Genomic Analysis of Heptosyltranferase I Predicting Pathogenicity of Vibrio Genus

Jeffery Shen[1], Erika A. Taylor[2], Philip Arevalo[3,4]

[1] Hazel Quantitative Analysis Center, Wesleyan University; [2]Department of Chemistry, Wesleyan University; [3]Department of Biology, Wesleyan University; [4]College of Integrative Science, Wesleyan University

Abstract

Lipopolysaccharide is a key component of the outer membrane of gram-negative bacteria, and Heptosyltransferase I (HepI) is a key enzyme in synthesizing LPS inner core [4, 5]. Recent research has identified HepI as a potential target for novel antibiotics [3]. Previous studies have analyzed HepI in *Escherichia coli* and *Haemophilus influenzae* [1], but the HepI of the genus *Vibrio,* which includes several human pathogens, remains largely uncharacterized. We hypothesized that since the outer membrane is important for interactions between host and pathogen, pathogens may have specific HepI adaptations that allow them to succeed in the host environment. In this study, we first examined the relationship between pathogenic potential and HepI amino acid sequence among over 200 species within the genus *Vibrio*. We also narrowed our search to look for key mutations in HepI that differentiated pathogenic and non-pathogenic strains of *Vibrio cholerae*. We found that a single mutation in the C-terminal domain of HepI (T330M) may be more associated with pathogenic strains of *V. cholerae*. This mutation occurs in a region of HepI does not have a structural homolog in *E. coli*, and suggests that the region may facilitate pathogenesis in *Vibrio*.

Several key residues are conserved in both *E.coli* and *Vibrio cholerae*

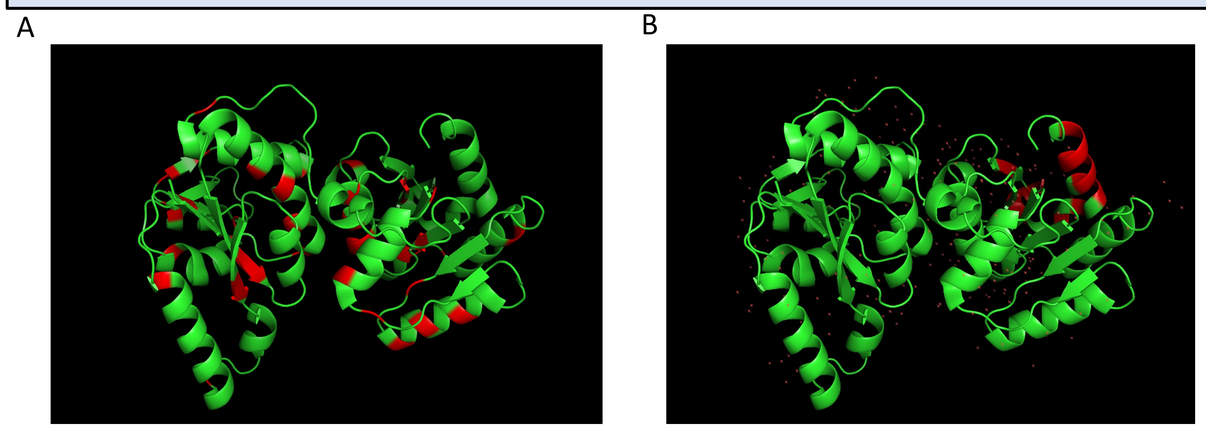

Figure 1: (A) We filtered out the positions of sequences that totally conserved in *E. coli* and *V. cholerae* and mapped these positions on the HepI structure of *E. coli* in red. (B) We filtered out the positions of sequences that conserved in *V. cholerae* but not in *E. coli* and mapped these positions on the HepI structure of *E. coli* in red.

At *Vibrio* genus level, there is no significant difference between pathogens and non-pathogens; Within *V. cholerae*, pathogen HepI is generally more conserved among pathogens

