

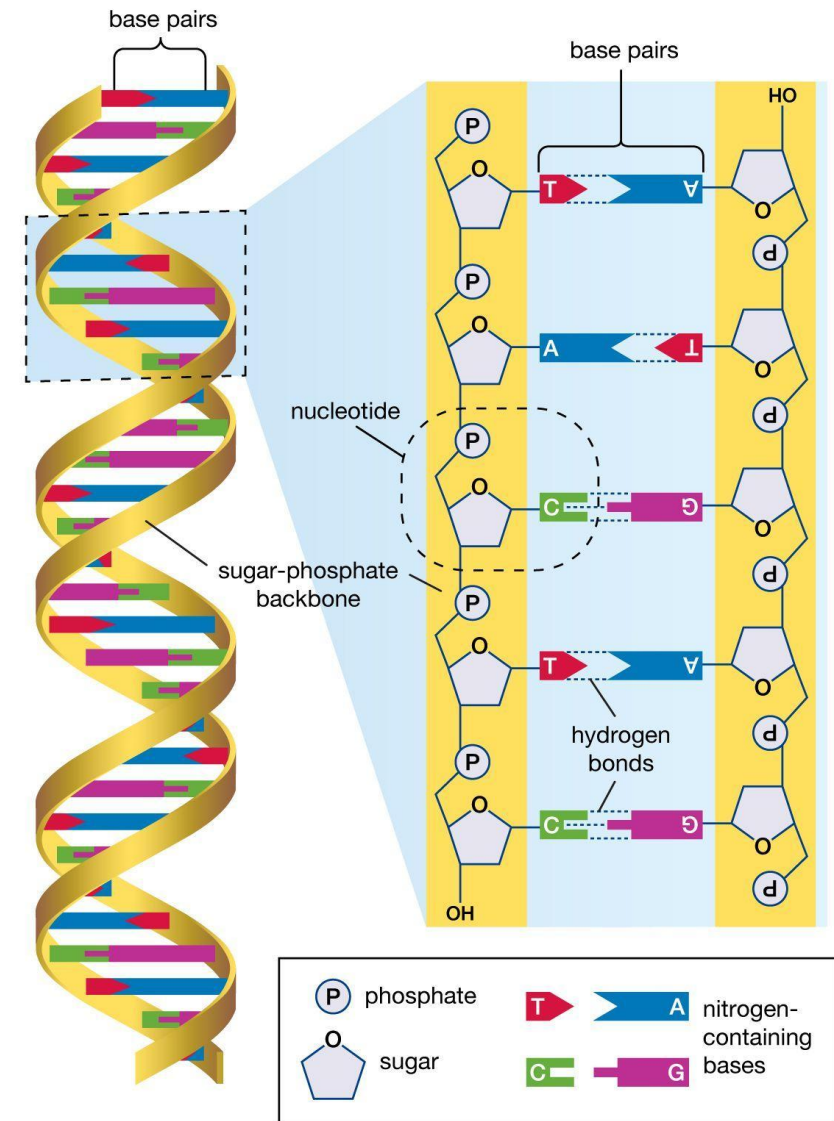
Human genome structure, function and clinical considerations

