



Some Basics of Molecular Biology

Hubert Rehrauer



**University of
Zurich^{UZH}**

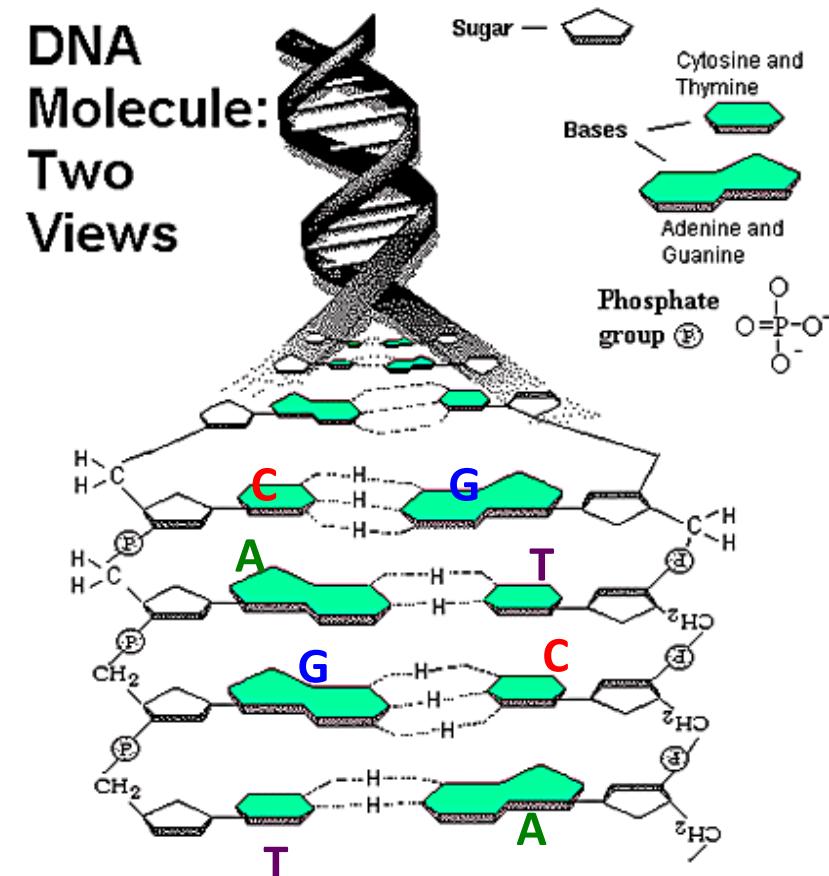
ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich



What is DNA?

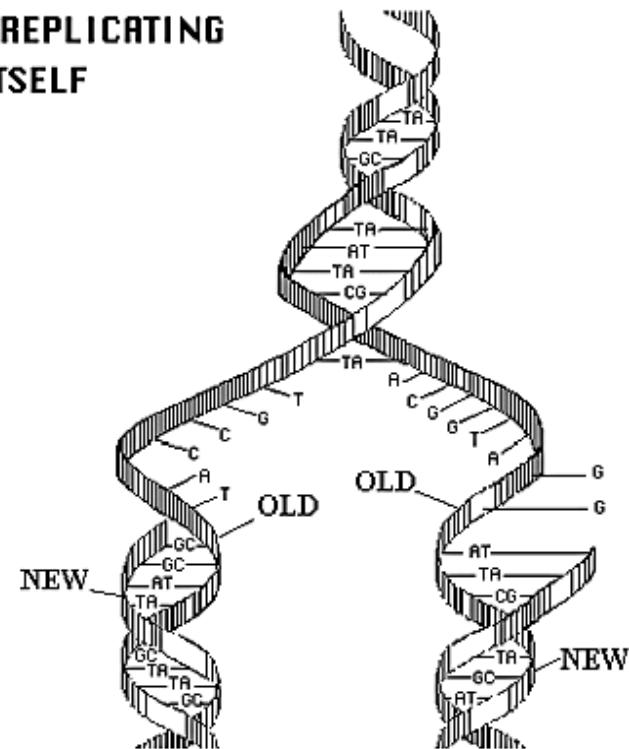
- A long backbone of sugars with nucleotides attached
 - Adenine (A)
 - Guanine (G)
 - Cytosine (C)
 - Thymine (T)
- It can form a self-complementary **double helix**
- In living organisms, the DNA is the carrier of the hereditary information, it is the **source code** of life



