Ecology and population genetics of clinical and environmental *Vibrio parahaemolyticus* in the U.S. Pacific Northwest

M. S. Strom*, E. D. Landis, R. N. Paranjpye, and J. W. Turner
NOAA Fisheries Service, Northwest Fisheries Science Center and West Coast Center for Oceans and Human Health, Seattle, WA, USA.

Vibrio parahaemolyticus (Vp) is a marine bacterium capable of causing severe gastroenteritis in humans, usually through the consumption of raw shellfish. Before 1995, Vpvibriosis was sporadic world-wide and caused by a relatively heterogeneous population of the bacterium. Since then, outbreaks have become more epidemic, with foci of infections traced to seafood harvested from single or geographically-linked sites. While initial outbreaks in India and Asia, and later in South America and the U.S. Gulf Coast region, have been attributed to a single serotypically-related pandemic clonal complex (O3:K6), other serotypes have been implicated in distinct geographical areas, including O4:K12 (U.S. Pacific Northwest, PNW) and O6:K18 (Alaska). Current risk assessment models are based on the presence of the virulence-associated genes *tdh* and *trh*, yet illnesses have been attributed to *tdh*- and/or *trh*- isolates. Previous phylogenetic studies have shown that Vp, like most *Vibrio spp*., is genetically diverse, and as yet there has been no definitive conclusion as to what genes are essential for virulence.

Using phenotypic, genetic, and genomic comparison methods such as Multi-Locus Sequence Typing (MLST) and Suppressive Subtractive Hybridization (SSH), we are examining the hypothesis that a set of highly-virulent clones of Vp with increased pathogenic potential have recently emerged in the PNW, and determining whether the emergence can be correlated with specific environmental parameters. MLST and other genotyping analyses of clinical and environmental Vp isolates from PNW sources demonstrate the extensive patterns of diversity as seen elsewhere. However, the majority of PNW strains obtained from human infections form a distinct clonal complex separate from most environmental isolates. Interestingly many environmental isolates obtained from PNW sources are phylogenetically related to the pandemic clonal complex, but this group has not been associated with clinical infections in the region. To further examine the genetic diversity between strains, SSH was used to identify specific genetic differences between the pandemic strain and individual clinical and environmental isolates from the PNW. In addition to identifying variable genetic features within and between the MLST groups, SSH analysis has expanded our understanding of the potential content of the Vp genome. In addition we are examining the phylogeny of specific virulence-associated genes that encode bacterial surface proteins, to determine if any of these genes are subject to selective adaptation in the environment.

An additional key finding includes data showing that Vp strains possessing the *tdh* and *trh* genes make up a significantly higher percentage in the PNW than what is found elsewhere, yet only a small fraction of these belong to the clonal group that appear capable of causing disease. Therefore reliance on the presence of *tdh* as a predictor of pathogenic Vp is not appropriate for all geographical areas. This finding shows that there is a need to carry out genome sequencing of a variety of strains in order to identify appropriate genetic markers that can be used to rapidly identify strains with higher pathogenic potential. Population structure analysis formed the decision basis for strain selection for Vp genome sequencing, projects that are now in progress.

The application of comparative genomics analysis should lead to a better understanding of the genetic differences between the relatively few strains possessing a higher pathogenic potential and other lesser virulent or avirulent environmental strains. The implications of this information for the development of improved risk assessment and health early warning systems, including *in situ* biosensors, will also be discussed.