Phuc Loi Luu, PhD

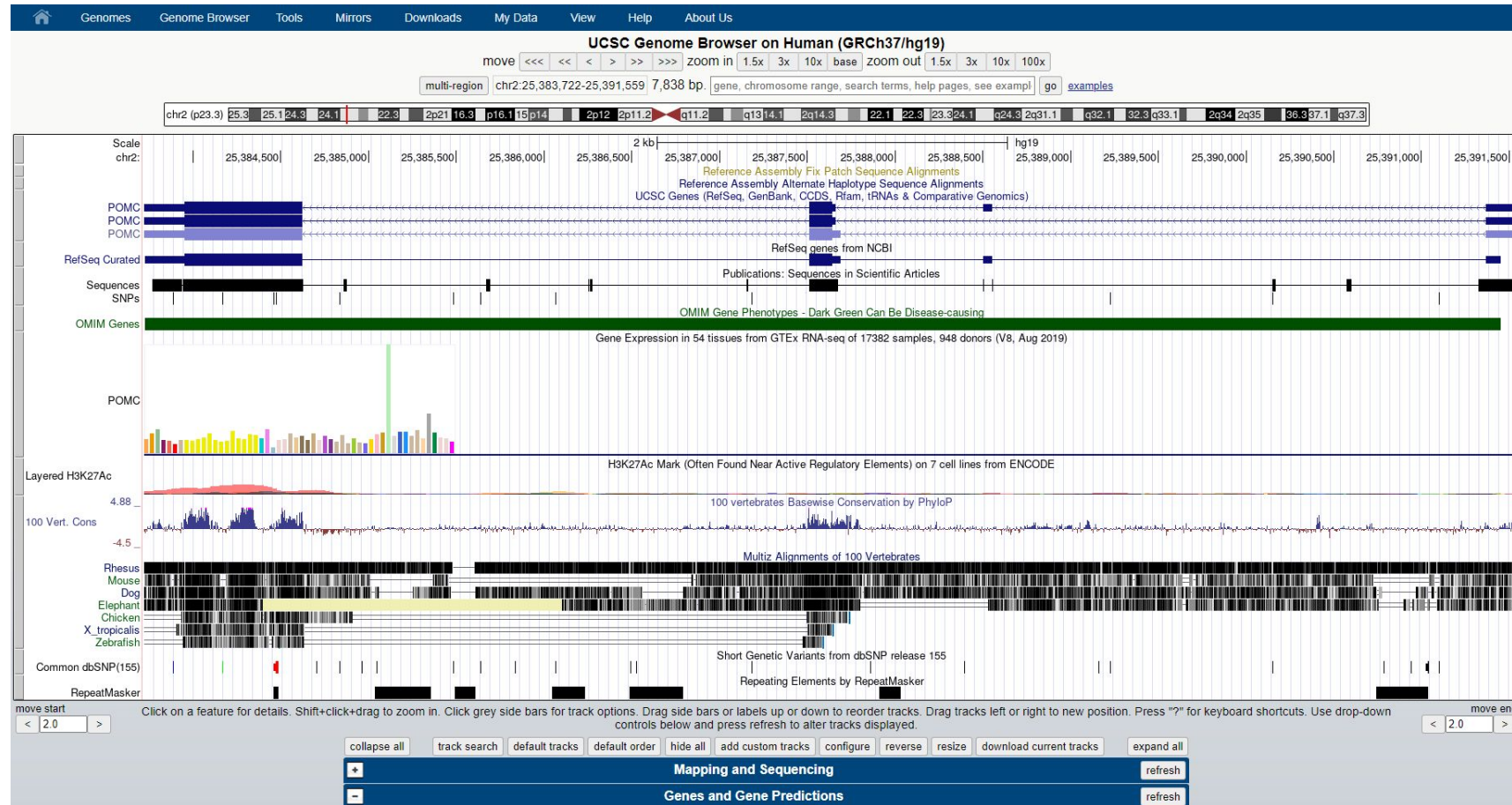
09/04/2023

Content

- Human reference genome and gene annotation
- DNA structure
- Human genome project (HGP)
- Human genome in a nutshell
- Genome components
- Genome 3D structure
- Gene structure



Understanding Human reference genome through the UCSC Genome Browser (<https://genome.ucsc.edu/>)



Understanding Gene Annotation through GENCODE (<https://www.gencodegenes.org/>)

[Human](#)[Mouse](#)[How to access data](#)[FAQ](#)[Documentation](#)[About us](#)

HUMAN

GENCODE 43 (08.02.23)



MOUSE

GENCODE M32 (08.02.23)



The goal of the GENCODE project is to identify and classify all gene features in the human and mouse genomes with high accuracy based on biological evidence, and to release these annotations for the benefit of biomedical research and genome interpretation.

GENCODE are [updating the annotation](#) of human protein-coding genes linked to SARS-CoV-2 infection and COVID-19 disease.

The [Long-read RNA-seq Genome Annotation Assessment Project](#) (LRGASP) Consortium for systematic evaluation of different methods for transcript computational identification and quantification using long-read sequence data **has launched**.

GENCODE are [supporting the annotation](#) of non-canonical human ORFs predicted by Ribo-seq data.

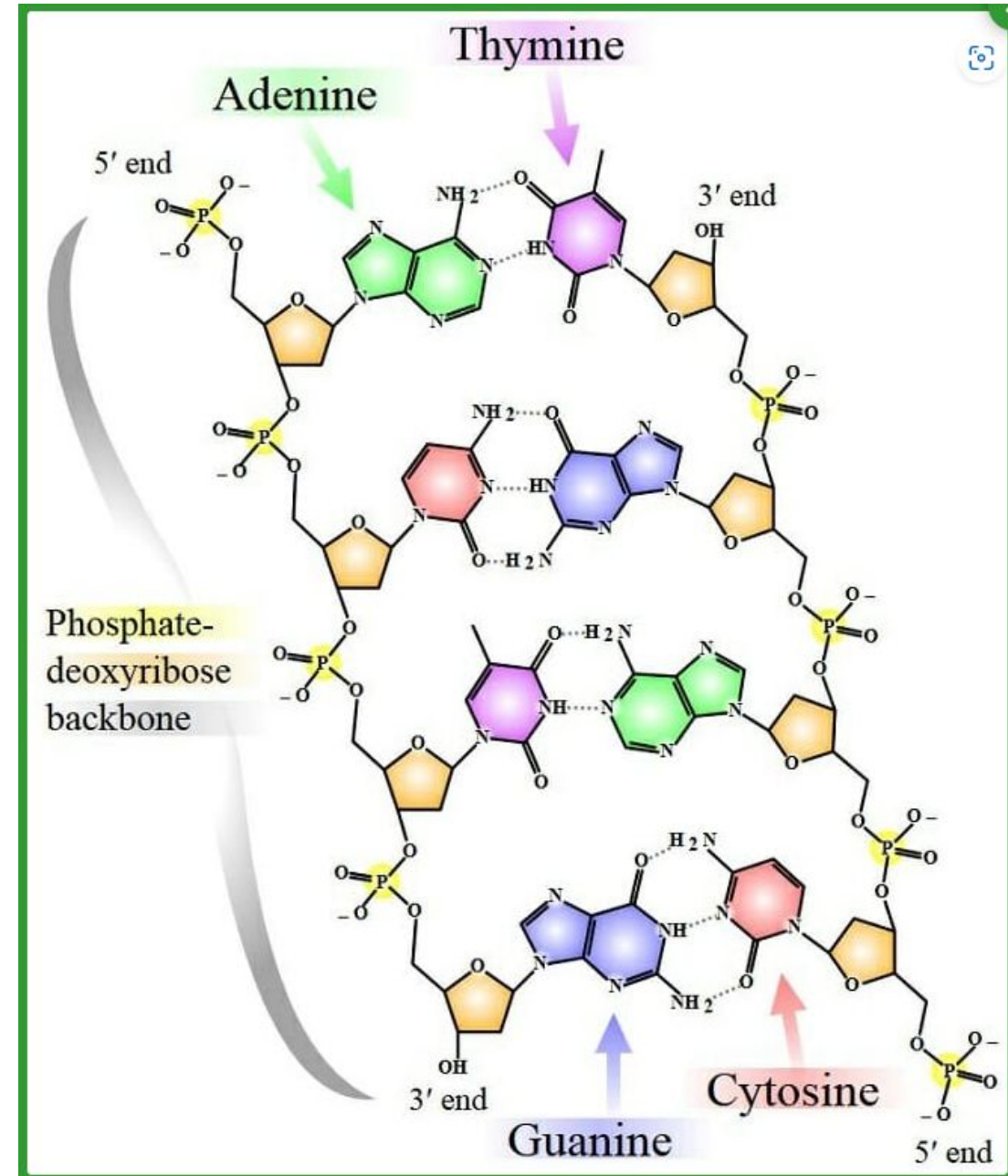
Race to discover the 3D structure of DNA

Figure 3. Rosalind Franklin at the Cabane des Evettes in the French Alps (c 1951). Photo credit: Vittorio Luzzati.



