Computer genomics

Yonathan Berith Jaramillo Ramirez

September 27, 2021

1 Rembering old topics on DNA

What is DNA made of?

- Pentose sugar (the *sides* of the ladder)
- Phosphate group (the *sides* of the ladder)
- Nitrogenous bases (the rungs of the ladder)

What is a polymer:

Any of a class of natural or synthetic substances composed of very large molecules, called macromolecules.

What is a Nitrogenous base made of?

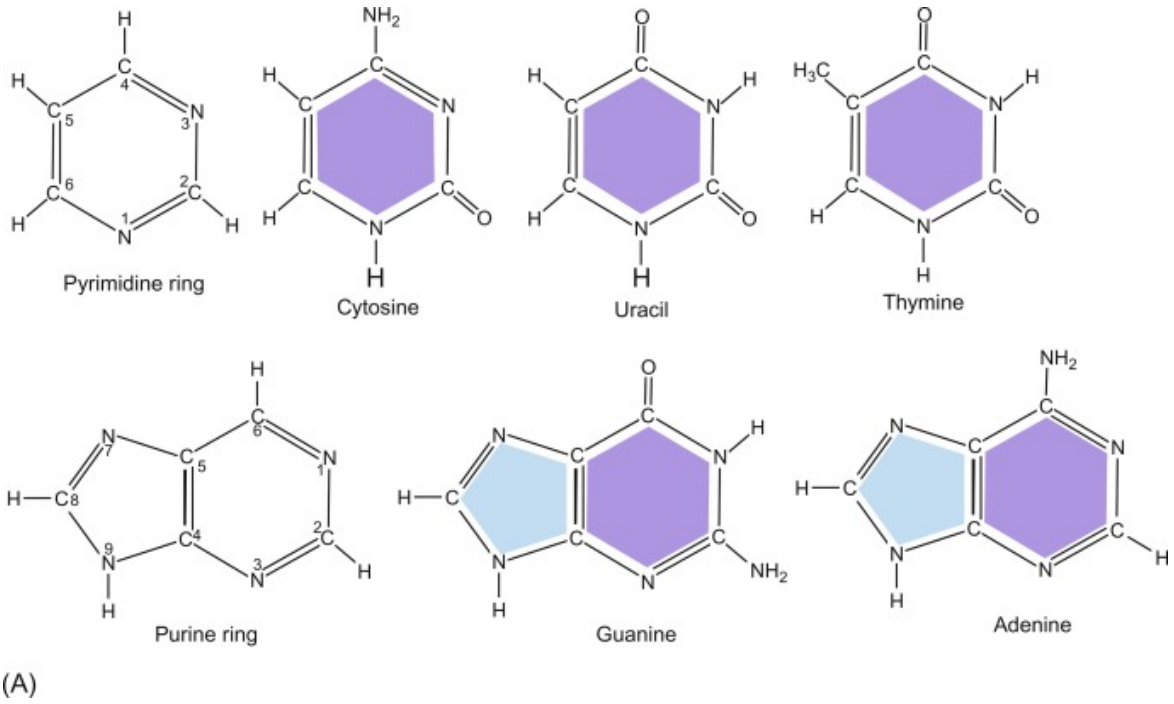

Figure 1: Nitrogenous bases are made of...

What process can you do with DNA?

- Translation (Protein Synthesis)
- Transcription (Synthesizes RNA)
- Central Dogma (Transcription + Translation)

Where on earth can I find DNA?

The nucleus of a cell contains Chromosomes and DNA is contained within the Chromosomes.

