

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 5, 18.11.2021

Exercise 1 (DCJ halving)

(8 pts)

For each of the following duplicated genomes, compute the DCJ halving distance $h_i = h_{\text{DCJ}}(\mathbb{D}_i^f)$ and find a perfectly duplicated genome \mathbb{P}_i^f with a matching between the genes of \mathbb{P}_i^f and \mathbb{D}_i^f , giving a halving sorting scenario with h_i optimal DCJ operations that transform \mathbb{D}_i^f into \mathbb{P}_i^f .

1. $\mathbb{D}_1^f = [3\ 5\ \bar{4}\ 2\ \bar{5}] [2\ 1] [3\ 4\ 1]$

2. $\mathbb{D}_2^f = (3\ 5\ \bar{4}\ 2\ \bar{5}) (2\ 1\ \bar{1}\ 3\ 4)$

Exercise 2 (Solution space of DCJ halving sorting)

(4 pts)

Denote by $\rho\mathbb{G}$ the genome obtained after applying a DCJ operation ρ to a genome \mathbb{G} .

Now consider the duplicated genome:

$$\mathbb{D}^f = [\bar{4}\ 1\ \bar{4}\ \bar{3}\ 2] [\bar{2}\ 3\ 1] [5\ \bar{5}].$$

List all possible optimal (1st step) DCJ halving operations that could be applied to \mathbb{D}^f , that is, the set of DCJ operations $R = \{\rho : h_{\text{DCJ}}(\mathbb{D}^f) = h_{\text{DCJ}}(\rho\mathbb{D}^f) + 1\}$.

Exercise 3 (DCJ double distance)

(6 pts)

Let \mathbb{S}^f be a singular circular genome and \mathbb{D}^f be a duplicated circular genome. Show that finding common adjacencies greedily between some perfectly duplicated genome $\mathbb{P}^f \in 2 \cdot \mathbb{S}^f$ and \mathbb{D}^f as a first step is an optimal approach towards the exact computation of the DCJ double distance $d_{\text{DCJ}}^2(\mathbb{S}^f, \mathbb{D}^f)$ (although in general the computation of the DCJ double distance is NP-hard).

Hint: the ambiguous breakpoint graph might be an useful tool for this task.