Complete genome analysis of highly pathogenic non-O1/O139 *Vibrio cholerae* isolated from *Macrobrachium rosenbergii* reveals pathogenicity and antibiotic resistance related genes

周一凡，顾舒文，李杰，纪鹏，张颖杰，吴聪聪，姜群，高晓建，张晓君 a*

扬州大学动物科技学院 扬州 225009

Abstract

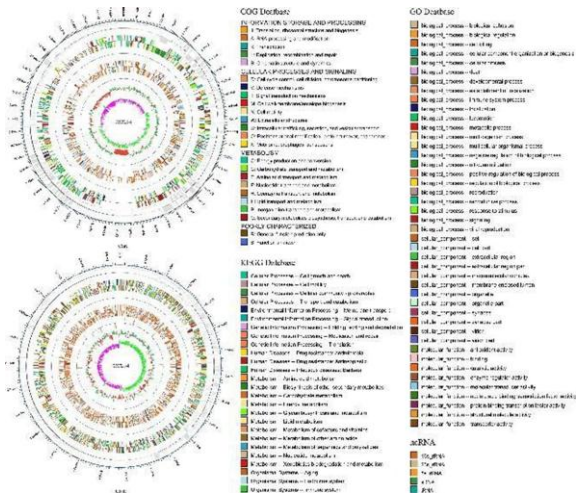
Non-O1/O139 *Vibrio cholerae* is highly virulent pathogen that causes mass mortalities of various aquatic animals. In the present study, we sequenced the whole genome of non-O1/O139 *V. cholerae* GXFL1-4 isolated from *Macrobrachium rosenbergii* for revealing the pathogenicity and antibiotic resistance. The result showed its genome contained two circular chromosomes and one plasmid with a total size of 4,282,243 bp, which harbored 3,869 coding genes. Among them, 3047, 2659 and 3661 genes were annotated in COG, GO and KEGG databases, respectively. In addition, 372 potential virulence genes were predicted based on the VFDB database, including type II, III, IV and VI secretion systems related genes, flagella genes and pilus formation or motility related genes, etc. Blast results in the CARD database showed that the strain contained 27 categories of 148 antibiotic resistance related genes, including efflux pump complex antibiotic resistance genes and antibiotic resistance gene cluster genes, etc. The PHI database annotated 320 genes related to pathogen-host interaction, including T3SS function genes, virulence regulatory factors, transcriptional regulators, two-component response regulator related genes, etc. The whole-genome analysis suggested that the pathogenic non-O1/O139 *V. cholerae* strain GXFL1-4 might have a complex molecular mechanism of pathogenicity and antibiotic resistance.

Materials and methods

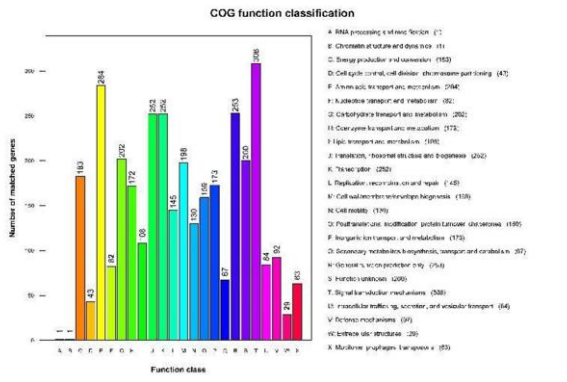
- Genomic DNA was extracted with the SDS method. Sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit for Illumina. The whole genome was sequenced using Illumina NovaSeq PE150 at the Beijing Novogene Bioinformatics Technology Co., Ltd.
- Clean data is assembled with Short Oligonucleotide Aignment Program SOAP denovo software, SPAdes software and Assembly By Short Sequences software.
- We used different databases to predict gene functions, which were GO, KEGG and COG. For the pathogenicity and antibiotic resistance, we used the PHI, VFDB, CARD to perform the analyses.

