Algorithms in Genome Research
Winter 2025/2026

Exercises

Number 11, Discussion: 2026-January-30

1. Consider the following canonical pair of genomes:

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

   (a) Draw the relational graph $RG(\mathbb{A}, \mathbb{B})$ and compute the DCJ distance between $\mathbb{A}$ and $\mathbb{B}$.

   (b) Give the optimal DCJ sorting scenario from $\mathbb{A}$ to $\mathbb{B}$ that follows the greedy algorithm of first creating the adjacencies of $\mathbb{B}$ from left to right, then the left telomere of $\mathbb{B}$ and finally the right telomere of $\mathbb{B}$. Name the operations in your sorting scenario (fusions, fissions, inversions, translocations, etc).

   (c) If this optimal sorting scenario contains an intermediate genome with more than one circular chromosome, find an alternative optimal scenario in which each circular intermediate chromosome is immediately reintegrated. (No need to worry about running time of your procedure.)

   (d) Can you find yet another alternative optimal scenario without circular intermediates?

2. Solution space of sorting by DCJ:

   (a) 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?

   (b) Given two canonical genomes $\mathbb{G}_1$ and $\mathbb{G}_2$, let C be a k-cycle (with $k \geq 4$) in $RG(\mathbb{G}_1, \mathbb{G}_2)$. Denote by $E_\alpha(\mathbb{G}_1, C)$ the set of edges corresponding to adjacencies of genome $\mathbb{G}_1$ that are in C.

   What is the number of distinct DCJ operations that modify genome $\mathbb{G}_1$ and split C into two cycles?

   (*Hint: For each pair of edges in $E_\alpha(\mathbb{G}_1, C)$, there is exactly one DCJ splitting C into two cycles.*)

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

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

   (a) Give an example of a pair of mutually distinct genomes showing that this bound is tight.

   (b) Determine a tight upper bound for the SCJ distance with respect to the DCJ distance and explain your answer.