

## **Project 3: Read Mapping Algorithm**

### **System Configuration:**

- 15.6 GB memory
- 2.6 GHz processor (x8)

### **Summary of Results**

Genome Length: 5362496bp  
Number of reads: 500000 reads  
Construction time: 27s  
Prep time: 1.2s  
mapRead time: 937s  
output time: 215ms  
Total time: 965s  
Alignments per read: 1.07625

### **Justification**

We were able to run the read mapping algorithm on the largest possible dataset due to our time and space complexity in the Suffix Tree construction, prep, and findPath operations. However, our alignment code was flawed in calculating the percent identity, which prevented us from getting hits for the vast majority of the reads. When we lowered the minimum percent identity to 50%, we were able to detect local alignments with sufficient percent identity. This indicates to us that the local alignment algorithm runs, but fails to retrace with the correct optimal path. The findPath algorithm returns the (approximately) correct location candidates given the name of the read.