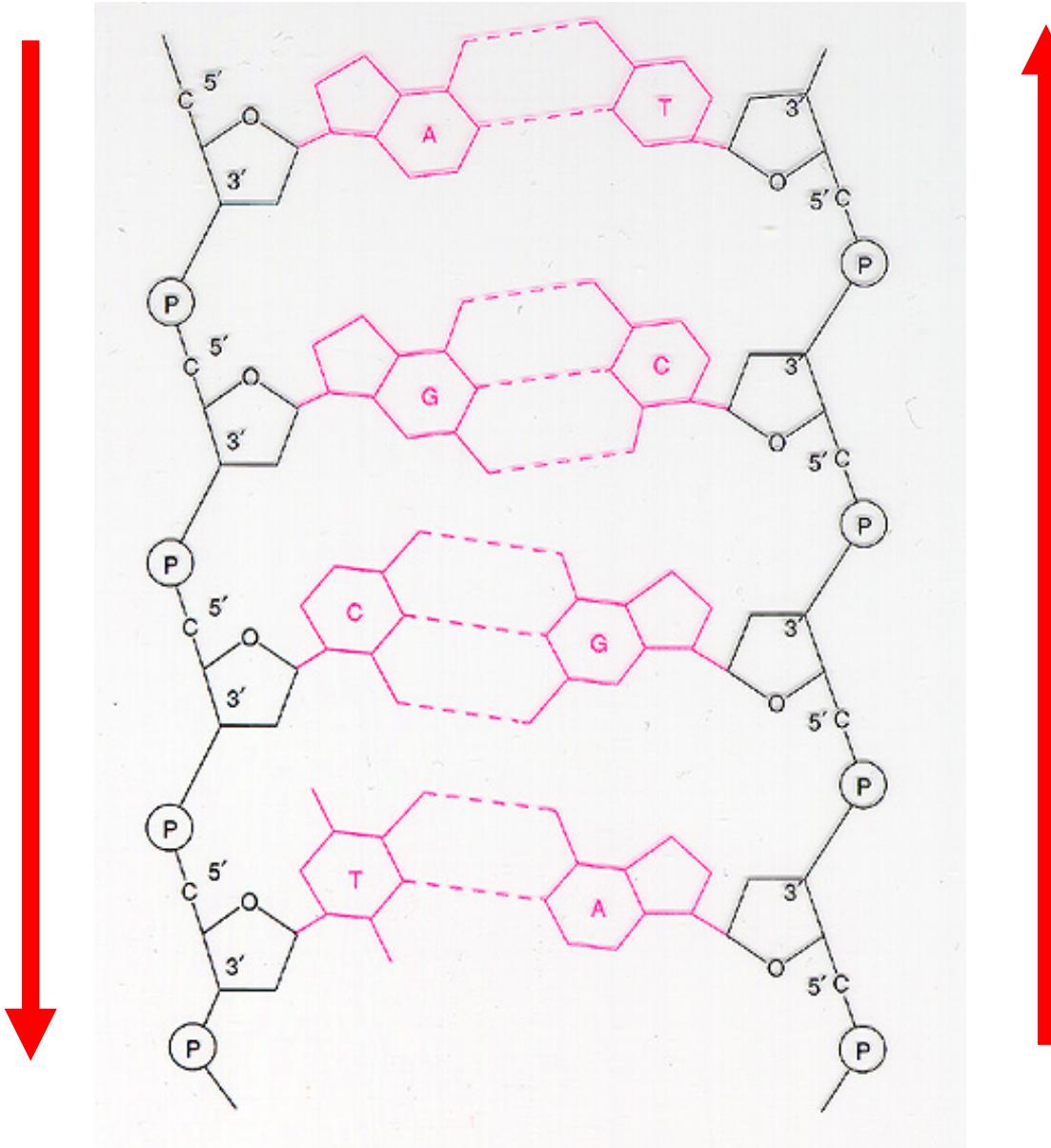
DNA (single stranded)



A strand of DNA has a direction. The asymmetric ends of DNA strands are called the 5' (five prime) and 3' (three prime) ends, with the 5' end having a terminal phosphate group and the 3' end a terminal hydroxyl group.