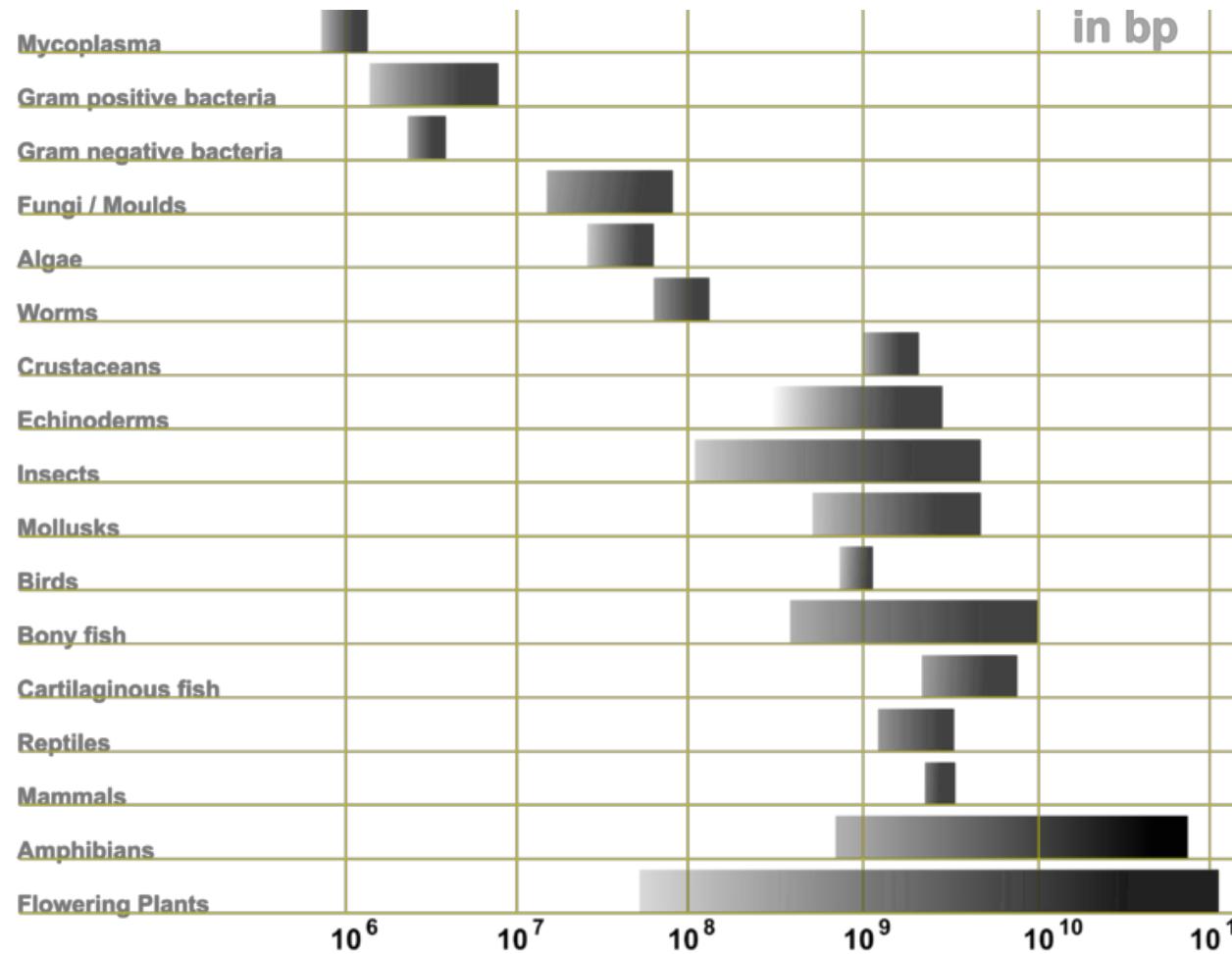
DNA replication

- The helix becomes unzipped and each strand acts as a template for a new complementary strand of DNA

