4. Research report (Follow the guideline on the next page)

Title of Research Project: Integrating a Genomic Perspective towards Biological Control Approaches: The Case in Arthropod Vectors

BACKGROUND

Mosquitoes are known vectors that transmit pathogenic microorganisms contributing to the toll of morbidities and mortalities worldwide. The Philippines considers mosquito-borne diseases such as dengue, chikungunya, and filariasis as an economic and health burden to its people. Because of the extent of insecticide resistance in mosquitoes, the use of biological control approaches is now recommended in controlling the vector. Genomic tools have been critical in mosquito-borne disease surveillance and control because they provide important genetic information on vector populations as a guide to the choice of the best tools that maximize the success of control interventions. With limited data or literature on these mosquito vectors in the Philippines, especially on the application of biological control agents, the proposed project aims to utilize these genomic tools to provide the necessary biological and ecological information on its control. The following objectives are: (a) To investigate the population genomics of *Ae. aegypti* in Metropolitan Manila to further elucidate their gene flow pattern and detect insecticide resistance using a Pool-based sequence analysis and (b) To determine the relationship between the genetic diversity and microbial community composition of *Culex quinquefasciatus*.

METHODOLOGY

Mosquito samples collected and deposited in the Molecular Ecology and Health (MEcoH) of Ehime University and the Biological Control Research Unit (BCRU) of De La Salle University were used. *Ae. aegypti* and *Cu. quinquefasciatus* samples of MEcoH were collected in Metropolitan Manila, Philippines from 2014 – 2017. All samples had been DNA extracted using the QIAGEN Blood and Tissue DNEasy Kit (Qiagen, Hilden, Germany) following the modified protocol of Crane (2011) and stored in -20 Ultra freezer.

The amplification of the CO1 gene in *Cu. quinquefasciatus* was carried out using the specified primers and thermal conditions using the protocol of Regilme et al (2021). PCR products were purified using the QIAquick 96 PCR Purification Kit (QIAGEN) following the manufacturer's instructions and sequenced in both directions by Eurofin Genomics, Inc, Tokyo, Japan. For microbiome analysis in *Cu. quinquefasciatus*, library preparation is carried out through PCR using modified primer sequences targeting the V4 hypervariable region of the 16S SSU rRNA gene as detailed in the Earth Microbiome Project (Capoaso et al., 2012). Afterwards, library-quality control will be performed by checking the library size distribution via the High-Sensitivity DNA chip (Agilent Bioanalyzer), and the libraries will be purified/size selected using SPRI beads (AmpureXP; Beckman Coulter Genomics) with an amplicon size of ~400-bp. Triplicate quantitative PCR reactions at appropriate dilutions will be performed to quantify the amplicon

libraries with the KAPPA Illumina Library qPCR Quantification kit (Kappa Biosystems). The normalized and pooled amplicon library will be sequenced using the Illumina MiSeq platform with paired-end reads of 300-bp per read.

