

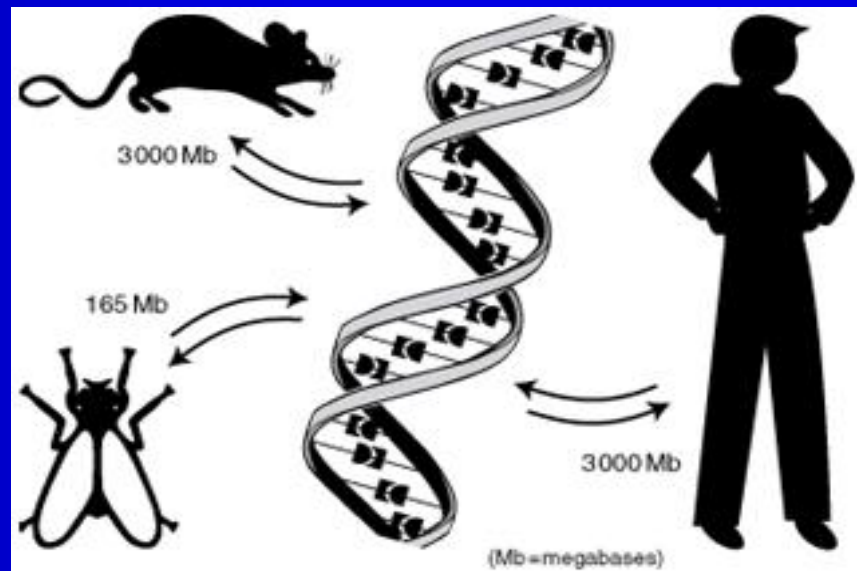
KB8019 Comparative Genomics

Teachers:

- Erik Sonnhammer

Assistants

- Deniz Secilmis
- Miguel Castresana
- Stefanie Friedrich



Roll Call

Registered students 2018:

Axelsson, Linnea
Bindu Suresh, Aishwarya

Boey, Zhong Hao Daryl
Challoori Mounika

Hallberg, Olaf
Hesselman, Maria Carmen

Juszczak, Kajetan
Kaldhusdal, Vilde

Kiik, Helen
Kimler, Kyle

Krali, Olga
Kumar, Sharmishtaa

Lautenbach, Maximilian Julius
Martinez Hernandez, Marina

Pérez Gómez, Fernando
Rodriguez, Lucie
Ropat, Maryia

Rosa, André
Senftleben, Maximilian Lukas

Valiukonyte, Milda
Xi, Yuanyuan

Xu, Fuqi
Xu, Shuhan

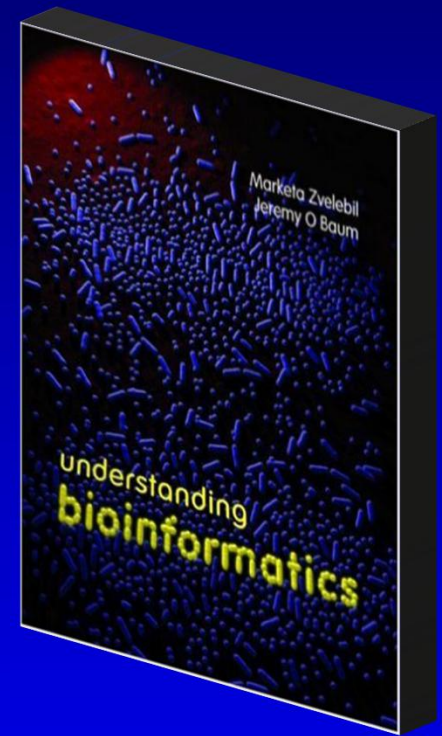
Yang, Ke
Zhang, Youcheng

Comparative Genomics

- Prerequisites ~ Applied Bioinformatics, DD2404
- Course dates: 2/5-1/6
- 4x2 Lectures
- 8 Practicals
- 3 Final project assignments.
- Practicals and final project assignments in groups of 2 persons
 - Analyse 5 anonymous genomes

