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 10, discussion: 04.07.2025

Exercise 1 (Family-free DCJ and DCJ-indel distances)
Given the following two circular genomes and gene similarities:

$$\mathbb{A} = (\mathtt{X}_1\ \mathtt{X}_2\ \mathtt{X}_3\ \mathtt{X}_4\ \mathtt{X}_5)$$

and

$$\mathbb{B} = (\mathtt{Y}_1\ \overline{\mathtt{Y}}_2\ \overline{\mathtt{Y}}_3\ \overline{\mathtt{Y}}_4\ \mathtt{Y}_5\ \mathtt{Y}_6)$$

Genome $\mathbb{A}$	Genome $\mathbb{B}$	Similarity (σ)
$\mathtt{X}_1$	$\mathtt{Y}_1$	0.8
$\mathtt{X}_1$	$\mathtt{Y}_2$	0.8
$\mathtt{X}_1$	$\mathtt{Y}_4$	0.4
$\mathtt{X}_2$	$\mathtt{Y}_2$	0.7
$\mathtt{X}_2$	$\mathtt{Y}_3$	0.3
$\mathtt{X}_3$	$\mathtt{Y}_2$	0.7
$\mathtt{X}_3$	$\mathtt{Y}_4$	0.6
$\mathtt{X}_4$	$\mathtt{Y}_3$	0.6
$\mathtt{X}_4$	$\mathtt{Y}_5$	0.2
$\mathtt{X}_5$	$\mathtt{Y}_5$	0.7
$\mathtt{X}_5$	$\mathtt{Y}_6$	0.6

1. Draw the gene similarity graph of $G_\sigma(\mathbb{A}, \mathbb{B})$. Recall that not only the edges, but also the vertices have weights.

2. Find a maximal matching M_1 of $G_\sigma(\mathbb{A}, \mathbb{B})$ such that (i) M_1 includes the edge from $\mathtt{X}_2$ to $\mathtt{Y}_2$ and (ii) M_1 has the maximum possible weight. Give the weight $w(M_1)$ and the complement $\widetilde{M_1}$ with its corresponding weight $w(\widetilde{M_1})$. Draw the (weighted) relational graph that can be derived from M_1 and calculate the weighted DCJ-indel distance $\mathrm{wd}^{\mathrm{ID}}_{\mathrm{DCJ}}(\mathbb{A}, \mathbb{B}, M_1)$.

3. Find a maximal matching M_2 of $G_\sigma(\mathbb{A}, \mathbb{B})$ such that (i) M_2 does not include the edge from $\mathtt{X}_2$ to $\mathtt{Y}_2$ and (ii) M_2 has the maximum possible weight. Give the weight $w(M_2)$ and the complement $\widetilde{M_2}$ with its corresponding weight $w(\widetilde{M_2})$. Draw the (weighted) relational graph that can be derived from M_2 and calculate the weighted DCJ-indel distance $\mathrm{wd}^{\mathrm{ID}}_{\mathrm{DCJ}}(\mathbb{A}, \mathbb{B}, M_2)$.

4. Draw the family-free multirelational diagram of $\mathbb{A}$ and $\mathbb{B}$. Recall that every gene has an indel edge and that extremity and indel edges have weights.