HMM for Bioinformatics

Paweł Błażej
Department of Genomics, Faculty of Biotechnology,
blazej@smorfland.uni.wroc.pl

27 lutego 2019

piece of literature

- W.H. Majoros, “Methods for Computational Gene Prediction”, Cambridge University Press 2007;

piece of literature

- W.H. Majoros, “Methods for Computational Gene Prediction”, Cambridge University Press 2007;
- T. Koski, “Hidden Markov Models for Bioinformatics”, Kluwer Academic Publishers 2001.

piece of literature

- W.H. Majoros, "Methods for Computational Gene Prediction", Cambridge University Press 2007;
- T. Koski, "Hidden Markov Models for Bioinformatics", Kluwer Academic Publishers 2001.
- `http://smorfland.uni.wroc.pl/~blazej/HMM/`

Getting started

1. Mathematically trained individuals prefer their problems to be stated simply and precisely;

Getting started

1. Mathematically trained individuals prefer their problems to be stated simply and precisely;
2. To persons with a mathematical training, biology appears to be bewildering array of terminology and conventions;

Getting started

1. The students are advised not to treat everything which will be said here as being "biologically true".

Getting started

1. The students are advised not to treat everything which will be said here as being "biologically true".
2. All the biological statements that will appear during this lecture are "probably approximately correct".

The first

Statement

All life on this planet depends on three types of molecule DNA, RNA and proteins,

The second

Statement

DNA is a vast library describing how the cell works

Statement

RNA acts to transfer certain short pieces of this library to different places in the cell at which point those smaller volumes of information are used as a templates to synthetize proteins.

Statement

Proteins form enzymes that perform biochemical reactions, send signals to other cells, form the body's major component and otherwise perform the actual work of the cell.

Short history of DNA and RNA

1. 1869 – DNA was discovered by Johann Friedrich Miescher;

Short history of DNA and RNA

1. 1869 – DNA was discovered by Johann Friedrich Miescher;
2. By the eaerly 1900s it was shown that DNA was a long molecule consisting of four types of bases;

Short history of DNA and RNA

1. 1869 – DNA was discovered by Johann Friedrich Miescher;
2. By the eaerly 1900s it was shown that DNA was a long molecule consisting of four types of bases;
3. By the 1920s, nucleid acids were grouped into two classes called DNA and RNA, that differ slightly intheir base composition.

Short history of DNA and RNA

fact 1.

Ironically, for a long time biologists paid a little attention to DNA.

Important date

1940 – Oswald Avery and colleagues proved that genes indeed reside on DNA.

What is the structure of DNA?

1. 1950 – Erwin Chargraff discovered one-to-one ratio of the adynine-to-thymine and guanine-to-citosine in DNA;

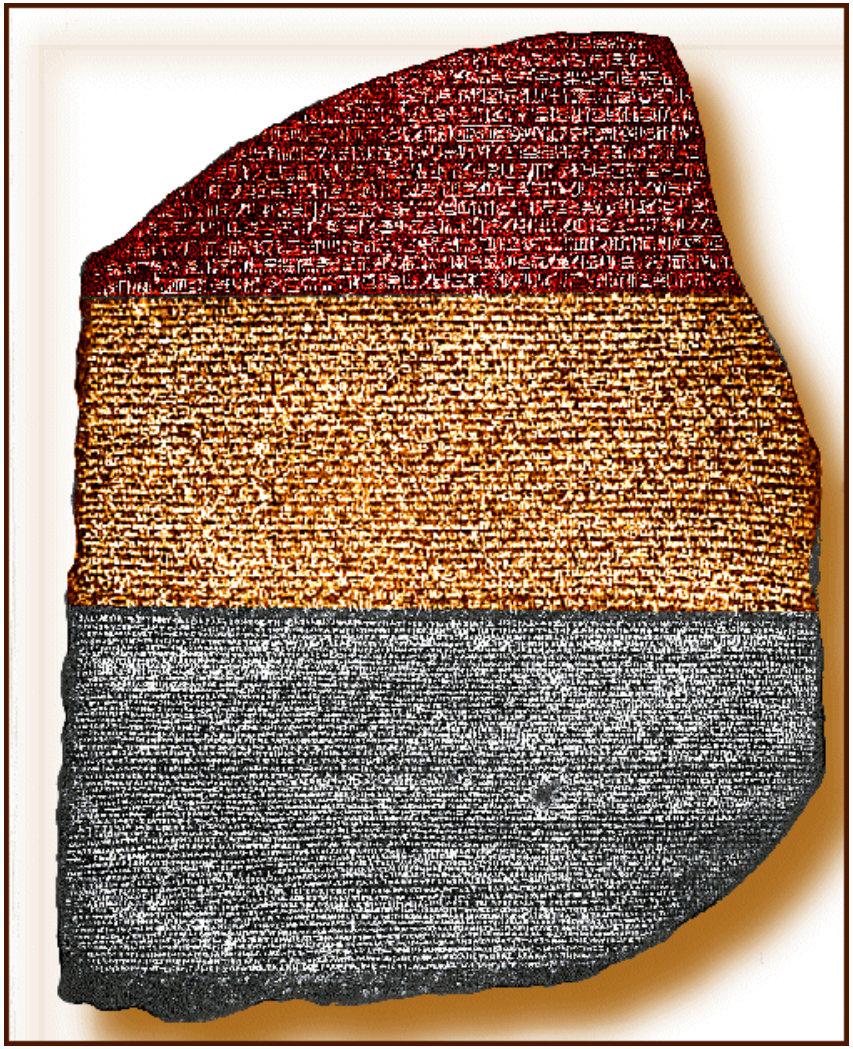
What is the structure of DNA?

