

The logo for VELSERA, featuring the word in a white, serif, all-caps font inside a black rectangular box.

VELSERA

SevenBridges

**Applied
Bioinformatics**



reveals

Agenda

- Introductions
- Course overview
- Bioinformatics intro
- Platform registration

Applied Bioinformatics

Introductions: Lecturers



Different backgrounds - bioinformatics engineers

- Bojana Smiljanić, BF



- Boris Majić, ETF



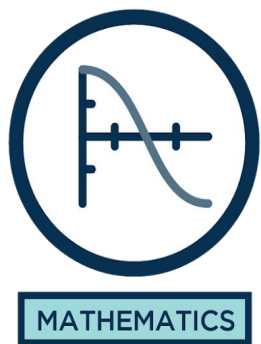
- Marko Zečević, ETF



- Milan Kovačević, PMF



Bioinformatics analyst



Applied Bioinformatics

Course overview



Course logistics (1/2)

- 2 classes each week
- A mixture of lectures and hands-on exercises
- Exercises will be done in IPython notebooks on the CGC platform
 - We will provide help with the Python syntax if needed
- The course is not covered by a single textbook
- Lessons are mostly linked

Course logistics (2/2)

- Attending classes is mandatory and highly recommended
- Practice test in the end of semester
- For all course related questions contact
 - milan.kovacevic@velsera.com

Course topics

- Introduction to biological background and sequencing (2 weeks)
- DNA analysis (4 weeks)
- RNA analysis (2 weeks)
- Structural variation detection (1 week)
- Cancer genomics (1 week)
- Methylation/epigenetics analysis (1 week)
- Exam practise (1 week)

Applied Bioinformatics

Bioinformatics intro



Bioinformatics intro

During this course you will learn/hear:

- More about
 - Human cells
 - DNA
 - RNA
- Why is the DNA so important?
- What can we find out from the DNA?
- More about certain diseases, like cancer
- Genome digitalisation
- How to analyse genomic data
- Interesting algorithms
- Optimisations for huge data

What is bioinformatics?



What is bioinformatics?



Bioinformatics is an interdisciplinary field of study that combines biology, computer science, and statistics to develop methods and tools for analyzing and interpreting biological data.

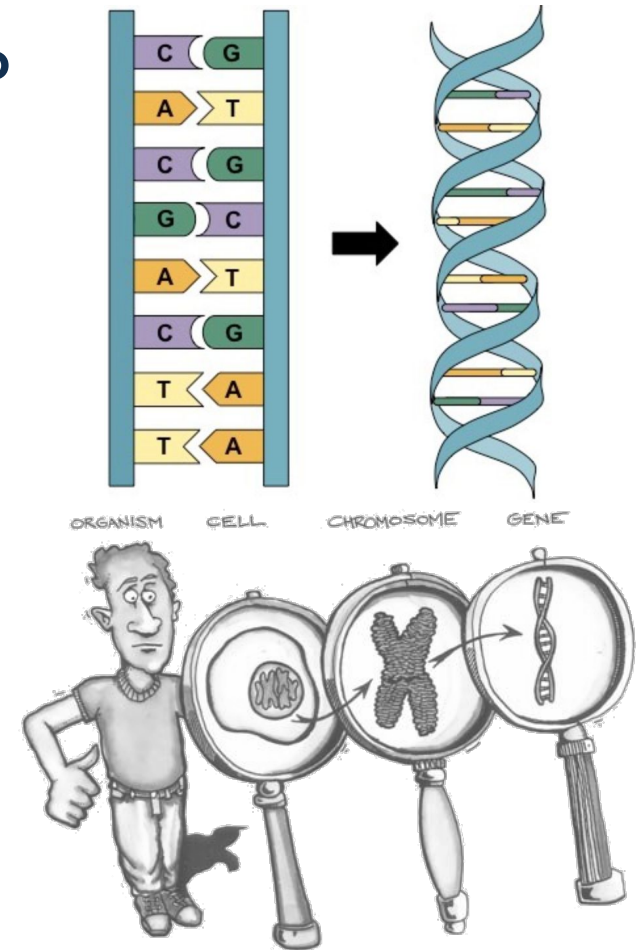


In bioinformatics, researchers use computational techniques to process and analyze large amounts of biological data, such as genomic and proteomic data. This includes tasks such as sequence alignment, genome annotation, and gene expression analysis. Bioinformatics also involves developing algorithms and software programs for analyzing and visualizing biological data.

