

## **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: 26s  
Prep time: 1.3s  
Map Read time: 963s  
Output time: 2.3s  
Total time: 990s  
Alignments per read: 1.07625  
Hit Rate: 77%  
Peak Memory Usage: 620MB

### **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. The findPath algorithm returns the (approximately) correct location candidates given the name of the read. When performing alignments, we encountered slightly over 1 alignment per read, and had a "best hit" rate of about 75%. The part of execution that took the most amount of time was mapRead, which is likely due to the exceptionally high number of reads. Despite the long amount of time taken, space complexity remained relatively low throughout testing, with peak memory usage at 620 MB.