1. 1950 – Erwin Chargraff discovered one-to-one ratio of the adynine-to-thymine and guanine-to-citosine in DNA;
2. 1951 – Maurice Wilkins and Rosalind Franklin obtained sharp x-ray images of DNA;
3. 1953 – James Watson and Francis Crick determined the double helical structure of a DNA molecule.

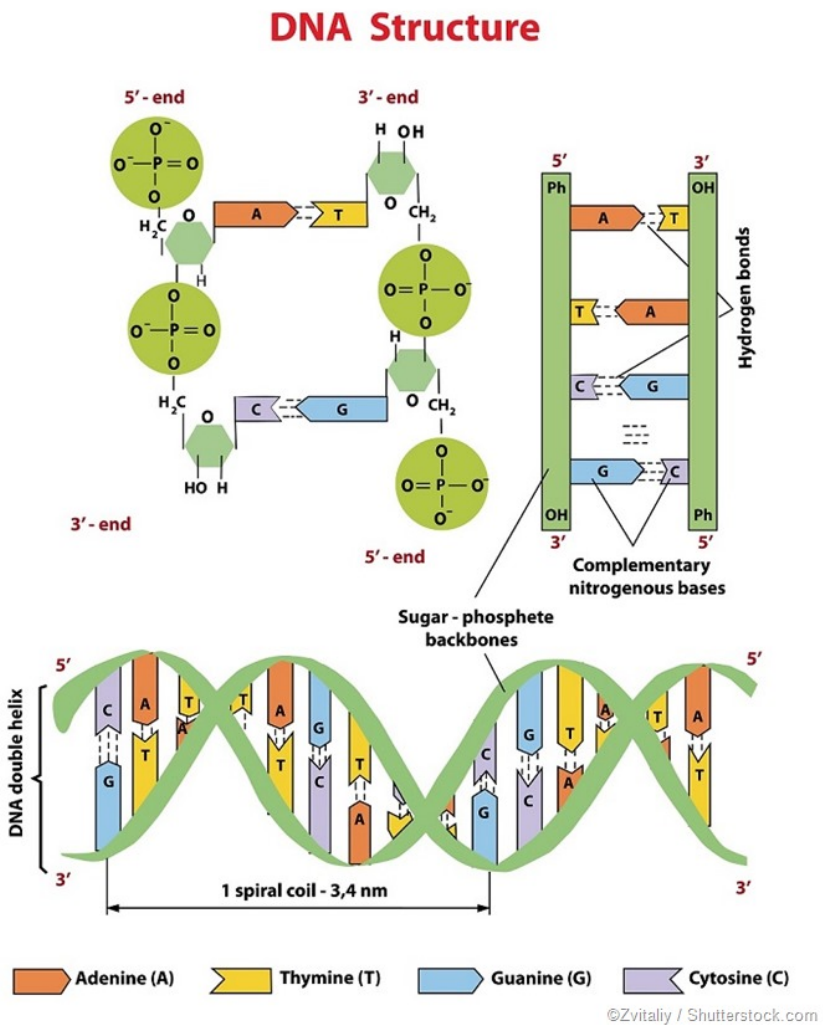
Human Genomic Project – 2001

The Human Genome Project originally aimed to map the nucleotides contained in a human haploid reference genome (more than three billion). The "genome" of any given individual is unique; mapping the "human genome" involved sequencing a small number of individuals and then assembling these together to get a complete sequence for each chromosome. Therefore, the finished human genome is a mosaic, not representing any one individual.

