

Algorithms in Comparative Genomics

Universität Bielefeld, WS 2021/2022

Dr. Marília D. V. Braga · Leonard Bohnenkämper

<https://gi.cebitec.uni-bielefeld.de/teaching/2021winter/cg>

Exercise sheet 1, 21.10.2021

Exercise 1 (Algorithm for breakpoint distance)

(5 pts)

Devise a linear time algorithm for computing the breakpoint distance $d_{BP}(\mathbb{A}_{\triangleright}^f, \mathbb{B}_{\triangleright}^f)$, where $\mathbb{A}_{\triangleright}^f$ and $\mathbb{B}_{\triangleright}^f$ are a pair of canonical genomes and can contain multiple linear or circular chromosomes.

Exercise 2 (Set of doubled genomes)

(3 pts)

For each of the following singular genomes, give the corresponding set of possible doubled genomes

1. $\mathbb{A}_{\triangleright}^f = \{ [1 \bar{2} 3] [4 5] \}$
2. $\mathbb{B}_{\triangleright}^f = \{ [\bar{1} \bar{3} 2] (4 5) \}$
3. $\mathbb{C}_{\triangleright}^f = \{ [\bar{2} 3 1] (\bar{4} 5) (\bar{6}) \}$

Exercise 3 (Breakpoint double distance)

(6 pts)

For the following pairs with singular genome $\mathbb{A}_{\triangleright}^f$ and duplicated genome \mathbb{B}_{\diamond}^f , compute the breakpoint double distance $d_{BP}^2(\mathbb{A}_{\triangleright}^f, \mathbb{B}_{\diamond}^f)$ and give the doubled genome \mathbb{A}_{\diamond}^f and an optimal matching of the genes of \mathbb{A}_{\diamond}^f and \mathbb{B}_{\diamond}^f .

1. $\mathbb{A}_{\triangleright}^f = \{ [1 2 3 4 5 6 7] \}$
 $\mathbb{B}_{\diamond}^f = \{ (1 2 3 2 3) [1 5 6 \bar{5} \bar{4} 7] (4 6 7) \}$
2. $\mathbb{A}_{\triangleright}^f = \{ (1 \bar{2}) (\bar{4} 3) \}$
 $\mathbb{B}_{\diamond}^f = \{ (\bar{4} 3 4 1 \bar{2} 3) (1 2) \}$

Exercise 4 (Maximal matching)

(4 pts)

Given two annotated genomes

$$\mathbb{A}^f = (1 3 1 8 4 7 7 8 \bar{7} 5 9) \text{ and}$$

$$\mathbb{B}^f = (2 1 \bar{5} 7 8 5 6 7 8 \bar{9} \bar{5} 3),$$

1. Give the number of markers per family in each genome.
2. Give the number of distinct pairs of genomes $(\mathbb{A}^{f_m}, \mathbb{B}^{f_m})$ that can be obtained by fixing a maximal matching of the genes within each family.