Genomic Analysis of Reproduction and Lifespan Evolution in *Drosophila Melanogaster*

Anwita Molaka – BIOL 4559: Computational Evolutionary Biology

Department of Biology, University of Virginia, Charlottesville VA

Background Information

The relationship and trade-off between reproduction and lifespan has been a topic of great interest over the past few decades. Genomic characteristics of life-history traits, such as life span, reproductive age and rate, and age-specific health differ even between closely related populations of eukaryotic organisms. Thus, it is important to study the genetic basis of variation to better comprehend this trade-off, and analyze what factors influence this relationship.

Previous studies have shown that increased selection on late-life activity leads to increased lifespan and reproduction late in life, which results in decreased early fecundity. One particular study investigated the genomic basis of aging and life-history evolution in *D. melanogaster,* one of the best model organisms to study the function of conserved genes important for human survival. Secondary analysis was done on this work of Remolina et. al. for this research.

Questions

1. Is there observed repeatability of evolution across the sample data?
2. What genes/chromosomes contributed to the response to selection favoring late-age fertility?
3. How are key genes observed enriched for aging and fertility around the world?

Experimental Approach

The ancestral population of flies for this study was derived from ~8000 offspring of females caught in New Jersey in 1998. Since fecundity was the major factor of analysis, the ancestral population only consisted of female flies. 50 generations of selection was done on age at first reproduction, favoring an older age. The dataset consisted of 3 treatment groups (delayed reproduction) and 3 control groups, which each group containing 6 biological replicates.

To explore the first question, principal component analysis (PCA) was done on the study's samples to observe patterns among the data.Then, using a linear model, the p-values of the PCA were shown in a sliding window. For the second question, an analysis of mean FST values was conducted to isolate genes of interest regarding effects on life-span and age-specific fecundity. A sliding window of single nucleotide polymorphisms (SNPs) was created as well to view areas of potential selection. To explore the third question, principal component analysis was done once again, but plotted the study's samples and *D. melanogaster* populations across the world. A chart containing predicted z-scores of biological processes the genes of interest are generally involved in was obtained from FlyEnrichr as well.

Results

Figure 1 (below). Principal component analysis highlight overlaps and differences of the study's control and experimental groups.

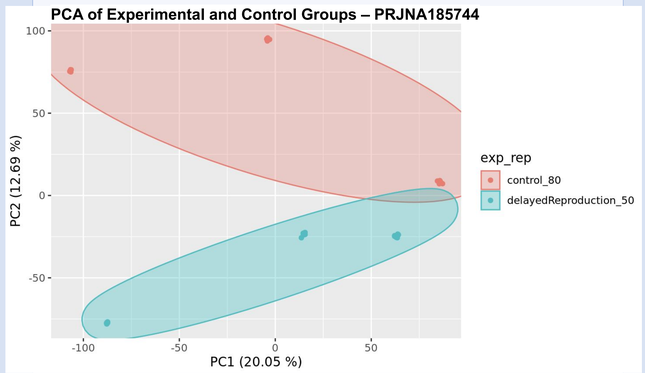


Figure 3 (right). Important genes were chosen based on if their mean FST values were over 0.15 (boundary for the samples), and they were present in the original paper's chart of genes showing strong evidence of selective sweep.

Figure 4 (below). Sliding window displaying SNPs across windows of various chromosomes, highlighting those specifically in chromosome 3L.

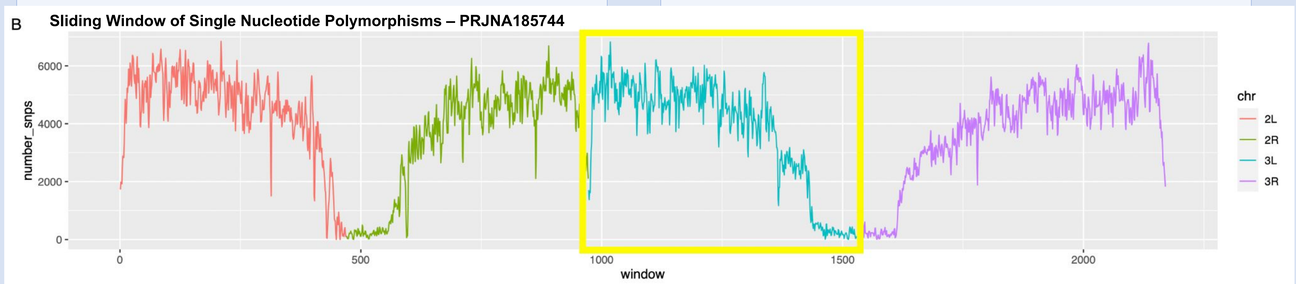


Figure 5 (below). Principal component analysis highlight the overlaps and differences of the study's samples and *D. Melanogaster* populations across the world by continent.