DNA REPLICATING ITSELF



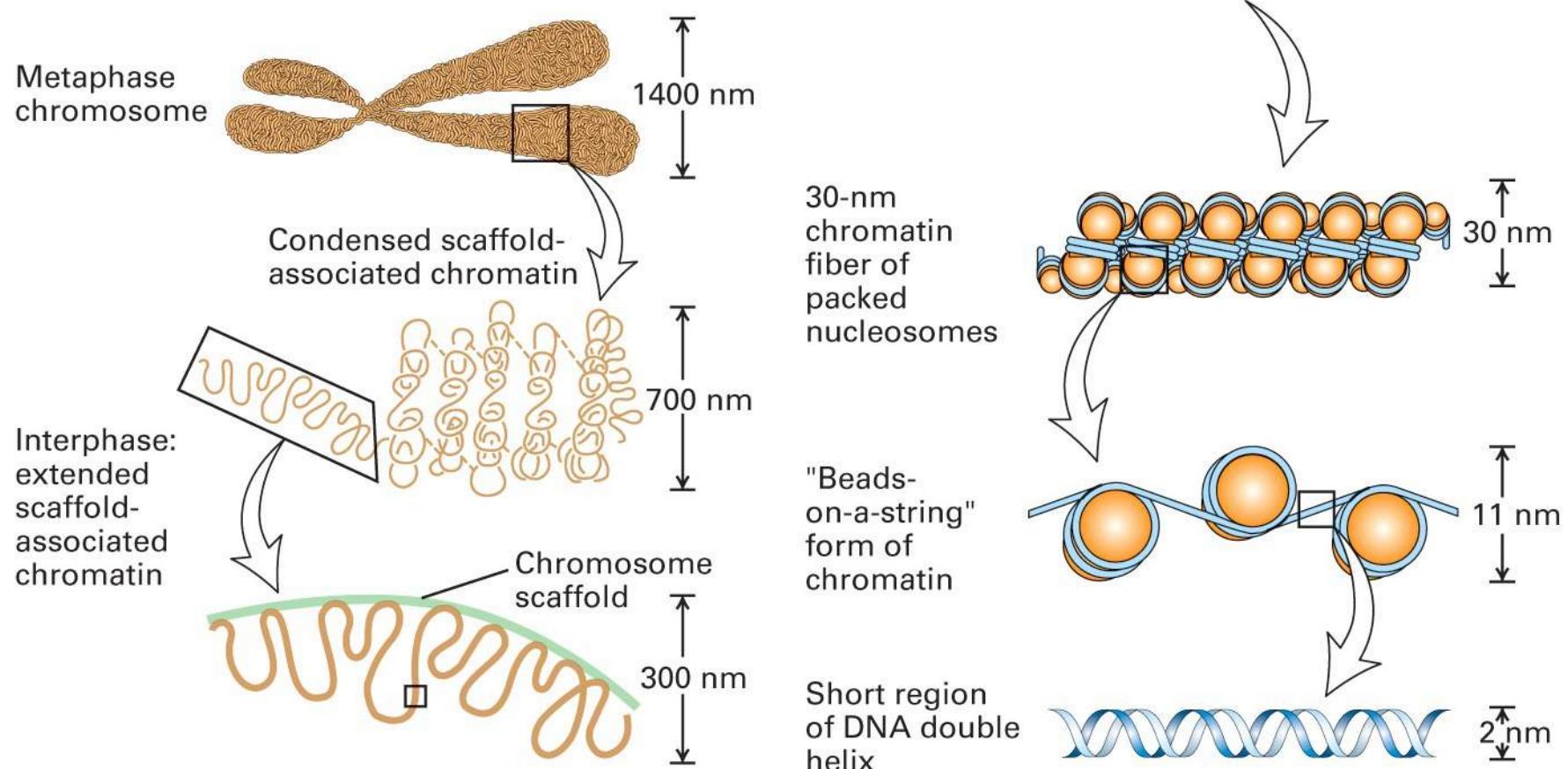
Genome Sizes



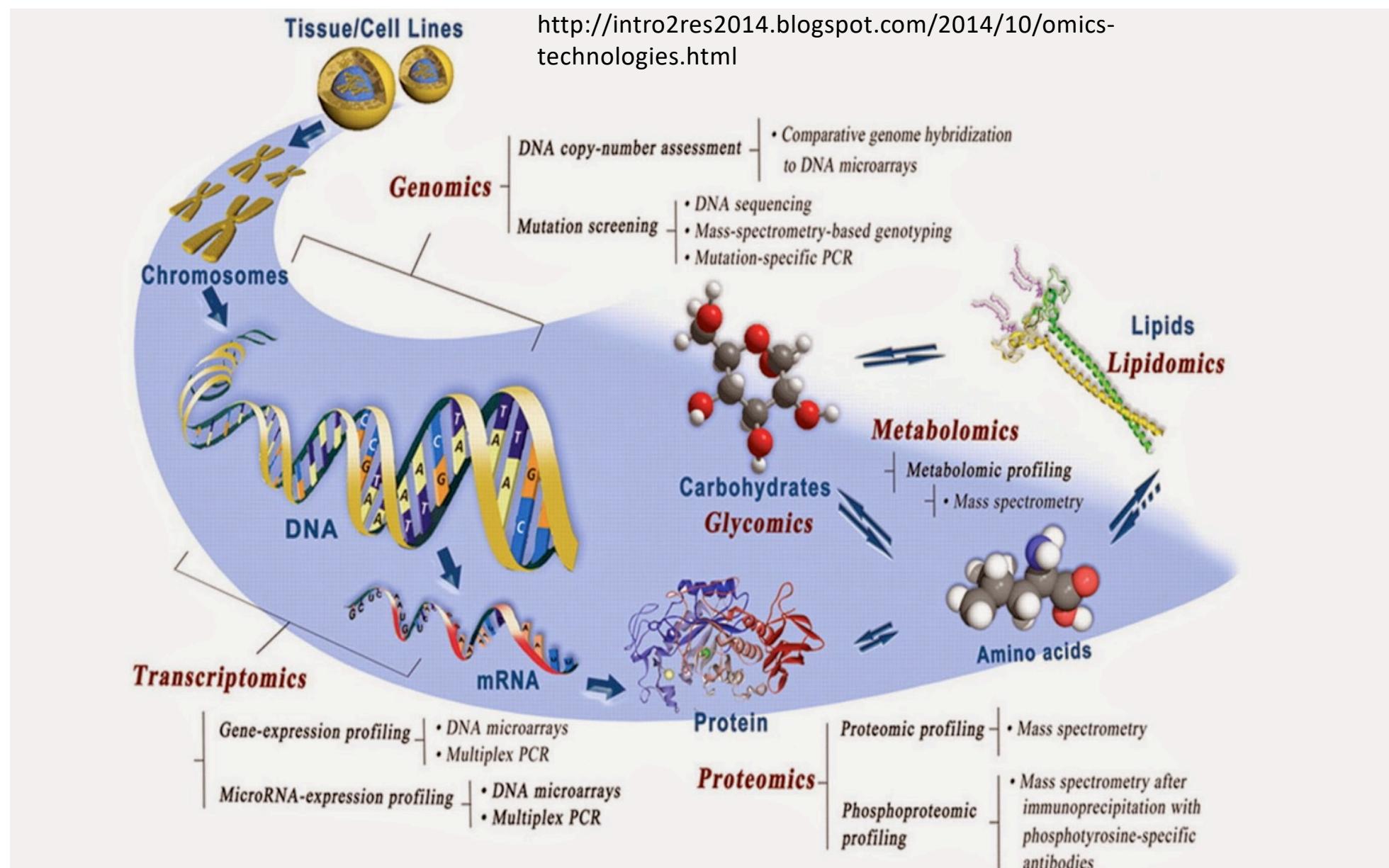
The size of the human genome is 3.2 billion base pairs. The length of this DNA string is approx. 2m.

http://en.wikipedia.org/wiki/Genome_size

DNA Superstructure

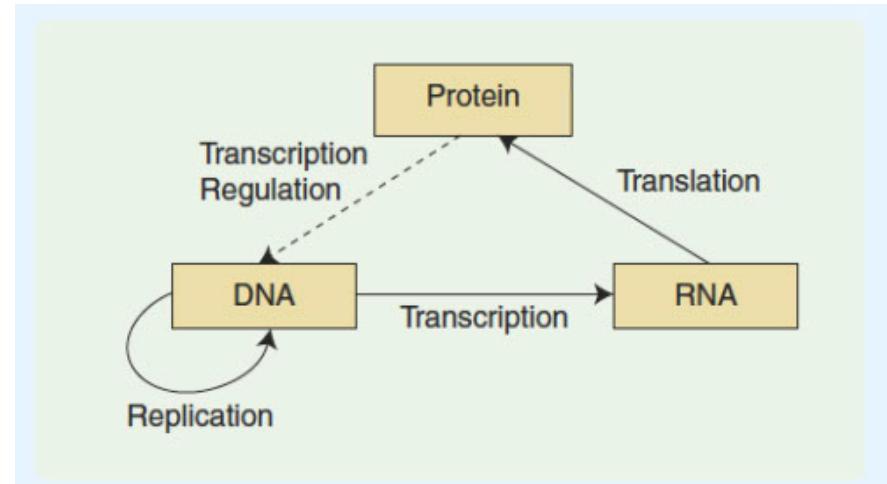


Lodish et al. *Molecular Biology of the Cell* (5th ed.). W.H. Freeman & Co., 2003.



