**Course KB8030 Comparative Genomics, 7.5 hp**

Preliminary schedule for 2024, version 3/27/24. _Do not make a copy of this file_ as it may change.

Hosted by Stockholm University, DBB.
Course responsible: Prof. Erik Sonnhammer, erik.sonnhammer@scilifelab.se

Course goals: to learn current techniques for analysing genomes and how comparative genomics can be used to understand the organisation, evolution, and function of genomic sequences.

Course literature:
•   Web resources.
•   Zvelebil and Baum, Understanding bioinformatics. Not as up to date as the web resources but has more in-depth explanations of many concepts and algorithms.

Course begin/end: May 2 – June 2, 2024
•   Classes at Science for Life Lab, in Air&Fire on Gamma 2.
•   The listed literature should be read before each class. Time is reserved for this in the morning of the class day.
•   Practicals are done via remote login on a server or locally on your own computer (not recommended). Accounts on server will be provided at the course start. Teacher assistants will be available in the Gamma 7 lunch room on non-class days 13.30-16.
•   Reports for practicals should be submitted during the week they are listed, but at the latest the Monday after.  Practicals are graded passed / not passed.
•   Slides will be provided before the class.
•   Teacher assistants: Davide Buzzao and Emma Persson
•   Course information on Google drive
•   Course information and reporting via Athena; log in with your SU account on https://athena.itslearning.com

**Week 1. The structure of prokaryotic and eukaryotic genomes; Gene prediction**
May 2, 10.00-12.00: Roll call, Introduction to course, and start of practicals
        Practical 1: Basic genome analysis. Briefing May 2, ~11.00
        Practical 2: Gene prediction. Briefing May 2, ~11.00

May 3, 14-16:
        Class 1. Genome organisation
                Intended Learning Objectives:
                •   Introduction to genomes
                •   Genome properties in different species
                •   Gene content of genomes
                •   Regulatory sequences
                •   Non-coding sequences
                •   Metagenomics
        Class 2. Gene prediction
                Intended Learning Objectives:
                •   Gene structure in different species
                •   Gene prediction in different species
                •   Promoter prediction in different species

Literature:
   https://en.wikipedia.org/wiki/List_of_biological_databases
   https://en.wikipedia.org/wiki/Domain_(biology)
   https://en.wikipedia.org/wiki/Genome
   https://en.wikipedia.org/wiki/C-value
   https://en.wikipedia.org/wiki/Translation_(biology)
   https://en.wikipedia.org/wiki/Genome_evolution
   https://en.wikipedia.org/wiki/Gene_prediction
   https://www.ncbi.nlm.nih.gov/blast/tutorial/
   Optional: https://en.wikipedia.org/wiki/Introduction_to_genetics
   Optional: https://en.wikipedia.org/wiki/Human_genome
   Optional: https://en.wikipedia.org/wiki/Bioinformatics
   Optional: https://en.wikipedia.org/wiki/Genetic_code
   Optional: https://en.wikipedia.org/wiki/Non-coding_DNA
   Zvelebil:
      Chapter 3 Dealing with Databases
      Chapter 9 Revealing Genome Features
      Chapter 10 Gene Detection and Genome Annotation

## Week 2. Evolution of genes and genomes
May 6, 14-16:
   Class 3. Phylogenetics
      Intended Learning Objectives:
      • Understand principles of phylogenetic trees
      • Understand methods of phylogenetic tree building
      • Understand methods to assess the quality of phylogenetic trees
   Class 4. Phylogenomics
      Intended Learning Objectives:
      • Understand principles of phylogenomics
      • Understand applications of phylogenomics
      • Understand phylogenetic profiles, metaphylogeny, and tree of life
      • Understand reasons for phylogenetic disagreement

Practical 3: Phylogenetic reconstruction. Briefing May 6, 16.00
Practical 4: Phylogenomics. Briefing May 6, 16.00

