Whole genome sequencing analysis of aquatic pathogenic bacterium Flavobacterium tructae

Hao Li，Liqun Lv

Introduction

The genus Flavobacterium, part of the family Flavobacteriaceae, has seen a rapid increase in species, with 416 described species currently. Members of the genus have been isolated from soil, water, sludge, plants, food, and human samples. Some strains of Flavobacterium are considered to be pathogens of fish diseases. In 2024, a new strain named CF-01 was isolated from diseased carp in a farm in Tangshan.

Microbial whole genome sequencing is an important method for genome localization, complete genome assembly, and comparative analysis. Sequencing can generate accurate genomes in terms of function and structure, support microbial identification and comparative genomics research, and help analyze and construct gene interactions.

Methods

1. Isolation and identification of strains

Tructae Flavobacterium CF-01 was isolated from carp in a breeding farm in Tangshan, Hebei Province by our laboratory. Preserved at the Strain Collection Center of Shanghai Ocean University. Select Tryptone Yeast Extract Salt (TYES) medium (containing 4.0g of tryptone, 0.4g of yeast extract, 0.5g of magnesium sulfate heptahydrate, 0.5g of anhydrous calcium chloride, and 15g of agar per 1000mL).

2. Whole genome sequencing of Tructae Flavobacterium CF-01

Genomic DNA → Split DNA → Repair Ends → Ligate Adaptors → Purify DNA → Sequencing
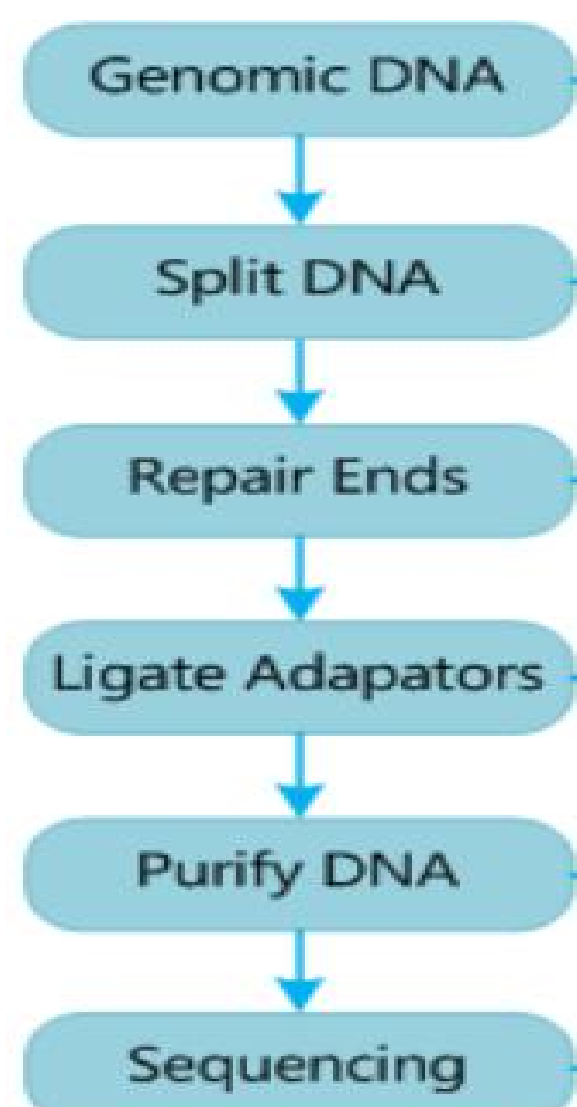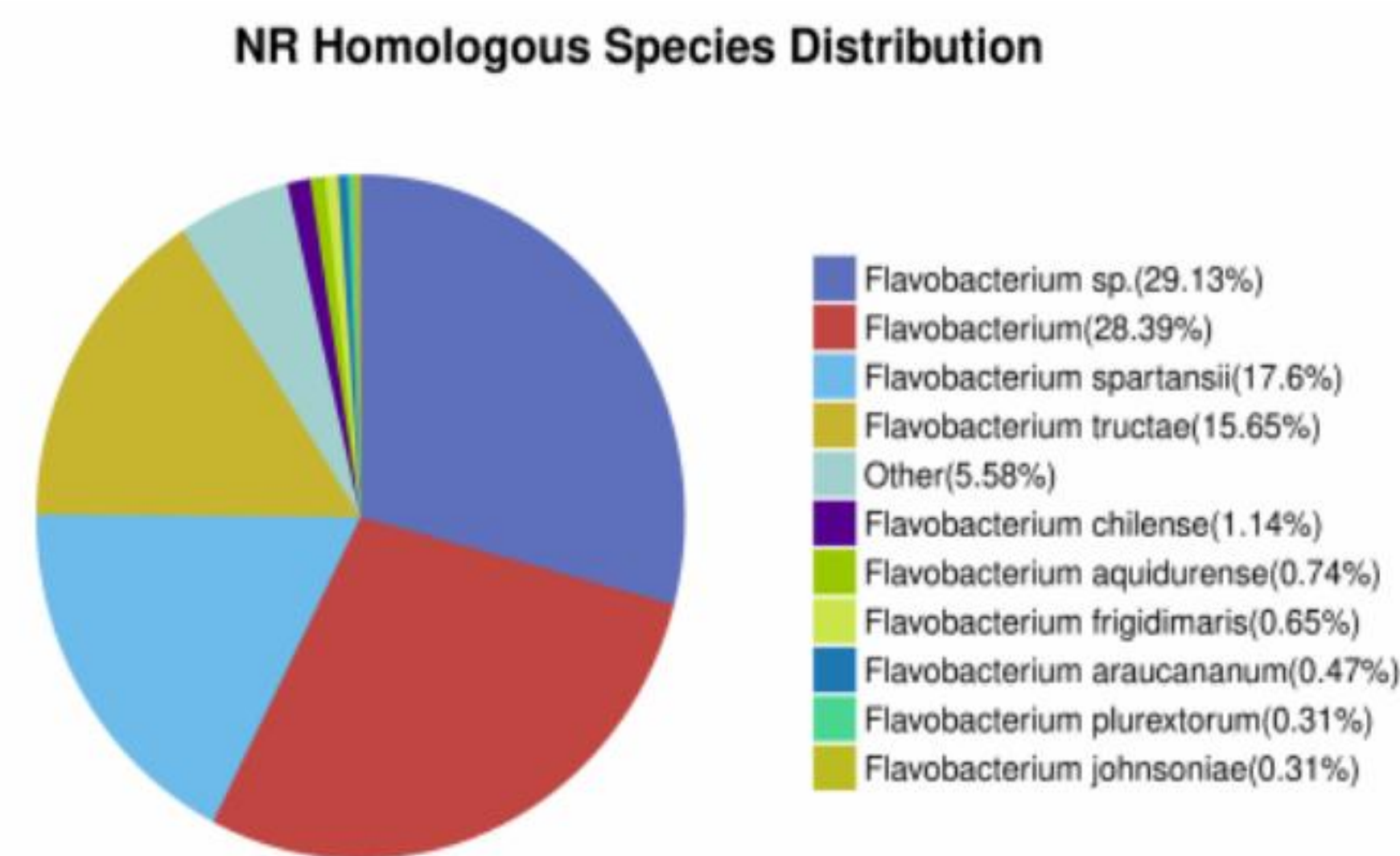
3. Genome annotation

Based on the predicted protein sequences of the coding genes, BLAST (version 2.3.0) was used to perform proportional alignment between the software and the NR database (Non Redundant Protein Database), KEGG database (Kyoto Encyclopedia of genes and genomes, KEGG), COG database (Cluster of Orthologous Groups of proteins), GO database (gene ontology) The Evalue of the sequence alignment was set to 1e-5, and the optimal matching result was selected as the annotation result for this gene.

