

Esame 08/11/2018

Andrea Passerini
passerini@disi.unitn.it

Informatica

Programma python

Scrivere un programma python che:

- prenda in ingresso un file con una sequenza genomica in formato FASTA, ed uno con una tabella di percentuale di utilizzo dei codoni nel genoma corrispondente.
- stampi per ogni codone le percentuali di utilizzo nella sequenza di ingresso e nel genoma.

- File FASTA

```
>lcl|NC_000067.6_cds_XP_006495613.1_1 ...  
ATGGCCGCTAAGTCAGACGGGAGGCTGAAGATGAAGAA...  
TTCGGGCTCTGTGCAAGGACTGGCTCCAGGCTTGCCGT...  
GCTGCCCGGACGGCGGTGGCTGCTCGCGCTGCTGCTGC...  
...
```

- Tabella utilizzo codoni

```
TTT 0.0172  
TCT 0.0162  
TAT 0.0122  
TGT 0.0114  
TTC 0.0218  
TCC 0.0181  
TAC 0.0161  
...
```

Esempio esecuzione

```
> python codon_usage_difference.py
Usage table file: codon_usage.txt
Sequence file: gene.fasta
GCC      0.0448  0.0260
GCT      0.0139  0.0200
AAG      0.0185  0.0336
TCA      0.0077  0.0118
GAC      0.0216  0.0260
GGG      0.0201  0.0152
AGG      0.0185  0.0122
CTG      0.0355  0.0395
AGC      0.0370  0.0197
GTG      0.0340  0.0284
...
```

Programma python: suggerimento

Si possono implementare 5 funzioni separate:

- 1 Una che legga il file FASTA e restituisca la sequenza del gene
- 2 Una che legga il file di utilizzo dei codoni e restituisca una mappa codone → percentuale di utilizzo
- 3 Una che data una sequenza restituisca il dizionario con le percentuali di utilizzo di ciascun codone (attenzione ai possibili modi di utilizzo della funzione `range`)
- 4 Una che dati due dizionari stampi per ogni chiave del primo i valori sia del primo sia del secondo
- 5 una che realizzi il programma richiesto usando le funzioni di cui sopra

Shell: esercizio #1

Ordinare e stampare a video le sequenze nel file `sequences.fasta` che non cominciano con una metionina (M), ma contenenti almeno una fenilalanina (F).
Esempio: **E****F**CD**F**SGNQAAGGVMVMDT...

Risultato atteso

```
AEETAALS...  
AESSDKLY...  
...
```

Shell: esercizio #2

Dato il file `leukemia.txt`, contare quanti valori di espressione negativi sono presenti nel file, mettendo insieme tutti i livelli di espressione di tutti i pazienti.

Risultato atteso

56401

Modalita' di esecuzione e consegna

- 1 Avviare la macchina in modalita' `ESAME`
- 2 Autenticarsi con nome utente `sci-esame` e password fornita dal docente
- 3 Il testo del compito ed i file necessari si trovano in una cartella `Testo` sul Desktop
- 4 Realizzare il programma python come file `programma.py` e scrivere gli esercizi da linea di comando in un file di testo `linea_di_comando.txt`
- 5 Creare sul Desktop una cartella con *nome_cognome* e metterci i due file realizzati.
- 6 Eseguire il logout ma NON spegnere la macchina