Assignment 2

The first task was to find the Open Reading Frames for the human, pantroglodytes, and the rat’s DNA. It was also to determine what fraction of the gene represents a potential protein codifying sequence.

Human (NC\_001807): 55%

Pantroglodytes (NC\_001292): 53%

Rat(NC\_005089): 61%

2. The second task was to randomize the gene sequence and list the bigger ORF size. The randomization I chose was to read the gene backwards. The book listed couple examples how to permute it, suck as, shifting each nucleotide individually or shitting the codons.

Open Reading Frame Size:

Human(NC\_001807): 360

Pantroglodytes(NC\_001292): 360

Rat(NC\_005089): 228

3.

The code is posted on Github on the following link: [orf.py](https://github.com/ICMC/Computation-Biology/blob/master/modulo2/ORF.py)