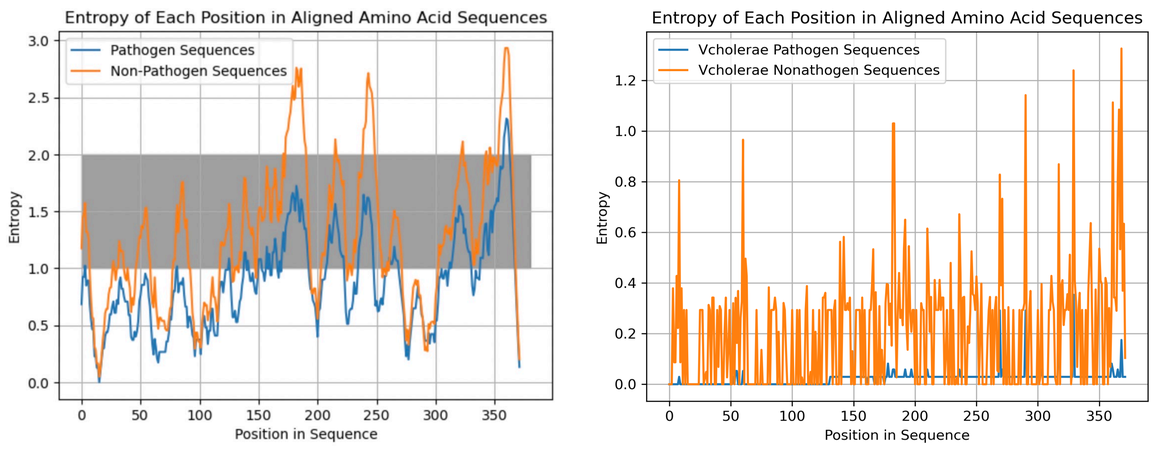

Figure 2: We calculated the entropies of pathogen and non pathogen sequences at each position. All data come from NCBI dataset portal and HepI sequences are retrieved from each genome. Here, the higher the entropy, the more variable types of amino acids is presented at this certain position across strains. We find that there is no significant difference in variability of Vibrio genus, but pathogen is generally more conservative in pathogen sequences.

References

[1] Gronow, S., Brabetz, W., Lindner, B., & Brade, H. (2005). OpsX from Haemophilus influenzae Represents a Novel Type of Heptosyltransferase I in Lipopolysaccharide Biosynthesis. *Journal of Bacteriology*, *187*(17), 6242–6247.

[2] Cote, J. M., Hecht, C. J. S., Patel, K. R., Ramirez-Mondragon, C. A., Sham, Y. Y., & Taylor, E. A. (2020). Opposites Attract: Escherichia coli Heptosyltransferase I Conformational Changes Induced by Interactions between the Substrate and Positively Charged Residues. *Biochemistry*, *59*(34), 3135–3147.

[3]: Milicaj, J., Hassan, B. A., Cote, J. M., Ramirez-Mondragon, C. A., Jaunbocus, N., Rafalowski, A., Patel, K. R., Castro, C. D., Muthyala, R., Sham, Y. Y., & Taylor, E. A. (2022). Discovery of first-in-class nanomolar inhibitors of heptosyltransferase I reveals a new aminoglycoside target and potential alternative mechanism of action. *Scientific Reports*, *12*(1).

[4]: Frirdich, E., & Whitfield, C. (2005). Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the Enterobacteriaceae. *Journal of Endotoxin Research*, *11*(3), 133–144.

[5]: Blaukopf, M., Worrall, L., Kosma, P., Strynadka, N. C., & Withers, S. G. (2018). Insights into Heptosyltransferase I Catalysis and Inhibition through the Structure of Its Ternary Complex. *Structure*, *26*(10), 1399-1407.e5.

The major HepI allele is associated with pathogenic strains of *V. cholerae*

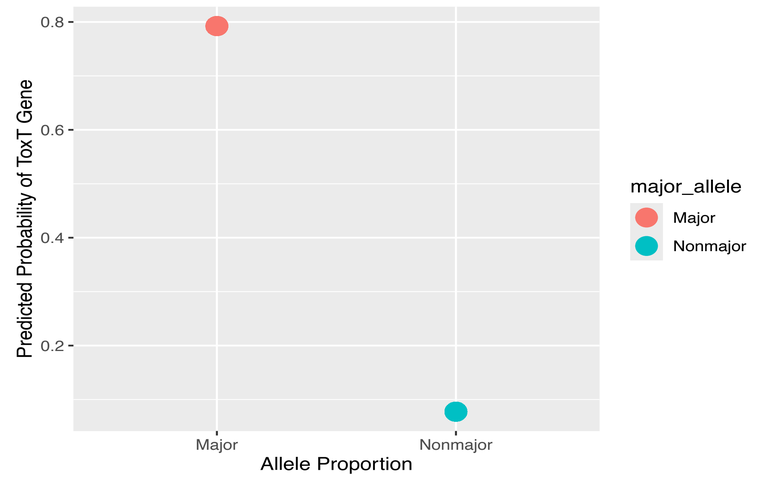

Figure 3: We used logistic regression to see whether being the major allele of HepI (WP_001173306.1) serves as a good predictor of whether the strain of bacteria carry ToxT, a key gene indicating pathogenicity. The logistic regression shows that ToxT = 2.4779 - 3.8174 major allele with p-value <2e-16. This suggests that carrying the major allele can possibly associate with pathogencity.