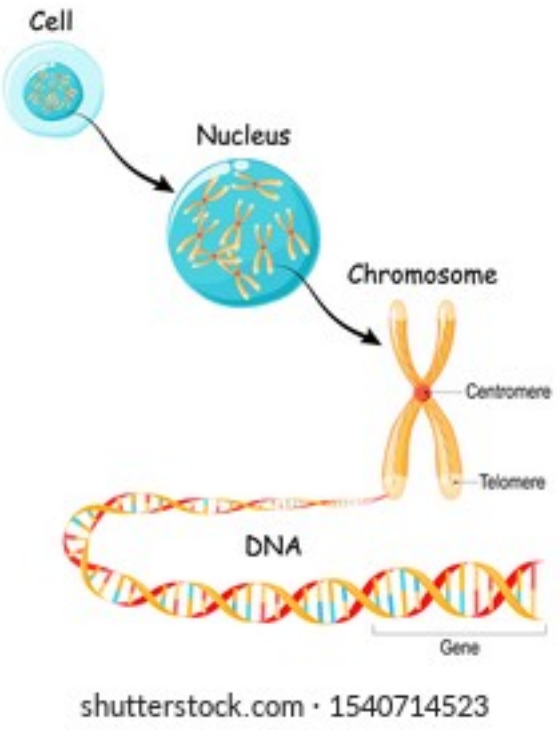

Figure 2: DNA...

Nucleotide composition

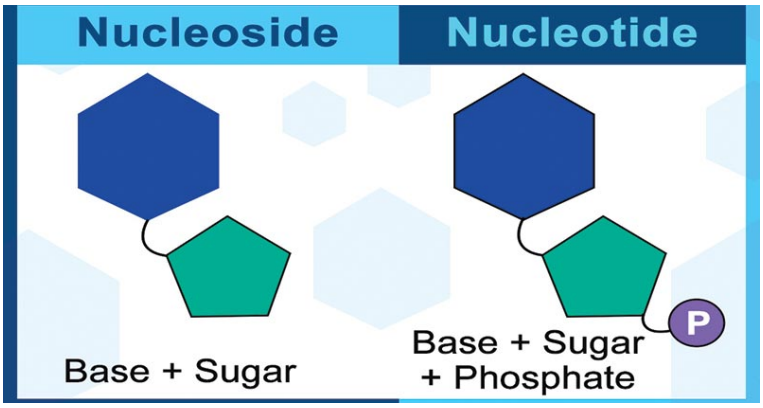

Figure 3: Here we see a nucleotide and a nucleoside

Remember!

