

Rearrangement-based fragment assembly

Nadiya Sitdykova

Advisors: Max Alexeev

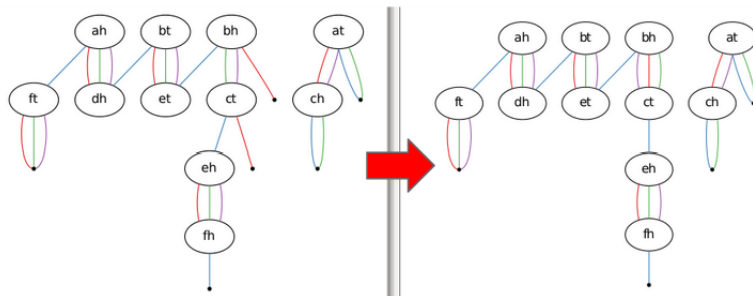
05.04.2014

Project Goal

Assemble high fragmented genomes of mosquitoes using rearrangement-based approach.

Simple idea:

For every connected component in breakpoint graph we try to find irregular edges. If we have exactly two irregular edges of the degree 1 of the same color, we glue respective “chromosomes”, creating one regular edge of the same color.



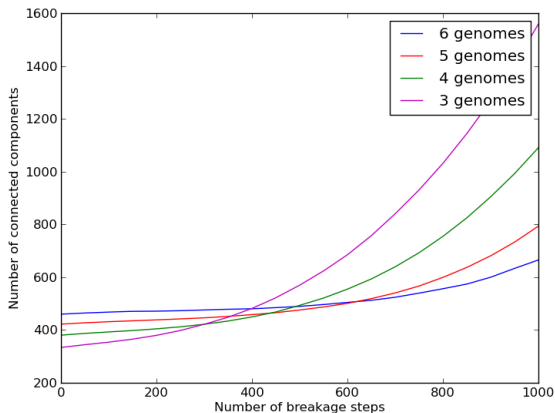
Number of gluing operation in different genomes in gambiae complex subtree:

Genome	number of gluings	change in number of fragments
A	3637	4649 → 1012
B	4	315 → 311
C	0	102 → 102
D	43	821 → 778
E	0	6 → 6

Seems not bad, but can we trust it?

cc and breakage steps correlation

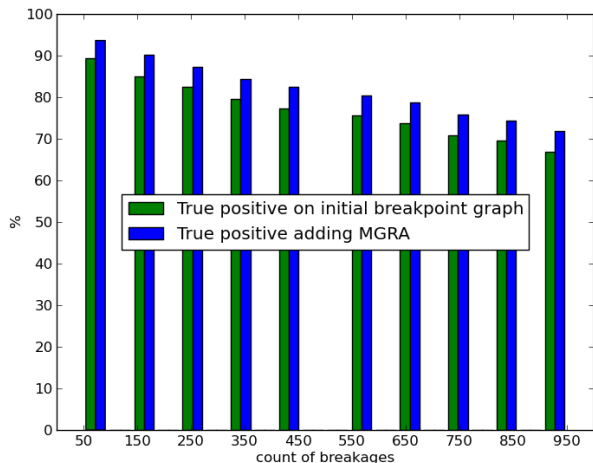
We simulated breakages in different subsets of set of 6 assembled mammal genomes and evaluated number of connected components in breakpoint graph.



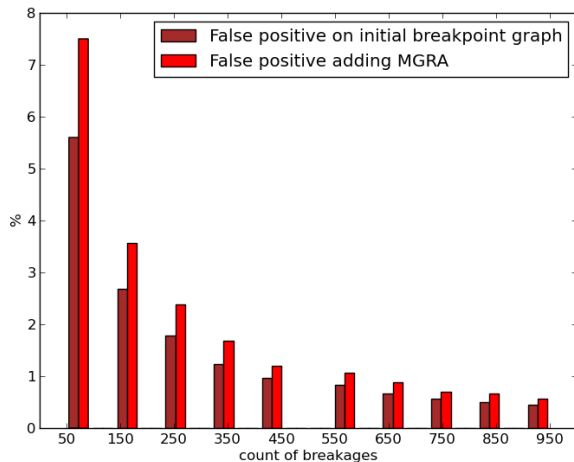
Each breakage step performs random breakage in each genome.

True positives

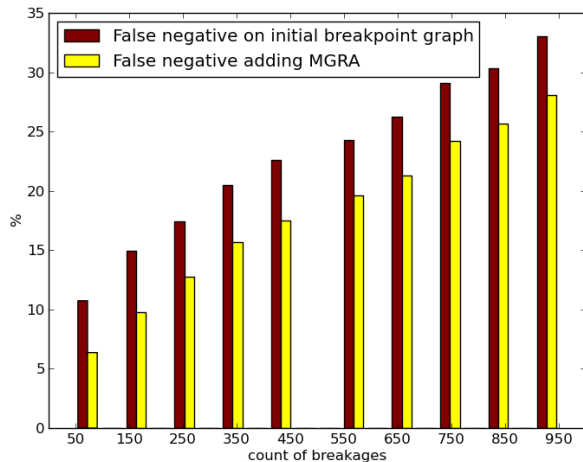
Also we applied MGRA with Step12 to mammal genomes after random breakages and evaluated it's assembly:



False positives



False negatives



Thank you for attention!