

Progress Report: MotifSeeker

Justin Hii, A16682841

Tusha Karnani, A17339806

Numa Yazadi, A16837717

Name of Tool: MotifSeeker

Our tool for Motif Enrichment Analysis aims to take peak files (.bed) and reference genome files (.fa) and return information about whether certain motifs are found and enriched in the genome or not.

We are currently trying to write code for searching the reference genome for the motifs (in the HOMER database linked [here](#)).

After gaining sequences of motifs that are present, we wish to use position-weight matrices (PWMs) and analyze them to gauge whether these motifs are enriched or not using p-values. We are done writing all the functions for these analyses; we took inspiration from the Lab 5 exercises.

Here is our [Github repository](#).

Our README contains relevant information about the tool.

In the coming weeks, we wish to increase our tool's functionality and test it out with several different data sets to optimize its performance.

Once we are done fine tuning our code, we will proceed to benchmark it using HOMER's findMotifsGenome.pl functionality. We will analyze and compare both tools based on time and memory efficiency.