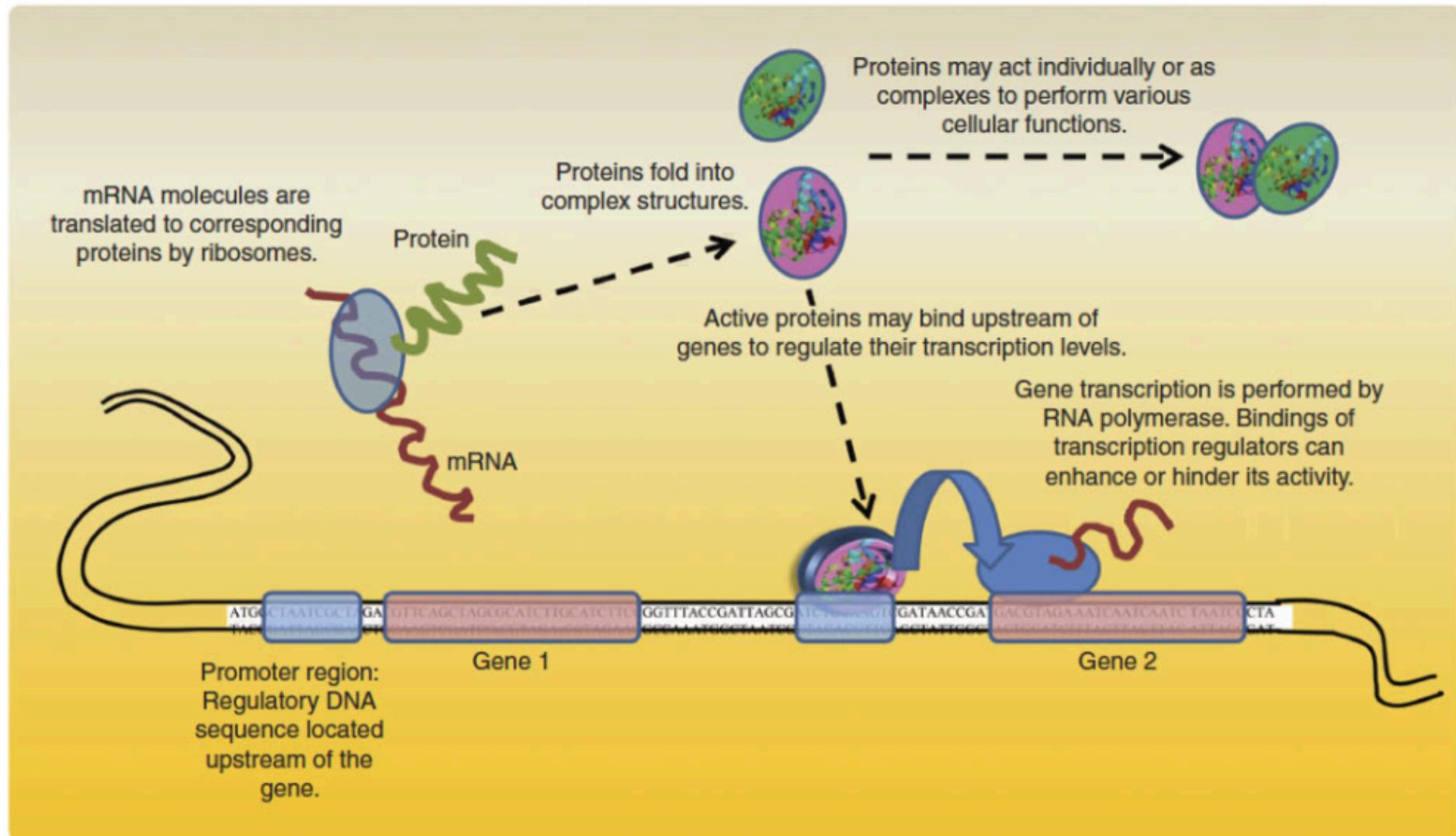
Genes

- A gene is a region of DNA that controls a hereditary characteristic
- Usually a gene is transcribed into a messenger RNA which is then translated into a protein.
- In humans genes constitute only ~3% of the human genome



10
01
1010
1
01
101
10
0
01
1

The Central Dogma



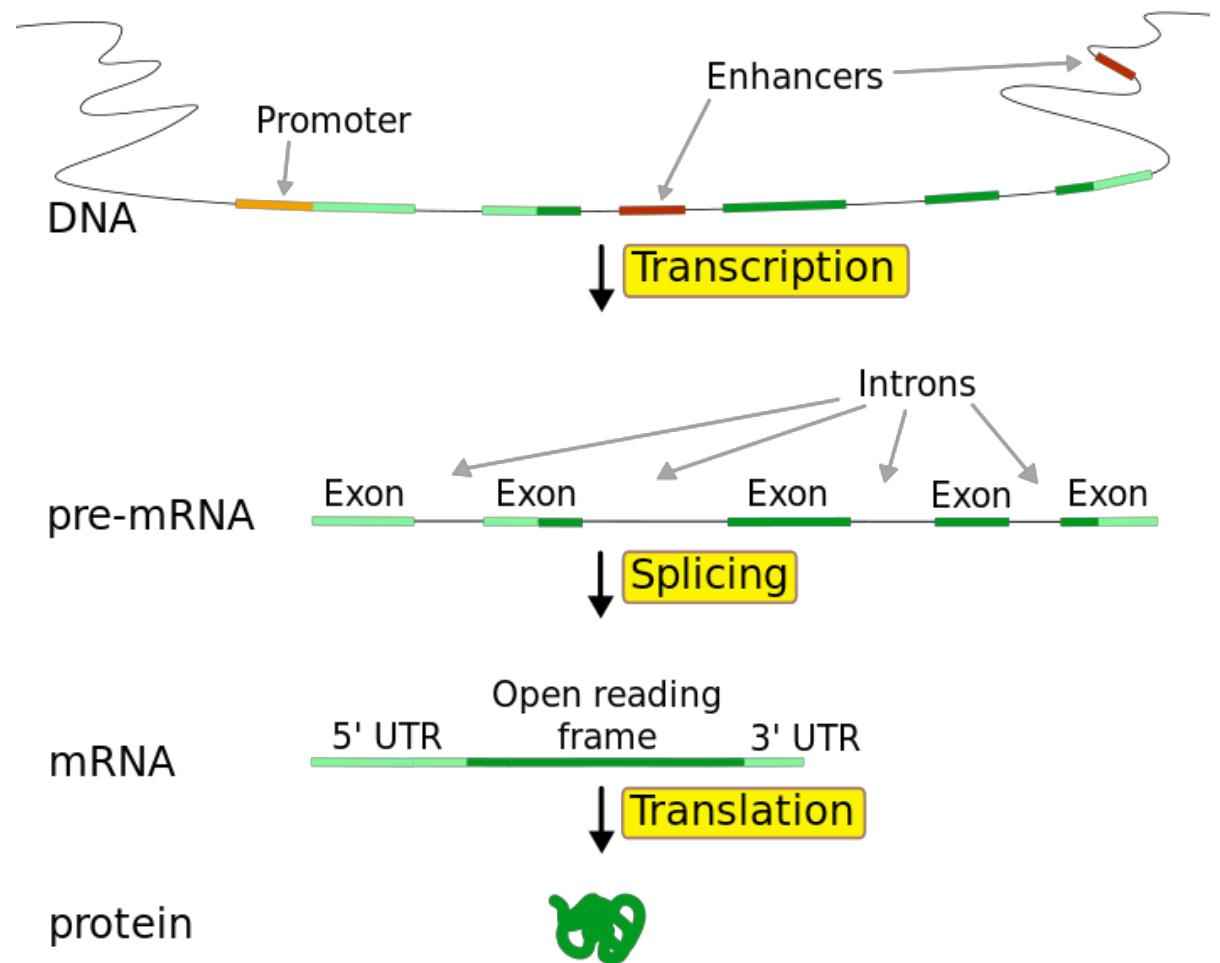
Transcription

The transcription process generates a messenger RNA molecule from a gene region.

RNA is like DNA but

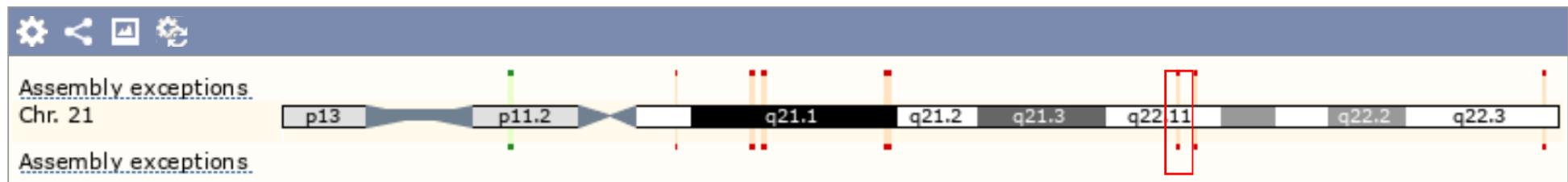
- the sugar-phosphate is different: ribose instead of deoxyribose
 - In all places where the DNA has a T the RNA has a U (uracil)

In higher organisms the protein coding sequences (exons) are interspersed by non-coding sequences (introns) which are spliced out.