Results

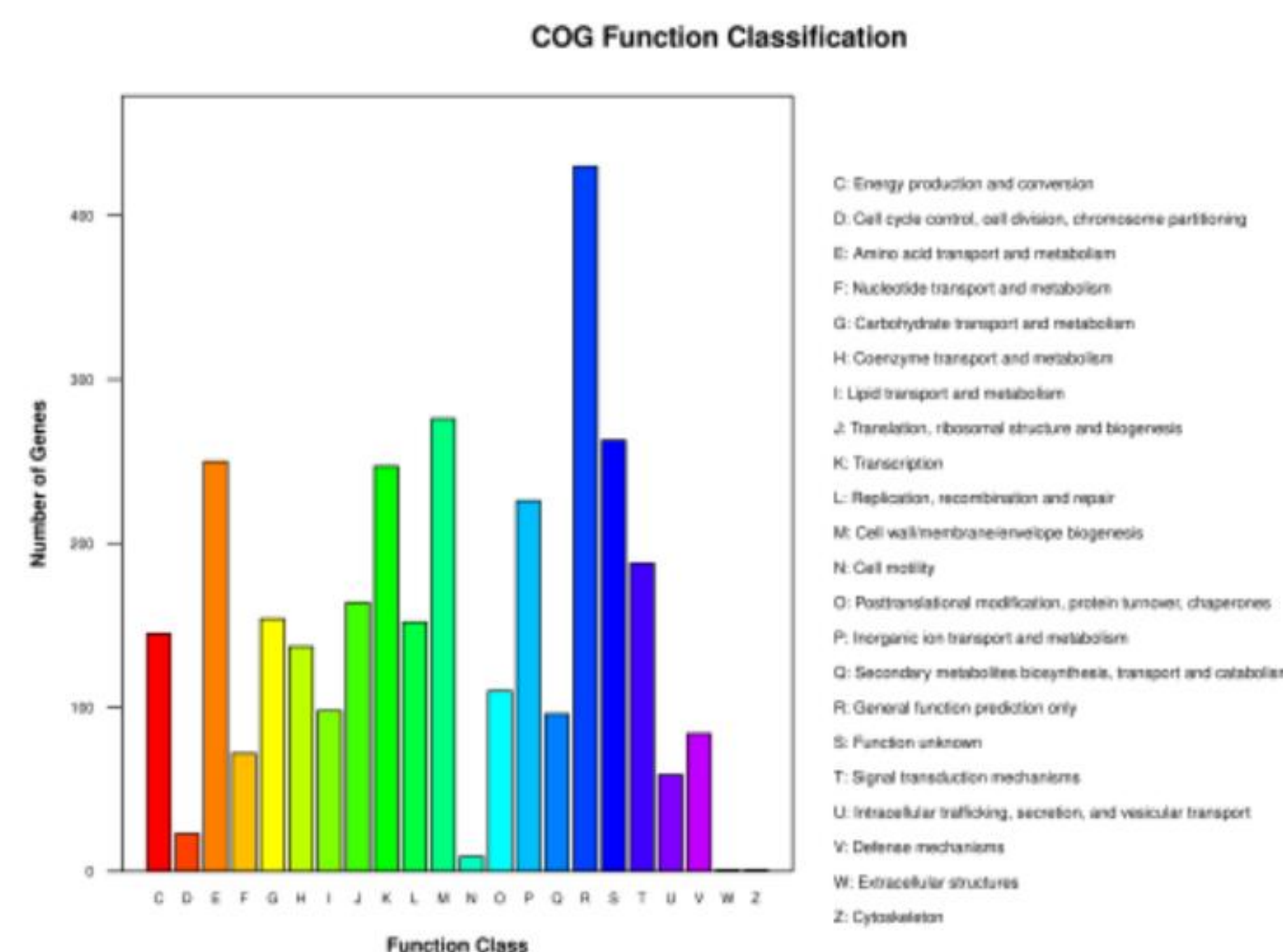