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 1, discussion: 02.05.2025

Exercise 1 (Algorithm for breakpoint distance)

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

Exercise 2 (Number of string representations)

1. Write down all string representations for the following singular genomes.
   - $\mathbb{A} = \{\ [1\,\bar{2}\,3]\ \ [4\,5]\ \}$
   - $\mathbb{B} = \{\ [\bar{1}\,\bar{3}\,2]\ \ (4\,5)\ \}$
   - $\mathbb{C} = \{\ [\bar{2}\,3\,1]\ (4)\ \}$
2. Develop a general formula for the number of string representations of singular genomes.

Exercise 3 (Breakpoint and SCJ distances)

Given two canonical genomes
$\mathbb{A} = [3\ 4\ \overline{8}]\ \ [2\ 1\ 7\ 5\ 6]\ \ (11\ 12\ \overline{13}\ 9\ 10)\ \ [\overline{15}\ \overline{14}]$ and
$\mathbb{B} = [1\ 2\ 3\ 4\ 5]\ \ (6\ 7)\ \ [8\ 9\ 10\ 11\ 12\ 13\ 14\ 15]$,

1. Draw the relational graph of $\mathbb{A}$ and $\mathbb{B}$.
2. What is the breakpoint distance between $\mathbb{A}$ and $\mathbb{B}$?
3. What is the SCJ distance between $\mathbb{A}$ and $\mathbb{B}$?
4. Explain the difference between the two distances (if any).

Exercise 4 (Bounds for SCJ distance)

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

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

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 breakpoint distance and explain your answer.