

GC Base Counts and Their Effects On Genomes

Ignas Ramanauskas 2023

Introduction

Using different techniques in Bash scripting and with the help of Vinogradov's published paper we analyzed GC content in yeast genomes to understand how their genetic makeup affects their functionality, mainly thermostability and bendability.

Methods

- ▶ Using multiple scripts like fasta-unfold, find-orfs and gc-count, aswell as python script, were used to analyze and modify data from genome fasta files and provide images with useful data.
- ▶ In A. E. Vinogradov's paper the genome data was collected from GeneBank. Bendability and curvature of nucleotide sequences were determined using the trinucleotide table of consensus values obtained from DNase I digestion and nucleosome positioning studies. The thermostability was determined using a unified dinucleotide table for free energy of melting [Vinogradov (2003)].

Results

- ▶ According to A. E. Vinogradov's research paper, it was found that higher GC percentages correlated with a slight decrease of thermostability
- ▶ With increased elevation of GC percentages it was seen that bendability also increased

Results

In Figure 1, the average percentages of GC bases in CLIB215 yeast species are presented. The GC content is consistently around 38 percent.

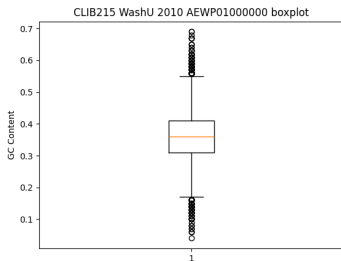
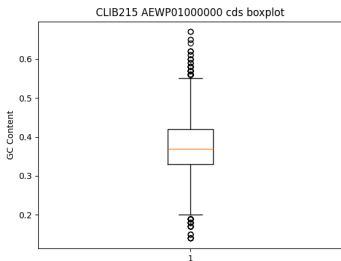


Figure: Average GC content percentages in CLIB215 yeast species from different databases.

Results

In this image we can see percentages of EC9-8 yeast species. This image provides us with a similar result as the first one.

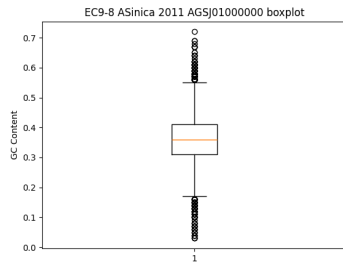
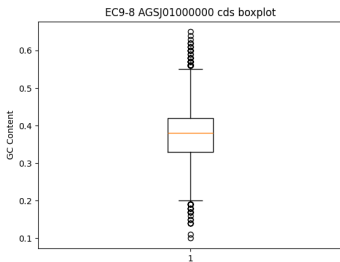


Figure: Average GC content percentages in EC9-8 yeast species from different databases.

Conclusions

- ▶ To summarise our results, according to A. E. Vinogradov, we saw that increased GC percentages in genomes resulted in slightly less thermostability, but with increased DNA's bendability, which could hint us that places with higher GC percentages are more likely to be coding regions.
- ▶ When comparing GC counts in different yeast species, we saw a similarity of GC percentages across all yeast species, which was expected. The differences in outliers between same species of yeast could be explained by having data from different databases.

References

Vinogradov, A. E. (2003). Dna helix: the importance of being gc-rich. *Nucleic Acids Research*, 31(7):1838–1844. Received January 21, 2003; Revised and Accepted February 12, 2003.