

Algorithms in Comparative Genomics

Universität Bielefeld, WS 2020/2021

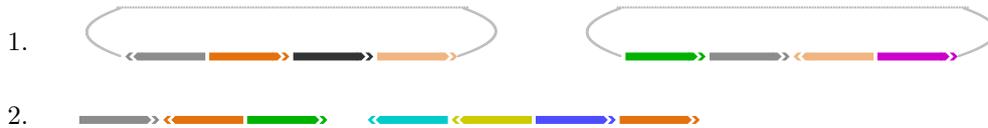
Dr. Marília D. V. Braga

<https://gi.cebitec.uni-bielefeld.de/teaching/2020winter/cg>

Exercise sheet 1, 29.10.2020

Exercise 1 (Genome representation)

Represent the following genomes in each pair as lists of gene orders (one per chromosome) and as list of (unambiguous) adjacencies, assigning arbitrarily numbers from 1 to n to the gene families (where n is the number of families in each genome pair).



Exercise 2 (Algorithm for breakpoint distance)

Devise a linear time algorithm for computing the breakpoint distance $d_{BP}(A, B)$, where A and B are canonical genomes on a given set of gene families $\mathcal{F} = \{1, \dots, n\}$ and can contain multiple linear or circular chromosomes.

Exercise 3 (Breakpoint double distance)

For the following duplicated genome D and singular genome S , compute the breakpoint double distance $d_{BP}^2(S, D)$ and give an optimal matching of the genes of D and $S \oplus S$.

$$D = (12323) [156\bar{5}\bar{4}7] (467)$$

$$S = [1234567]$$