Bioinformatics plays an important role in many areas of biological research, including genetics, molecular biology, evolutionary biology, and drug discovery. It is particularly useful for understanding complex biological systems and for identifying new targets for drug development.

Bioinformatics intro

- A, T, C, G
- 3.000.000.000 letters in a DNA molecule
- 46 chromosomes
- Almost everything in the body is predefined by the order of those letters

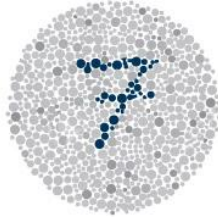


Bioinformatics intro

Some interesting characteristics:



LONGER EYELASHES



DALTONISM



LESS SLEEPING



SUPER STRENGTH

And some not that interesting things:

- Cancer
- Rare diseases
- Autoimmune diseases
- Neurological diseases

DNA digital representation

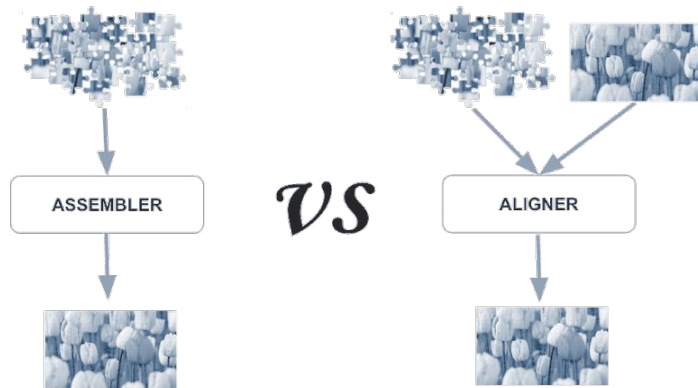
- DNA is stored as a string
- Human genome project and reference genome
- 3 billion As, Ts, Cs and Gs translates into ~ 3 GB of data