A strand of RNA is similar, only composed of ribonucleotides (one more oxygen atom), and uses a modified version of thymine (called **uracil**)

DNA (double stranded)



In a double helix
the direction
of the nucleotides
in one strand
is opposite
to their direction
in the other strand:
the strands
are **antiparallel**

C pairs only with G
A pairs only with T

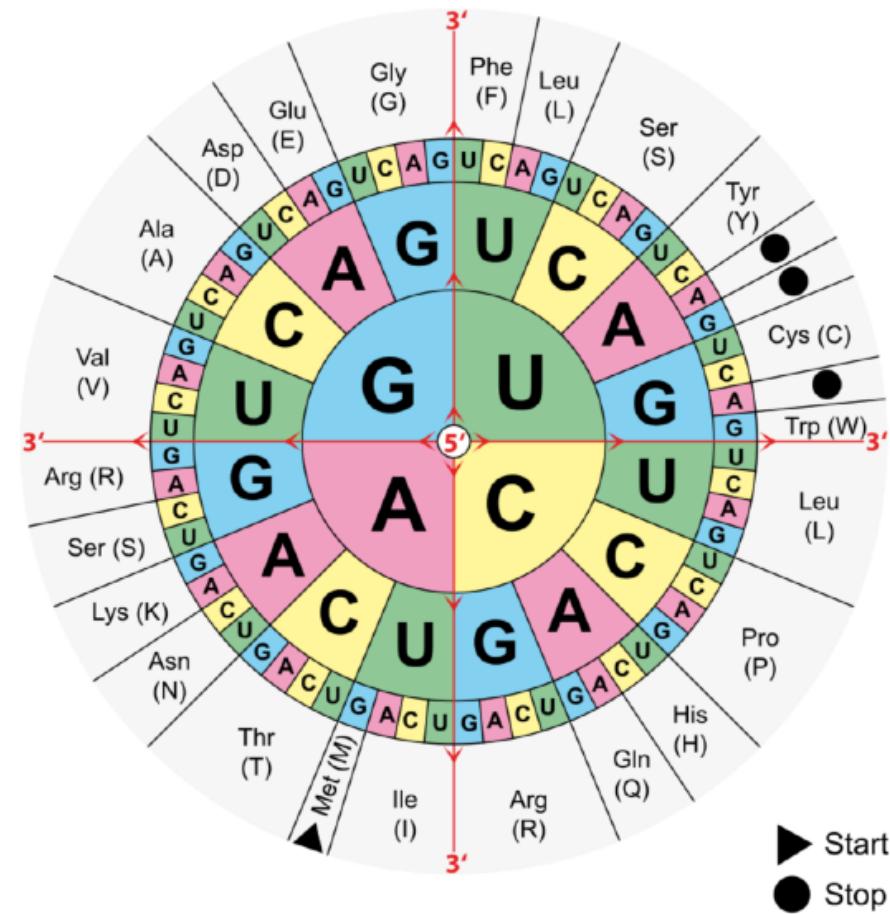
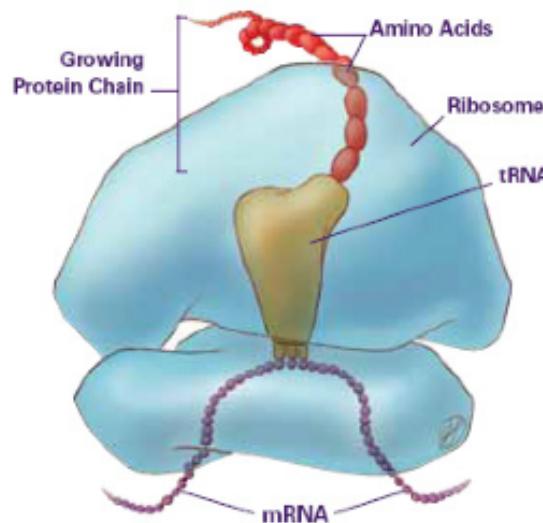
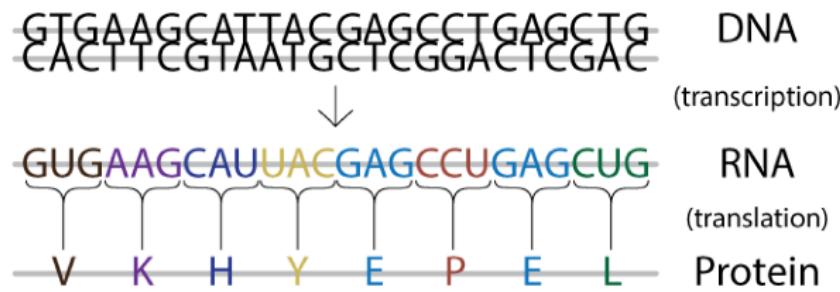
DNA inside living cells

