

Laboratorul 2

În acest laborator vom lucra cu fișiere.

Fișierele în python se deschid asemănător cu fișierele din C sau C++. Funcția pentru deschidere a unui fișier este `open()` iar parametri acestora sunt fișierul și modul în care se face citirea. Avem patru moduri de deschidere a unui fișier:

- 'r' – Acest mod deschide fișierul doar pentru citire;
- 'w' – Acest mod deschide fișierul doar pentru scriere;
- 'a' – Acest mod deschide fișierul pentru adăugare la sfârșitul fișierului;
- 'r+' – Acest mod deschide fișierul atât pentru citire cât și pentru scriere.

Pentru citirea informației din fișier folosim funcția `read()`. Pentru scrierea în fișier folosim funcția `write()`.

Exemplu

```
infile = open("sequence.txt", "r")
outfile = open("DNA.txt", "w")
firstLine = infile.read()
outfile.write(firstLine)
```

Exerciții

1. Intrați pe Ensembl și descărcați secvența fasta pentru gena FOXP2. Deschideți fișierul cu python. Extrageți secvențele aferente exonilor, excluzând intronii, folosind un model simplificat (începutul unui exon este ATG iar finalul lui poate fi TAA, TAG sau TGA). Extrageți toate cele șase combinații posibile (reading frames). Secvențele de exoni extrase se vor salva pe rânduri diferite în fișierul `exoni.txt`.
2. Preluati secvențele exonilor de la exercițiul anterior și generați secvențele de aminoacizi aferente. Salvați cele șase secvențe într-un fișier denumit `proteine.txt`.
3. Schimbați fișierele cu secvențele exonilor și aminoacizilor cu unul dintre colegi. Verificați dacă secvențele corespund.

Resurse:

<http://www.cbs.dtu.dk/courses/27619/codon.html>

https://en.wikipedia.org/wiki/Open_reading_frame