For SNP calling in *Ae. aegypti*, library preparation for the 100 ng of genomic DNA from each pooled sample was done using Rasic et al. (2014) protocol with slight modifications. The pooled library was sequenced in Illumina HiSeq2000 platform to obtain 100 bp paired-end reads. Following a quality control in FASTQC v.0.11.5 (Andrew, 2010), sequence reads were demultiplexed, filtered, for quality, and trimmed of 10bp MID sequences using STACKS v 1.42 (Catchen et al., 2013) using default settings. The reference genome of Ae. aegypti (50_AaegyptiLVP, vectorbase.org) was used to identify RAD loci from the denovo_map.pl in STACKS (m = 5, M = 2, n = 0). Afterward, SNP calling was done in POPOOLATIONS2 (Kofler et al., 2010). We mapped all our RAD reads against the reference catalog created in STACKS using the bwa mem algorithm of BWA v 0.7.12 (Li and Durbin, 2010). Mapping results were filtered for a minimum Phred quality score of 20 and converted into mpileup format in SAMTOOLS v 0.1.19 (Li et al., 2009). For each population we then calculated the allele frequency difference at each position with a minimum coverage of 10 and a minimum minor allele count of 2 using the snp-frequency-diff.pl script of POPOOLATION2 (Kofler et al., 2010). The microsatellites of previously studied *Ae. aegypti* were used and re-analyzed. Pairwise F_{ST} values were calculated using Arlequin v3.5.1.3 (Excoffier and Lischer, 2010) with 10,000 permutations to determine the magnitude of the genetic differentiation among mosquito populations. On the other hand, pairwise F_{ST} values were estimated in POPOOLATION2 only considering positions with a minimum coverage of 10 and minimum allele count of two per RAD locus. Using the pairwise F_{ST} values in microsatellites and SNPs, we generated a genetic dendrogram based on UPGMA in the fastcluster package (Müllner, 2013) in R version 3.5 (R Development Core Team, 2018).

RESULTS

Population genomics of *Ae. aegypti*

A total of 272,828,044 raw reads with 150 bp lengths were generated from the ddRAD sequencing, ranging from 12,697,800 to 29,508,570 reads per population. After trimming the adapters and filtering reads for quality, 335,280 reads were removed, retaining 272,492,764 reads. The total number of mapped reads to the reference genome of *Ae. aegypti* was 258,100,457 reads. These covered 29.72 % of the entire *Ae.aegypti* reference genome (1,278,715,314 bp). We identified 12,771 SNPs, with 1,116 SNPs privately found only among female populations and 1,312 SNPs found only among male populations. NMDS plots showed an unclear separation between the female and male populations (Figure 1). The male populations showed a more diverse structure than the female populations, with the North Manila and South Manila populations being genetically isolated. Ten populations from Metropolitan Manila showed a higher pairwise FST (0.029 - 0.069 (average = 0.043)) generated from SNPs data than microsatellites data (0 - 0.043 (average = 0.015)). The dendrogram generated

from microsatellite markers had shallow branch lengths (= 0.010) from the terminals to the common ancestor and very small genetic divergence (pairwise FST = 0 - 0.010). In contrast, in the dendrogram of neutral SNP markers, the overall branch length from the terminal to the common ancestor was longer (= 0.030), indicating the unique genetic structure of individual populations.

