Algorithms in Comparative Genomics

Universität Bielefeld, SS 2023
Dr. Marília D. V. Braga · Dr. Roland Wittler

https://gi.cebitec.uni-bielefeld.de/teaching/2023summer/cg

Exercise sheet 8, 2.6.2023

Exercise 1 (Optimal cover of component tree) (8 pts)

Given canonical circular chromosomes

$$\mathbb{A} = (0\ 2\ \overline{7}\ 6\ \overline{5}\ 3\ 4\ 8\ 10\ \overline{12}\ 9\ \overline{11}\ 13\ \overline{1}\ 14\ 16\ 21\ 17\ 19\ 18\ 20\ 22\ 27\ 23\ 25\ 24\ 26\ 28\ 15)$$

$$\mathbb{B} = (0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20\ 21\ 22\ 23\ 24\ 25\ 26\ 27\ 28)$$

1. Draw the relational (or the breakpoint) diagram of $\mathbb{A}$ and $\mathbb{B}$.

   (You can use the Java program InversionVisualization provided on the course website: enter the values for chromosome $\mathbb{A}$, without the first value (0) and assume that the first vertex of the output diagram is 0^h and the last vertex is 0^t.)

2. Construct both the chained component tree $\Upsilon_{\blacksquare}(\mathbb{A}, \mathbb{B})$ and the component tree $\Upsilon_{\circ}(\mathbb{A}, \mathbb{B})$.

3. Find an optimal cover (i.e. a cover with minimum cost) for $\Upsilon_{\circ}(\mathbb{A}, \mathbb{B})$.

4. Compute the inversion distance $d_{\text{INV}}(\mathbb{A}, \mathbb{B})$.

Exercise 2 (Bounds for the SCJ distance) (6 pts)

Theoretical bounds for the SCJ distance with respect to the DCJ and the inversion distances are

$$d_{\text{DCJ}}(\mathbb{A}, \mathbb{B}) \leq d_{\text{SCJ}}(\mathbb{A}, \mathbb{B}) \leq 4\, d_{\text{DCJ}}(\mathbb{A}, \mathbb{B})$$

$$2\, d_{\text{INV}}(\mathbb{A}, \mathbb{B}) \leq d_{\text{SCJ}}(\mathbb{A}, \mathbb{B}) \leq 4\, d_{\text{INV}}(\mathbb{A}, \mathbb{B})$$

For each one of these four bounds:

Show that it is tight by giving an example of pairs of mutually distinct genomes that fulfill it.