

Algorithms in Genome Research
Winter 2025/2026

Exercises

Number 12, Discussion: 2026-February-6

1. Consider the following canonical genomes

$$\mathbb{G}_1 = \{ [1 \ 2 \ 3 \ 4 \ 5] \}, \mathbb{G}_2 = \{ [1 \ 2 \ \bar{3} \ 5 \ 4] \}, \mathbb{G}_3 = \{ [\bar{2} \ 3 \ 1 \ \bar{4} \ \bar{5}] \} \text{ and } \mathbb{G}_4 = \{ [4 \ 5 \ 3 \ 1 \ 2] \}$$

and let $\mathcal{S}^3 = \{\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3\}$ and $\mathcal{S}^4 = \mathcal{S}^3 \cup \{\mathbb{G}_4\}$. For each of the two sets \mathcal{S}^3 and \mathcal{S}^4 :

- (a) Compute a general SCJ median $\mathbb{M}_{\text{SCJ}}^k$ of \mathcal{S}^k and give its score.
- (b) Is there another SCJ median of \mathcal{S}^k that is distinct from $\mathbb{M}_{\text{SCJ}}^k$?
(Justify your answer by giving a distinct median or explaining why it does not exist.)

Is $\mathbb{M}_{\text{SCJ}}^3$ also a breakpoint median of $\mathbb{G}_1, \mathbb{G}_2$ and \mathbb{G}_3 ?

If *no*: Compute a breakpoint median of $\mathbb{G}_1, \mathbb{G}_2$ and \mathbb{G}_3 .

If *yes*: Is there another breakpoint median of $\mathbb{G}_1, \mathbb{G}_2$ and \mathbb{G}_3 that is distinct from $\mathbb{M}_{\text{SCJ}}^3$?
(Justify your answer by giving a distinct median or explaining why it does not exist.)

2. Breakpoint median.

- (a) Given three canonical linear genomes $\mathbb{L}_1, \mathbb{L}_2$ and \mathbb{L}_3 , let \mathbb{M} be the general breakpoint median of $\mathbb{L}_1, \mathbb{L}_2$ and \mathbb{L}_3 . Let \mathbb{M}_L be a genome obtained by breaking, for each circular chromosome of \mathbb{M} , one adjacency with smallest weight into two telomeres. Give an example showing that the obtained \mathbb{M}_L is not necessarily a linear breakpoint median of $\mathbb{L}_1, \mathbb{L}_2$ and \mathbb{L}_3 . (Hint: It is enough to consider canonical genomes with four genes.)
- (b) Can you design an exact polynomial time algorithm to compute the linear breakpoint median of three canonical linear genomes? Justify your answer by describing the algorithm or explaining why it cannot be designed.

3. Let **UNICIRCBPMEDIAN** be the problem of computing a unichromosomal circular breakpoint median of three given canonical unichromosomal circular genomes.

- (a) Show how **UNICIRCBPMEDIAN** can be reduced to the NP-hard Travelling Salesman Problem (TSP).
- (b) What does this reduction mean for the complexity of **UNICIRCBPMEDIAN**?