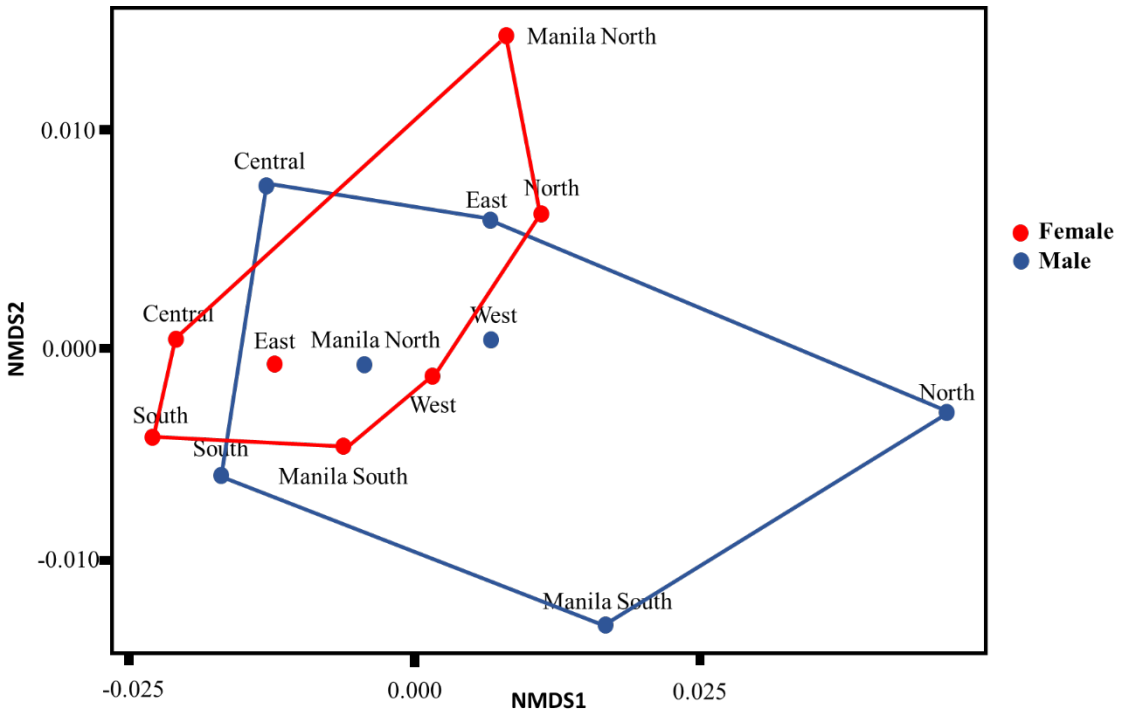

Figure 1. Population genetic structure of male (blue) and female (red) Ae. aegypti populations according to Non-metric Multidimensional Scaling (NMDS).

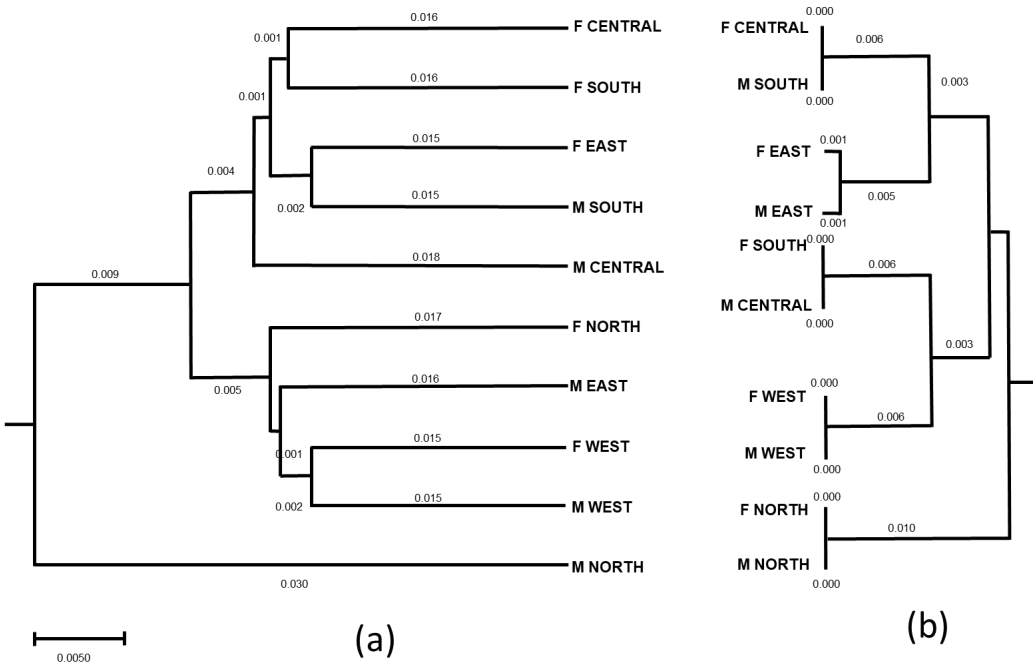

Figure 2. Comparison of dendrograms generated by hierarchical clustering (UPGMA) based on genetic dissimilarity using (a) SNP loci and (b) Microsatellite loci dataset from

Metropolitan, Philippines.

