Metagenomic sequencing of *Culex tarsalis* from the field

Hanna Retallack[1], Leslie Goo[2], Amy Kistler[2], Eric Haas-Stapleton[3], Joseph DeRisi[1,2]

[1]*University of California, San Francisco, San Francisco, CA*
[2]*Chan Zuckerberg, Biohub, San Francisco, CA*
[3]*Alameda County Mosquito Abatement District, Hayward, CA*

Eric.Haas@mosquitoes.org

INTRODUCTION

Culex tarsalis is a vector of West Nile virus and other arboviruses that can breed in landscapes with pooled rain water and in urban settings that have uncovered water containers such as fouled swimming pools or flood control canals. The lack of a complete genome or even subgenomic mitochondrial genome sequence for this vector has hampered our ability to monitor insecticide susceptibility. Moreover, its full microbial payload has not been extensively examined. To address these gaps, we carried out a metagenomic next-generation sequencing (mNGS) pilot study to capture both the *Cx. tarsalis* sequences and non-host microbial sequences from *Cx. tarsalis* collected under standard conditions at a single site during peak abundance in fall of 2017.

METHODS

Mosquitoes were captured at a marsh habitat located in Fremont CA (USA) using EVS CO_2-baited traps and frozen using dry ice or immobilized using triethylamine (TEA). Mosquitoes were sorted on a chill table, female *Cx. tarsalis* collected, and separated into head and abdomen pools (n=5 and n=20) before storage at -80 ^{o}C. DNA and total RNA were extracted from homogenized *Cx. tarsalis* specimens using ZR duet DNA/RNA miniprep plus kit. No significant differences in RNA or DNA yields were detected between specimens frozen on dry ice or treated with TEA. NGS libraries were prepared for each pool and sequenced on an Illumina NextSeq 550 sequencer. An average of 30M reads (range 17M – 41M) were obtained for RNA samples, and an average of 38.6M reads (range 26M-50M) were obtained for DNA samples. *Cx. tarsalis* sequences were identified through nucleotide sequence similarity to a compilation of all publicly available non-redundant mosquito sequences present in the NCBI GenBank repository. A near complete draft of the 14,850bp mitochondrial genome and transcriptome has been assembled for *Cx. tarsalis*. For each specimen pool, we assessed the presence of mutations in genes known to confer insecticide resistance such as the voltage-gated sodium channel (kdr). DNA sequence and RNA transcriptome sequences obtained from these studies will be made available to the community for further study.

RESULTS

Between 0.6 % - 1.2 % of the total RNA reads corresponded to non-host sequences, which were investigated further for the presence of mosquito food sources, commensals, and potentially pathogenic microbes. Alignment to the NCBI set of non-redundant nucleotide (nt) and protein (nr) databases revealed the presence of known and potentially novel viral and non-viral microbes of mosquitoes. These microbes partitioned differentially across the head and abdomen fractions. Our results highlight the potential for mNGS approaches to accelerate our understanding of vector dynamics and insecticide resistance in California. Whether and how these results may vary according to environmental context and time of year has yet to be determined.