

Algorithms for Bioinformatics (Autumn 2015)

Study Groups 2 (Tue 15.9. 12-14 B222)

Group A: Lastname beginning with A–J

- Read about *branch and bound* algorithms for the Motif Search and Median String problems in last year's lecture notes (Lecture 2, primarily slides 61–68):

https://www.cs.helsinki.fi/juha.karkkainen/opetus/14s/AfB/AfB_lecture2_20140904.pdf

Note that some definitions are different from this year's slides. In particular, Score is defined as the number of characters matching the consensus, and the goal of Motif Search is to maximize rather than minimize Score.

- At the study group discuss the branch and bound algorithms and try to estimate if they are a big improvement over the brute force algorithms.

Group B: Lastname beginning with K–P

- Read the the sections *Results* and *Discussion* in the paper

Tompa et al.: **Assessing computational tools for the discovery of transcription factor binding sites.**

Nature Biotechnology 23, 137 - 144 (2005)

<http://dx.doi.org/10.1038/nbt1053>

- At study group, discuss what are the most important performance measures and which are the best tools.

Group C: Lastname beginning with R–Ö

- Read the sections *Suffix trees* and *The WEEDER algorithm* in the paper

Pavesi, Mauri, Pesole: **An algorithm for finding signals of unknown length in DNA sequences.**

Bioinformatics (2001) 17 (suppl 1): S207-S214.

http://dx.doi.org/10.1093/bioinformatics/17.suppl_1.S207

Slides 77–83 in last year's Lecture 2 (see Group A above) may be helpful.

- At the study group, discuss the WEEDER algorithm based on the pseudocode in Fig. 4.