Algorithms for Bioinformatics (Autumn 2015)
Study Groups 1b (Tue 8.9. 12-14 B222)

Each student should belong to exactly one of the groups C, D and E. The groups are assigned at the lecture on Thu 3.9. If you were not present at the lecture, please contact the lecturer and ask for your group.

Group C: Random assignment at lecture

- Read the subsection titled *Probabilities of patterns in a string* on pages 35–40 in the course book Compeau & Pevzner.

- At the study group solve the "exercise break" problems and discuss the "STOP and Think" questions in the subsection.

Group D: Random assignment at lecture

- Study the paper

  Luo, Zhang, Gao: **Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes**.
  Front Microbiol. 2014; 5: 482.
  `http://dx.doi.org/10.3389/fmicb.2014.00482`

  to obtain an overview of the methods used by Ori-Finder 2.

- At study group, compare the methods in the paper to the methods described at the lecture.

Group E: Random assignment at lecture

- Read the subsections *Enhanced suffix arrays*, *Enumerating k-mers and their occurrence counts*, and *Analysis of time and space requirement* (pages 5–6) in the paper

  Kurtz, Narechnia, Stein, Ware: **A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes**.
  BMC Genomics 2008, 9:517
  `http://dx.doi.org/10.1186/1471-2164-9-517`

  Note that the pdf-version of the paper displays the letter "ℓ" incorrectly as the symbol "$<$". This has been corrected in the handouts distributed at the lecture. The HTML-version does not have this problem.

- At the study group, discuss how to use the enhanced suffix array to find the most frequent k-mers.