

Introduction to Bioinformatics (autumn 2005)

Excercise 1

Group	time	place
Riikka Kaven	Tuesday 4.10 at 12.15–14.00	BK106

- (Problem 2.17 in course book J & P): There are n bacteria and one virus in a Petri dish. Within the first minute, the virus kills one bacterium and produces another copy of itself, and all of the remaining bacteria reproduce, making 2 viruses and $2(n - 1)$ bacteria. In the second minute, each of the viruses kills a bacterium and produces a new copy of itself (resulting in 4 viruses and $2(2(n - 1) - 2) = 4n - 8$ bacteria); again, the remaining bacteria reproduce. This process continues every minute. Will the viruses eventually kill all the bacteria? If so, design an algorithm that computes how many steps it will take. How does running time of your algorithm depend on n ?
- Define the following concepts. A sentence or two on each is enough:
 - genome
 - gene
 - transcription
 - codon
 - translation
 - polymerase chain reaction
 - restriction enzyme
 - gene expression
- Simulate the greedy reversal sorting algorithm ("kääntöjärjestämisalgoritmi", SimpleReversalSort on page 129 in J&P), that increases $prefix(\pi)$ at each step, on input $\pi = 10\ 8\ 6\ 4\ 2\ 9\ 7\ 5\ 3\ 1$.
- Let $\pi = \pi_1\pi_2 \cdots \pi_n$ be a random permutation. Hence, $\mathbb{P}(\pi_i = j) = 1/n$ for all $1 \leq j \leq n$ at each fixed position i . Calculate the probability that $prefix(\pi)$ increases only by 1 at the first step of the greedy reversal sorting algorithm. What does this indicate on the estimated overall amount of reversals needed by the algorithm (no exact calculations needed here, just an educated guess)?
- (Warmup for exhaustive motif search) Write down a pseudo-code that prints out all DNA sequences of length n .