Class 1

Dr. Giuliano Colosimo & Giulia Maiello

Welcome

Class overview

- Your teachers
- Opening quiz
- Course overview
- Bioinformatics
- Setting up your computers

Figure 1: Dr. Giuliano Colosimo

- Research Associate working @ University of Rome "Tor Vergata"
- I was a PostDoc for 5 years with the San Diego Zoo Wildlife Association
- I got a PhD in Conservation Genetics @ Mississippi State University
-
  giuliano.colosimo@uniroma2.it

Figure 2: Giulia Maiello

- PhD Student
- giulia.maiello@uniroma2.it

Opening quiz

Bioinformatics

What is Bioinformatics?

From Wikipedia

Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data, in particular when the data sets are large and complex. […] bioinformatics combines biology, chemistry, physics, computer science, information engineering, mathematics and statistics […] using computational and statistical techniques.

Important

Bioinformaticians deal with a variety of different type of data! **We are not bioinformaticians!!** We are evolutionary biologists!! We will focus on nucleotide sequence data.

- Back in the days, biologists could analyze their data on personal computers. Sanger Sequecing produced anywhere between 50 and 1000 nucleotides at a time to be analysed on a regular computer.

Why Bioinformatics

- Back in the days, biologists could analyze their data on personal computers. Sanger Sequecing produced anywhere between 50 and 1000 nucleotides at a time to be analysed on a regular computer.
- Since 1994 scientist have begun developing new technologies to parallelize DNA sequencing and produce high throughput data.

- Back in the days, biologists could analyze their data on personal computers. Sanger Sequecing produced anywhere between 50 and 1000 nucleotides at a time to be analysed on a regular computer.
- Since 1994 scientist have begun developing new technologies to parallelize DNA sequencing and produce high throughput data.
- As these technologies improve, biologists need to harness the tools to design meaningful experiments to produce data and analyse them.
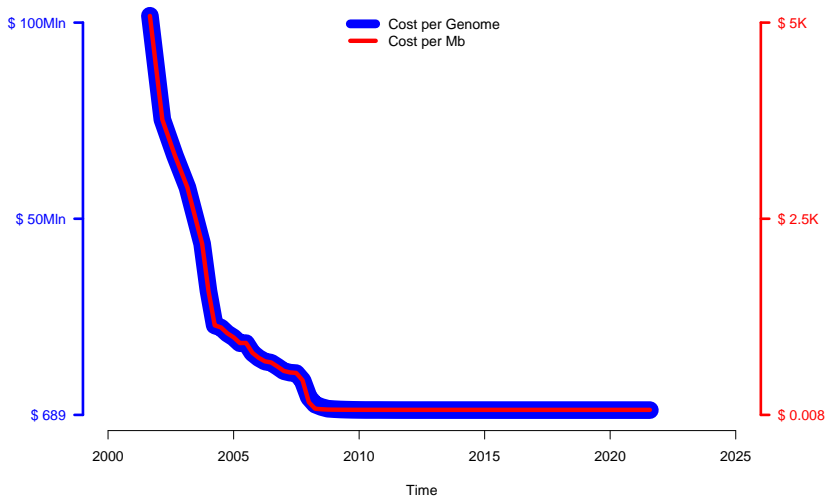
Figure 3: Drop of sequencing cost for Genomes and Nucleotides. Data from https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

Setting up your computer

To visualize and analyse data and to produce reports on all the activities of this course we will be using **R** and **RStudio**. These are free cross-platform software.

- Visit the **R** (https://cran.r-project.org) page and download the **R** version compatible to your OS. Install the software following the instructions.

- Visit the **RStudio** (https://www.rstudio.com/products/rstudio/download/) page and download the **RStudio** Desktop version compatible with your OS. Install the software following the instructions.