DNA Structure- Double Helical Watson & Crick Model

Figure 4. James Watson and Francis Crick with their DNA double helix model, 1953. Photo credit: A. Barrington Brown/Science Photo Library



Structure of DNA

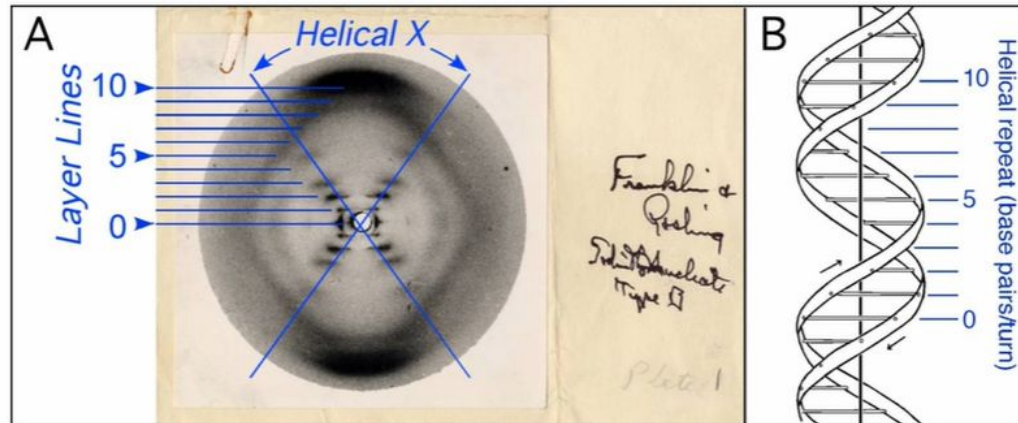


Figure 5. (A) Photo 51, the diffraction pattern from DNA in its so-called B configuration. The dimensions of DNA are: pitch $P = 3.4$ nm, radius $R = 1$ nm. Several important features include the characteristic X-shape and the ten diffracted orders per X. (B) A schematic diagram of DNA helical repeat structure. Photo is adapted from (Watson and Crick, 1953).

