Computational Biology Homework Guide 2021

# Overview

You will be asked to complete four homework assignments throughout the semester. Here are the directions for each of them. I have written the instructions in R with all the information you need to decipher them in this document.

####HW 1####

paper = list of computational biology papers

*for (i in 1:4) do {summary(paper[[i]]) }* ***#Note:*** *Each paper must fall into a different discipline of biology.*

####HW 2####

paper = list of biological modelling papers

*for (i in 1:2) do {analyze(paper[[i]]) } #****Note:*** *Each paper must use a different model, but they can be about the same topic in biology.*

####HW 3####

paper = list of computational ecology papers

*for (i in 1:2) do {analyze(paper[[i]]) }*

####HW 4####

paper = list of bioinformatics papers

*for (i in 1:2) do {analyze(paper[[i]]) } #Note: For our purposes, bioinformatics concerns the analysis or processing of molecular biology data.*

You can (and should) reuse the same papers for multiple assignments. For example, in HW 1 you can try to pick computational ecology and bioinformatic papers (ideally ones that include questions about modelling). You can also choose to present one of these papers later in the semester.

# summary()

Write a bulleted summary for each paper you address. The summary should include the following information.

1. The main biological question/problem addressed. (1 sentence per main question)
2. A brief description of the computational elements (even if superficial) utilized in the research and how it relates to addressing the main problem. (no more than 3 sentences for each computational element)
3. How the authors shared their computational elements. Did they write a software package? Did they share their code on GitHub? Did they only give the algorithms/models in the paper? Search for them and link to them.
4. How the findings of the research contribute to science or society (i.e., broad conclusions of the paper)

# analyze()

For each paper you address, answer the following questions.

1. What coding language was used to build/execute the computational elements of the paper?
2. Describe (in detail) the computational elements in the paper and how they address a main question/problem. Describe software packages, algorithms, models, and basic tools utilized.
3. What computational issues had to be overcome? (for example: was computational time, memory, data availability, or other resources a limiting issue?)
4. Was this a theoretical or empirical paper? If theoretical only, how can the authors test their theory in the real world? If empirical and/or theoretical, how do the findings match with what is expected from the underlying theory (e.g. simulations)?
5. Did this study address “big data”? If so, how did they scale their method? If not, how could they scale their method?

# Where to find papers?

Ultimately, any peer-reviewed scientific journal publishing biology papers could have research about computational biology. However, there are some that are more likely to have them than others. Here are some excellent journals where you can find papers.

*NOTE: Don’t feel pressure to cover papers that are strictly within the discipline of computational biology (e.g., those that develop new software, tools, or algorithms). As long as there is a strong computational element to the research presented and research paper in biology is acceptable.*

* Nature Methods
* Genome Research
* Nature Structural and Molecular Biology
* Genome Biology
* Nucleic Acids Research
* EMBO Journal
* Systematic Biology
* Molecular Biology and Evolution
* Bioinformatics
* Bioinformatics Advances
* BMC Bioinformatics
* Applied Bioinformatics
* Briefings in Bioinformatics
* Molecular Ecology
* PLoS Computational Biology
* Journal of Computational Biology
* Ecological Informatics
* Nature Ecology and Evolution
* Global Change Biology
* Conservation Letters
* Journal of Ecology
* International Society for Computational Biology Community Journal