

## Bio Informatics

### Assignment 1

The first part of the assignment was to download the gene sequence from the *Bacteriophage Lambda* (# NC\_001416) from GenBank, write a program to count the “GC” count, and plot a chart. The code to accomplish this task was written in Python and is located on Github on the following address: [sequence.py](#)

On Figure 1, the chart shows that the “C” count increases when the “G” count increase as well. It also shows that the “GC” count increases with the increment of the window size.

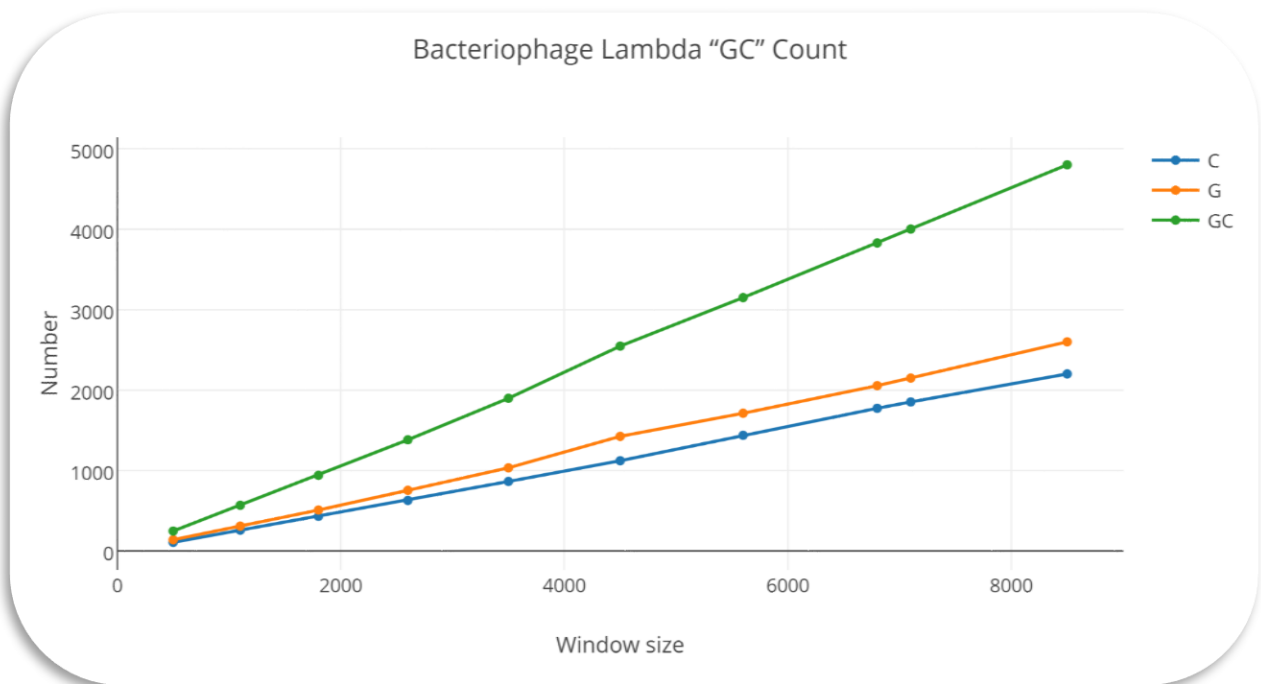


Figure 1. Bacteriophage Lambda “GC” count chart

The task for the following charts was to analyze the gene sequence of the Mitochondria of two different species. The first specie is Homosapiens Sapiens, the second species is Pantroglodytes.

Figure 2 shows the percentage of the nucleotides A,T,C, and G of both species. It is possible to observe that both genes have almost an identical percentage of those nucleotides on the mitochondria gene sequence. Meaning that the mitochondria that belongs to the Homosapiens and Pantroglodytes function alike.

The source code for Figure 2 is located at: [ACGTprobability.py](#)

