

## PacBio's SMRT Sequencing Provides Unprecedented Insight into Genome of Disease-causing Mosquito

November 14, 2018

### Improved Reference Genome for *Aedes aegypti*, Carrier of Zika Virus and Other Public Health Threats, Published in Nature

MENLO PARK, Calif., Nov. 14, 2018 (GLOBE NEWSWIRE) -- Pacific Biosciences of California Inc., (Nasdaq:PACB), the leading provider of high-quality sequencing of genomes, transcriptomes and epigenomes, today announced that an international team of scientists have published the highest quality genome assembly to date of the *Aedes aegypti* mosquito genome using its Single Molecule, Real-Time (SMRT®) technology. The peer-reviewed paper appears online today in the journal [Nature](#).

Analysis of the *Ae. aegypti* mosquito genome could help scientists combat the pest and the infectious diseases it spreads, including Zika, dengue, chikungunya, and yellow fever. Generation of a high-quality reference genome for this species is critically important for exploring the genomic basis of insecticide resistance, disease transmission, blood-feeding host preference, and development of novel repellents.

"The high-quality genome assembly and annotation described here will enable major advances in mosquito biology and has already allowed us to carry out a number of experiments that were previously impossible," the authors write. The effort was led by Leslie B. Vosshall and colleagues at Rockefeller University, along with researchers at several other institutions. The sequencing was performed by Pacific Biosciences at the company's headquarters in California.

Previous efforts to sequence the 1.3 gigabase, highly repetitive *Ae. aegypti* genome suffered from shorter segment sequences that had to be pieced together, resulting in more than 30,000 gaps. Using SMRT Sequencing data, the team produced an assembly that is highly contiguous, representing a 93 percent decrease in the number of contigs.

Dr. Vosshall commented: "A lot of the most important questions in mosquito biology are being attacked with molecular biology and genetics. Without a properly assembled genome, it is extremely difficult to make any progress because we lack information about the total number of genes and their arrangement on chromosomes. The new genome assembly – powered by extremely high-quality PacBio data—has already transformed the field. Because the mosquito genome is highly repetitive, only long-read sequencing produces the high-quality data necessary to bridge the 'oceans' of repetitive DNA to find the 'islands' of meaningful DNA containing genes and gene-related sequences."

"The completion of this mosquito genome provides an enormous contribution to future global health efforts and may address the devastating diseases carried by this species, such as the Zika virus," said Jonas Korlach, Chief Scientific Officer of Pacific Biosciences. "SMRT Sequencing has been employed extensively in the study of infectious disease, particularly for the smaller genomes of viral and microbial pathogens. Recent performance improvements in the technology have also made it more cost effective to resolve the larger, more complex genomes of host species, providing a more complete picture of disease lifecycle. The methods used in this reference genome assembly create a path for producing high-quality reference assemblies for other challenging species."

As the field of genomics evolves, there is a growing awareness in the scientific community of the importance of long-read sequencing data. The ability to read DNA in long, uninterrupted segments provides researchers with a more complete and accurate picture of gene elements and their positioning. PacBio sequencing offers the most comprehensive view of genomes — including the full spectrum of genetic variation — of any sequencing technology on the market today by providing very long average read lengths, very high consensus accuracy, and uniform coverage. SMRT Sequencing technology consistently produces some of the longest average read lengths available in the industry, averaging greater than 20,000 base pairs, with some reads more than 60,000 base pairs.

More than seventy scientists collaborated to create this new assembly, AaegL5, which has already revealed some of what biologists have been missing. For instance, one gene crucial for sex determination—the *Nix* gene—was entirely missing from previous *Ae. aegypti* assemblies. Being able to locate where this gene and the corresponding male-determining factor (M-factor) appear on the chromosome is important for scientists who wish to breed and release males into the wild as a strategy to bias the population towards males rather than disease-transmitting females over multiple generations.

The new genome has also provided critical clues into how the insects sense a vast array of chemical cues, such as the carbon dioxide and human body odor that attracts female mosquitoes. Another important find was the location of markers that differentiated mosquitoes capable of transmitting certain viruses, like dengue. Other discoveries enabled by the new assembly could provide insight into insecticide resistance, as well as identify potential new candidate insecticides. By tracing ion channels, for example, the team identified a type of insecticide widely used in agricultural and veterinary applications that could be re-purposed to kill mosquito larvae.

The genome was assembled into contigs using PacBio's FALCON-Unzip assembler, followed by scaffolding to the three *Ae. aegypti* chromosomes end-to-end using Hi-C technology. The researchers also generated a dramatically improved gene set annotation (AaegL5.0) using the NCBI RefSeq pipeline.

An infographic about how genome references like this mosquito accelerate science is available [here](#). More information about the benefits of PacBio sequencing for plant and animal sciences is available [here](#).

### About Pacific Biosciences

Pacific Biosciences of California, Inc. (NASDAQ:PACB) offers sequencing systems to help scientists resolve genetically complex problems. Based on its novel Single Molecule, Real-Time (SMRT®) technology, Pacific Biosciences' products enable: *de novo* genome assembly to finish genomes in order to more fully identify, annotate and decipher genomic structures; full-length transcript analysis to improve annotations in reference genomes, characterize alternatively spliced isoforms in important gene families, and find novel genes; targeted sequencing to more comprehensively

characterize genetic variations; and real-time kinetic information for epigenome characterization. Pacific Biosciences' technology provides high accuracy, ultra-long reads, uniform coverage, and the ability to simultaneously detect epigenetic changes. PacBio® sequencing systems, including consumables and software, provide a simple, fast, end-to-end workflow for SMRT Sequencing. More information is available at [www.pacb.com](http://www.pacb.com).

### **Forward-Looking Statements**

All statements in this press release that are not historical are forward-looking statements, including, among other things, statements relating to future availability, uses, accuracy, quality or performance of, or benefits of using, products or technologies, the suitability or utility of methods, products or technologies for particular applications, studies or projects, the expected benefits of sequencing projects, and other future events. You should not place undue reliance on forward-looking statements because they involve known and unknown risks, uncertainties, changes in circumstances and other factors that are, in some cases, beyond Pacific Biosciences' control and could cause actual results to differ materially from the information expressed or implied by forward-looking statements made in this press release. Factors that could materially affect actual results can be found in Pacific Biosciences' most recent filings with the Securities and Exchange Commission, including Pacific Biosciences' most recent reports on Forms 8-K, 10-K and 10-Q, and include those listed under the caption "Risk Factors."

Pacific Biosciences undertakes no obligation to revise or update information in this press release to reflect events or circumstances in the future, even if new information becomes available.

### **Contacts**

#### Media:

Nicole Litchfield

415.793.6468

[nicole@bioscribe.com](mailto:nicole@bioscribe.com)

#### Investors:

Trevin Rard

650.521.8450

[ir@pacificbiosciences.com](mailto:ir@pacificbiosciences.com)



Source: Pacific Biosciences of California, Inc.