



Biological Sequence Analysis

2022

Shilpa Garg
Anders Krogh
Amelie Stein



Learning outcomes for the course

Knowledge

know and understand the main algorithms and methods used in

- pairwise and multiple alignment
- searching of sequence databases
- hidden Markov models of biological sequences
- finding and discovery of motifs in biological sequences
- mapping and assembly of DNA sequencing data

understand the biological contexts in which to apply biological sequence analysis

Skills:

- derive simple probabilistic methods and algorithms for biological sequence analysis
- explain the algorithms and methods covered in the course
- interpret results of standard methods used in biological sequence analysis
- select the proper sequence analysis methods and tools for a given biological problem
- run some standard computer programs for biological sequence analysis

Learning Outcomes cont'd

Competences:

The student will obtain these general competences:

- Will be able to **discuss and explain** methods in biological sequence analysis with researchers in bioinformatics and related sciences
- Will be able to contribute to interdisciplinary projects involving biological sequence analysis
- Will be able to read, understand, and discuss scientific literature in biological sequence analysis

Topics

- Pairwise alignment
- Data structures for efficient searching
- Search heuristics (BLAST)
- Hidden Markov Models (HMMs)
- Multiple sequence alignments
- Co-evolution
- Motif discovery
- Next-gen sequencing (NGS)

See also the “Schedule” entry on Absalon

Exam: Nov 7-10

- If you have a conflict on any of these dates, let me know as soon as possible
- Oral exam
- Random selection of topic
- 30min preparation time, open books
- 25min actual exam, starting with a presentation on that selected topic
- Then questions/discussion following up on the presentation
 - may also cover other topics from class

ERDA

Current members of COURSES/BIO/BSA:

Actions	members
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=David Rodriguez Morales/emailAddress=gsn480@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Luca Robinson/emailAddress=ntp570@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Weikun Liu/emailAddress=bht586@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Michaela Válková/emailAddress=jrh448@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Weizhe Zhang/emailAddress=cwt988@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Isabel Marleen Pöttsch/emailAddress=fnj847@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Jianpeng Zheng/emailAddress=cfv939@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Yinghan Jiang/emailAddress=wfz471@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Tobias Svend-Aage Beyer-Pedersen/emailAddress=tobias.pedersen@sund.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Conor David Ohare/emailAddress=bfr679@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOLOGI/OU=NA/CN=Daniel Hansen/emailAddress=bcw849@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Javier Sanchez/emailAddress=qrd871@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Finnja Hendrike Becker/emailAddress=zkw116@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Rebecca Distefano/emailAddress=fwh492@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Shanshan He/emailAddress=wkq953@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Jaime Moreno Martínez/emailAddress=zvx353@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Pablo Atienza Lopez/emailAddress=cgr232@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Taraneh Strunk/emailAddress=jxn241@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Cristina Rodríguez Sanjuán/emailAddress=swm300@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIO/OU=NA/CN=Sven Larsen-Ledet/emailAddress=sven.larsenledet@bio.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Jonas Blomquist Jørgensen/emailAddress=pcf255@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Alexander Michael Villadsen/emailAddress=mwn735@alumni.ku.dk
    	 /C=DK/ST=NA/L=NA/O=BIOINFORMATIK/OU=NA/CN=Gerardo José Rodríguez Alarcón/emailAddress=xgm987@alumni.ku.dk

Introductions

- Your name
- Educational background
- One boring fact about you

ILOs for today

- Variables, indices
- Sum, product notation
- Probabilities
- Logarithms
- Algorithmic complexity in time & space
- Dotplots
- (sequence comparison, if time)

There will be a quiz on Absalon at the end of the first week, please help us by filling it out

Questions?

Please ask - during class, in breaks, after class

Please use the Absalon Discussion boards if you have questions outside of class time, so that others in the class can also benefit from the answer :)