Literature:
   https://en.wikipedia.org/wiki/Phylogenetic_tree
   https://en.wikipedia.org/wiki/Substitution_model
   https://en.wikipedia.org/wiki/UPGMA
   https://en.wikipedia.org/wiki/Neighbor_joining
   https://www.nature.com/articles/nrg1603.pdf, boxes 1 and 2 (Delsuc et al., 2005)
   https://genome.cshlp.org/content/8/3/163.long (Eisen, 1998)
   https://en.wikipedia.org/wiki/Phylogenetic_profiling
   https://en.wikipedia.org/wiki/Phylogenetic_network
   Optional: https://en.wikipedia.org/wiki/Models_of_DNA_evolution
   Optional: https://en.wikipedia.org/wiki/List_of_phylogenetics_software
   Optional: https://en.wikipedia.org/wiki/Phylogenetic_tree_viewers
   Optional: https://en.wikipedia.org/wiki/Bootstrapping_(statistics)

Optional: https://en.wikipedia.org/wiki/Phylogenomics
Optional: https://en.wikipedia.org/wiki/Phylogenetics
Zvelebil:
Chapter 7: Recovering Evolutionary History
Chapter 8: Building Phylogenetic Trees

## Week 3. Synteny and orthology analysis
May 13, 14-16:
Class 5. Gene order
Intended Learning Objectives:
- Introduction to gene order conservation (synteny)
- Analysing synteny with dotplots
- Using synteny for inferring phylogeny
Class 6. Orthology
Intended Learning Objectives:
- Definitions & assumptions
- Applications
- How to identify orthologs
- Tree reconciliation
- Graph-based methods; Tree-based methods; Other methods
- Comparison of ortholog databases

Practical 5: Orthology. Briefing May 13, 16.00
Final project assignment: Briefing May 13, 16.00

Literature:
Inferring Orthology and Paralogy (Altenhoff et al., 2019)
https://en.wikipedia.org/wiki/Sequence_homology
https://en.wikipedia.org/wiki/Synteny
https://en.wikipedia.org/wiki/Dot_plot_(bioinformatics)
https://genomevolution.org/wiki/index.php/Synteny:_Getting_the_Big_Picture
InParanoiDB 9: Ortholog Groups for Protein Domains .. (Persson & Sonnhammer, 2023)
Optional: https://questfororthologs.org/
Optional: https://orthology.benchmarkservice.org/
Optional: https://www.ncbi.nlm.nih.gov/pubmed/11934753
Zvelebil:
Chapter 7.2 Molecular Evolution and its Consequences

## Week 4. Interaction networks
May 20, 14.00-16.00:
Class 7. Interaction networks
Intended Learning Objectives:
- Introduction to network biology
- Discovering protein interactions
- Network properties
- Network databases (STRING, FunCoup)
Class 8. Applications of network analysis
Intended Learning Objectives:
- Network analysis for gene identification
- Pathway analysis

- Gene regulatory networks

Practical 6: Interaction networks. Briefing May 20, 16.00

Literature:
    https://en.wikipedia.org/wiki/Biological_network
    https://en.wikipedia.org/wiki/Interactome
    https://en.wikipedia.org/wiki/Network_science
    https://en.wikipedia.org/wiki/Scale-free_network
    https://en.wikipedia.org/wiki/Clustering_coefficient
    https://en.wikipedia.org/wiki/Systems_biology
    https://en.wikipedia.org/wiki/Pathway_analysis
    The Architecture of Biological Networks (Wuchty et al., 2006)
    FunCoup 5: Functional Association Networks... (Persson et al., 2021)
    https://www.youtube.com/@sonnhammerlab
    Optional: https://www.ncbi.nlm.nih.gov/pubmed/27151197 (Ogris et al., 2016)
Zvelebil:
    Chapter 17: Systems Biology

**Week 5. Final project assignment**
(Briefing in week 3)

May 31, 10.00-13.00, Air&Fire Gamma 2: exam