Scott Bogen
Project Report

McCreight's linear-time algorithm was used to build a suffix tree for various input strings, and the time elapsed to build these trees were recorded. The tests for this algorithm were done on an AMD Ryzen 5 1600 3.2GHz processor with 16GB of DDR4 RAM. The results of the tests are listed in the below table.

Name	Input size (n)	Time elapsed (ms)
Opsin-Human	3302	3359
Opsin-Mouse	3843	4051
BRCA2	11382	11150
Tomato	155461	355709

Table I - Results from 4 separate inputs

These results point to a time complexity best described as linear, with the initial three trials consuming approximately 1 millisecond(ms) of time for every byte of the input string. The time elapsed for the *Tomato* trial spent closer to 2ms, but this result is still best approximated as linear. On the test computer, the trials did not consume any meaningful amount of memory (0.0%) at runtime. These results are indeed what was expected from the trials, as the algorithm claims a linear time relationship with the input size, and such a relationship was observed.

The BWT index for the *Tomato* sequence took the form "AAAGGAA…TCGA". Compared to sample sequences and BWT indices, these results do appear correct.