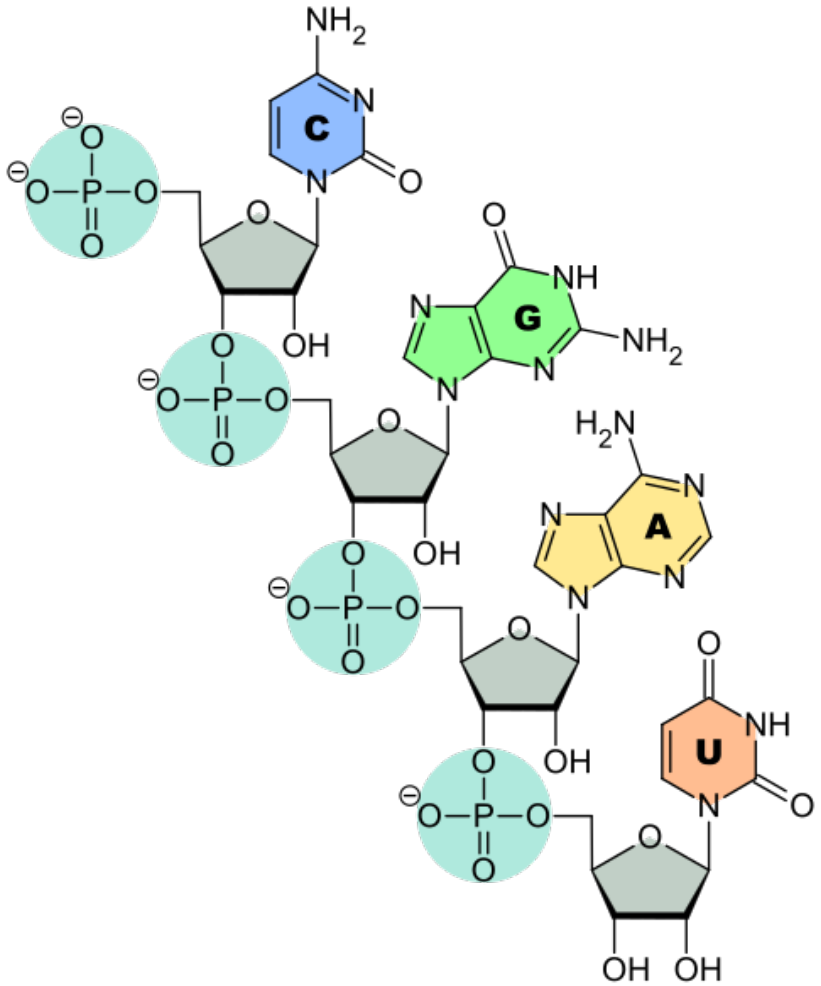
Rosetta Stone



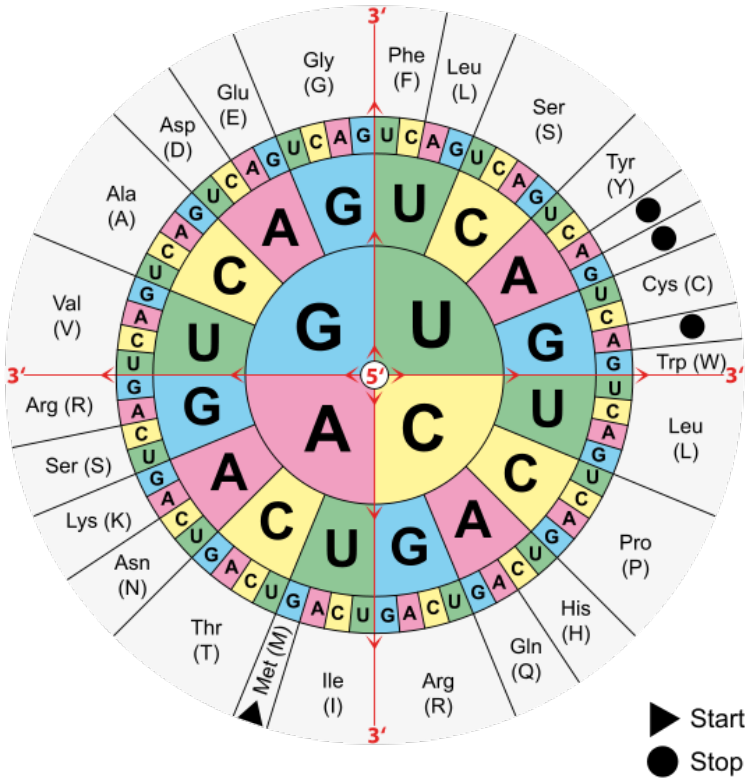
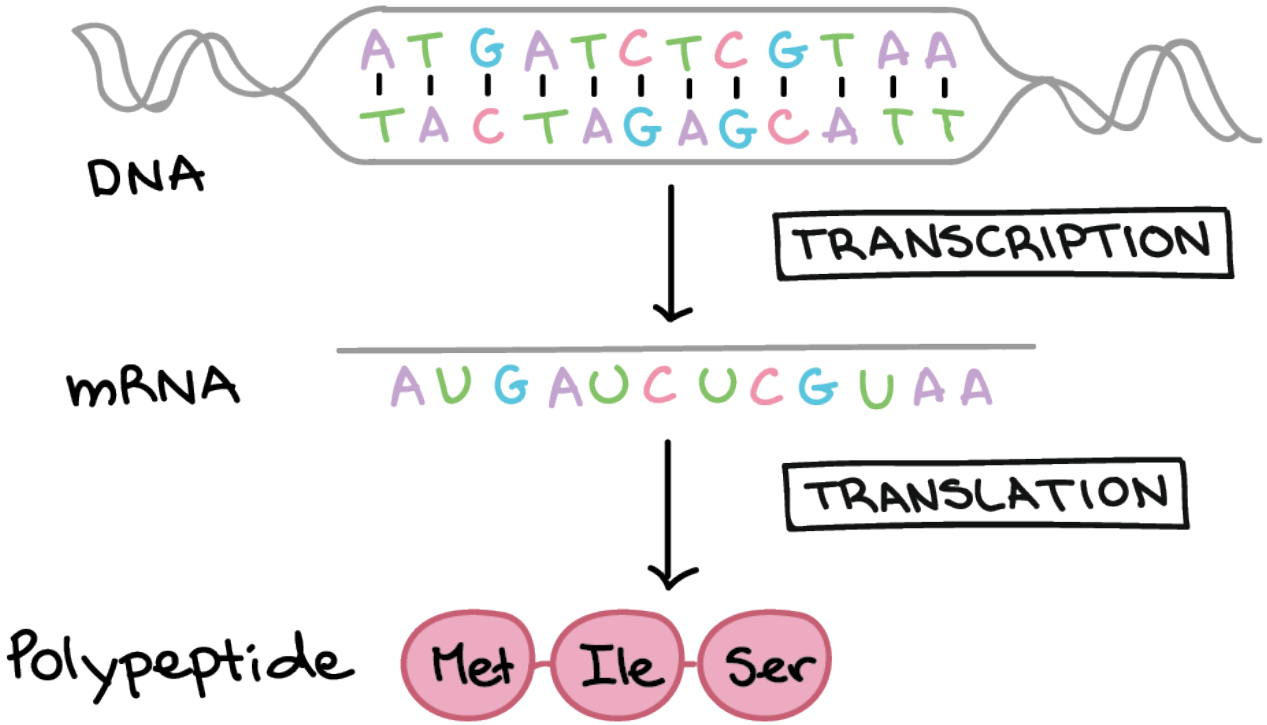
DNA and RNA


DNA Structure
5′ - end
3′ - end
3′ - end
5′ - end
Hydrogen bonds
Complementary nitrogenous bases
Sugar - phosphete backbones
DNA double helix
1 spiral coil - 3,4 nm
Adenine (A)
Thymine (T)
Guanine (G)
Cytosine (C)
©Zvitaliy / Shutterstock.com

DNA and RNA



The standard genetic code



The Central Dogma of Genetics