Structure of DNA

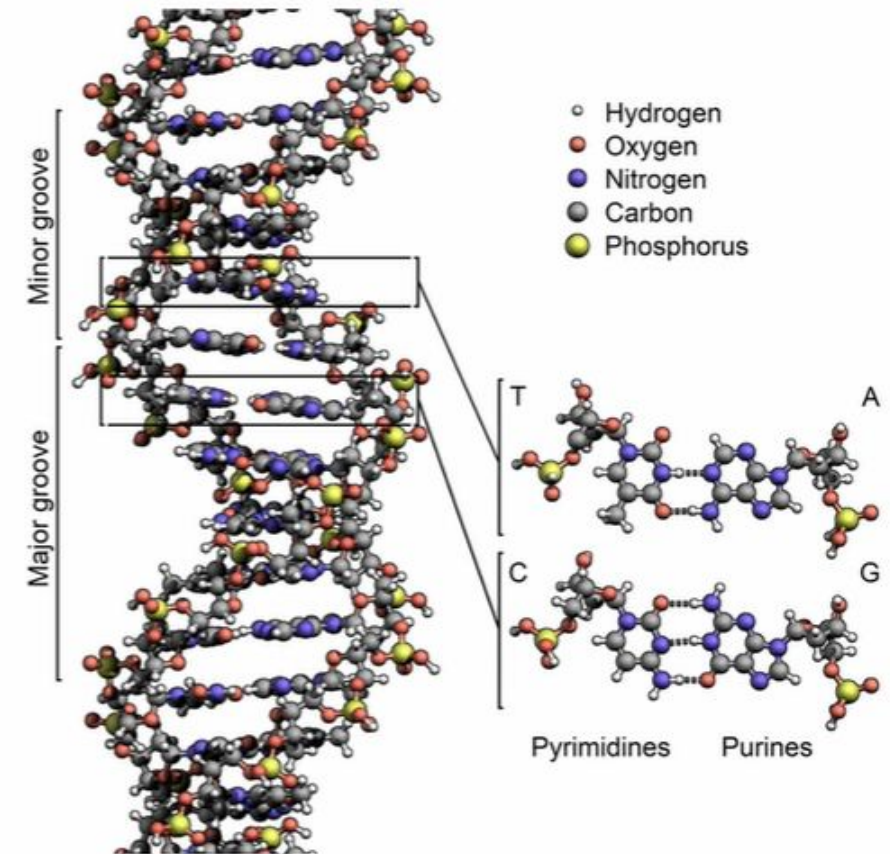
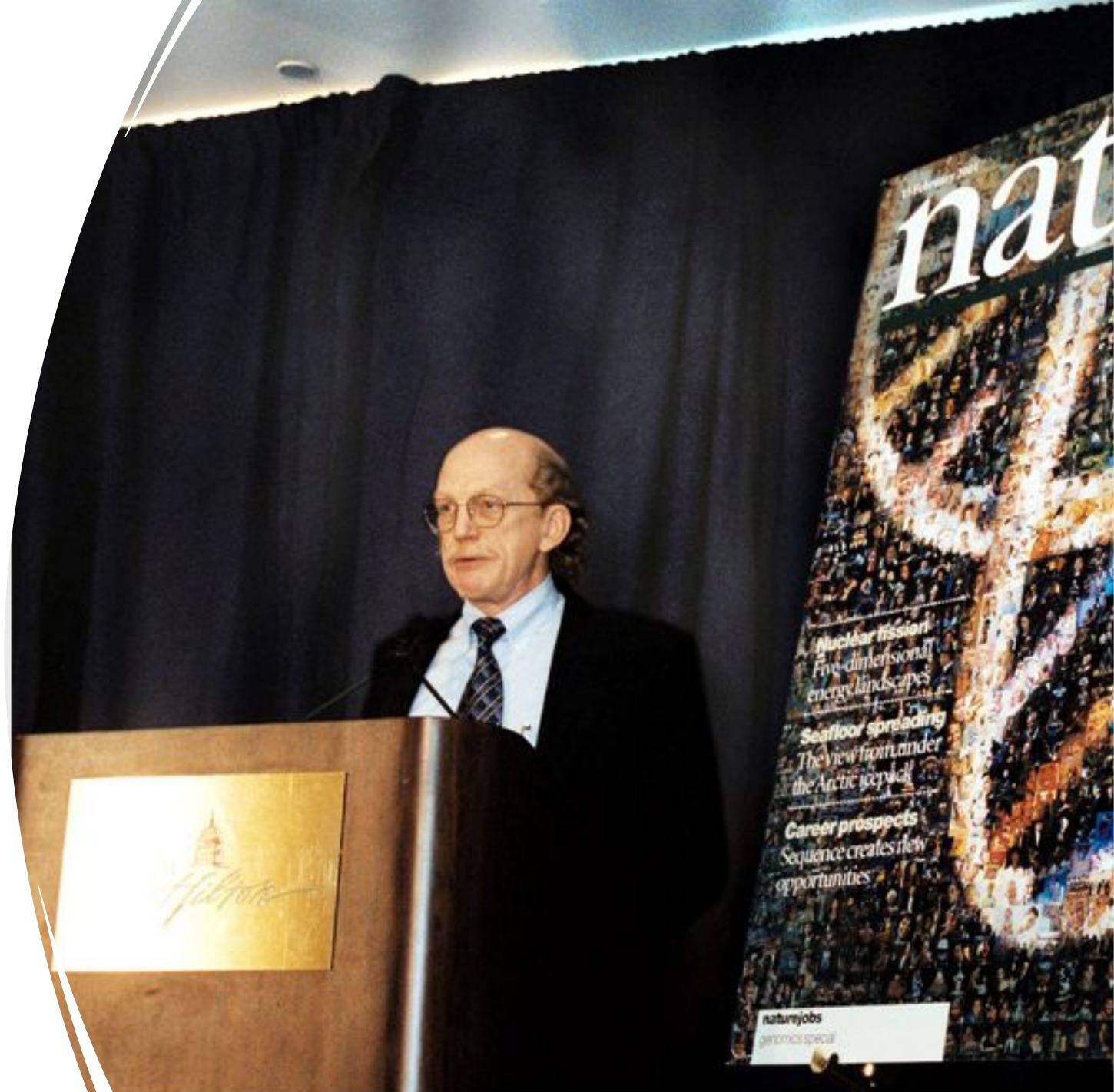


Figure 6. The structure of DNA. Image adapted from Zalli et al (2020).

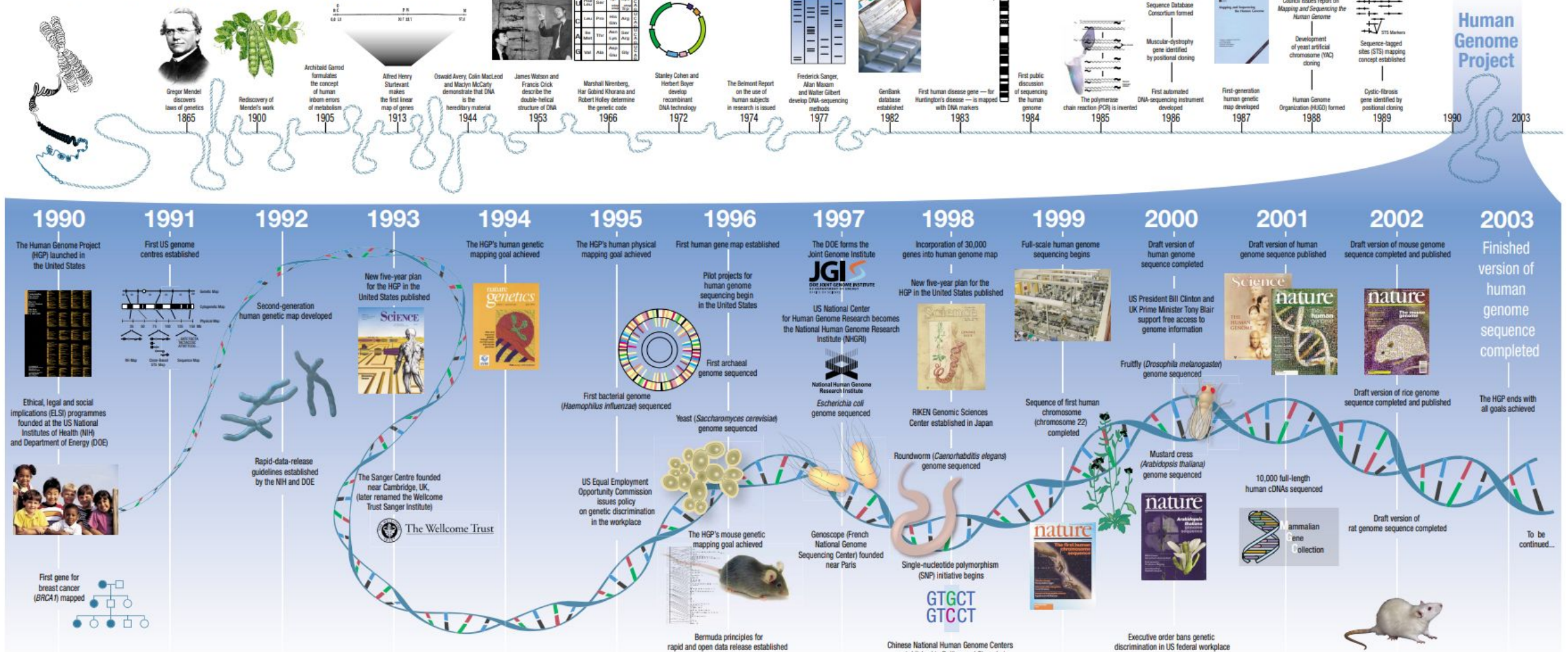
Human genome project (HGP)