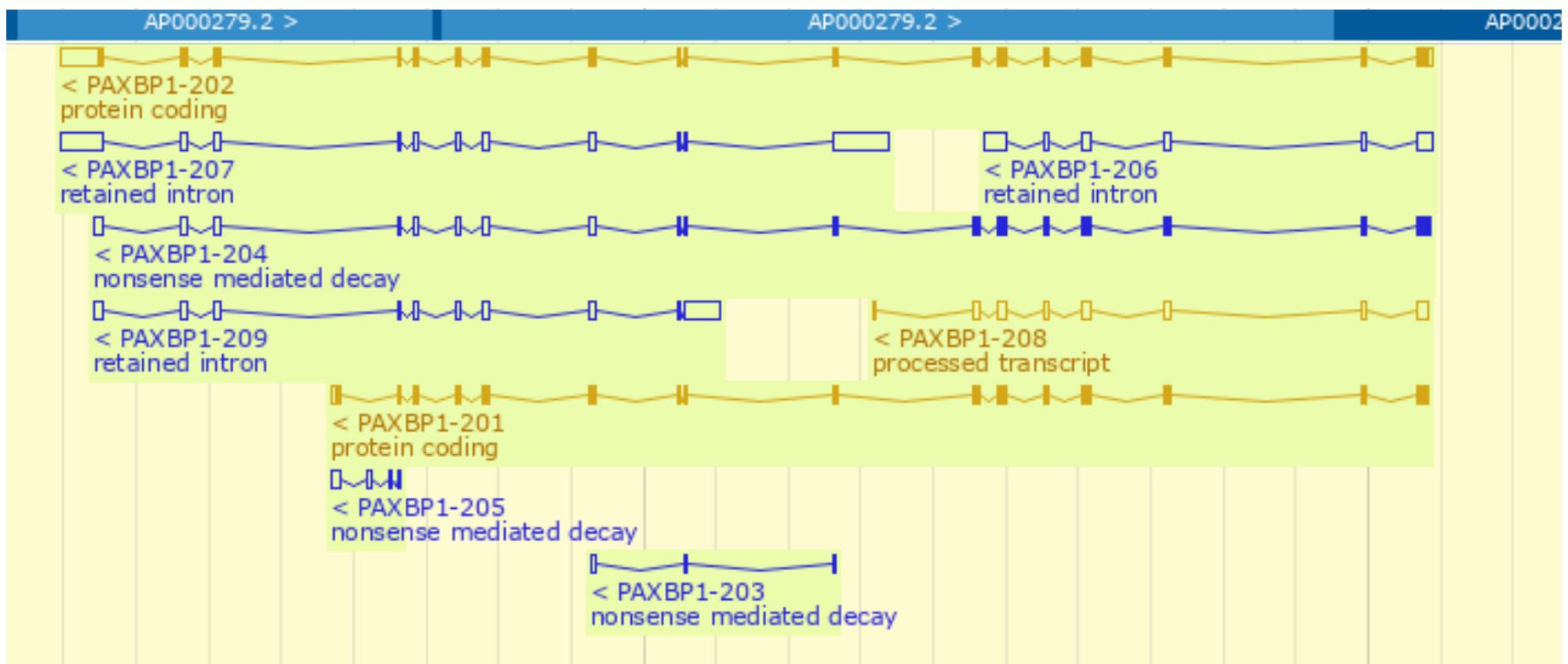
Gene: PAXBP1

Chromosome 21: 32,733,142-32,772,549



Region in detail ?





