Annotation of the Genomic Landscape on Contig29 of the D. eugracilis Dot Chromosome

Kristina Zudock
Washington University in St. Louis
Annotation of the Genomic Landscape on Contig29 of the D. eugracilis Dot Chromosome

Kristina Zudock

Mentor: Sarah Elgin

Most Drosophila have a small chromosome known as the dot chromosome (F element). This chromosome is unusual because it appears to be almost entirely heterochromatic, yet its 1.3 Mb-long arm has a normal gene density and three times the number of repetitive elements on euchromatic chromosome arms. The ~80 genes found on the F element are expressed and function in heterochromatin at the same levels as is typical of euchromatic genes. The genomic landscape of the D. eugracilis dot chromosome was examined in this study. D. eugracilis was selected because it is more recently diverged (10-15 million years) from a common ancestor with D. melanogaster than other Drosophila previously analyzed, meaning its dot chromosome regulatory motifs likely share enough similarities with D. melanogaster’s to be recognizable. This study focuses on the annotation of contig29, an ~56 kb region of the D. eugracilis fourth chromosome. One goal of this project is to identify the most parsimonious gene model for the two features on contig29 as compared to the D. melanogaster ortholog. Each gene model was required to have appropriate stop and start codons, coding exon boundaries, and splice sites that do not lead to phase shifts across an intron while being in congruence with transcription data. Using a mixture of bioinformatics tools and genomic databases along with D. eugracilis expression data (including RNA-Seq analysis in embryos, adult male, and adult female flies), the coding spans for two genes and their isoforms were identified—MED26 (two isoforms differing in their 5’ UTRs) and bt (five unique isoforms). Additionally, genomic data for D. melanogaster, D. biarmipes, and D. eugracilis was examined for evidence of transcription start sites for the two features on contig29. Ultimately, annotation of these conserved regulatory motifs will provide insight into how gene regulation functions on the dot chromosome.