Algorithms for Genome Rearrangement
Summer 2016

Exercises

Number 3, Submission deadline: 2016 June 02, 10:00 a.m.
(Send your solutions in pdf format by email to `klamkiew@cebitec.uni-bielefeld.de`)

1. Consider the linear genomes $A = (\circ, 1, a, -2, b, 3, \circ)$ and $B = (\circ, 1, c, 2, d, 3, \circ)$.   (3 Pt)

   (a) Draw the adjacency graph $AG(A, B)$. What is the DCJ-indel distance between A and B?

   (b) Find a scenario with 3 indels, applying only optimal DCJs. Redraw the adjacency graph after the first DCJ.

   (c) Find a scenario with 2 indels. Redraw the adjacency graph after the first DCJ.
   (Hint: Find a neutral DCJ that reduces the number of runs.)

2. Prove the formula for the DCJ-indel potential λ of a cycle C of the adjacency graph $AG(A, B)$,   (4 Pt)

$$\lambda(C) = \left\lceil \frac{\Lambda(C) + 1}{2} \right\rceil,$$

   where $\Lambda(C)$ is the number of runs of indels in C.

   How does the corresponding formula for a path P with $\Lambda(P)$ runs look like?

3. Consider the following two genomes $A = (\circ, 1, 2, 3, 4, 5, 6, \circ)$ and $B = (\circ, -3, 1, 2, -6, 5, \circ)$.   (3 Pt)

   (a) Find a third genome C such that the triangle inequality for the uncorrected DCJ-indel distance does not hold, i.e.

$$d_{\text{DCJ-indel}}(A, B) > d_{\text{DCJ-indel}}(A, C) + d_{\text{DCJ-indel}}(C, B).$$

   (b) Show how the distance can be corrected such that the triangle inequality holds again.