Washington University in St. Louis

Washington University Open Scholarship

Volume 13

Washington University Undergraduate Research Digest

Spring 2018

# Comparative Analysis of Transciption Start Sites and Conserved Motifs in Drosophila F and D Elements

Ben French
*Washington University in St. Louis*

Joe Zielke
*Washington University in St. Louis*

# Comparative Analysis of Transcription Start Sites and Conserved Motifs in *Drosophila* F and D Elements

*Ben French and Jon Zielke*

*Mentor: Sarah C. R. Elgin*

The *Drosophila melanogaster* Muller F element is unusual in that this chromosome is packaged mostly as heterochromatin, but contains ~80 protein-coding genes. Past studies have shown that classical markers of heterochromatin (e.g., HP1a) are depleted at the transcription start sites (TSS) of active F element genes, which suggests that the factors that regulate F element gene expression are located near the TSS. To define search regions for efforts to identify potential regulatory sites, we manually annotated the TSS positions of genes on the *Drosophila biarmipes* F element (118 TSS) and on a euchromatic region at the base of the D element (258 TSS). These TSS annotations are based on multiple lines of evidence (e.g., sequence homology, RNA-seq data, RNA polymerase II ChIP-seq data). We analyzed promoter shapes (i.e., peaked, intermediate, broad) and the distributions of transcription factor binding sites (TFBS) for these TSS. We found that a substantially greater proportion of D element promoters are classified as peaked compared to F element promoters in both species. To further characterize core promoters of F element genes, we partitioned the *D. biarmipes* F element promoters based on known TFBS observed in the *D. melanogaster* orthologs. We then analyzed each sub-population (partition) using the MEME suite to identify known and novel motifs in the promoters of *D. biarmipes* F element genes. For the dl, twi, Med, da, Udx, zfh1, hb, and Med + dl partitions, no significant motifs (E-value < 1E-05) were found. The most significant motif discovered by MEME was for the zfh1 partition, with an Expect value of 4.3E-04. A Tomtom search of this motif against the FlyFactorSurvey database did not show any significant matches to known TFBS (false discovery rate < 0.01). Hence this motif might be novel.