Volume 12

Washington University
Undergraduate Research Digest

Spring 2017

Comparative Annotations of Transcription Start Sites and Identification of Conserved Motifs on the Drosophila F Element

Ben French
Washington University in St. Louis

Comparative Annotations of Transcription Start Sites and Identification of Conserved Motifs on the *Drosophila* F Element

Ben French

Mentor: Sarah C. R. Elgin

The Muller F element is an unusual domain within the *Drosophila melanogaster* genome because it is packaged mostly as heterochromatin, but also contains approximately 80 protein-coding genes. Past studies have shown that Heterochromatin Protein 1 (HP1a) and the di- and tri-methylation of histone 3 at lysine 9 (both associated with heterochromatin) are depleted at the transcription start sites (TSS's) of active F element genes. This evidence suggests that the factors regulating the expression of F element genes are likely located near the TSS's. We are currently in the process of manually annotating the TSS's of *Drosophila biarmipes* F element TSS's. Manual annotation is a meticulous process that requires the interpretation of multiple lines of evidence (e.g. sequence homology, RNA-seq data, and RNA polymerase II ChIP-Seq data) to define TSS positions and search regions. The resulting high quality TSS annotations will be used in subsequent analyses. An initial investigation of a set of 11 *D. melanogaster* F element genes expressed in fly fat bodies has been conducted using the motif finding tool, MEME (http://meme-suite.org/). The three most significant motifs identified by MEME were similar to the known transcription factor binding motifs for *Trl*, nub, and *lrbp18*. Additionally, we performed comparative analyses of the known *Drosophila* core promoter motifs present in the F elements of *D. melanogaster* and *D. biarmipes*. The most common of these motifs are BRE^{d} (found in 6.6% of annotated TSS), Inr (9.8%), and DPE (4.9%). The presence of these known motifs was noted in the annotations of TSS's in *D. biarmipes* as part of the annotation protocol. With this analysis of the conservation of known motifs and the identification of possible novel motifs, we hope to uncover factors that enable F element genes to be expressed in a heterochromatic environment.