- Was large, well-organized, and highly collaborative international effort that generated the **first sequence of the human genome** and that of several additional well-studied organisms.
- Carried out from 1990–2003, it was one of the most ambitious and important scientific endeavors in human history.

Human Genome Project Timeline



Landmarks in genetics and genomics



PEAS COURTESY J. BLAMIRE, CITY UNIV. NEW YORK; WATSON & CRICK COURTESY A. BARRINGTON BROWN/SP; SCIENCE COVERS COURTESY AAAS

Human genome project (HGP)

- The sequence of the human genome generated by the Human Genome Project was not from a single person.
- Rather, it reflects a patchwork from multiple people whose identities were intentionally made anonymous to protect their privacy.
- The project researchers used a thoughtful process to recruit volunteers, acquire their informed consent, and collect their blood samples.
- Most of the human genome sequence generated by the Human Genome Project came from blood donors in Buffalo, New York; specifically, 93% from **11 donors**, and 7% from one donor.

Human genome project (HGP)

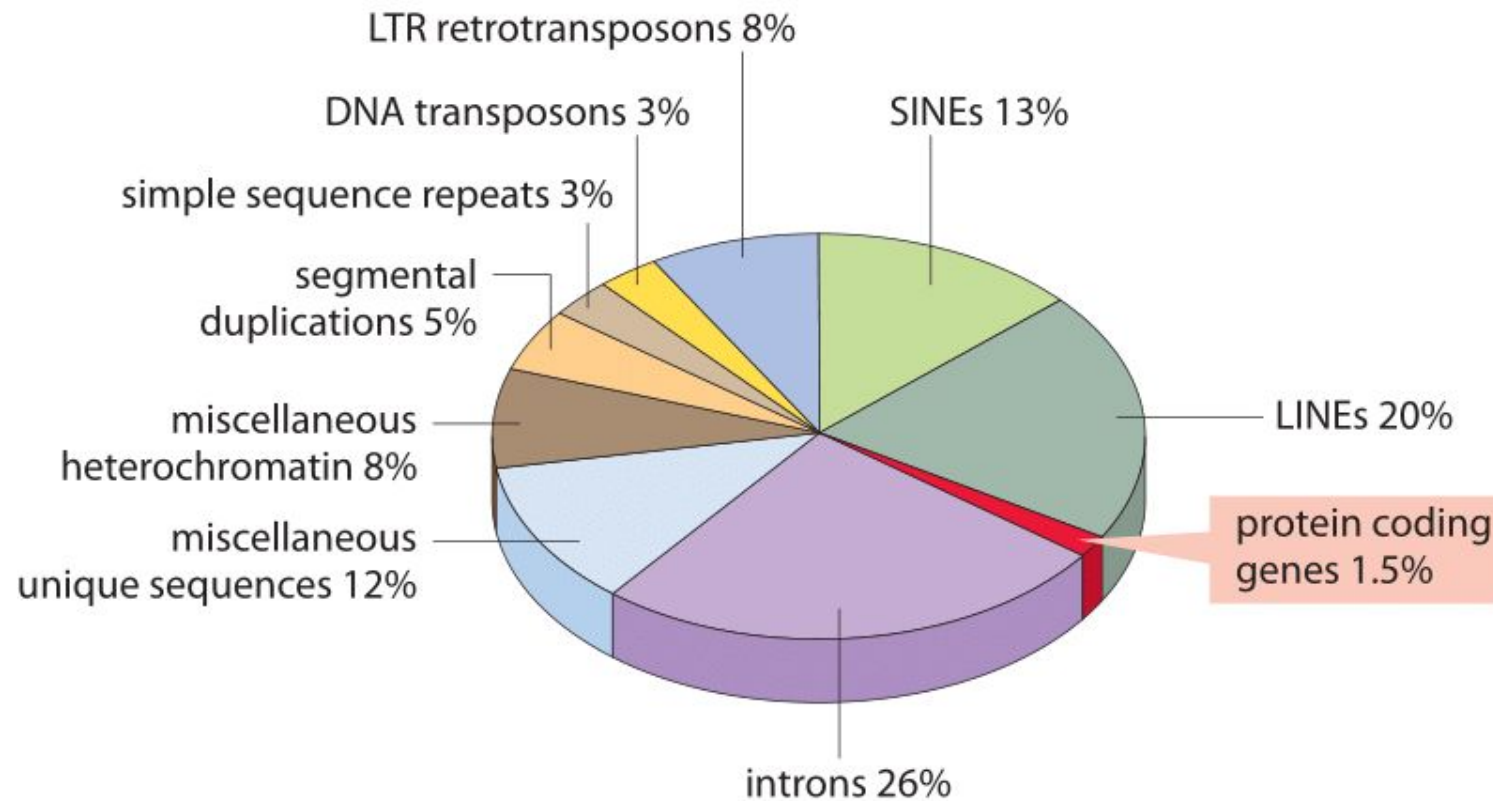
- The Human Genome Project could not have been completed as quickly and effectively without the dedicated participation of an international consortium of thousands of researchers.
- In the United States, the researchers were funded by the Department of Energy and the National Institutes of Health, which created the Office for Human Genome Research in 1988 (later renamed the National Center for Human Genome Research in 1990 and then the National Human Genome Research Institute in 1997).
- The sequencing of the human genome involved researchers from 20 separate universities and research centers across the United States, United Kingdom, France, Germany, Japan and China.
- The groups in these countries became known as the International Human Genome Sequencing Consortium.