Video

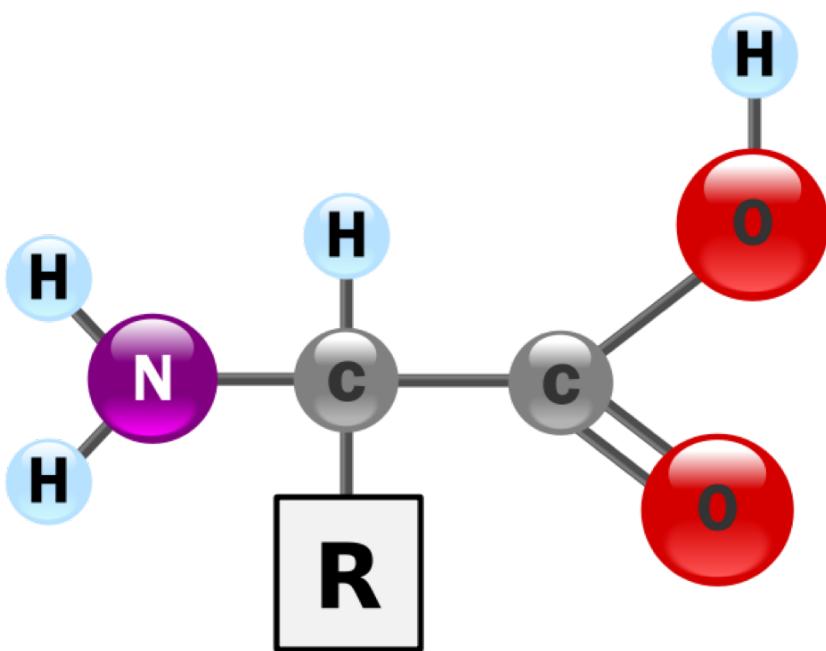
Information Science and Biology

- Information science applied to:
 - analysis (encoding and decoding),
 - classification,
 - manipulation,
 - storage,
 - retrieval,
 - transmission.
- of biological information **stored inside living cells**

Cells reading and decoding information: the GENETIC CODE



Amino acids



Twenty-One Amino Acids			Positive Side chain charge at physiological pH 7.4	
A. Amino Acids with Electrically Charged Side Chains			Positive	
			Lysine (Lys) K	
			Histidine (His) H	
			Aspartic Acid (Asp) D	
			Glutamic Acid (Glu) E	
			Negative	
			Aspartic Acid (Asp) D	
			Glutamic Acid (Glu) E	
			Glutamic Acid (Glu) E	
B. Amino Acids with Polar Uncharged Side Chains			C. Special Cases	
			Cysteine (Cys) C	
			Selenocysteine (Sec) U	
			Glycine (Gly) G	
			Proline (Pro) P	
			Glycine (Gly) G	
			Proline (Pro) P	
D. Amino Acids with Hydrophobic Side Chain			D. Amino Acids with Hydrophobic Side Chain	
			Alanine (Ala) A	
			Valine (Val) V	
			Isoleucine (Ile) I	
			Leucine (Leu) L	
			Methionine (Met) M	
			Phenylalanine (Phe) F	
			Tyrosine (Tyr) Y	
			Tryptophan (Trp) W	

From amino acids to proteins

[Video](#)

From amino acids to proteins

- The genetic code encodes 20 amino acids + three codons used for «stop translation»
- But, we have 21 amino acids incorporated in eukaryotic proteins
- Where is the 21st codon?