Course materials

- Wikipedia pages, see schedule in Mondo
- Zvelebil & Baum book (2007), chapters:
 - 3 Dealing with Databases
 - 7 Recovering Evolutionary History
 - 8 Building Phylogenetic Trees
 - 9 Revealing Genome Features
 - 10 Gene Detection and Genome Annotation
 - 17 Systems Biology
- Basic bioinformatics knowledge from **chapters 4, 5, 6, 7, 8, 11, 12, 15, 16 covered in the introduction course** KB7004 is assumed.
- Everybody must read the listed material before the lecture. Please email me questions that you want to discuss at the lecture.
- Compulsory attendance on lectures, but **no exam**.



Final project assignments

Analyse 5 genomes by:

1. DNA genome statistics tool, protein statistics tool
Compute GC content and nucleotide (dinucleotide) frequency in a genome
Compute amino acid (diamino acid) frequencies in a proteome
2. ORF finder
Compute Open Reading Frames (ORFs) in a genome.
3. Distance matrix tool
Computes the distance between two genomes from a DNA statistic above.
Use distance matrix to create a species tree.

Write report and give 15 minute presentation on last day

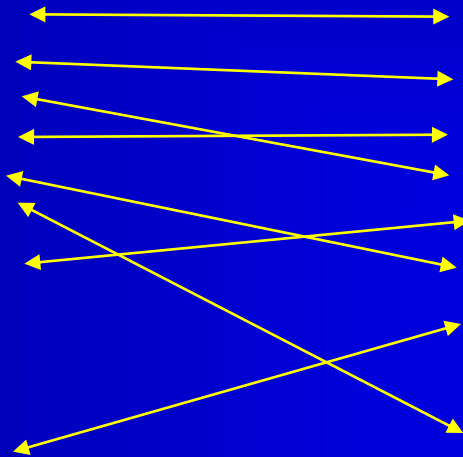
Comparative Genomics

- Week 1 The structure of genomes; Gene prediction
- Week 2 Evolution of genes and genomes
- Week 3 Orthology analysis
- Week 4 Function and interaction prediction
- Week 5 Project work, report writing
- Last day: Project presentations

Plagiarism

- Plagiarism means claiming or implying original authorship of (or incorporating material from) someone else's written or creative work, in whole or in part, into one's own without adequate acknowledgement. It concerns e.g. pictures, text, graphs or diagrams.
- Reports are checked for plagiarism.
- A student who is considered guilty of plagiarism will receive a warning and risks being suspended.

Why Comparative Genomics?



What is Comparative Genomics?

Google [define:Comparative Genomics]:

- The study of human genetics by comparisons with the genetics of other organisms.
- Study of the relationship of genome structure and function across different biological species.
- Comparative studies of whole genomes help researchers understand what parts of the genome in one organism are similar to those in another, how the overall structures of genes and genomes have evolved, and what these findings tell us about gene expression, gene regulation, and how to interfere ...
- The study of relationships between the genomes of different species or strains. Comparative genomics is an attempt to take advantage of the information provided by the signatures of selection to understand the function and evolutionary processes that act on genomes. ...

Course Goals

After this course, the student will:

- Be able to describe the contents of different types of genomes
- Be able to predict genes in a genome
- Understand the principles of sequence analysis and be able to apply such methods on complete genomes and proteomes
- Understand the principles for phylogenetic analysis and to apply phylogenetic methods
- Understand the principles for orthology analysis and to apply orthology identification methods
- Understand the principles for protein function analysis and to predict protein function.

Course Grades

A = excellent

B = very good

C = good

D = satisfactory

E = sufficient

Fx = insufficient

F = completely insufficient

Course Grades

Weighting scheme:

- Weight of all practicals' reports = 0.5
(each report has $1/8$ weight)
- Weight of project report/presentation = 0.5