



# Вовед во биоинформатика

Лабораториски вежби

# Подготовка

- ◆ Инсталација на **Biopython** во **Google colab**.
- ◆ Користејќи ја документацијата дадена на линкот да се реализираат следните задачи дадени во презентацијата.
- ◆ <http://biopython.org/DIST/docs/tutorial/Tutorial.html>

# Задачи

1. Да се реализираат примери за парсирање на FASTA и GenBank формати, онака како што е објаснето во поглавје 2.4 од туторијалот.
2. Да се напишат 3 примери за поврзување со различни биолошки бази, како што е описано во поглавје 2.5 од туторијалот.
3. Да се напише пример за работа со Seq објекти, какошто е описано во поглавје 3, и да се обрне внимание на процесите транскрипција и транслација.

## Задачи

4. Да се преземе комплетната ДНА секвенција на *Yersinia pestis* (бактерија која ги инфектира белите дробови и предизвикува пневмонија). До неа се пристапува со употреба на идентификациониот број **NC\_005816** во GeneBank базата (работете со SeqIOобјекти, имате детали во документацијата).
5. Имплементирајте репликација на секвенцијата потпомогната од функции во `biopython`.

## Задачи

6. Со помош на `biopython`, пронајдете ги секвенциите на различните кодни региони означени како (CDS). CDS се регионите добиени после процедурата на отсекување на инtronите.
7. Одберете еден CDS и најдете ги сите старт и стоп кодони во неговата РНА.
8. Направете транслација за да ги добиете сите можни протеински секвенции.

## Задачи

9. Одберете еден од кодните региони и направете мутација со поместување на рамката за 2 нуклеотиди. Анализирајте ги новите протеински секвенции. Внимавајте, сега има нови старт и стоп кодони.
10. Да се разработи поглавје 5.5 за креирање на сопствени записи во fasta формат.
11. Да реализира примерот за предвидување на структурата на оперонот кај бактеријата *Bacillus subtilis* описан во поглавје 16 од туторијалот. Да се употреби методот на Логистичка регресија и Машини со Поддржувачки Вектори (SVM) и да се споредат резултатите.

## Задачи

12. Да се испрограмира алгоритмот на Nussinov како функција во python која на влез ќе добива стринг од нуклеотиди, а на излезе ќе ја дава нивната секундарна структура.

Влез нека биде следната tRNA молекула:

ACCGCGGGCGUCCGGCUCAGCUUGCCGCUGCAGCGCCGCAGUAGCCGAAACU  
UCCGGCUCGGGCCGUGGGCGGUCUCUAGUCGCCCG

13. Да се истражи можноста на biopython за одредување на секундарна структура описана во поглавје 11.6.8.