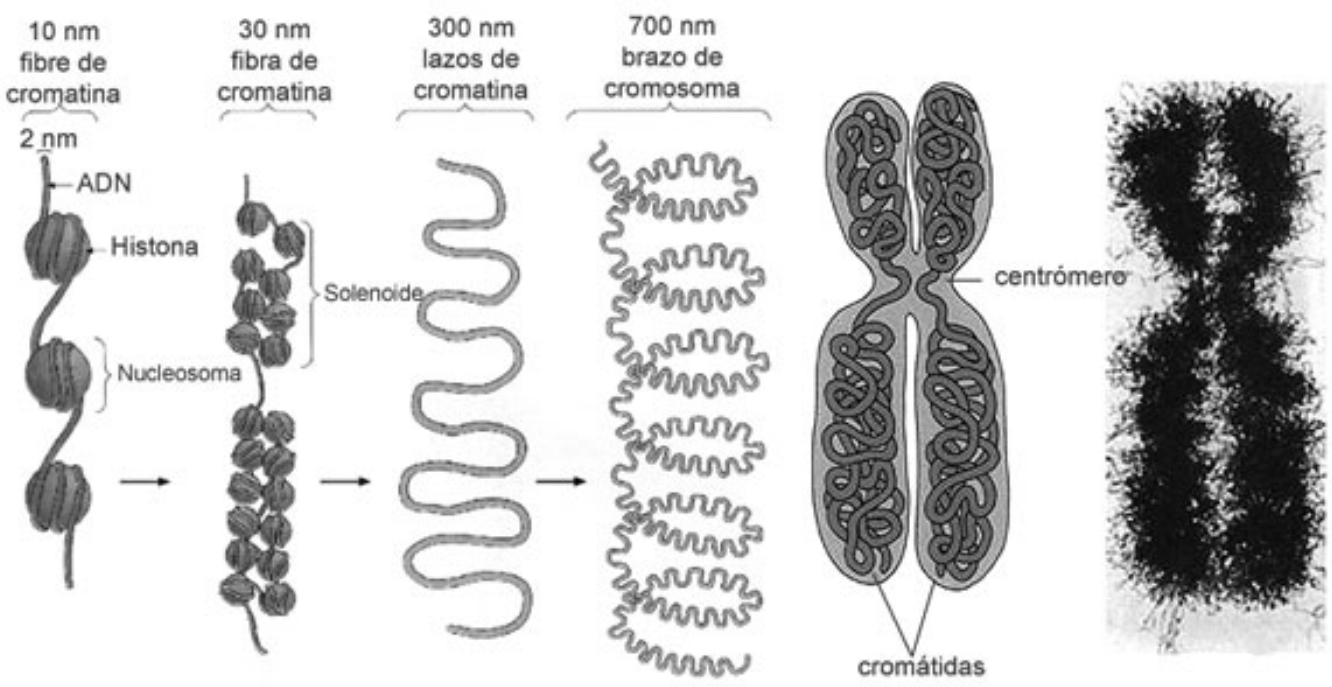

Figure 4: A deeper look into Chromosomes

2 Beggining and struggles in the age of genomics.

Notes about the video:

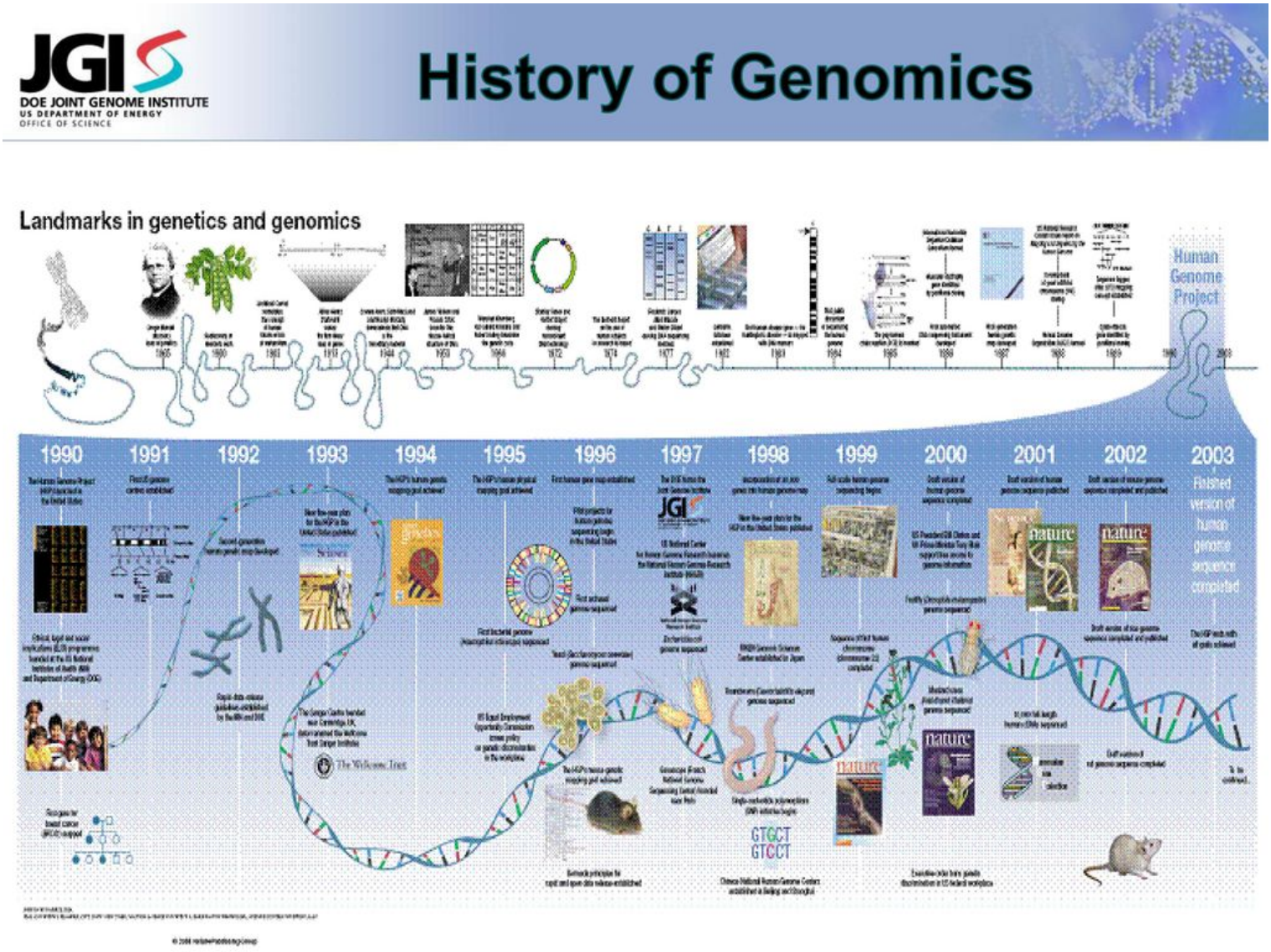

Figure 5: History of Genetics

History

3 Important years

1865: Laws of genetics (Mendel laws).
1900: Darwin theory + Mendel laws join together
1913: Alfred Henry first linear genetic maps
1953: Watson and Crick discover the structure of DNA

1966: Determination of the genetic code
1977: Creating of the first methods to sequence DNA
1982: GenBank database was established
1985: Discovery of the chain reaction polymerase
1988: They start to talk about the human genome project
1990: The human genome starts and the ELS(Ethical, legal and social implications)
1993: The wellcome Trust Sanger Institute
1997: Joint Genome institue and National Humans Genome Research Institute (JGI and NHGRI)
2003: Finished version of the human genome sequence completed

Challenges Biology: Concrete the structure and function of the genomes.
Health: Find benefits based on the human genome knowledge.
Society: Use genomics to maximize the benefis and minimize dangers in society.

- Humans share 99 of their DNA.
- Thymine gets along with Adenine. (hydrogen bonding)
- Guanine gets along with Cytosine. (hydrogen bonding)
- Model organisms: yeast, bacteria.

The biggest challenges:

- Indentify the functional and structural components in the HG.
- Understand the organization of the genetic networks and protein routes to see how they co-relate to the phenotype in organisms.
- To develop a detailed hereditary variation in the HG.
- To comprehend the genetic variation between species as a factor for evolution.
- To develop laws that allow the generalized use of the information about the HG for both research purposes and clinic purposes.
- Find diseases caused by pecific gens.
- Use probability to predict genetic diseases and drug side effects.
- Kill mosquitos using genetics haha
- Become super humans?
- Relationship (gens) - (human behavior).
- Find genetic variations that contribute on your helath and resistance against diseases.
- Ethics boundaries because of course we are mad scientists.

The solutions:

1. **Technological Developments:**
   - Sanger sequencing
   - Marker based genetics.
   - Cloning.
   - Chain reaction of the polymerase
2. **Making resources available:**
3. **Computational Biology**
4. **Education at different levels**
5. **ELSI (Ethic Legal Social implications)**

4 A brief history of bioinformatics lecture

Some cool stuff

- Hemoglobin, helps showing a degree of similarity over long evolutionary time using the gen variations.
- Pairwise protein sequence alignments:
  (dot-matrix methods, dynamic programming and word methods)
- Needleman–Wunsch algorithm.
  (computationally impractical)