Human genome project (HGP)

- In June 2000, the International Human Genome Sequencing Consortium **announced** that it had produced a draft human genome sequence that accounted for 90% of the human genome. The draft sequence contained more than 150,000 areas where the DNA sequence was unknown because it could not be determined accurately (known as gaps).
- In April 2003, the consortium **announced** that it had generated an essentially complete human genome sequence, which was significantly improved from the draft sequence. Specifically, it accounted for 92% of the human genome and less than 400 gaps; it was also more accurate.
- On March 31, 2022, the Telomere-to-Telomere (T2T) consortium announced that had filled in the remaining gaps and produced the **first truly complete human genome sequence**.

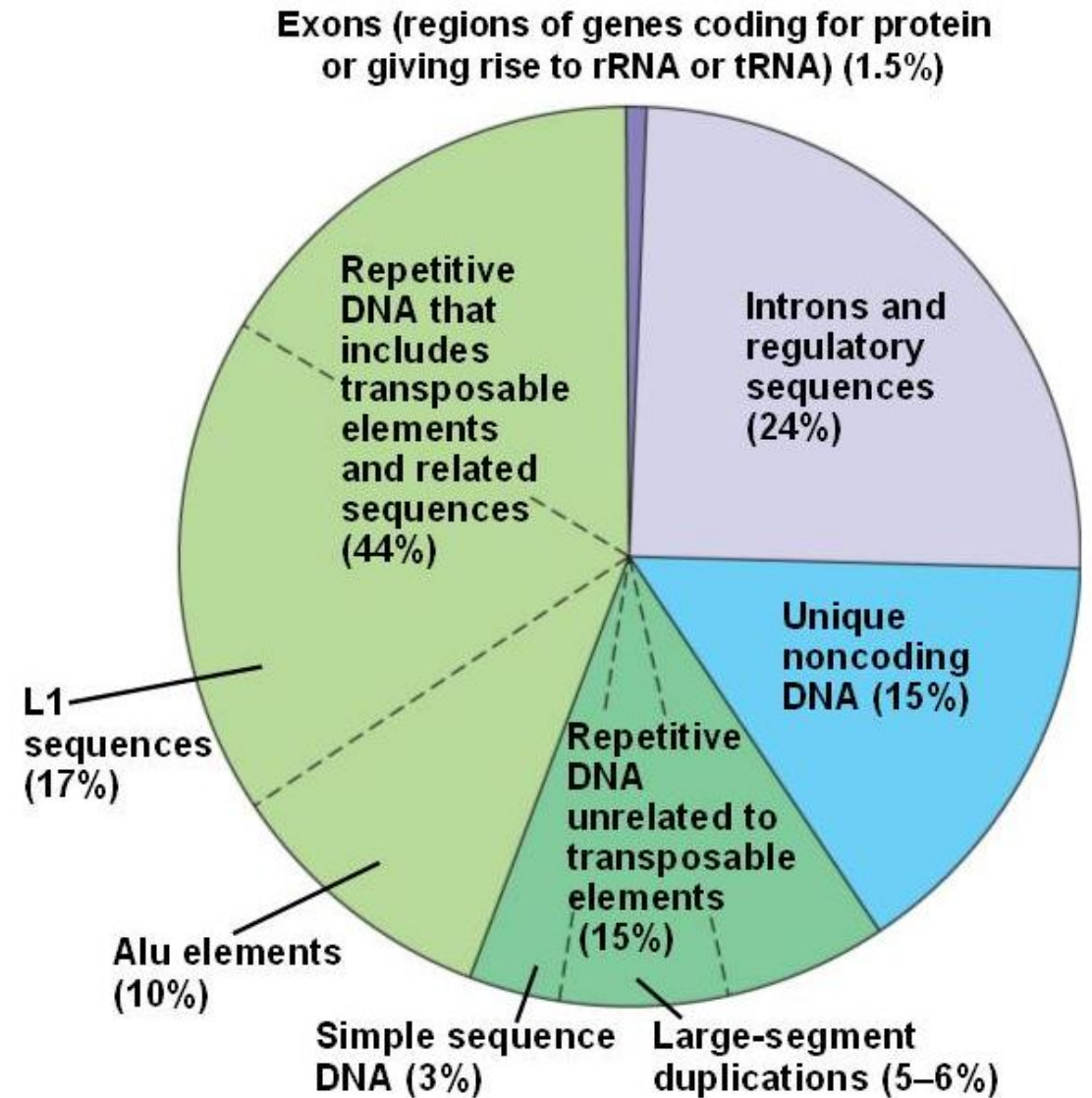
Genome components

main components of the human genome



Genome components

- The human genome contains less than 2% coding exons within genes.
- The remaining DNA consist of:
 - introns and regulatory sequences such as enhancers
 - unique noncoding DNA contains many pseudogenes (genes that have accumulated mutations and became nonfunctional)
 - repetitive DNA are sequences that are repeated many times; much of these are
 - transposable elements that can "jump" between chromosomes, leading to transpositions.

