

Analysis report for the "gag" gene in 9 supposedly alike sequences from a database

Luis Jimenez

March 3, 2025

1 Introduction

The analyzed sequence extracted from GeneBank is KY574515.1, and the sequences used for the comparison by %GC content were extracted from a database only with genomes that have the target gene, which are as follows: P860667.1, KP860668.1, KP860669.1, KP860670.1, PP483791.1, PP483792.1, PP483793.1, PP483794.1, PP483795.1. Additionally, a comprehensive BLASTN analysis against the database was executed to calculate the percentage identity, suggesting potential validation of the target sequence's similarity to those from the database.

2 Graphic comparison via GC content

2.1 Density of % GC Content and % GC Content for the randomized sequences

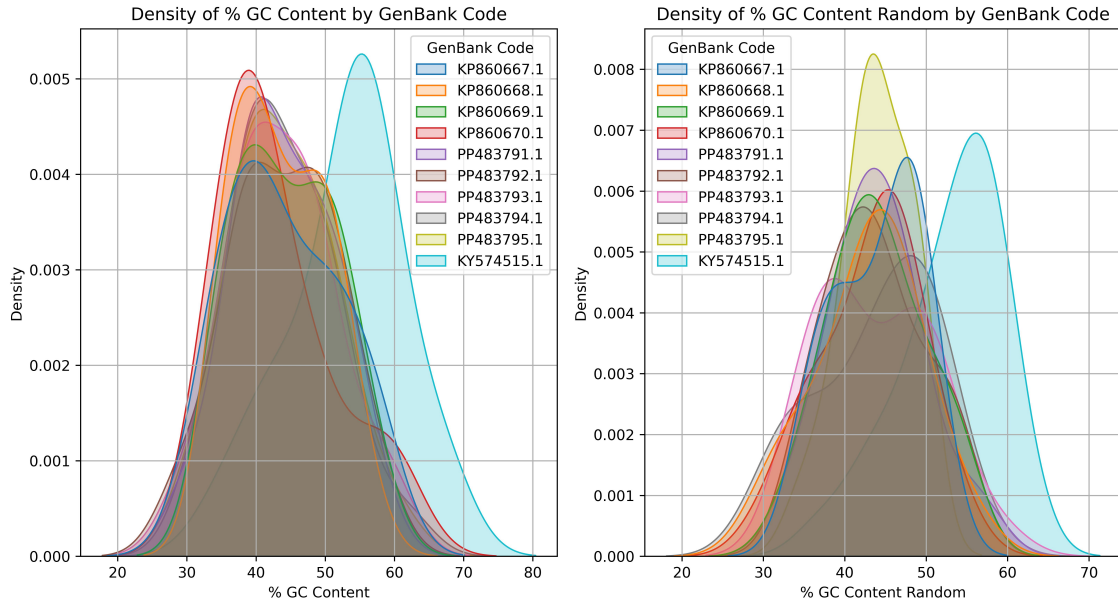


Figure 1: Density of % GC content in the sequences analyzed by fragments at a time.

2.2 Scatterplot of % GC Content by position and GenBank Code with regression lines

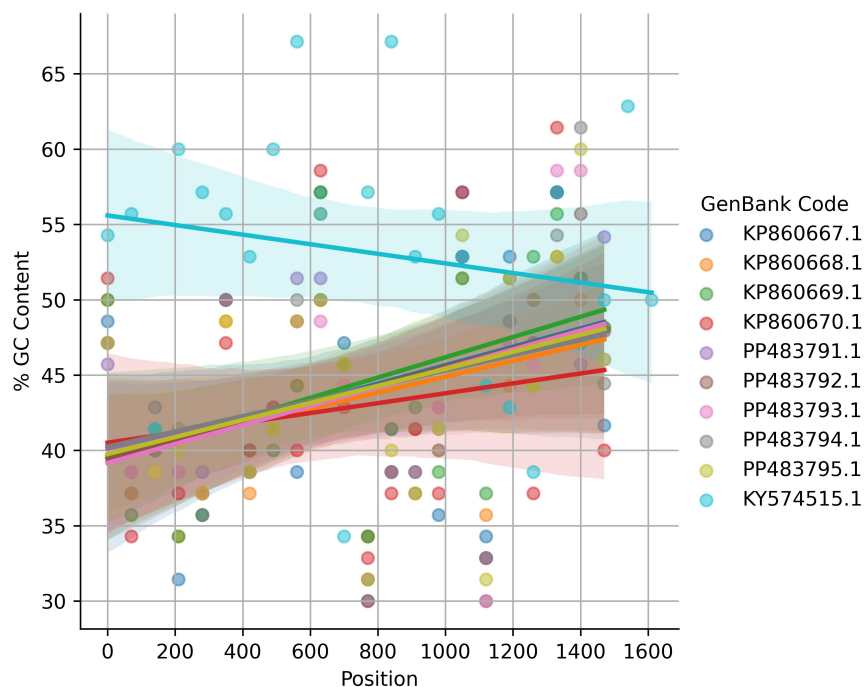


Figure 2: Dispersion and regression lines for % GC Content by position

3 BLASTN Analysis Results

The BLASTN analysis was performed to identify homologous sequences to the target sequence within the NCBI nucleotide database. The results revealed no significant matches, indicating a complete absence of detectable sequence similarity. This suggests that the target sequence does not share significant homology with any sequences currently deposited in the database.