Algorithms in Comparative Genomics

Exercise sheet 6, 19.5.2021

Exercise 1 (DCJ halving) (8 pts)

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

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

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

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:

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

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

Exercise 3 (DCJ double distance) (6 pts)

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