Protein Coding
 Ensembl protein coding
 merged Ensembl/Havana

Non-Protein Coding
 processed transcript
 pseudogene
 RNA gene

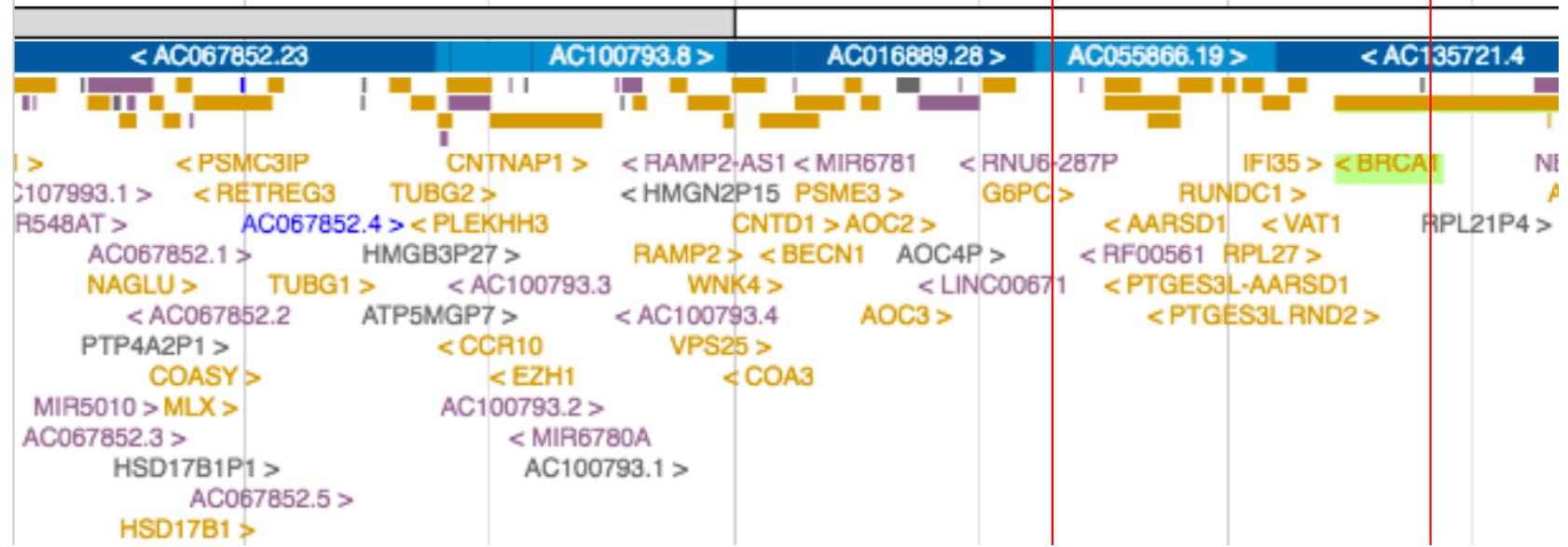
PAXBP1

CTCF
 Promoter Flank

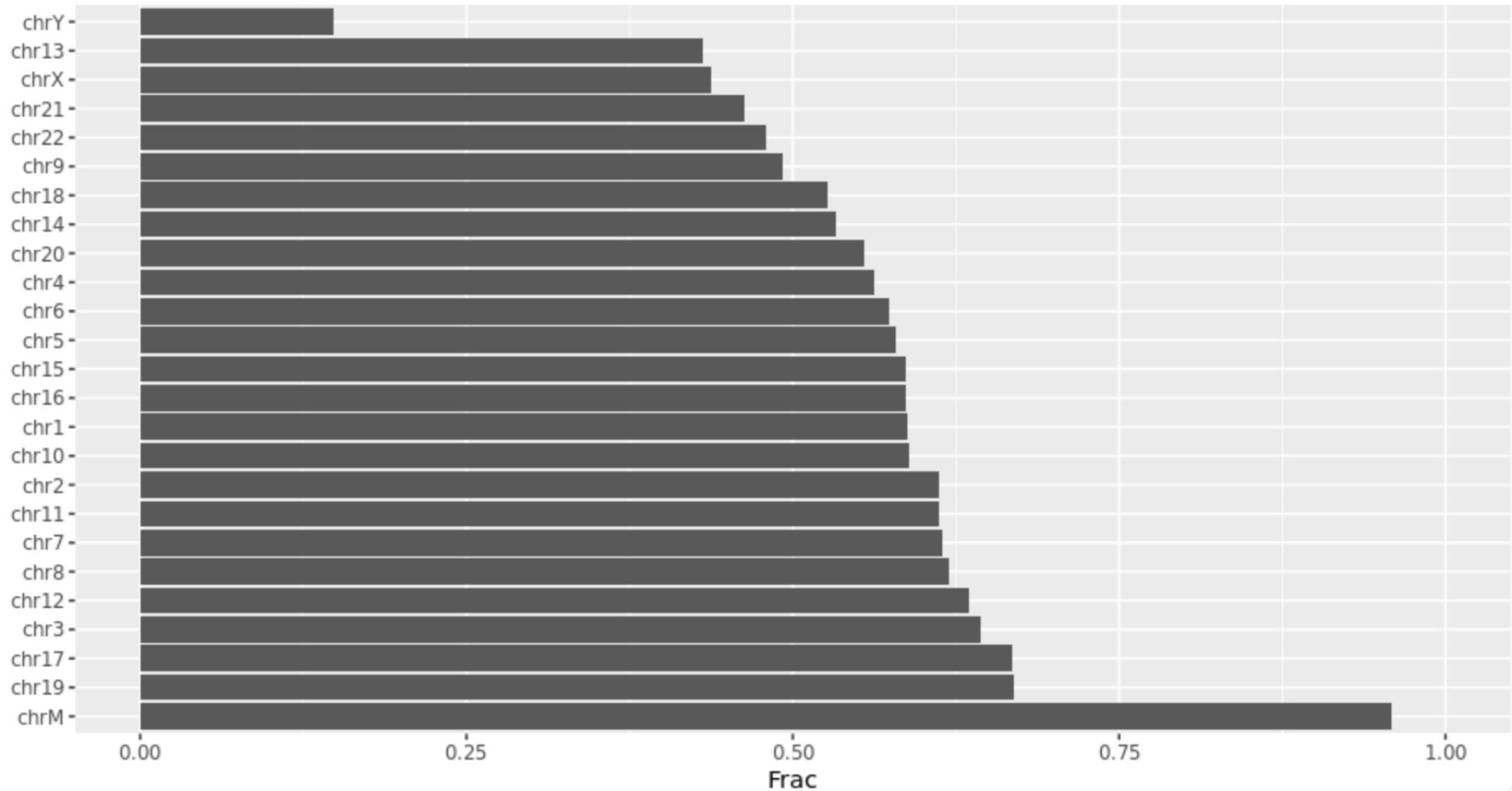
Enhancer
 Transcription Factor Binding Site

Promoter

Chromosome bands
Contigs
Genes
(Comprehensive set
from GENCODE 31)



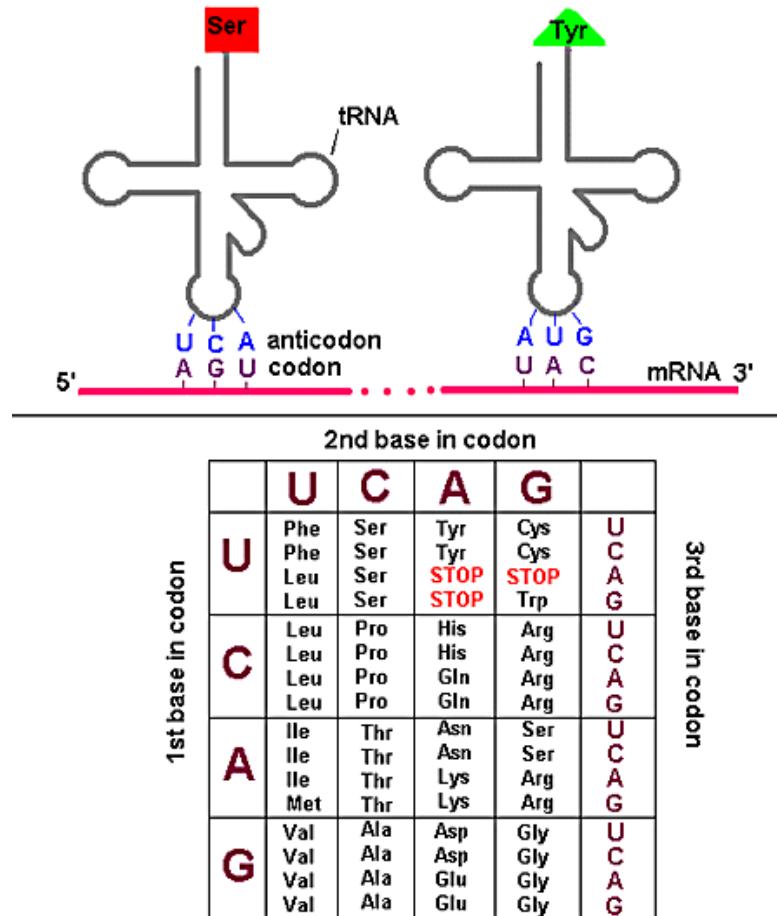
Fraction of Chromosomes that can be transcribed