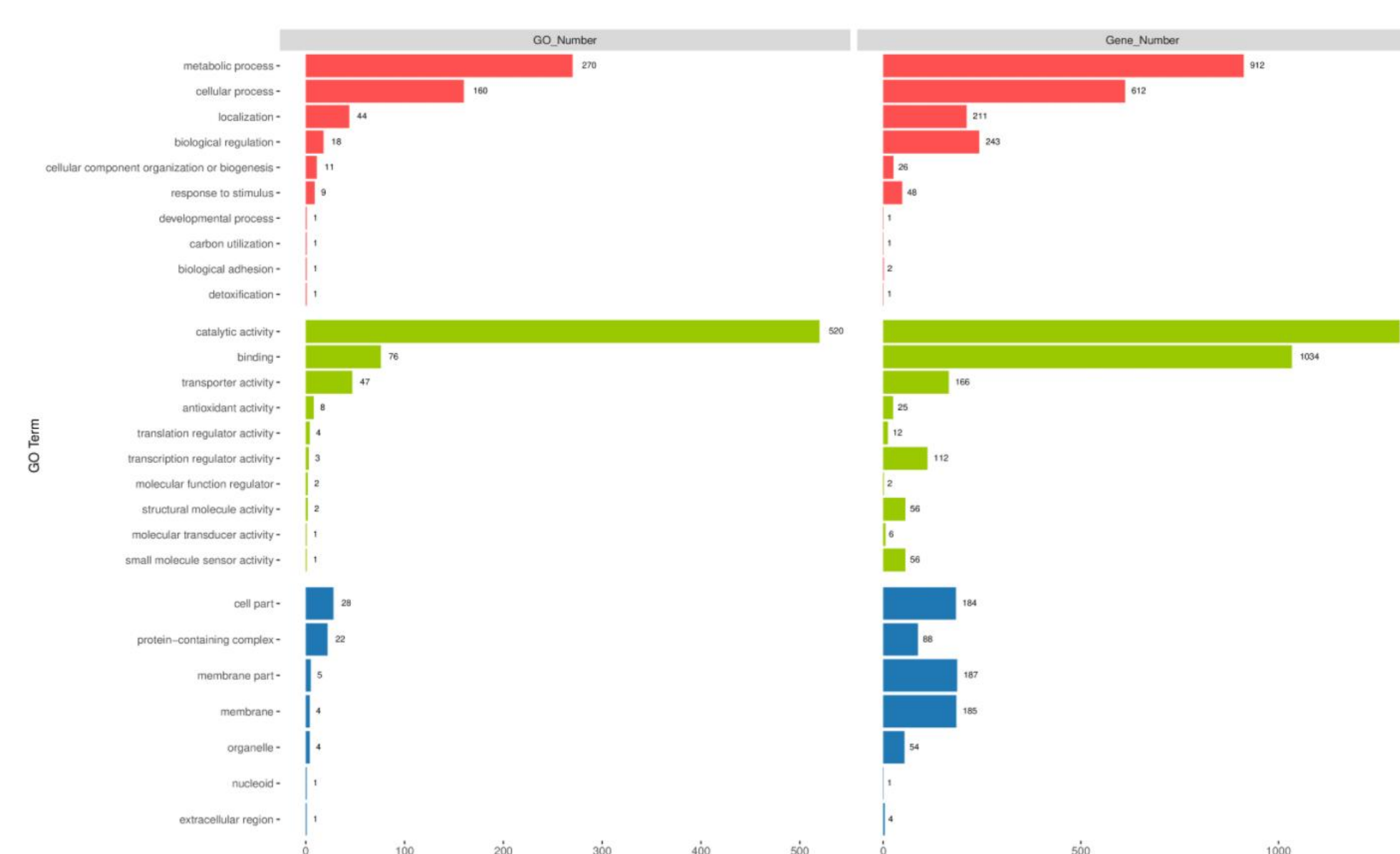### 1. NR database annotation

