PRELIMINARY SURVEY OF TRANSPOSABLE ELEMENTS FROM THE FIRST SPECIFIC PATHOGEN FREE (SPF) SHRIMP *Penaeus vannamei* PRODUCED IN THE UNITED STATES

Gladys Zuniga*, Acacia Alcivar-Warren

The Shrimp Epigenome (ShrimpENCODE) Project
Environmental Genomics Inc., Southborough, MA, 01772 USA, and
Fundación para la Conservación de la Biodiversidad (FUCOBI), Quito, Ecuador
fucobi@gmail.com

As part of our efforts to understand the epigenetic mechanisms associated with susceptibility of *Penaeus vannamei* to bacterial and viral diseases, we proposed that horizontal transfer (HT) and methylation of transposable elements (TEs) are potential mechanisms involved in susceptibility to Acute Hepatopancreatic Necrosis Disease (AHPND), white spot syndrome virus (WSSV), and Infectious hypodermal and hematopoietic necrosis virus (IHHNV). To test this hypothesis, we needed access to a fully assembled *P. vannamei* genome sequence (expected size is 2.8 Gb), but none is currently available in the NCBI genome databases. The only genome sequence available for *P. vannamei* is a 1.6 Gb draft prepared with muscle DNA of male shrimp of Kehai isolate from China (GCA_003789085.1; breed Kehai No. 1) originally from the United States.

A pilot genome sequencing effort was initiated through 'The Shrimp Epigenome (ShrimpENCODE) Project' of the FUCOBI Foundation of Ecuador and Environmental Genomics Inc., MA, USA, with help from Amplicon Express and Pacific Biosciences collaborators. High molecular weight DNA from offspring of SPF Kona Line *P. vannamei* domesticated by the breeding program of the U.S. Marine Shrimp Farming Program (USMSFP) maintained at the Oceanic Institute in Kona and Oahu, HI were used for sequencing. TEs were classified as either Class I [Long terminal repeat (LTR) retrotransposons and non-LTR retrotransposons] or Class II (DNA transposons).

A total of 312 TEs were identified in a partial genome sequence (~470 Mb). These included: 105 DNA transposons [EnSpm(1), Harbinger(11), hAT(13), Kolobok(2), Mariner/Tc1(10), Merlin(12), MuDR(1), P(1), piggyBac(8), Polinton(3), Transib(2), unclassified(41)]; 119 LTR retrotransposons [BEL(15), Copia(2), Gypsy(92), unclassified(10)]; 76 non-LTR retrotransposons [CR1(7), Daphne(7), Ingi(4), Jockey(4), Nimb(7), Penelope(18), Proto2(2), RTE(9), R4(2), Vingi(2), SINE2(3), unclassified(11)]; 1 integrated Nimavirus, *Nimav-1_LVa* (279,905 bp), and 11 unclassified. All sequences are deposited in Repbase (https://www.girinst.org/censor/index.php). The four most abundant TE families are: *RTE-3_LVa*, *RTE-2_LVa*, *NonLTR-1_LVa*, and *Nimb-1_LVa*. The complete genome of *Nimav-1_LVa* was found in the genome of *P. vannamei* Kehai isolate from China. This nimavirus seems to insert exclusively into the telomeric pentanucleotide microsatellite (TAACC/GGTTA)n (https://pubmed.ncbi.nlm.nih.gov/31947590/). Future research will focus on (a) sequencing a continuous, fully assembled reference genome from either brood stocks of the original SPF *P. vannamei* produced by the USMSFP or wild shrimp from Ecuador, and (b) understanding the epigenetic mechanisms associated with HT of retrotransposons.