Translation: The Genetic Code

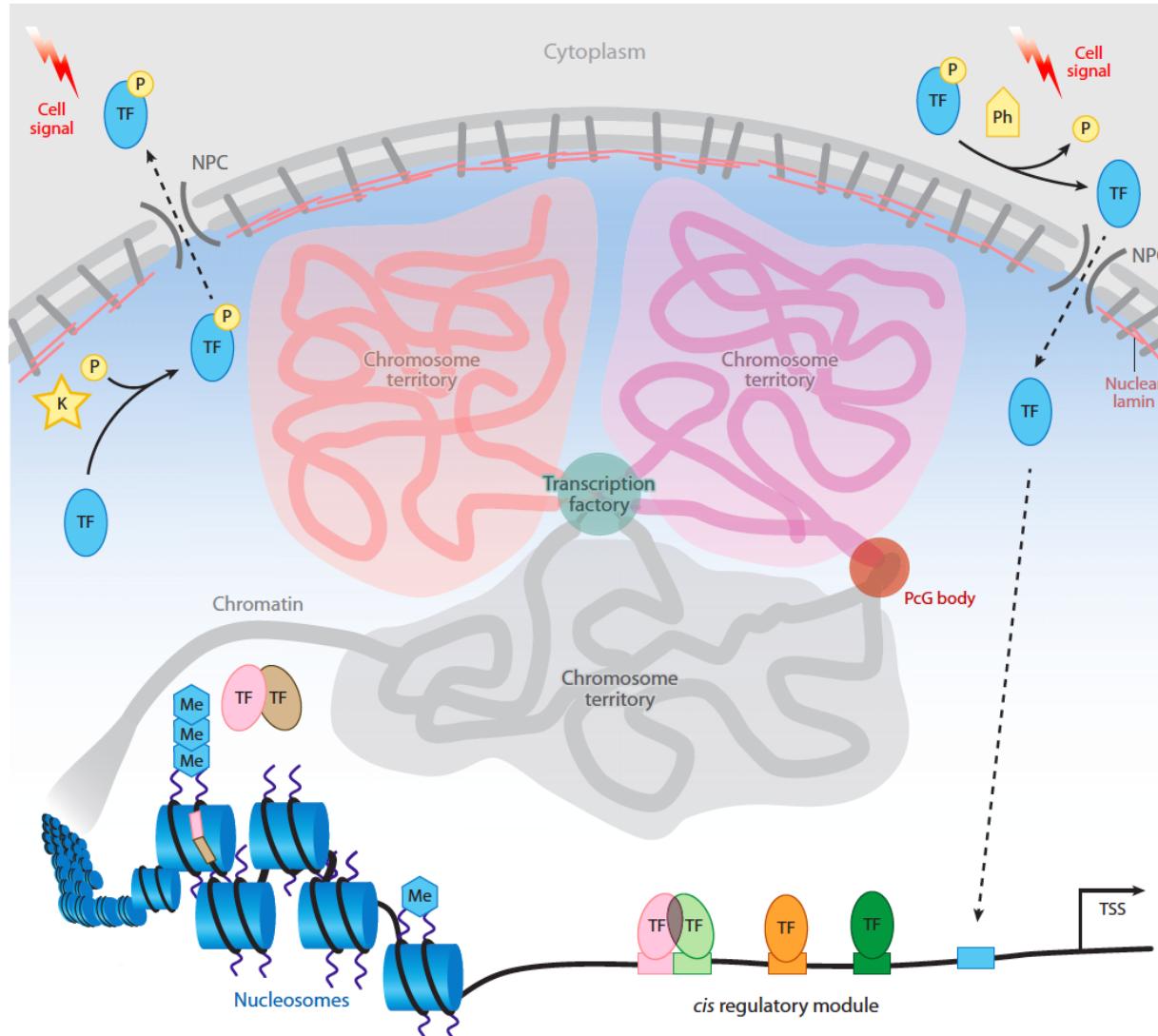
The translation process generates a protein based on the information in the messenger RNA

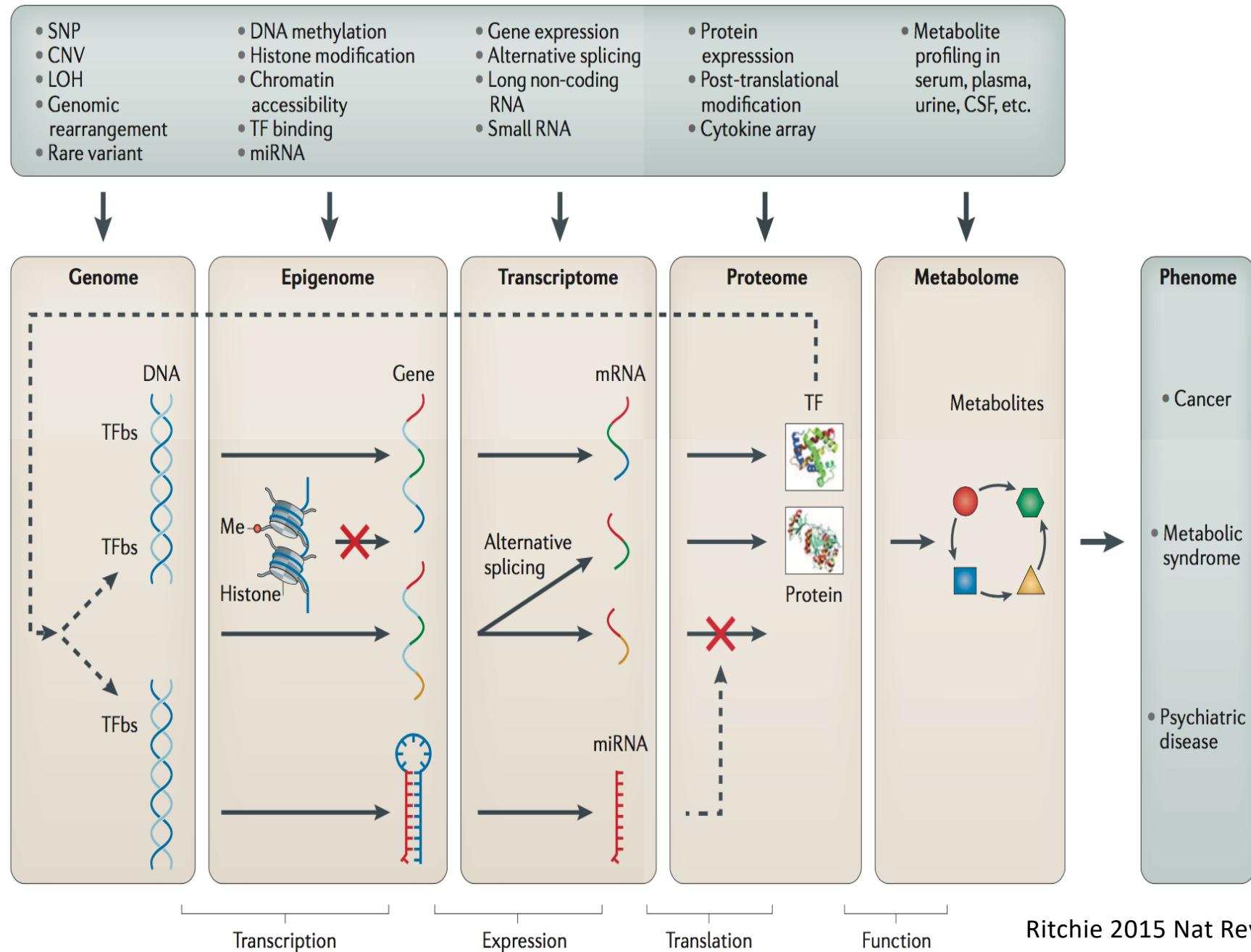
- A protein is a linear polymer of amino acids linked together by peptide bonds.
- Proteins are the main functional chemicals in the cell, carrying out many functions, for example catalysis of the reactions involved in metabolism.
- Proteins have a complex spatial structure



