

Algorithms for Bioinformatics (Autumn 2015)

Study Groups 4 (Tue 29.9. 12-14 B222)

Group A: Random assignment at lecture

- Read sections 1 and 2 and especially section 2.5 *Inexact matching: bounded traversal/backtracking* in the paper

Li, Durbin: **Fast and accurate short read alignment with Burrows–Wheeler transform.**

Bioinformatics. 2009 Jul 15; 25(14): 1754–1760.

<http://dx.doi.org/10.1093/bioinformatics/btp324>

- At study group, discuss the inexact matching method based on the pseudocode in Fig. 3. Hint: The four recursive calls to INEXRECUR correspond to insertion, deletion, substitution and match.

Group B: Random assignment at lecture

- Read about *general gap penalty model* and *affine gap penalty model* on pages 42–45 of the book:

Sung: **Algorithms in Bioinformatics: A Practical Introduction.**

CRC Press 2010.

Copies are distributed at the lecture.

Also watch the short lecture video by Nikolay Vyahhi covering the affine gap penalty:

https://youtu.be/Npv180dQ_4Y?list=PLQ-851Q1PqFNmbPEsMoxb5dM5qtRaVShn

- At the study group, discuss the basic ideas of the method. How do the tables in the book excerpt correspond to the levels in the video?

What is the best global alignment of CGAGAT and CAT using the affine gap model? Use cost +4 for a match, -2 for mismatch, -3 for gap opening, -1 for gap extension. What is the score of the alignment?

Group C: Random assignment at lecture

- Read about *gene prediction by spliced alignment* on pages 203–207 of the book

Jones & Pevner: **An Introduction to Bioinformatics Algorithms.**

MIT Press, 2004.

Copies are distributed at the lecture.

- At the study group, discuss the basic ideas based on Figure 6.28, and the derivation of the recurrences. What is the running time of the algorithm?