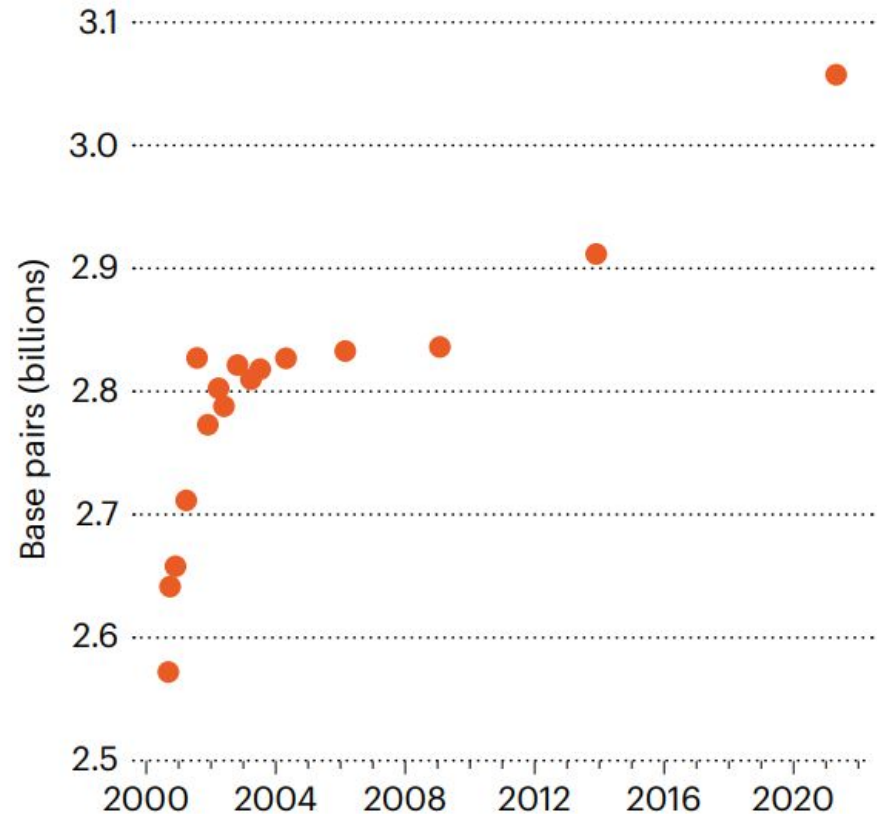


A COMPLETE HUMAN GENOME IS CLOSE: HOW THE GAPS WERE FILLED

Researchers added 200 million DNA base pairs and 115 genes – but they've yet to finish the Y chromosome.

COMPLETING THE HUMAN GENOME

Researchers have been filling in incompletely sequenced parts of the human reference genome for 20 years, and have now almost finished it, with 3.05 billion DNA base pairs.



0.3% of sequence might still have errors. Includes X but not Y chromosome. Count excludes mitochondrial DNA.

SOURCE: ADAM PHILLIPPY

(35) Lessons from the Human Genome Project - YouTube

- <https://www.youtube.com/watch?v=qOW5e4BgEa4>

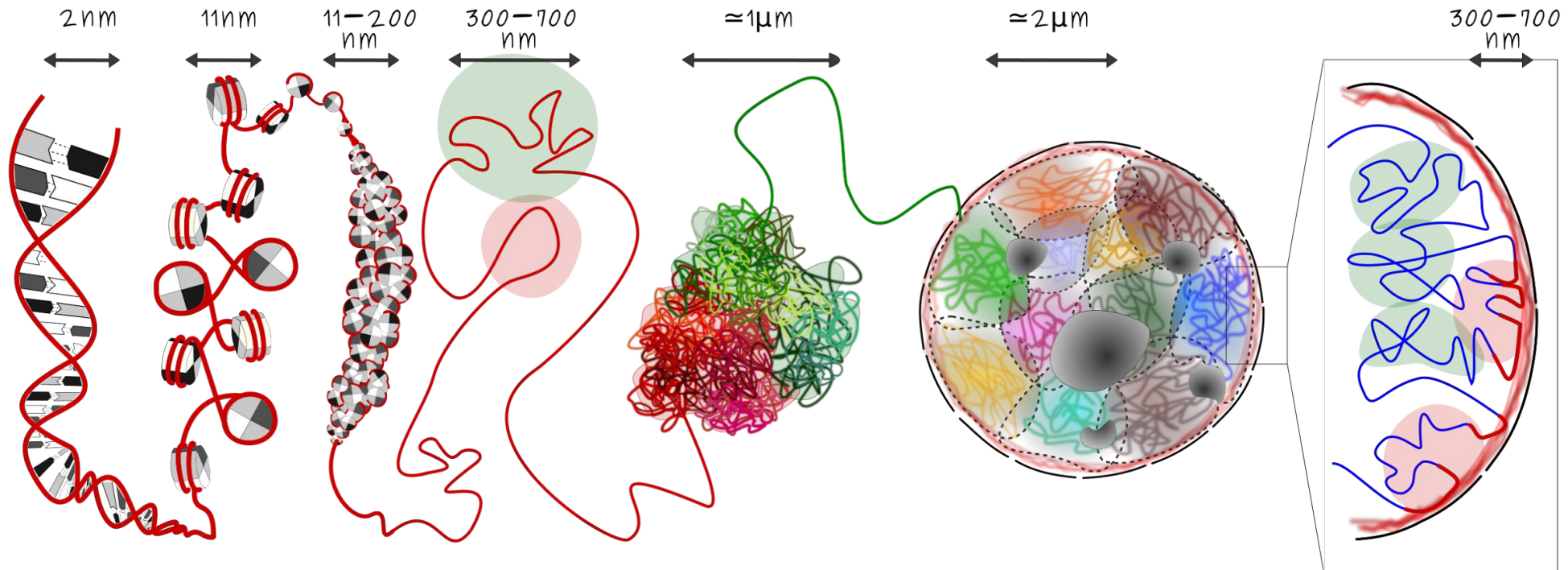
How did the Human Genome Project affect biological research in general?











