

BIN506 Assignment 2

2024-2025 Spring

Due Date: 18 March, 23.59

Late Submission Policy: After the due time, 10 points of deduction will be applied for each extra hour. No submission will be allowed after 19 March 09.40.

Uploading your assignment as a PDF is mandatory.

Following supplementary readings are highly recommended:

- [BLAST: Program Selection Guide](#) (Section 6.1 will be especially helpful.)
- [BLAST: Frequently Asked Questions](#)

Assignment Questions

1. What are the differences between **local**, **global**, and **global with no end-gap penalty** alignment methods?
2. Between **local** and **global** alignment methods, which one provides biologically more relevant information? Explain your reasoning.
3. Use [bl2seq](#) to align **NP_000509.1** to **NP_000549.1**
 - a. Which BLAST program did you choose? Why?
 - b. What is the sequence type in the final alignment?
 - c. What are the significance value and the total score of the alignment? Define them, and comment on the result you get.
4. Use [bl2seq](#) to align **NM_000558.3** to **NP_000549.1**
 - a. Which BLAST program did you choose? Why?
 - b. What is the sequence type in the final alignment?
 - c. What are the significance value and the total score of the alignment? Comment on the result you get.
5. Use **seq1.fasta** to **seq2.fasta** to answer the following questions.
 - a. Use [dotmatcher](#) to align the sequences with the given parameters and interpret the graph: Window size = 10, Threshold = 50.
 - b. Use [EMBOSS Needle](#) and [EMBOSS Water](#) to align the sequences. Compare and elaborate on the results. Explain why the resulting alignment is shorter in length in EMBOSS Water.
 - c. Look at the dotplot and the alignment results. Do the alignment results satisfy the information provided by the dotplot? Explain briefly.

During all alignments, please consider which sequence you use as the database sequence, and which sequence you use as the query sequence.