Our Final Goal is

To design the Synthetic Minimal Genome of a bacterium, that the information on essential genes, such as direction, position, length and function, is essential. In addition a new analyzing method which calculates the distance between replication origin and each essential gene (DTO: Distance to origin) in each species, and provides the number of essential genes within 10 percent of total genome size are required. Using this method, we confirmed a distribution of essential genes in each organism.

In this study the information of essential genes will be obtained from DEG (Database of Essential Genes). We will re-group essential genes by CDG distribution for construction of our database which is connected to a software named GGD (Genome Organization Database & Designer).

Analysis

To determine the appropriate analytical method, basic analysis was performed. We classified group by rep, rep depend on Equality of strand & direction. We observed the frequency of entire and essential genes into rep and scale.

Analysis of Frequency (1st Step) - SAS

- Rep=1, same code both strand and direction are same, has more frequency at almost scales

Analysis of Frequency (2nd Step) - SAS

- As a result, with the strand and direction of the same sign, the specific features tended to be concentrated in to the distribution.

Analysis of Similar Genes

- Similar Genes have Not only same CDG number (function), but also same strand, direction and scale(position).
- Sorting by CDG Alphabet group
- Sorting by dendrogram

Analysis of Essential Genes

We needed a work that inference essential gene, so as to cover unidentified genes together.

- Using NCBI Protein BLAST

We chose a method to give a CDG number and to assume as essential gene by doing blast for 1079 species of genes and essential genes together. However, the method costs a lot of times, so we needed another method.

Analysis of Similar Genes - Java

- Similar genes have Not only same CDG number (function), but also same strand, direction and scale(position).
- Sorting by CDG Alphabet group
- Sorting by dendrogram

Results

Our main goal is to develop the Browser program having both Viewer and Designer. About Designer, it wasn’t completed yet because it was supposed to apply to the program after establishing Genome Organization Database. It only has the function of Viewer, made with Java.

Future Work

We try to discover Principle of genome organization of Essential genes. But we can’t search for the useful principle. We think that is attributable to target of analysis limited in 15 species and few of analysis too. Therefore, We need to find and develop new methods that be able to analyze the essential genes out of a essential gene pool we made. Also, we try to analyze the essential genes using various methods above DTO and frequency. If we find the Principle, we could apply to design minimal genome or chromosomal.

Reference