■ font 4

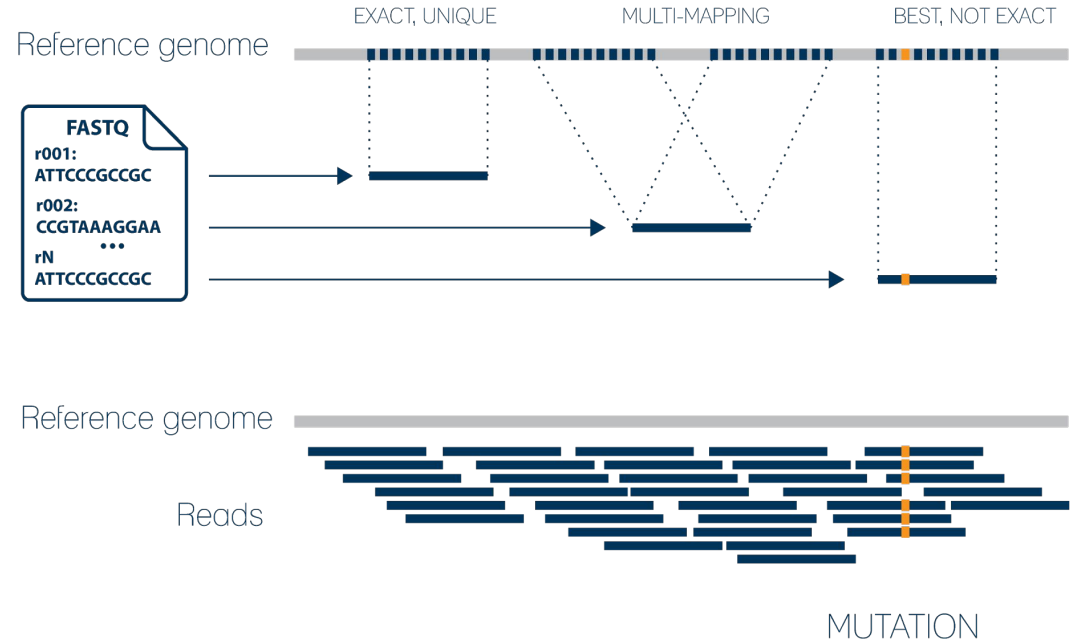


Bioinformatics intro

Common 1st step:
Reconstruct personal genome



$$\Theta(n^2) \text{ vs } \Theta(nm)$$



Bioinformatics intro

Single nucleotide variant



Deletion



Insertion



Inversion



Copy number variant



Translocation



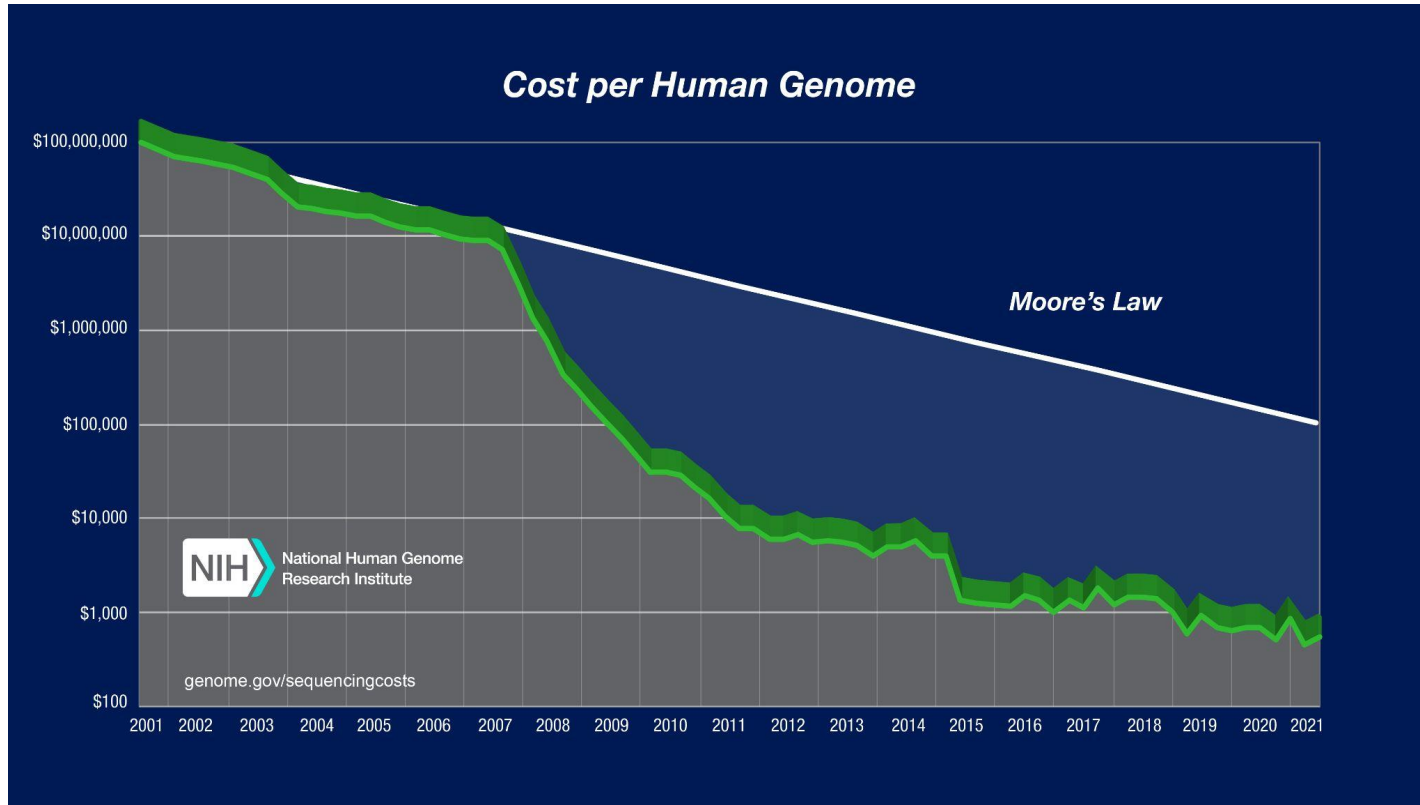
Whole genome duplication



Duplication



Why bioinformatics



source: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>

Why bioinformatics



My Full DNA: Whole Genome Sequencing with mtDNA

€259.00 EUR ~~€850.00 EUR~~ You save €591.00 EUR

Type of report: My Full DNA...

1

+

-

ADD TO CART

DESCRIPTION

My Full DNA by Dante Labs gives you unparalleled access to your health, your family traits, and predispositions. Having your full genome sequenced means:

Applied Bioinformatics

Platform registration



CGC registration

- Exercises are going to be done on CGC platform
- **cgcbioinformatics.com**
- CGC = Cancer Genomics Cloud
 - Funded by NCI - National Cancer Institute (NIH)
 - Powered by Seven Bridges
 - For academic use
 - Many researchers/ institutes / labs are using it for their analysis
- More information about the project available [here](#).
- Registration

Course materials

<https://github.com/milankovacevicsbg/MATF-2023-Primenjena-Bioinformatika>

Thank you!