NR Homologous Species Distribution

- Flavobacterium sp.(29.13%)
- Flavobacterium(28.39%)
- Flavobacterium spartansii(17.6%)
- Flavobacterium tructae(15.65%)
- Other(5.58%)
- Flavobacterium chilense(1.14%)
- Flavobacterium aquidurense(0.74%)
- Flavobacterium frigidimaris(0.65%)
- Flavobacterium araucananum(0.47%)
- Flavobacterium plurextorum(0.31%)
- Flavobacterium johnsoniae(0.31%)

2. KEGG database annotation

3. COG database annotation

COG Function Classification

4. GO database annotation

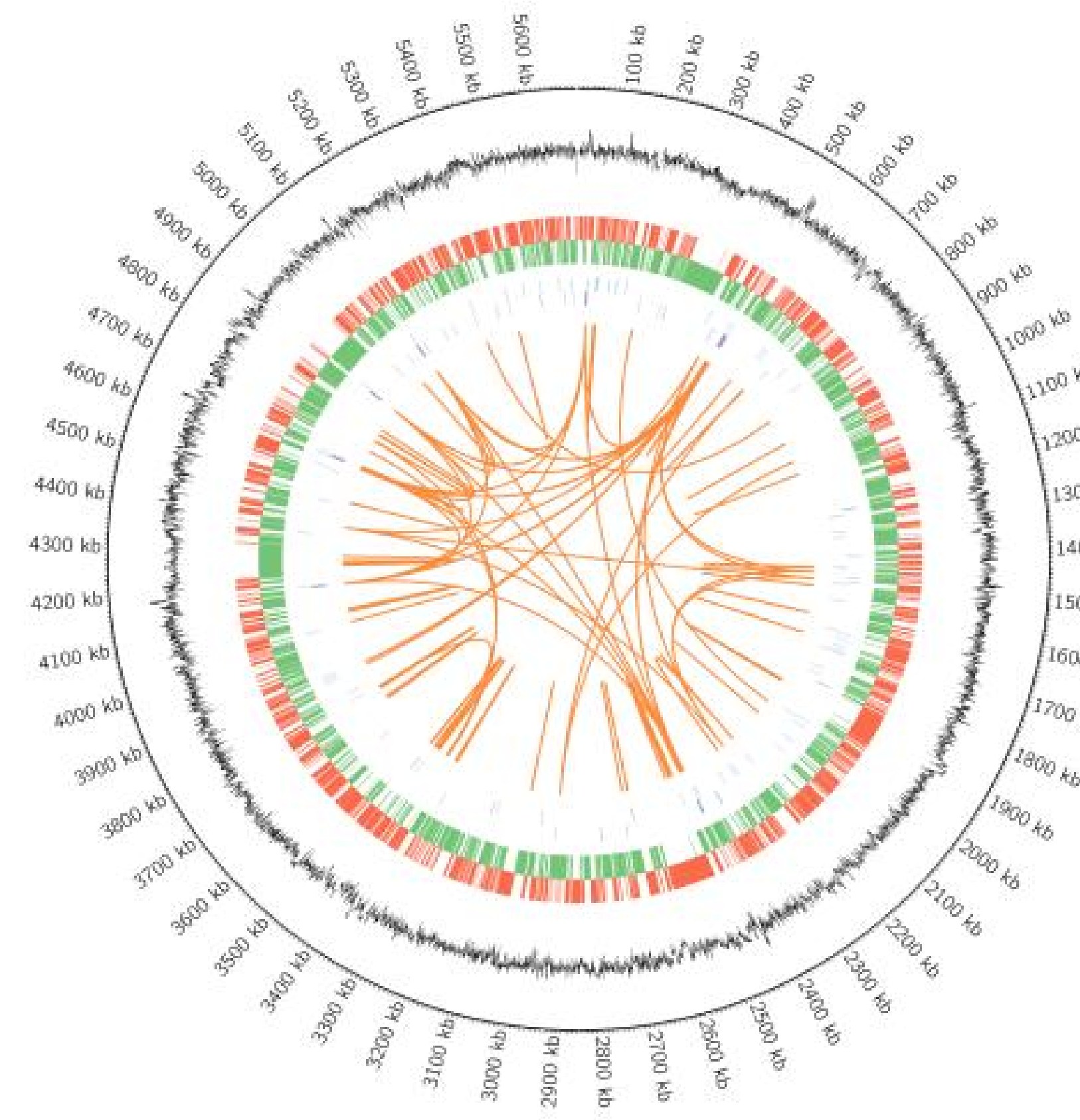### 5.Genomic Circle Map Display

Figure annotation: The circle diagram displays seven types of information, from outside to inside: the first circle is genome location information, the second circle is GC content information, the third circle is coding genes on the positive strand (marked in red), the fourth circle is coding genes on the negative strand (marked in green), the fifth circle is ncRNA information on the positive strand (marked in blue), the sixth circle is ncRNA information on the negative strand (marked in purple), and the seventh circle is long repeated sequence information within the genome (marked in orange).

Conclusion

This study provides information on the genome structure, gene function, and gene interactions of CF-01, which helps to understand the biological characteristics of bacteria.The NR database results show that the most similar species to CF-01 are Flavobacterium spartansii and Flavobacterium tructae. Flavobacterium spartansii was classified as Flavobacterium tructae a few years ago. So CF-01 can be identified as Flavobacterium tructae.In COG clustering analysis, conventional functions predict the most.The development of genomics has provided broader prospects for research on biological characteristics, population evolution, gene analysis, pathogenic mechanisms, and new drug development. Through this study, the genome sequence of Flavobacterium tructae CF-01 has been obtained, providing a deeper understanding of the genomic information of Flavobacterium tructae. This information will help clarify the pathogenic mechanism of Flavobacterium tructae and develop new disease resistance strategies in the future.