

# Principles in Bioinformatics I

## Instructions for the final project

Each student received an unlabeled sequence from different origins (viral, eukaryote, or prokaryote). In addition, each student must have their skills in the use of bioinformatics tools and programming to:

- 1) Design and execute a strategy to evaluate which organisms the sequences are from.
- 2) Search the genes encoded in each sequence (try with more than one gene search strategy if necessary).
- 3) Design and execute a strategy to obtain functional information from those genes.
- 4) Discuss the relationship between the available information concerning the organism's biology, physiology, or lifestyle and some features obtained from their sequences.
- 5) Present an example of a study of the genes, metabolic pathways, or genetic components found in their respective genomes that can be studied as a whole (e.g., a genomic island, a metabolic process, or an extended paralog family). Use comparison with related organisms or models to complement your study.

## Evaluation criteria

The project will be evaluated by considering three components: a written paper-like manuscript (corresponding to 50% of the final qualification) and an oral presentation (the other 50%) presenting the strategy and its results.

The manuscript is required to have the following components (with their respective scores):

- Format (10%): It must follow the *Frontiers* format for *Original Research* papers. *LaTeX* and Word templates can be found at <https://www.frontiersin.org/about/author-guidelines> (section 1.2).
- Content (30%): The manuscript must contain a clear, straightforward narrative. It must be composed of the following sections:
  - A title.
  - An abstract (100-250 words long).
  - An introduction/background section focused on the studied organism, showing the required antecedents for a proper understanding of the proposed problem and how you establish the aims of the manuscript.
  - A methodology section focused on explaining the principles of the tools utilized and the logic justifying their use. In addition, you must also discuss how scalable your method is.
  - A “Results and Discussion” section explores the interpretation of the results, the biological significance of the results obtained from the strategy, how those results address the question you originally proposed, and what kind of future studies or approaches can address open questions.

- A “References” section. Use the “Frontiers” format for those references.
- A Supplementary Material section. This part must contain the commands or the scripts generated and utilized during the analysis. They can be sent in a ZIP file.
- Narrative - Strategy (30%): The manuscript narrative must be focused on both the utilized bioinformatics strategies and the biological significance of the obtained results. The evaluation of this section depends on the method’s relevance.
- Narrative - Biological significance (30%): Consider the questions arising from your exploration of the unlabeled sequences and develop a biological exploration of the data.

The oral presentation must have PowerPoint slides (or a similar tool). The time limit is 25 minutes. The presentation must have a coherent narrative, focusing on both the strategy and the significance of the results. After the presentation, the teacher group will ask questions about the presentation's content. The quality of the answers will also be considered in the final oral presentation qualifications.

The components to be evaluated in the oral presentation are:

- Clarity and briefness (20%): Respect the 20-minute length. Your exposition must be clear and concise, containing well-separated sections.
- Content (60%): Develop your content considering the aforementioned aspects in the manuscript's evaluation.
- Questions (20%): quality of the answers to the questions from the teachers.

## Proposed deadlines

**Send the manuscript before June 13, 2025, 11:59 pm.**

**Oral presentations on June 18, 2025, 2 pm - 4:40 pm.**

# Manuscript Evaluation Guidelines

Category	Level: Low (<10%)	Level: Medium (50%)	Level: High (100%)
Format (10%)	<ul style="list-style-type: none"> <li>* No Frontiers document format was used</li> <li>* Absence of proposed manuscript sections</li> <li>* No page lines were used</li> <li>* No Frontiers citation format was used</li> </ul>	Partial accomplishment of the following components: <ul style="list-style-type: none"> <li>* Frontiers document format</li> <li>* Clear manuscript sections</li> <li>* Use of page lines</li> <li>* Use of Frontiers citation format</li> </ul>	Total accomplishment of the following components: <ul style="list-style-type: none"> <li>* Frontiers document format</li> <li>* Clear manuscript sections</li> <li>* Use of page lines</li> <li>* Use of Frontiers citation format</li> </ul>
Content (30%)	<ul style="list-style-type: none"> <li>* Absence of manuscript sections</li> <li>* No connection between the introduction and the aim of the manuscript</li> <li>* No connections between elements from Methods and Results/Discussion sections</li> <li>* Incomplete Discussion section, with no final conclusions</li> </ul>	Partial accomplishment of the following components: <ul style="list-style-type: none"> <li>* Presence and boundaries of manuscript sections</li> <li>* Effective connection between the introduction and the aim of the manuscript</li> <li>* Effective connection between elements from Methods and Results/Discussion sections</li> <li>* Degree of completeness of the Discussion section, including final remarks</li> </ul>	Total accomplishment of the following components: <ul style="list-style-type: none"> <li>* Presence and boundaries of manuscript sections</li> <li>* Effective connection between the introduction and the aim of the manuscript</li> <li>* Effective connection between elements from Methods and Results/Discussion sections</li> <li>* Degree of completeness of the Discussion section, including final remarks</li> </ul>
Narrative - Strategy (30%)	<ul style="list-style-type: none"> <li>* Absence of connection between the introduction and the formulation of the strategy</li> <li>* No explanation about why the strategy is useful for a certain biological problem</li> <li>* Absence of refe</li> <li>* No analysis or discussion of the results</li> <li>* No discussion or evaluation of changes in the scalability</li> </ul>	Partial accomplishment of the following components: <ul style="list-style-type: none"> <li>* The connection between the introduction and the formulation of the strategy is established</li> <li>* Available explanation about why the strategy is useful for a certain biological problem</li> <li>* Available combination between different tools and/or scripts to generate the strategy</li> <li>* Available comparison between 2 or more tools for at least 1 step of the overall strategy</li> <li>* Available analysis or discussion of the results</li> <li>* Available discussion or evaluation of changes in the scalability</li> </ul>	Total accomplishment of the following components: <ul style="list-style-type: none"> <li>* The connection between the introduction and the formulation of the strategy is established</li> <li>* Available explanation about why the strategy is useful for a certain biological problem</li> <li>* Available combination between different tools and/or scripts to generate the strategy</li> <li>* Available comparison between 2 or more tools for at least 1 step of the overall strategy</li> <li>* Available analysis or discussion of the results</li> <li>* Available discussion or evaluation of changes in the scalability</li> </ul>
Narrative - Biological Significance (30%)	<ul style="list-style-type: none"> <li>* In the introduction, the justification of strategy is not presented at all</li> <li>* The presented strategy does not contain any step focused on obtaining special or specific data</li> <li>* In Methods, the strategy description is insufficient to be reproduced</li> <li>* In Methods, no information about pre-requisite software/dependencies are shown</li> <li>* In Results and/or Discussion, the strategy could not show its</li> </ul>	Partial accomplishment of the following components: <ul style="list-style-type: none"> <li>* In the introduction, the justification of strategy is presented</li> <li>* The presented strategy does contain steps focused on obtaining special or specific data</li> <li>* In Methods, the strategy description allows the reproduction</li> <li>* In Methods, there is available information about pre-requisite software/dependencies</li> <li>* In Results and/or Discussion,</li> </ul>	Total accomplishment of the following components: <ul style="list-style-type: none"> <li>* In the introduction, the justification of strategy is presented</li> <li>* The presented strategy does contain steps focused on obtaining special or specific data</li> <li>* In Methods, the strategy description allows the reproduction</li> <li>* In Methods, there is available information about pre-requisite software/dependencies</li> <li>* In Results and/or Discussion,</li> </ul>

	utility or significance	the strategy is discussed in its utility or significance	the strategy is discussed in its utility or significance
--	-------------------------	--	--