Establishing a Modular Pipeline for Identifying the Cis-Regulatory Elements and Transcription Factor Binding Sites in *C. elegans*

D. K. Adrian Williams\(^1,2\), Kee-Myoung Nam\(^1\), and Jeremy Gunawardena\(^1\)

\(^1\)Department of Systems Biology, Harvard Medical School, Boston, MA and \(^2\)Wright State University, Dayton, OH

The long-term goal of this project is to perform a comparative analysis of how regulatory mechanisms of neuron type specification have evolved across the animal phylogeny. As a first step, we created a modular computational pipeline that would identify cis-regulatory sequences and transcription factor binding sites that regulate the expression of *C. elegans* neuron type–specific genes.

**Introduction**

For the TFs with no PWMs found in CisBP we searched for putative homologues using the HMMER tool and found their corresponding PWMs.

We used to database Wormbase to find all of the transcription factors for each marker gene.

The inputs to our pipeline are the effector genes you want the CREs and TF binding sites for. In this example, we are inputting the TPH-1, CAT-2, and CAT-4 marker genes.

**What’s Next?**

The immediate goal in the future would be to use this pipeline to compare the binding profiles of these genes to those of negative control genes, along with examining the question of cooperativity: which pairs of transcription factors appear to co-bind more often than expected by chance?

**Acknowledgements**

A huge thank you to the Gunawardena Laboratory, Chris Nam, and the Harvard Systems Biology Department, in addition to the National Science Foundation Grant #1462629.

**References**

Figures 1 and 2 show examples of individual transcription factor binding sites for the marker genes. Investigating why these differences between transcription factors arise would be an ideal next step in this project.