## Algorithms in Comparative Genomics

Bielefeld University, Summer Semester 2025 Dr. Marília D. V. Braga · Prof. Dr. Jens Stoye https://gi.cebitec.uni-bielefeld.de/teaching/2025summer/cg Exercise sheet 2, discussion: 09.05.2025

## Exercise 1 (DCJ model)

Given the following pair of canonical genomes:

 $\mathbb{A} = \{ [2 \ \bar{1} \ \bar{4} \ \bar{3}], [5 \ 7 \ 6 \ 9 \ 8 \ 10] \} \text{ and } \mathbb{B} = \{ [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10] \} \}$ 

- 1. Draw the relational graph  $RG(\mathbb{A}, \mathbb{B})$  and compute the DCJ distance between  $\mathbb{A}$  and  $\mathbb{B}$ .
- 2. Give an optimal DCJ sorting scenario from A to B. Name the operations in your sorting scenario (fusions, fissions, inversions, translocations, etc).
- 3. If your optimal sorting scenario contains circular chromosomes in two consecutive intermediate genomes, find an alternative optimal scenario in which each circular intermediate chromosome is immediately reintegrated (no need to worry about running time of your procedure).
- 4. Can you find yet another alternative optimal scenario without circular intermediates?

## Exercise 2 (Solution space of sorting by DCJ)

- 1. Given genomes  $\mathbb{A} = \{ (1 \ 2 \ 3) \}$  and  $\mathbb{B} = \{ (1 \ 3 \ 2) \}$ , how many different optimal DCJ scenarios sorting  $\mathbb{A}$  into  $\mathbb{B}$  can you find?
- Given two canonical genomes G<sub>1</sub> and G<sub>2</sub>, let C be a k-cycle (with k ≥ 4) in RG(G<sub>1</sub>, G<sub>2</sub>). Denote by E<sub>α</sub>(G<sub>1</sub>, C) the set of edges corresponding to adjacencies of genome G<sub>1</sub> that are in C. What is the number of distinct DCJ operations that modify genome G<sub>1</sub> and split C into two cycles? (*Hint: For each pair of edges in E<sub>α</sub>*(G<sub>1</sub>, C), there is exactly one DCJ splitting C into two cycles.)

## Exercise 3 (Bounds for SCJ distance)

A theoretical lower bound for the SCJ distance with respect to the DCJ distance is

$$d_{DCJ}(\mathbb{A}, \mathbb{B}) \leq d_{SCJ}(\mathbb{A}, \mathbb{B}).$$

- 1. Give an example of a pair of mutually distinct genomes showing that this bound is tight.
- 2. Determine a tight upper bound for the SCJ distance with respect to the DCJ distance and explain your answer.