A mutation (T330M) in a region unique to *V. cholerae* is more prevalent among pathogenic strains

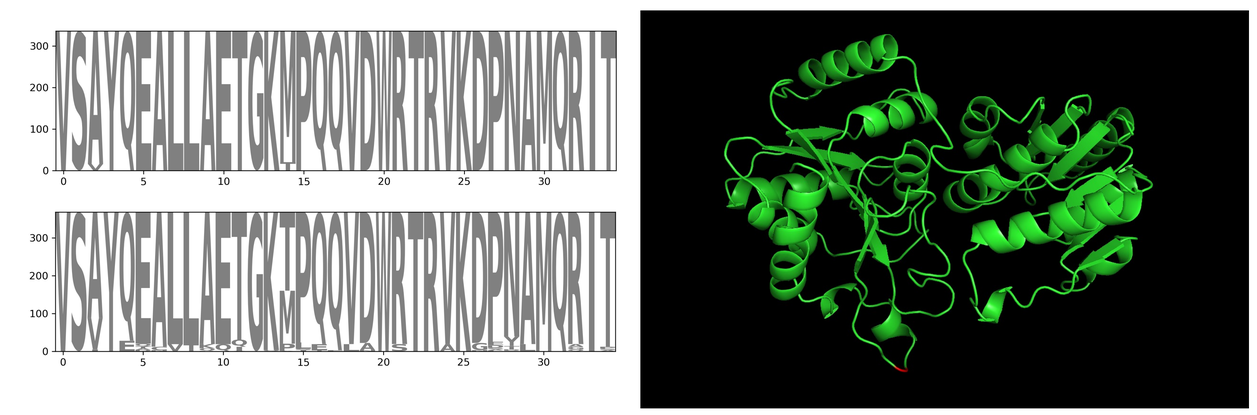

Figure 4: (A) We generated the logo plot for the positions 316-350 in both sequences of pathogen and non-pathogen of Vibrio Cholerae. The larger the character at a certain position indicating the larger probability of this amino acid appear at this position. We find that at 330th position, almost all pathogen is dominated by Methionine, but non-pathogen as approximately 50% Threonine and 40% Methionine. (B) We mapped the 330th position on Vibrio cholerae structure on PyMol software and found that this is located on a loop that is totally deleted in *E. coli*.

T330M has evolved at least twice among *V. cholerae* pathogens

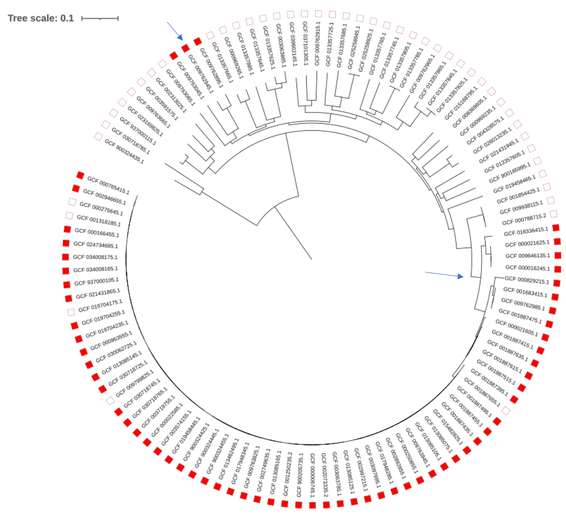

Figure 5: This tree is made using whole genome alignment of 109 Vibrio cholerae strains. The genomes with red dots are the genome of pathogen. The genomes pointed by the arrows are where the T330M mutation happens. This tree suggests that this mutation is not totally associated with one ancestor, and it at least evolved twice.

Next Steps

- **Expand our dataset to compare more *V.cholerae* genome.**
- **Refine the definition of pathogenicity and quantify the pathogenicity.**
- **Examine the biochemical properties of HepI and Hep II interactions.**