Genetic diversity and Microbial community composition of *Culex quinquefasciatus*

A total of 570 *Cu. quinquefasciatus* mosquito individuals were used wherein 302 were female while 268 were male. All samples were collected from 5 households within Metropolitan Manila from October 2016 til January 2016. In terms of its population genetics, 13 haplotypes were identified wherein the majority are from haplotype 1. Overall FST values in male and female populations are 0.021 and 0.013 respectively. For microbiome analysis, a total of 13,235,342 raw reads were generated. After trimming the adapters and filtering reads for quality, 6,754,824 were retained and assigned to 925 zOTUs using the UPARSE pipeline. Further analysis showed that the 925 zOTUs were classified into 20 bacterial phyla, 34 class and 196 genera. Figure 3a shows the shared and unique zOTUs between male and female *Cu. quinquefasciatus* mosquitoes. Both males and females yielded the same zOTU composition however their read abundances differ (Figure 3b and 3c). No significant difference in terms of microbial diversity between male and female populations was observed. Further analysis showed that there is no significant correlation between

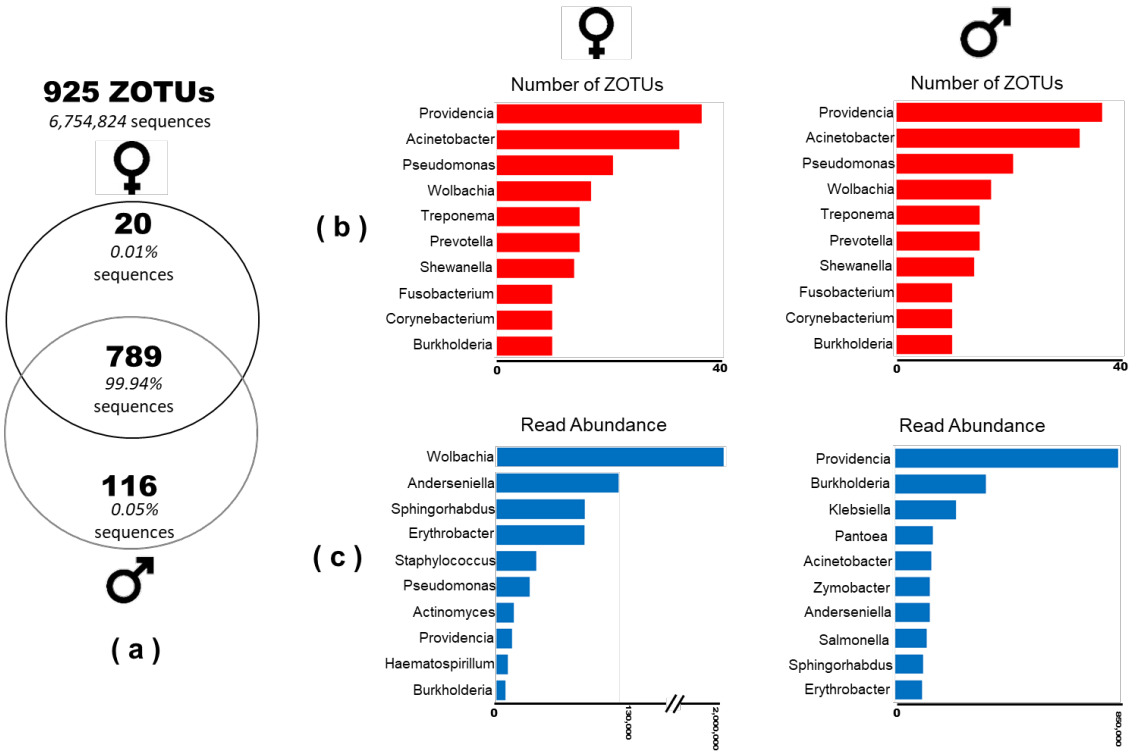

Figure 3. Microbiome analysis of *Cu. quinquefasciatus*. (a) Shared zOTUs between male and female mosquito populations; (b) zOTU composition and (c) abundance between male and female mosquito populations