# Fundamentals of computational biology

**Lecture notes**

Camilo García-Botero

2022-04-11

# Table of contents

# Preface

We started this book with the aim of compiling the lectures of the course Fundamentals of Computational Biology offered at Universidad EAFIT for undergrad students in Biology. The course has been taught from different perspectives from its creation, yet the last iteration was divided into three modules. i) introduction to Unix (4 lectures) ii) introduction to sequence analysis and genomics (7 lectures) and iii) principles of structural biology (4 lectures).

Lectures are focused on a theoretical-practical approach were basic concepts from biology, bioinformatics and computer science and interleave with the practice to solve challenges.

# Introduction

Here we present a course centered book of the Fundamentals of Computational Biology. We will cover several topics, from using the unix tools, the importance of package manager systems (such as homebrew and conda), sequencing technologies, sequence alignments, molecular phylogenetics, genome assembly and annotation, and variant calling analysis.

# Part I

# Unix

# 1 Welcome to the command line

In this chapter we will explore the fundamentals of the command line. That is the concepts of Unix based systems the command line (CLI) and how we can use it to access information programmatically.

# Part II

# Sequence analysis

# 2 Introduction to sequence analysis

In this chapter we will discuss several about several points of view about bioinformatics and computational biology and how to get started with the command line being a biologist, we will further consider several biological concepts that appear central to understand the manipulation of biological data.

## Endless debate: bioinformatics vs. computational biology

## Getting started with the command line

## The duality of DNA

## The central ~~dogma~~ theory of molecular biology extended

## Sequencing strategies

## Sequencing over time

## Some insights from sequencing genomes

# 3 Sanger analysis

This is a section about the first gen sequencing tech

## Databases exploration

## Sanger sequencing methods

### The chain termination method

### Sanger with capillary electrophoresis

### Strengths and limitations of Sanger methods

## Files from Sanger

## Sanger processing workflow

## The 16S rRNA and its relevance for sequencing

# Part III

# Challenges demostrations