Siddhant Khanna

DNA Assignment

StringStrand

dna length = 4,639,221

cutting at enzyme gaattc

Class	splicee	recomb	time	appends
StringStrand:	256	4,800,471	6.762	1290
StringStrand:	512	4,965,591	6.521	1290
StringStrand:	1,024	5,295,831	6.817	1290
StringStrand:	2,048	5,956,311	7.550	1290
StringStrand:	4,096	7,277,271	9.514	1290
StringStrand:	8,192	9,919,191	12.945	1290
StringStrand:	16,384	15,203,031	21.311	1290
StringStrand:	32,768	25,770,711	34.931	1290
StringStrand:	65,536	46,906,071	77.161	1290
StringStrand:	131,072	89,176,791	162.796	1290
StringStrand:	262,144	173,718,231	266.165	1290

StringBuilderStrand

Class	splicee	recomb	time	appends
StringBuilderStrand:	256	4,800,471	0.044	1290
StringBuilderStrand:	512	4,965,591	0.058	1290
StringBuilderStrand:	1,024	5,295,831	0.046	1290
StringBuilderStrand:	2,048	5,956,311	0.033	1290
StringBuilderStrand:	4,096	7,277,271	0.045	1290
StringBuilderStrand:	8,192	9,919,191	0.036	1290
StringBuilderStrand:	16,384	15,203,031	0.066	1290
StringBuilderStrand:	32,768	25,770,711	0.083	1290

StringBuilderStrand: 65,536 46,906,071 0.124 1290

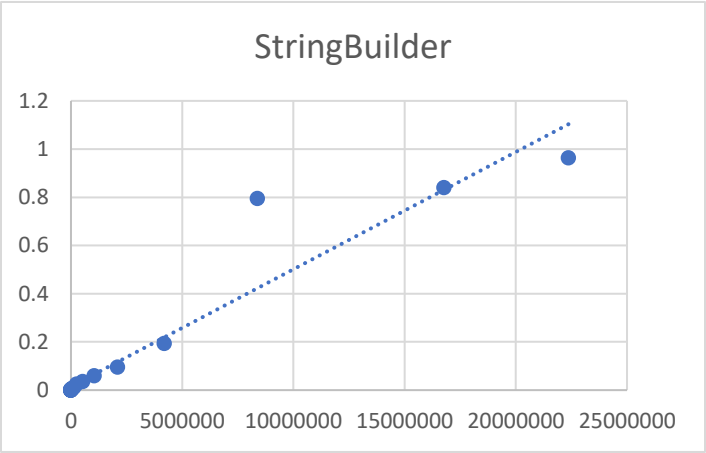
StringBuilderStrand: 131,072 89,176,791 0.242 1290

StringBuilderStrand: 262,144 173,718,231 0.365 1290

NLLH

StringBuilderStrand

BSize	Time
16	0
32	0.0011
64	0.0006
128	0.0003
256	0.0001
512	0.0014
1024	0.0026
2048	0.0015
4096	0.0012
8192	0.0025
16384	0.0023
32768	0.005
65536	0.0076
131072	0.0103
262144	0.0256
524288	0.0365
1048576	0.0597
2097152	0.0952
4194304	0.1934
8388608	0.796
16777216	0.840
22369621	0.964

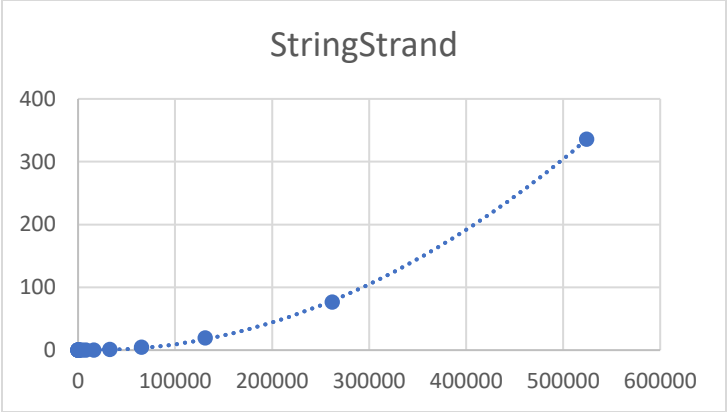


StringBuilder class has a runtime of o(bs) time. As we plot the graph we see the points form a linear regression. Thus our hypothesis was correct.

StringStrand

16	0.0001
32	0.0001
64	0.0002
128	0.0002
256	0.0011
512	0.0027
1024	0.0088
2048	0.0085

4096	0.0287
8192	0.0997
16384	0.3081
32768	0.9889
65536	4.5631
131072	19.6087
262144	76.4775
524288	335.849

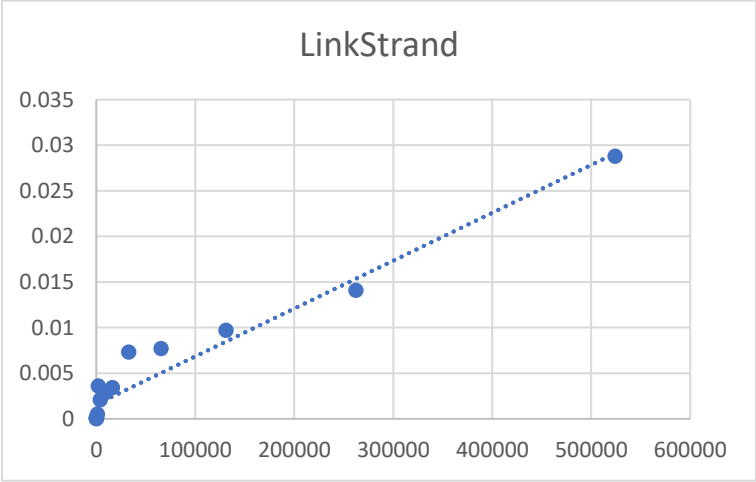


Here, StringStrand is not based on LinkedLists, thus has a big o of $o(b^2s)$. Thus our graph shows the hypothesis is correct.

LSH

LinkStrand

16	0
32	0
64	0.0001
128	0.0001
256	0.0001
512	0.0002
1024	0.0005
2048	0.0036
4096	0.0021
8192	0.0028
16384	0.0034
32768	0.0073
65536	0.0077
131072	0.0097
262144	0.0141
524288	0.0288
1048576	0.0540
2097152	0.4950
4194304	3.9775



The LinkStrand class is the Linked List method. It's big o is o(N-b). The linear curve is off by a constant equal to b, thus the program cannot catch small numbers thus values at the start are 0. Thus our hypothesis was correct.