The Genetic Code

Transcriptional Regulation

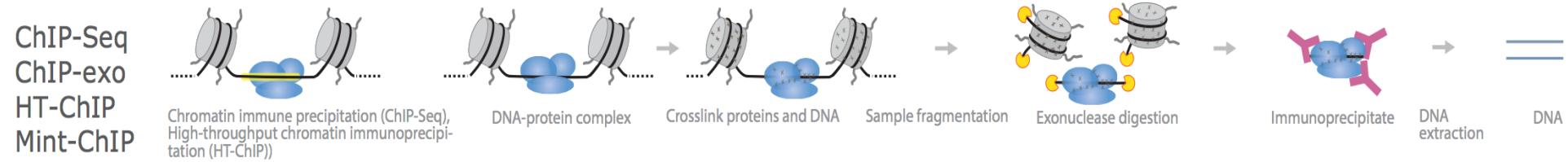




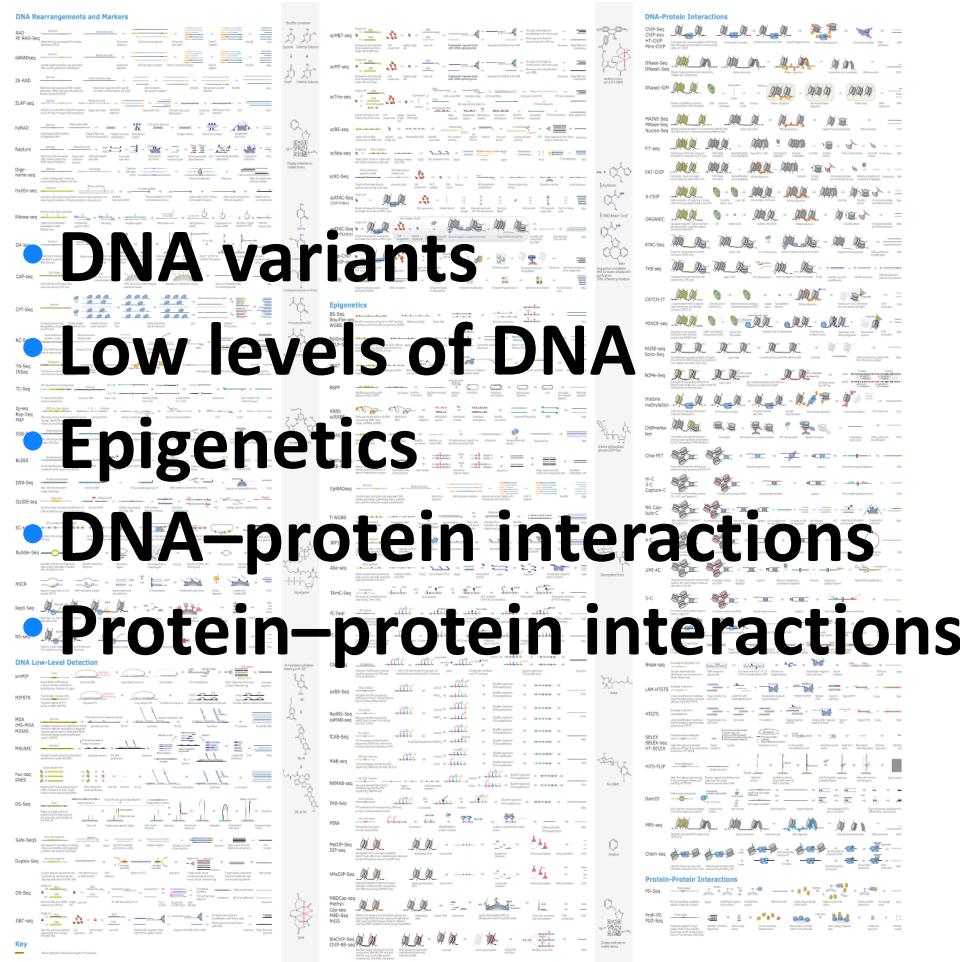
NGS Protocols

- Example: Preparation of DNA for a ChIP-seq experiment

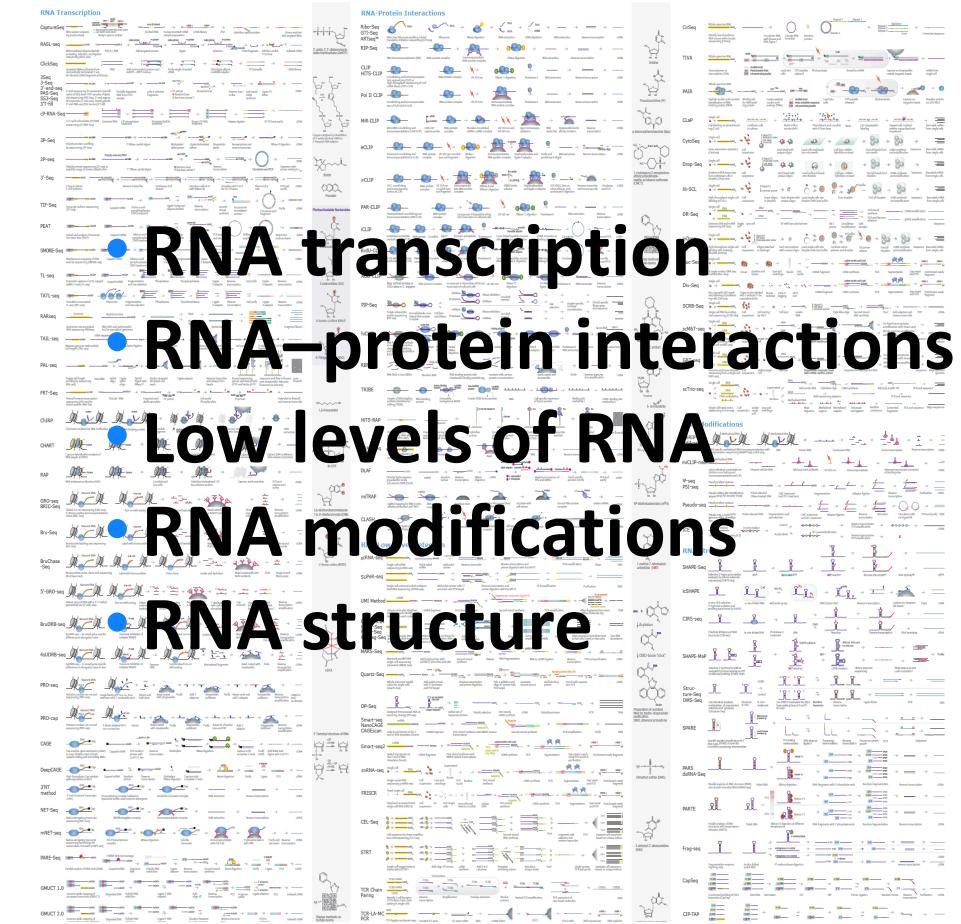
DNA-Protein Interactions



- The preparation determines how sequenced reads have to be interpreted



- DNA variants
- Low levels of DNA
- Epigenetics
- DNA–protein interactions
- Protein–protein interactions



- RNA transcription
- RNA–protein interactions
- Low levels of RNA
- RNA modifications
- RNA structure