

# Pràctica UF2.Pt (Òmiques)

**Institut:** Provençana (L'Hospitalet de Llobregat)

**Curs:** 2022-2023

**Grup:** DAWBIO-2

**Mòdul:** M14 Mòdul de Bioinformàtica

**UF:** UF2 – Bioinformàtica

**Instrument d'avaluació:** UF2.Pt

**Tema:** Ciències Òmiques

**Data:** 2023-02-10

## Notes:

- Grups:
  - La pràctica és fa en grups de dues o tres persones.
- Pràctica:
  - La pràctica consisteix en fer un estudi com el que faria un investigador interessat en comparar un gen present a diverses espècies.
  - Escolliu el següent:
    - Un organisme A que us interressi.
    - Diverses espècies de l'organisme A.
    - Almenys un gen G present en les espècies.
  - Descarregueu-vos informació dels gens, extraieu informació del gen, alineu les seqüències i mireu d'escriure alguna conclusió.
  - El codi, la taula resum i les conclusions han d'estar en fitxers .py.
  - Apunteu el vostre grup al full de càlcul compartit (PENDING).
  - És recomanat consultar els organismes i gens amb el professor.
- Tasques:
  - L'estudi està dividit en cinc tasques. Teniu els detalls a la pàgina següent.
- Puntuació:
  - Cada tasca val 1,00 punt. La puntuació màxima de la pràctica és 5,00 punts.
- Moodle:
  - Poseu el nom del grup i de tots els estudiants al principi del fitxer .py.
  - Comprimeu tots els fitxers addicionals (genbanks, etc) en un arxiu **.zip**.
  - L'arxiu ha de contenir tots els fitxers necessaris per poder-se executar **a Linux**.
  - El nom del zip ha de ser «grup-uf2-pt-omiques.zip».
  - Si el fitxer no segueix aquest format no es corregirà.
  - Pugeu l'arxiu .zip a la tasca del moodle abans de la data límit.
  - El .zip ha de contenir el codi de tothom. Només cal que ho pugui un membre del grup.

## Tasques

Escolliu el següent:

- Un organisme A.
- Diverses espècies de l'organisme A.
- Almenys un gen G present en les espècies.

Descarregueu-vos informació del gen G a cadascuna de les espècies, extraieu informació del gen, alineu les seqüències i mireu d'escriure alguna conclusió.

- L'important no és la conclusió, si no el procés realitzat.
- S'han de comparar unes 10 seqüències aproximadament.
- Si no teniu suficients espècies per arribar a les 10 seqüències, escolliu més d'un organisme.
- Cada grup ha d'estudiar un gen i organisme diferent.

### 1. Formular la hipòtesi (1,00 punt)

- Formuleu una hipòtesi que voleu validar sobre el gen i les espècies escollides.
- L'important és fer-ho de forma lògica i amb sentit.
- No importa la seva rellevància científica ni si ja es coneix la resposta o no.

### 2. Descarregar Genbanks utilitzant Bio.Entrez (1,00 punt)

- Feu la cerca d'informació utilitzant Bio.Entrez i escolliu els camps correctament. Podeu fer proves prèvies manualment a la interfície web del NCBI.
- Utilitzeu BioPython per descarregar-vos de forma automàtica les fitxes GenBank.

### 3. Extreure informació utilitzant Regexps (1,00 punt)

- Les fitxes Genbank contenen molta informació textual.
- Utilitzeu expressions regulars per extreure informació útil de les fitxes.
- Aquesta informació després l'haureu de posar a la taula resum final.

### 4. Alineament de seqüències (1,00 punt)

- Utilitzeu BioPython per alinear una de les seqüències amb la resta de seqüències.
- Hi ha d'haver al menys 10 comparacions.
- Guardeu-vos la puntuació de l'alineament i ordeneu les comparacions segons la similitud.

## 5. Mostrar els resultats (1,00 punt)

- Mostreu una taula resum feta en Pandas on es mostrin com a mínim les següents columnes:
  - Accession number de la seqüència.
  - Nom descriptiu de la seqüència.
  - Informació rellevant extreta amb regexps.
  - Puntuació de l'alineament.
- La taula ha d'estar ordenada per la puntuació de l'alineament, de major a menor puntuació.
- Estudieu la similitud de les seqüències amb puntuació major i les seves fitxes Genbank.
- Intenteu extreure alguna conclusió i validar o no la vostra hipòtesi. Un paràgraf o dos ja és suficient. L'important no és la conclusió si no veure que heu fet tot el procés correctament.
- El codi, la taula i les conclusions han d'estar en un JupyterLab notebook.

## Avaluació

- Tots els alumnes han d'haver programat codi Python.
- El codi ha de funcionar per ser avaluat, si no la pràctica es puntuarà amb un zero.
- Els alumnes hauran de fer una demostració i respondre les preguntes del professor. El professor pot demanar canvis en el codi.
- Els alumnes han de respondre correctament les preguntes sobre el seu propi codi i fer els canvis demanats (si és el cas), si no la pràctica es puntuarà amb un zero.

## Recursos

- **NCBI:** <https://www.ncbi.nlm.nih.gov/>
- **BioPython:** <https://biopython.org/>
- **EOL:** <https://eol.org/>