

Algorithms in Comparative Genomics
Summer 2018

Exercises

Number 2, return 2018 May 25

1. Given two signed genomes

$$A = [3 \ 4 \ -8] \ [2 \ 1 \ 7 \ 5 \ 6] \ (11 \ 12 \ -13 \ 9 \ 10) \ [-15 \ -14]$$

and

$$B = [1 \ 2 \ 3 \ 4 \ 5] \ (6 \ 7) \ [8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \ 15].$$

- (a) What is the breakpoint distance between A and B ?
(b) What is the SCJ distance between A and B ?
(c) Explain the difference between the two distances (if any).
2. The SCJ and breakpoint distances are given by

$$d_{\text{SCJ}} = 2n - 2c_2 - p \quad \text{and} \quad d_{\text{BP}} = n - a - t/2,$$

where n is the number of genes, p and c_2 are, respectively, the number of paths and the number of cycles of length 2 of the adjacency graph, and a and t are the number of common adjacencies and the number of common telomeres, respectively. The intuitive relationship between the SCJ and BP distance is $d_{\text{SCJ}} = 2d_{\text{BP}}$. Using the above equations, describe for which types of genomes this relationship is true.

3. Theoretical bounds for the SCJ distance with respect to the breakpoint distance and the DCJ distance are

$$d_{\text{BP}}(A, B) \leq d_{\text{SCJ}}(A, B) \leq 2 d_{\text{BP}}(A, B)$$

and

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

Give examples showing that these bounds are tight.