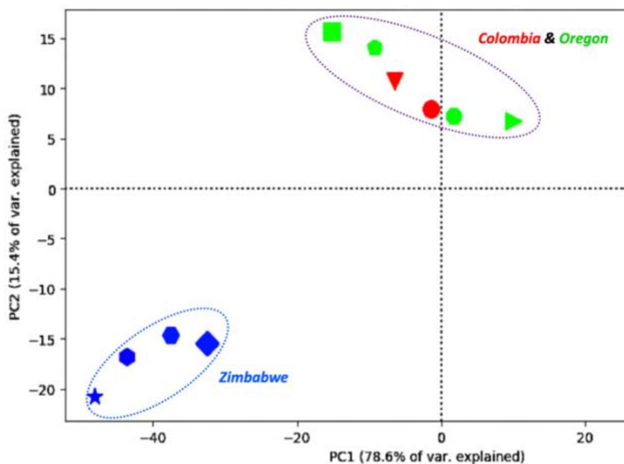
Using Reporter Transgenes to Verify Polymorphic Enhancers in Drosophila

Saivaahini Neelagiri
Michael Palopoli
Bowdoin College
Biology

Mutation is the fundamental precursor to evolution, without which species could not adapt to the selective pressures of a given environment. Genetic mutations (base substitutions, deletions, insertions) produce either structural or regulatory variation, both of which facilitate a species' evolution. Mutations within structural genes alter at least one amino acid in the polypeptide chain, resulting in a protein with a modified structure and, in some cases, function. However, for the current project we are focusing on regulatory variation. Regulatory gene mutations produce variation in gene expression–a tightly regulated process that controls what genes are transcribed and then translated into functional products.

The current research tests the hypothesis that regulatory variation is a crucial feature of phenotypic variation in nature. Specifically, we hypothesize that variation in the usage of cis-regulatory elements between and within Drosophila species is a major contributing factor to phenotypic variation between and within species.Palopoli lab has assessed the extent of enhancer function similarity within and between species of *Drosophila*.



The work the Palopoli lab has previously done supports the hypothesis. The Principal Component Graph above groups together populations based on similarity in chromatin conformation genome-wide. As shown, The Drosophila melanogaster population native to Africa had a vastly different chromatin conformation profile compared to that of its South American and North American counterparts. Moreover, the variations in chromatin conformation profiles were further magnified when comparing different Drosophila species to each other.

Because active enhancers tend to be "open" in structure while inactive enhancers tend to be folded tightly in a "closed" conformation, the lab examined and measured the chromatin conformation throughout the genome. Using ATAC-seq, the lab identified 30,000 high confidence peaks (open chromatin DNA stretches). Around 40% (1200) peaks are between genes or in introns, thus likely active enhancers. Furthermore, Visual scanning indicated that around 11% (1300) of intronic peaks are in some populations & species, but not others. In general, greatest peak (enhancer) variations were between species, but some variations between populations were present.

The project goals for summer 2023 was to (1) Verify presence of enhancer titled E1 in constructed plasmid created in Summer 2022 (2) Identify additional enhancers present in melanogaster genome (3) Amplify selected sequences and insert into pJFC28-UAS-GFP plasmid (4) Verify inserts by sequencing (5) Send constructed plasmids for insertion into Drosophila genome.

Tracks were browsed using position ranges generated by statistical analysis that identified the locations of significant polymorphic peaks. These peaks were then compared to NCBI RefSeq gene transcripts to ensure either intronic or intergenic location. If peaks were determined to be significantly polymorphic and intergenic or intronic, they were flagged as a potential cis-regulatory element. A peak's presence was also verified in D. melangoster populations from Oregon as these are the populations enhancer-plasmid constructs will be inserted into. Using Geneious, primer pairs for polymerase chain reaction amplification of selected enhancer sequences were designed. Enhancer PCR sequences were bioanalyzed and ran through a gel in order to accumulate linearized backbone for Gibson Assembly. To achieve this linearization, circular pJFRC28 plasmid was digested with HindiIII and NheI. After gel electrophoresis, bands of approximately 8 kb were cut from the gel and purified. The enhancer sequences were then integrated into pJFC28 plasmid via Gibson Assembly.

In the coming months the integrated constructs will be inserted in *drosophila melangastor* brain tissue, along with a reporter gene (GFP), and tracked for enhancer activity. Using molecular biology techniques, we will attach putative enhancers to a reporter gene, along with a promoter sequence to initiate the transcription process. The reporter gene we plan to use encodes for Green Fluorescent Protein. Since the protein is bioluminescent, we can monitor the gene activity by measuring how much fluorescent protein is produced. If the sequence is indeed an enhancer, the rate of transcription should exceed the basal rate and comparatively more fluorescent protein should be produced. If the open chromatin sequence does not impart a significant increase in protein, then the assumption will not hold.

The significance of this project goes beyond fruit flies. Understanding the extent and basis of regulatory variation will provide critical insight into processes governing gene expression, establish a flexible relationship between genotype and phenotype, and highlight the role selective pressures play in determining how that relationship evolves in a species. Each step in the flow of information from DNA to RNA to protein gives the cell an opportunity to self-determine its function by modifying the amount and type of protein it manufactures. Although structural variation can certainly contribute to adaptation, it is important to understand how regulatory variation can confer specific advantages by enhancing and inhibiting the production of certain genes.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration or of the Maine Space Grant Consortium.

Funded by the Maine Space Grant Consortium
Bowdoin College is an affiliate of the Maine Space Grant Consortium