

Comparative genomic analysis of clinical and environmental strains of *Vibrio parahaemolyticus* isolated in Brazil: insight into their virulence potencial

Leandro de Oliveira Santos, Paulo Mascarello Bisch, Wanda Maria Almeida von Krüger

Laboratório de Física-Biológica, Instituto de Biofísica Carlos Chagas Filho - UFRJ

Vibrio parahaemolyticus is a Gram-negative bacteria found in marine and estuarine environments. Some strains are human pathogens and can cause diarrhea, wound infections and septicemia. *V. parahaemolyticus* have been isolated in Brazil during outbreaks of gastroenteritis, from oyster farms, fish markets and restaurants, but their genomes have not been sequenced. The major factors involved in the pathogenesis of *V. parahaemolyticus* are the hemolysins (TDH and TRH) that form pores in the host membrane causing water and electrolytes efflux, and the Type 3 (T3SS) and Type 6 (T6SS) Secretion System that inject bacterial effectors directly into the host cytosol. The gene encoding the thermolabile hemolysin (*tlh*) is considered a signature molecular marker for the species. In this work we sequenced, assembled and annotated the genomes of 5 Brazilian isolates of *V. parahaemolyticus*: 3 (Cascavel, 17381 and 17384) from fecal samples of gastroenteritis patients and 2 (20173 and 20142) of oysters from an oyster farm. Genome DNA samples were prepared from overnight cells grown in LB medium at 37°C, using the Wizard Genomic DNA Purification Kit (Promega). Integrity and concentration of the samples were evaluated by agarose gel electrophoresis. The DNA was fragmented and the paired-end sequencing libraries were prepared using the Nextera DNA Library Kit. Sequencing was performed on the Illumina MiSeq, producing 75pb paired-end sequence data. Quality of the reads was evaluated by the FastQC program. Trimming and filtering were done (Phred quality score > 20) by the Trimmomatic algorithm. The filtered reads were assembled into contigs using Velvet Optimizer and Spades, and then mapped to the two chromosome of the clinical strain, *V. parahaemolyticus* RIMD2210633, with CONTIGuator. The GView tool was used to generate graphical representations of the chromosomes. The genomes were annotated with RAST server. All strains contain two circular chromosomes of approximately 3.1 and 1.7 Mbp, with a mean GC content of 45.2%. The *tlh* gene was identified in all the 5 genomes, in agreement to results of PCR analyses. Two copies of *tlh* were identified in the genomes of the strains Cascavel and 17381, and one in the 17384 genome. The *trh* gene wasn't detected in the 5 genomes. Moreover, two clusters of T3SS genes, whose sequences slightly differ from those of the reference genome, were detected in the 5 genomes. T6SS genes were also found in all genomes. These results demonstrate that the Brazilian strains, regardless of origin, have virulence-related genes and pathogenic potential.