

# THE COMPLETE GENOME OF AN ENDOGENOUS NIMAVIRUS (*Nimav-1\_LVa*) FROM SPECIFIC PATHOGEN-FREE SHRIMP *Penaeus vannamei* – THE NEED FOR REFERENCE GENOMES OF PENAEIDS AND OTHER CRUSTACEAN GENOMES

Acacia Alcivar-Warren<sup>1\*</sup> and Weidong Bao<sup>2</sup>

<sup>1</sup>Environmental Genomics Inc., P.O. Box 196, Southborough, MA 01772.

<sup>2</sup>Genetic Information Research Institute, 20380 Town Center Lane, Suite 240, Cupertino, CA 95014.

White spot syndrome virus (WSSV), the lone virus of the genus *Whispovirus* under the family *Nimaviridae*, is one of the most devastating viruses affecting the shrimp farming industry (<https://pubmed.ncbi.nlm.nih.gov/31947590/>). Knowledge about this virus, in particular its evolution history, has been limited, partly due to its large genome and the lack of other closely related free-living viruses for comparative studies.

A reconstructed full-length endogenous nimavirus consensus genome, *Nimav-1\_LVa* (279,905 bp) was found in the genome sequence of the Kehai isolate of *Penaeus vannamei* from China (GCA\_003789085.1; breed Kehai No. 1, cultured shrimp imported from USA; 1.6 Gb) in which ~12 copies of *Nimav-1\_LVa* exist. This endogenous virus seems to insert exclusively into the telomeric pentanucleotide microsatellite (TAACC/GGTTA)<sub>n</sub>. 117 putative genes are predicted. Sequence analysis of these genes indicates that there are four more recognizable nimaviruses core/ancestor genes, wsv112 (dUTPase), wsv206, wsv226 and wsv308 (nucleocapsid protein), making the total number to be 43. Some *Nimav-1\_LVa* contain introns, such as g012 (IAP), g046 (CHH), g155 (innexin), g158 (BI-1-like). More than a dozen *Nimav-1\_LVa* genes are involved in the pathogen-host interactions. We hypothesize that g046, g155, g158 and g227 (semaphorin 1A like) are recruited host genes for their roles in immune regulation.

Availability of *Nimav-1\_LVa* sequence will help understand the genetic diversity, epidemiology, evolution, pathogenicity, and virulence of WSSV. Future studies will focus on the possibility that *Nimav-1\_LVa* represents a free-living virus, yet unidentified but infecting crustaceans, because almost identical *Nimav-1\_LVa* sequences were also found in genome sequences from black tiger shrimp *P. monodon* isolate Shenzhen from China (GCA\_002291185.1, wild shrimp; 1.4 Gb) and *P. japonicus* Guanxi isolate from Japan (GCA\_002291165.1; wild shrimp; 1.6 Gb). In addition, three relatives of *Nimav-1\_LVa* are detected in *P. monodon*: *Nimav-1\_PMo*, *Nimav-2\_PMo* and *Nimav-3\_PMo*. So far, *Nimav-1\_LVa* sequences have not been found in the genomes of *P. monodon* isolate 26D (GCA\_007890405.1; cultured shrimp from Vietnam, originally from Australia; ~1.6 Gb), *P. monodon* isolate SGIC\_2016 from Thailand (GCA\_015228065.1; cultured shrimp from Surat Thani, Thailand; ~2.4 Gb), and *P. chinensis* isolate QD-2010 from China (GCA\_016920825.1; wild shrimp; ~1.6 Gb).

Future research should focus on (a) sequencing fully assembled reference genomes for all Penaeid shrimp, (b) determine, by FISH, the chromosome location of the integrated *Nimav-1\_LVa* or its relatives in crustacean genomes, and (c) survey the existence of Nimavirus in more Crustacean species. Search and isolate the free-living *Nimav-1\_LVa* virus, if possible.