  $$O(L^N)$$

  Where L is sequence length and N the amount of sequences

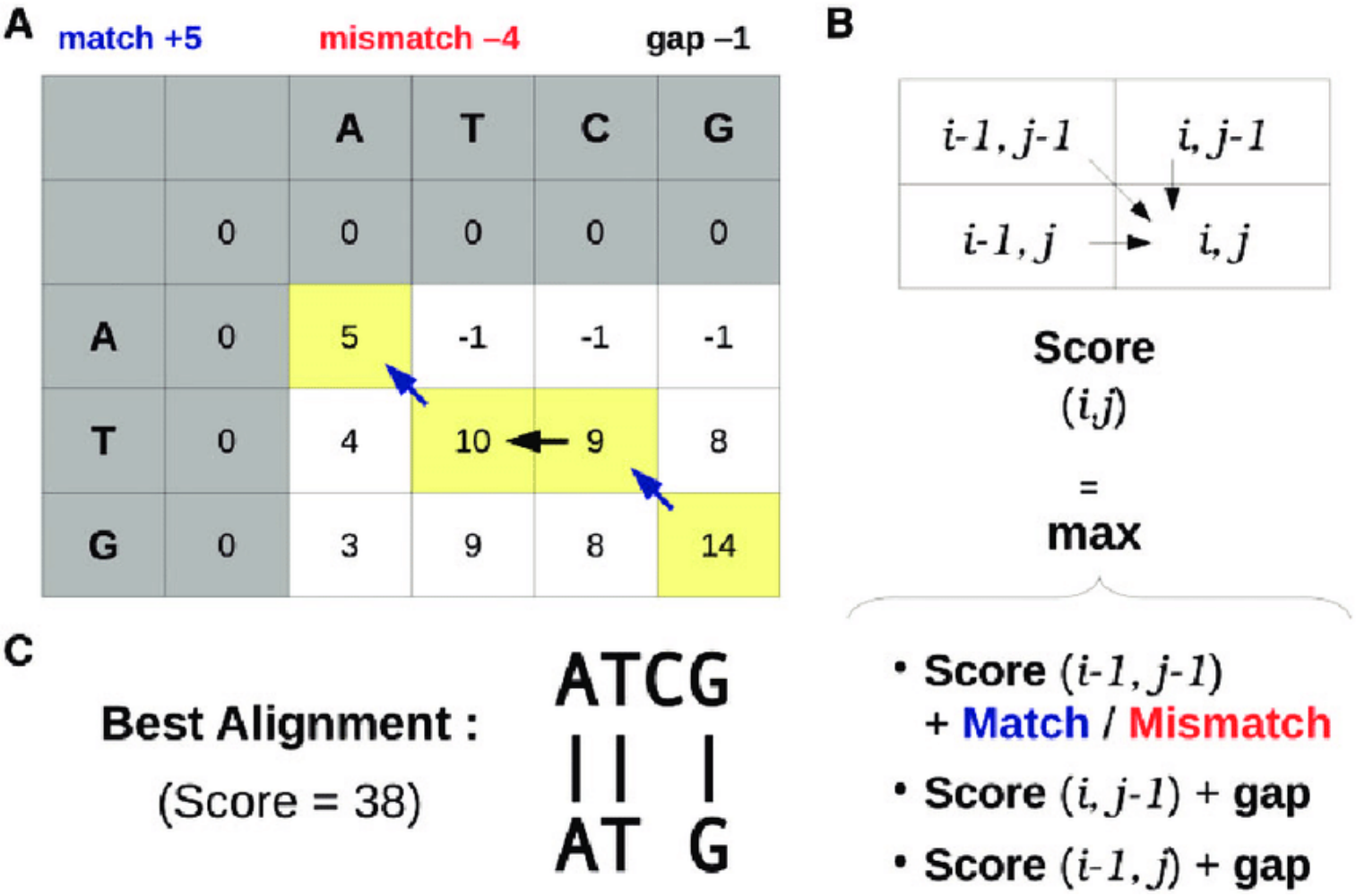

Figure 6: Needleman–Wunsch algorithm

Progressive sequence alignment

- Performing a Needleman–Wunsch alignment for all sequence pairs.
- Extracting pairwise similarity scores for eachpairwise alignment.
- Using those scores to build a guide tree and then...
- Aligning the two most similar sequences, and then the next more similar sequence, and so on, according to the guide tree.

CLUSTAL software (Feng–Doolittle algorithm)

Maxam–Gilbert sequencing method 1976