- Discuss and give an example!

How did the Human Genome Project affect medical research and clinical application in general?

- Discuss and give an example!

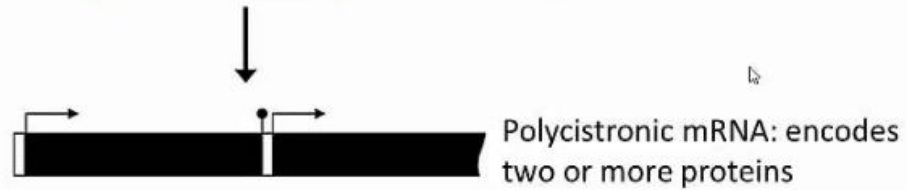
DNA Chromatin fiber DNA loop & Compartment, Chromatin/TAD domains Chromosome territory inside cell nucleus Lamina-associated domain(LAD)



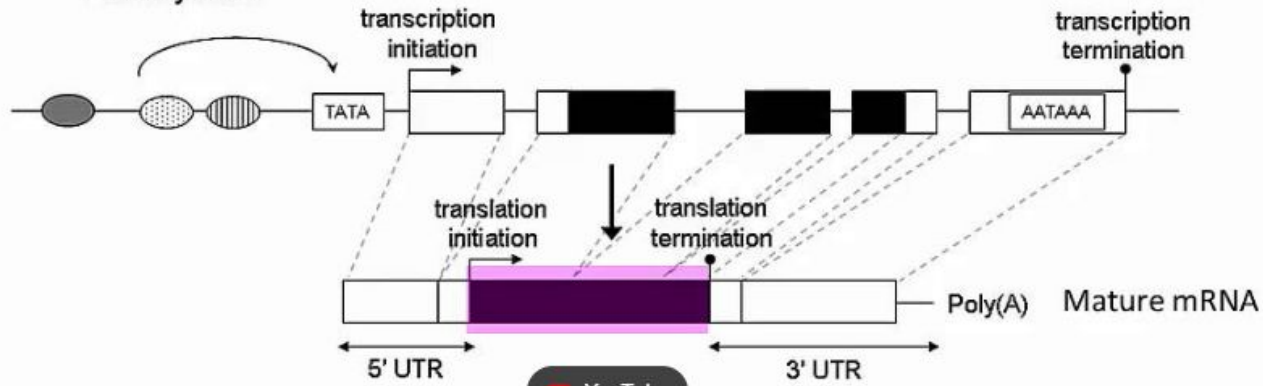
-  Nucleobases
-  Nucleosome, composed of 8 histones protéines
-  Repressive chromatin domain, mainly heterochromatin
-  Active chromatin domain, mainly euchromatin
-  Chromosome territory
-  Nucleolus
-  Nuclear speckles
-  Nuclear membrane, including schematic nuclear pores
-  Nuclear lamina
-  LAD

Gene structure

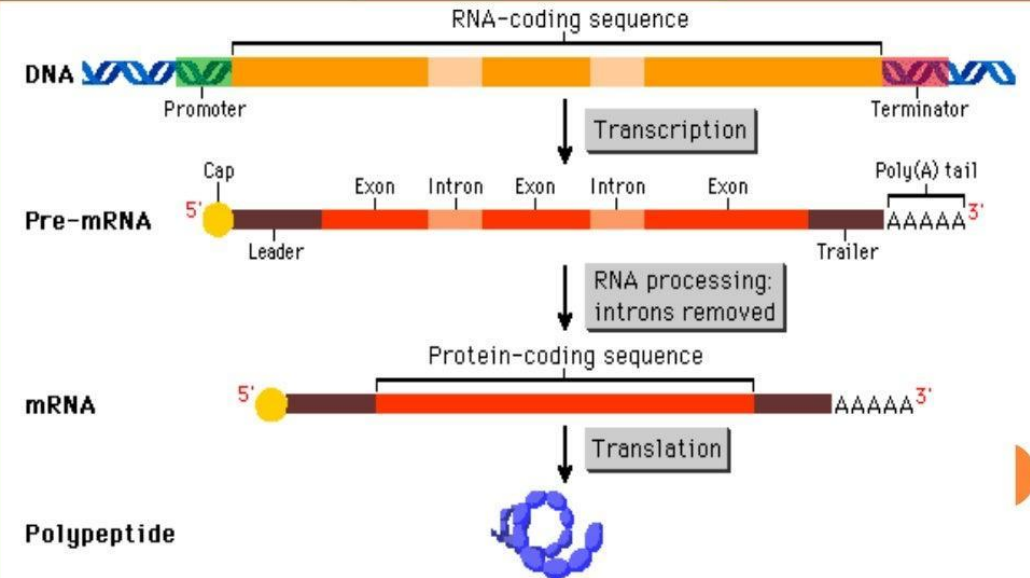
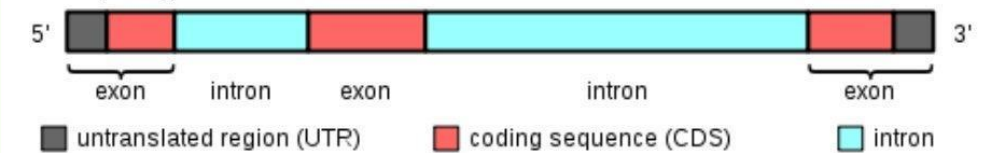
Prokaryotes:



Eukaryotes:



Eukaryotic gene.



Gene structure

