

# Algoritmi di Bioinformatica

**Zsuzsanna Lipták**

Laurea Magistrale Bioinformatica e Biotechnologie Mediche (LM9)  
a.a. 2014/15, spring term

Organisation

# Organisation

- **course times:** Wed 8:30 - 10:30, Thu 8:30 - 11:30
- **language:** English, but you can ask questions in English or in Italian
- **webpage:**  
<http://profs.sci.univr.it/~liptak/ALBioinfo.html>  
**coming soon!** (careful: at the moment there is last year's course there)  
will include: current info, materials, presentation topics etc.

## Organisation (cont.)

- **email:** `zsuzsanna.liptak@univr.it`  
Please put "corso Algoritmi di Bioinformatica" in the subject line
- **office:** CV 2, 1st floor, room 1.79
- **office hours:** Tue 8:30-10:30
- **attendance:** not obligatory but highly recommended  
(read what I wrote on the webpage of the course)

## Organisation (cont.)

- **final grade** (voto): 50% Algoritmi, 50% Linguaggi
- **grade Algoritmi:**
  - 50% **written**: 2 exams (2 hours each) in April+June
  - or** 1 exam (3 hours) June or any other exam session
  - 50% **oral**: short presentation of a topic in front of class (June or July)
  - or** oral exam on the whole course (any exam session)
- **presentations**: to be prepared in 1 or 2, based on material that I give you — **only** possible during this semester (June or July)

## Organisation (cont.)

- **final grade** (voto): 50% Algoritmi, 50% Linguaggi
- **grade Algoritmi:**
  - 50% **written**: 2 exams (2 hours each) in April+June
  - or** 1 exam (3 hours) June or any other exam session
  - 50% **oral**: short presentation of a topic in front of class (June or July)
  - or** oral exam on the whole course (any exam session)
- **presentations**: to be prepared in 1 or 2, based on material that I give you — **only** possible during this semester (June or July)

**Questions?**

## Goals of this course

1. to learn about some basic **problems and algorithms** behind common bioinformatics applications (sequence alignment, sequence similarity, phylogenetic reconstruction)
2. to get an idea of some basic **computational issues** involved (complexity, efficiency, limitations, problem specification)

## Goals of this course

1. to learn about some basic **problems and algorithms** behind common bioinformatics applications (sequence alignment, sequence similarity, phylogenetic reconstruction)
2. to get an idea of some basic **computational issues** involved (complexity, efficiency, limitations, problem specification)

The **differences in background** (biology/biotech vs. computer science) will be resolved

- by doing all algorithms in enough detail (comprehensible **and** interesting)
- different levels/difficulty in the project presentation part (also w.r.t. grades)

# Overview

- Part I: Sequence Analysis
  - Pairwise sequence alignment
  - **Detour:** Algorithm analysis
  - Multiple sequence alignment
  - String similarity and distance
  - Scoring matrices
  - Heuristics: FASTA, BLAST
- Part II: Phylogenetics
  - **Detour:** Trees and graphs
  - algorithms for distance-based data
  - character-based data, Perfect Phylogeny
  - Small Parsimony: Fitch's algorithm
  - Large Parsimony: heuristics



## Books

- **João Setubal, João Meidanis:** Introduction to Computational Molecular Biology (1997).—my favourite, 1 copy in library
- **Neil C. Jones and Pavel A. Pevzner:** An Introduction to Bioinformatics Algorithms (2004).—3 copies in library

## Books

- **João Setubal, João Meidanis**: Introduction to Computational Molecular Biology (1997).—my favourite, 1 copy in library
- **Neil C. Jones and Pavel A. Pevzner**: An Introduction to Bioinformatics Algorithms (2004).—3 copies in library
- **R. Durbin, S. Eddy, A. Krogh, G. Mitchinson**: Biological Sequence Analysis (1998).—oriented towards probabilistic models
- **Dan Gusfield**: Algorithms on Strings, Trees, and Sequences (1997).—the bible of string algorithms, a bit dated now
- **Hans-Joachim Böckenhauer and Dirk Bongartz**: Algorithmic Aspects of Bioinformatics (2010).—very formal, mathematical
- **Joseph Felsenstein**: Inferring Phylogenies (2004).—important book on phylogenetics
- **Cormen, Leiserson, Rivest (& Stein)**: Introduction to Algorithms (different editions, 1990-onwards).—the bible of algorithms, a must have for anyone interested in algorithms (buy second hand, old editions are also fine)

## Books (2)

More biologically oriented:

- **David M. Mount:** Bioinformatics: Sequence and Genome Analysis (2004).—good book but very detailed!
- **Arthur Lesk:** Introduction to Bioinformatics (2008).

Careful, in biologically oriented books, the treatment of algorithms may not be in-depth not enough! Positive side: more application-oriented.

## Books (2)

More biologically oriented:

- **David M. Mount:** Bioinformatics: Sequence and Genome Analysis (2004).—good book but very detailed!
- **Arthur Lesk:** Introduction to Bioinformatics (2008).

Careful, in biologically oriented books, the treatment of algorithms may not be in-depth not enough! Positive side: more application-oriented.

I also use (downloadable from the course webpage):

- Sequence Analysis I+II, lecture notes from Bielefeld University, by Jens Stoye and others
- Phylogenetics, lecture notes from Bielefeld University, by Jens Stoye and others

Mostly much too detailed, do not print! We will assign chapters from these for the projects.