
**ESTUDIO DE SECUENCIAS 5K EN VIBRIO PARAHAEMOLYTICUS Y
ANÁLISIS DE LA MICROEVOLUCIÓN DE ESTA BACTERIA EN LAS
COSTAS CHILENAS**

**DIEGO ALEJANDRO DÍAZ DOMÍNGUEZ
INGENIERO EN BIOINFORMÁTICA**

RESUMEN

Vibrio parahaemolyticus es una bacteria marina asociada a enfermedades gastrointestinales en humanos. En Chile, una población clonal de esta especie (clon pandémico) causó un brote de diarrea, con más de 10.000 casos sólo el verano del 2005. Estudios realizados en el Centro Nacional de Genómica y Bioinformática han descubierto que algunas cepas chilenas de *V. parahaemolyticus* poseen una secuencia de 5600 bp, denominada 5k, que no está presente en el genoma de referencia para esta población clonal. En las cepas chilenas, anterior a la región 5k, existe otra región de aproximadamente 5000 bp, llamada 5k', que sí está en el genoma de referencia. Dentro de la región que comprende 5k y 5k' existen

3 regiones repetidas, una en cada extremo y una tercera que separa 5k de 5k'. Se ha observado que en la región 5kk' hay segmentos que presentan similitud con profagos crípticos y algunas proteínas putativas que podrían jugar un rol en la virulencia. Este estudio tiene como objetivo determinar cual es la implicancia de las secuencias 5k y 5k', además de estudiar la microevolución de *V. parahaemolyticus* en las costas chilenas, para así poder entender mejor su comportamiento durante estos años. La metodología será un análisis in silico, con programas como BLAST, CLC, CELERA, SMALT, MAUVE, Bioperl, etc. Se espera corroborar que las secuencias 5k y 5k' se originaron a partir de elementos móviles, posiblemente adquiridos por transferencia horizontal.

ABSTRACT

Vibrio parahaemolyticus is a marine bacteria associated with gastrointestinal diseases in humans. In Chile, a clonal population of this bacteria (pandemic clon) caused a diarrhea infection with more than 10,000 cases only at the summer on 2005. Studies in “Centro Nacional de Genómica y Bioinformática” have discovered that some chilean strains of *V. parahaemolyticus* have a sequence of 5600 bp, called 5k, this one isn't presented in the reference genome of this population. In the chilean strains, previous to 5k region, there is another region of about 5000 bp, called 5k' which exists in the reference genome. Inside of the region between 5k and 5k', there are 3 repeat regions, each one in other side and a third one which separate 5k from 5k'. It has observed in the 5k-5k' region there are segments that show similitudes with cryptic prophages and some putative proteins that could play a role in the virulence. The aims of this study are to determine which is the implication of 5k and 5k' sequences and to study the microevolution of *V. parahaemolyticus* at the chilean coast, to understand its behavior during last years. The methodology that we are going to use will be an in silico analysis, with softwares like BLAST, CLC, CELERA, SMALT, MAUVE, Bioperl, etc. We hope to check that 5k and 5k' sequences correspond to mobile elements, possibly acquired by lateral transfer.