2024

Whole Genome Sequencing for the Millipede Cherokia georgiana

Elena Cruz
Georgia College & State University, elena.cruz@bobcats.gcsu.edu

Will Wittstock
Georgia College & State University, william.wittstock@bobcats.gcsu.edu

Daniel Hastings
Georgia College & State University, daniel.hastings@bobcats.gcsu.edu

Arnab Sengupta
Georgia College & State University, arnab.sengupta@gcsu.edu

Bruce A. Snyder
Georgia College & State University, bruce.snyder@gcsu.edu

Whole Genome Sequencing for the Millipede *Cherokia georgiana*

Out of thousands of known millipede species, only five sequenced genomes of species (in four of sixteen orders) are publicly available. No whole genomes and limited genetic information are available for incredibly diverse families such as Xystodesmidae. Our research goal is to sequence the whole genome of the millipede *Cherokia georgiana*. A *de novo* sequence of the complete genome of a North American species will facilitate future research in understanding gene expression under a variety of conditions. Many interesting biological processes in millipedes are poorly described, such as the production of a defensive hydrogen cyanide secretion found in the Polydesmida. While genes in this cyanogenic pathway have been identified, its regulation under stressed conditions is unclear. This may also aid in understanding this species' phylogenetic relationship to other North American species. Here, we present our research strategy for *de novo* sequencing using the next-generation sequencing platform Oxford Nanopore MinION. For this, DNA extraction kits were used to extract high-molecular weight (HMW) DNA to be used in sequencing. NGS library quality was evaluated using the Agilent TapeStation 4200 automated electrophoresis system. Long read sequencing data was collected using multiple flow cells to enhance read-depth and accuracy. Mitochondrial genome assembly was conducted using preliminary sequencing data and the mitogenomes of related species. Further genome assembly will utilize open-source assembly tools such as Shasta, Canu, and Flye. In future experiments, we will extract total RNA using the reagent TRIzol and conduct RNA-seq analysis to identify actively transcribed genes.