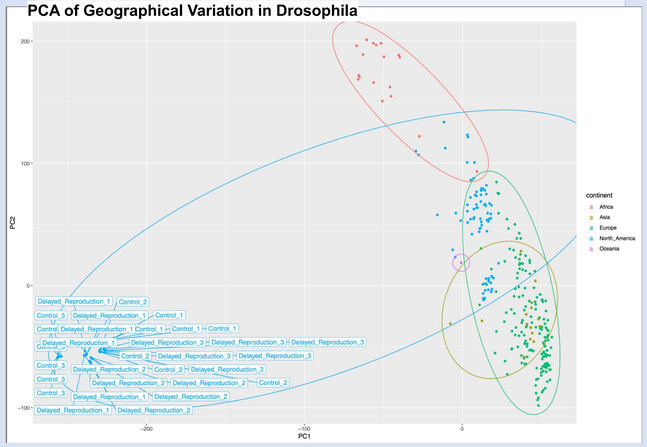
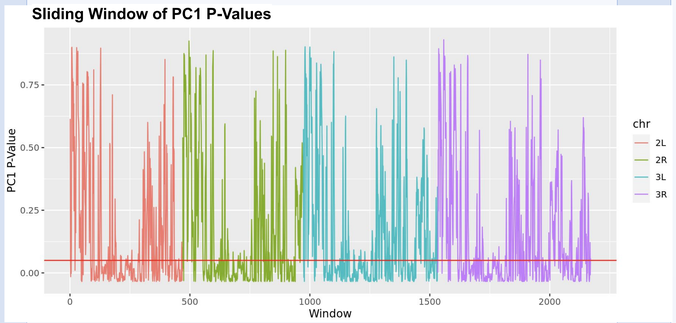


Figure 2 (below). Sliding window displays PC1 p-values across windows of different chromosomes.



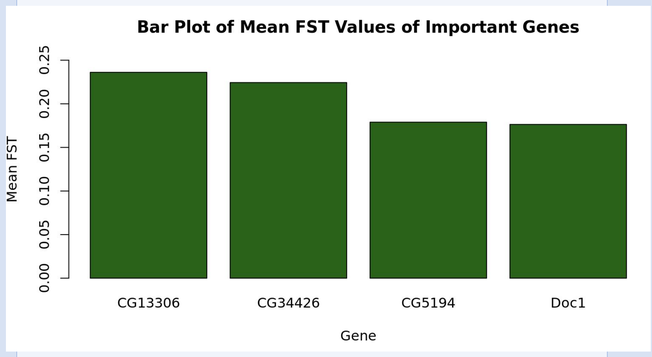
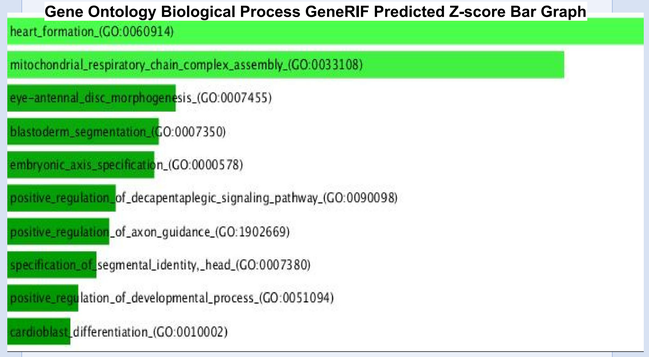


Figure 6 (below). Bar graph of the 10 most significant gene ontology biological processes that our 4 key genes from Figure 3 are a part of in *D. Melanogaster.*



Discussion

The PCA in Figure 1 shows two experimental and two control groups are close together, indicating similar evolutionary patterns among the data, including the biological replicates. The sliding window of PC1 P-values in Figure 2 showed that most of this similarity was not statistically significant. The red horizontal line represents a p-value of 0.05, and only a few windows displayed a statistically significant evolutionary pattern.

The bar plot in Figure 3 revealed 4 important genes with high mean FST values, indicating significant differentiation between samples. These are key genes involved in the response to selection on reproduction and lifespan evolution. The sliding window in Figure 4 showcases peaks in the number of SNPs across important chromosomes. These peaks, especially in chromosome 3L (highlighted), indicate increased genetic variation, which could be due to changes in important genes contributing to selection favoring late-age fertility.

The PCA in Figure 5 indicates that the study's data exhibits patterns much different from other *D. melanogaster* populations. However, the other samples across the different continents exhibit similar patterns, indicating that the evolutionary response to selection done in this study is significant enough to change the genome from what is normally observed. The bar graph of predicted z-scores in Figure 6 indicates that key genes affecting reproduction and lifespan evolution in this study are generally involved in conserved biological processes important for growth and fertility for all *Drosophila*. These include heart formation, regulation of development, and embryonic processes.

Conclusion

Based on three major questions explored, several conclusions were made:

1. There is some observed repeatability of evolution across the sample data, but most are not statistically significant.
2. CG13306, CG34426, CG5194, and Doc1 contribute to the response to selection favoring late-age fertility. These are all located on chromosome 3L.
3. Key genes identified are involved in conserved processes important for growth and fertility in *Drosophila* all over the world.

Acknowledgements

Remolina, S. C., Chan, P. L., Leips, J., Nuzhdin, S. V., & Hughes, K. A. (2012). Genomic Basis of Aging and Life-History Evolution in Drosophila melanogaster. Evolution, 66(11), 3390–3403. https://doi.org/10.1111/j.1558-5646.2012.01710.x

Thank you to Professor Alan Bergland, TA Connor Murray, and BIOL 4559 Remolina project group for guidance and support!