Results

- The chromosome I and II are 1,154,802 and 3,079,090 bp, with a G + C content of 45.99% and 47.97%, respectively (NCBI accession number CP090386 and CP090387). The plasmid is 48351 bp, with a G+C content of 41.3% (NCBI accession number CP090388).

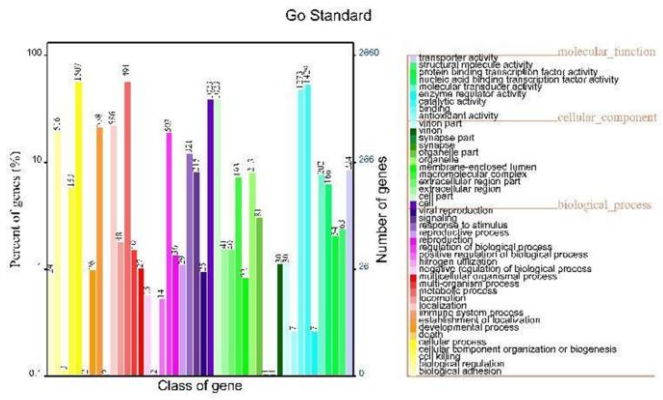


- COG annotation results showed that 3047 genes were annotated into 24 classes of genes, accounting for 78.75% of total genes in non-O1/O139 *V. cholerae* strain GXFL1-4.

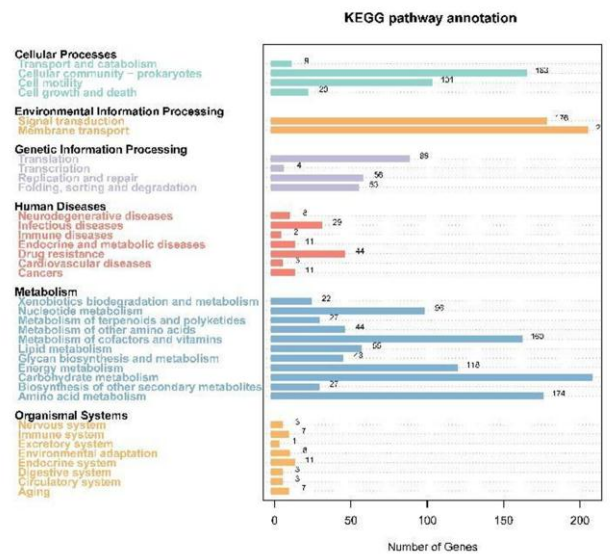


Results

- The functional annotation results in the GO database showed that 2659 genes were annotated into three classes of genes, which accounted for 68.72% of total genes of non-O1/O139 *V. cholerae* strain GXFL1-4.

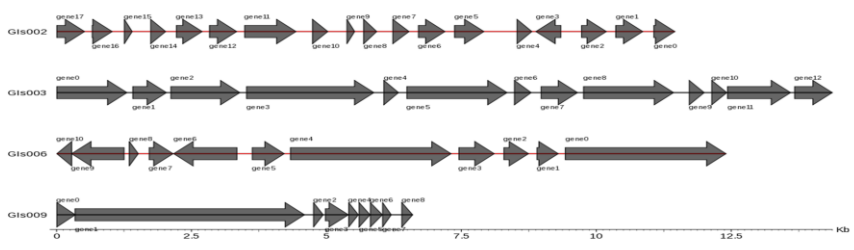


The results of the KEGG pathway analysis showed that 3661 genes were annotated into 189 known metabolic pathways.



Results

- Through the IslandPath-DIOMB online system, 9 genomic islands were predicted to be contained in the whole genome of the strain GXFL1-4. The longest genomic island was 81,279 bp, and the shortest one was 6,594 bp.



- According to the annotation results of the PHI database , the strain GXFL1-4 contained 320 mutant phenotypes related to pathogen-host interaction.