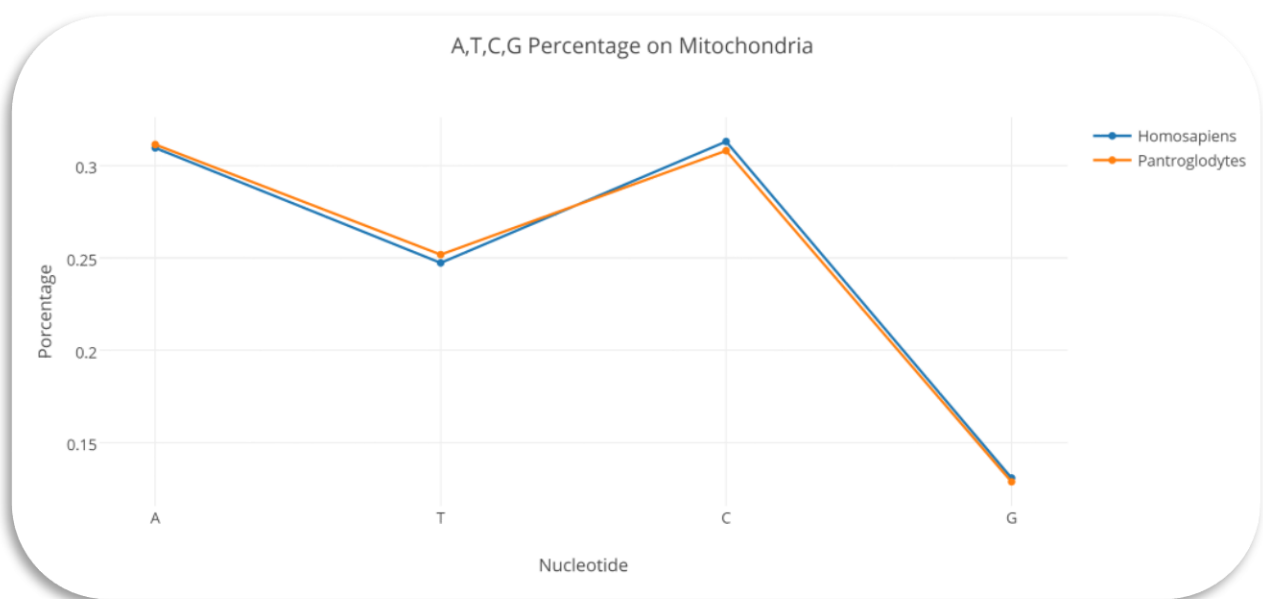


Figure 2. A,T,C,G Nucleotides Percentage

The last chart, Figure 3, shows the relation between the “GC” count between the Homosapiens and Pantroglodytes mitochondria. Again, the similarities are very high. In some of the windows the “GC” count is even the same. This way, we can assume that even though the genes come from different species they have a large number of similarities in the nucleotide count because they perform the same task.

The source code for the Figure 3 chart is located at: [mitochondria.py](https://github.com/anna-pawlicka-maule/mitochondria.py)

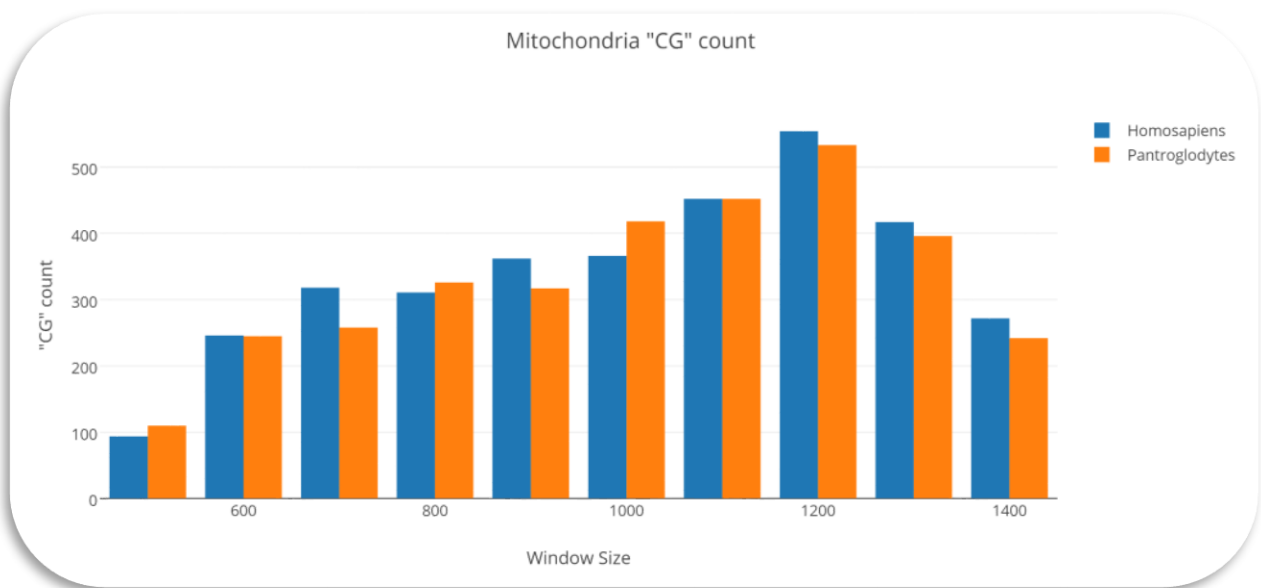


Figure 3. Mitochondria “GC” count